ROUTLEDGE LIBRARY EDITIONS: IRAN

Various Authors



AGRICULTURE, POVERTY AND REFORM IN IRAN

Mohammad Javad Amid

ROUTLEDGE LIBRARY EDITIONS:
IRAN



ROUTLEDGE LIBRARY EDITIONS:
IRAN

AGRICULTURE, POVERTY AND REFORM IN IRAN



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

AGRICULTURE, POVERTY AND REFORM IN IRAN

MOHAMMAD JAVAD AMID

Volume 25

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published in 1990

This edition first published in 2011
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Avenue, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 1990 Mohammad Javad Amid

Printed and bound in Great Britain

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 13: 978-0-415-57033-6 (Set)
eISBN 13: 978-0-203-83010-9 (Set)
ISBN 13: 978-0-415-61438-2 (Volume 25)
eISBN 13: 978-0-203-83022-2 (Volume 25)

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original copies may be apparent.

Disclaimer
The publisher has made every effort to trace copyright holders and would welcome correspondence from those they have been unable to trace.

Agriculture, poverty and reform in Iran

Mohammad Javad Amid



London and New York

First published 1990 by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
a division of Routledge, Chapman and Hall, Inc.
29 West 35th Street, New York, NY 10001

© 1990 Mohammad Javad Amid

Typeset by Pat and Anne Murphy, Highcliffe-on-Sea, Dorset
Printed and bound in Great Britain by
Mackays of Chatham PLC, Chatham, Kent

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

Amid, Mohammad Javad, *1949–*
Agriculture, poverty and reform in Iran.
1. Iran. Agricultural land. Land tenure. Reform
I. Title
333.3350955

ISBN 0-415-03561-9

Library of Congress Cataloging-in-Publication Data

applied for

Contents

List of figures vii

List of tables viii

Acknowledgements x

1 **Introduction** 1
The pattern of the study 2
Difficulties concerning the statistics 3
Iranian agriculture: a general view 4

2 **Agricultural backwardness in theory** 8
Traditional tenancy systems: major features and problems 9
Sharecropping and efficiency 11
Sharecropping, investment and innovation 16
Sharecropping as risk sharing 19
Additional motives for sharecropping 22
Concluding remarks 25

3 **Pre-reform Iranian village** 28
Land concentration 28
Sharecropping 33
Techniques of production and investment in agriculture 40
Agricultural production and rural income 42
Concluding remarks 47

4 **Theoretical analysis of land reform** 49
Definition of land reform 50
Political dimension of land reform 51
Economics of land reform 53
Concluding remarks 66

5 **An analysis of the Iranian land reform programme** 68
A review of the background of the land reform 69
The first stage of land reform 72
The second stage of land reform 74
The third stage of land reform 76
Analysis of the land reform legislation, and the government's agrarian attitude and policies 77
Concluding remarks 86

6 **The agrarian system after the land reform** 88
The implementation of the Iranian land reform 89
The distribution of land ownership and land tenure after the reform 102
Concluding remarks 109

7 **Economic consequences of the land reform** 111
The land reform and irrigation sources 112
The rural capital market 113
The rural labour market 119
Modernization of farming techniques after the reform 121
The new agrarian system and agricultural production 125
Concluding remarks 134

8 **Rural poverty and inequality after the reform** 135
Prevalence of subsistence farming 135
Persistence of low levels of living in rural areas 138
Conclusion 143

Glossary 146

Notes 147

Bibliography 165

Index 174

Figures

Map	Administrative division of Iran, 1966	xii
1.1	Annual precipitation in Iran	6
1.2	Sectional view of a *qanat*	7
2.1	Determination of labour intensity in sharecropped and non-sharecropped lands: the traditional approach	11
2.2	Sharecroppers' disincentive for innovation: an illustration	16
2.3	Choice of tenure system by a poor peasant	20
2.4	The impact of labour market dualism on the choice of tenure system	24
4.1	Choice of techniques of cultivation	58
8.1	Lorenz curve for rural consumption expenditure	140

Tables

1.1	Land utilization in Iran, 1977	5
3.1	Distribution of land ownership in 1,300 villages surveyed in Tehran and Demavand areas, 1949	30
3.2	Numbers and areas of Iranian holdings by type of tenure, 1960	34
3.3	Average number of plots in average holdings in five provinces, *c.* 1966	36
3.4	Kind of power used in agriculture in selected provinces, *c.* 1960	40
6.1	Progress made in obtaining villages subject to redistribution in the first stage, 1962–75	90
6.2	Villages purchased in different regions in the first stage up to September 1963	91
6.3	Completion of the first stage as of December 1975	93
6.4	Implementation of the second stage, February 1965–February 1966	94
6.5	Peasants affected by different options of the second stage up to October 1972	95
6.6	Implementation of the third stage of land reform	100
6.7	Results of the Iranian land reform, 1962–75	103
6.8	Distribution of land holdings by size, 1974	106
6.9	Growth of holdings and their areas, 1960–74	110
7.1	Relative importance of various sources of agricultural credit, 1963–72	114
7.2	Credit allowed by the Agricultural Bank, 1963–73	117
7.3	Average state loans per recipient by source of credit, 1968–74	118
7.4	State loans and grants to co-operative societies and farm corporations, 1968–71	118

7.5 Percentage share of wage labour in total labour employed by farms of different size-groups, 1974 120
7.6 Sales of tractors and combines, 1961–5 121
7.7 The application of machinery and fertilizer to agriculture, selected years 122
7.8 Distribution of machinery by selected provinces, 1972 122
7.9 Distribution of fallow land by farm size, 1974 124
7.10 Land under staple and cash crops, 1960–75 125
7.11 Production of staple and cash crops by source of information, 1960–75 127
7.12 Annual growth rates of crop production, selected periods 128
7.13 Production shares of selected provinces in the total production of principal cash crops, 1973 and 1978 129
7.14 Yield of major crops in Iran, 1961–78 130
7.15 Production indices of agriculture, 1967–77 131
7.16 Value of agricultural exports and imports, 1961–73 132

8.1 Distribution of farmers by share of crop sold and by farm size, 1974 136
8.2 Approximate size of farms required from $500 income per capita 137
8.3 Per capita rural income, 1972 138
8.4 Per capita consumption expenditures of the urban and rural households, selected years (1959 prices) 139
8.5 Housing in rural and urban areas, 1976 142

Acknowledgements

I began to write this book as my doctoral dissertation at Uppsala University, Sweden. Since its completion less than two years ago, I have received comments and suggestions and thought over certain sections. My subsequent revision of the dissertation has now resulted in this book.

I have benefited considerably from others in completing this study. I wish to acknowledge my great debts to Professor Mats Lundahl who provided me with his kind and encouraging supervision. His thorough readings of various drafts of the study and suggestions for the improvement of both form and content were extremely valuable. I am also heavily indebted to Olle Mellander, who read and commented on almost everything I wrote concerning this study and was kindly ready to advise me at any time.

I would like to express my gratitude to Professor Göran Ohlin who provided me with helpful supervision in the early stages of my research. Thanks are also due to Yngve Andersson, Christer Gunnarsson, Professor Fred Halliday, Rojelio Imable, Keith McLachlan, Christian Nilsson, Cyrus Salmanzadeh, Chris Scott and Bo Utas whose suggestions, encouragement and help at different stages of my study were of great importance. Moreover, my thanks go to the participants of the seminars at the Department of Economics, Lund University, where various chapters of this study were discussed.

I am also deeply indebted to Professor Ann K.S. Lambton who kindly read almost all of the chapters concerning Iran. Her valuable comments substantially improved this work. Furthermore, I owe a great debt to Professor Bo Gustafsson for his useful comments and encouragement and to Professor Bengt-Christer Ysander for stimulating me to speed up the completion of this research during its final stages, when it was really required.

Special thanks are also due to Alan Harkess for correcting and polishing my English and to Eva Holst, who very kindly and

skilfully typed many versions of this study. Moreover, I am very grateful to Jan Wallanders Stiftelse whose generous grants contributed much to the accomplishment of my research.

Last but not least, I have to express my deep appreciation to my wife and our daughter for their patience in putting up with my long preoccupation with working on this book.

Uppsala
M.J. Amid



Administrative division of Iran, 1966

Chapter one

Introduction

Iran, like many other developing countries, has experienced a land reform in recent decades. The Iranian land reform programme started in early 1962 accompanied by considerable rhetoric. Speeches and statements made by various government authorities gave the impression that eroding rural poverty and stimulating agricultural development were among the major objectives of the programme. The reform was also initiated without delay. The complete implementation of the land reform programme, however, lasted about a decade and went through three different stages during this period.

A relatively large number of Iranian peasants have received land as a result of the land reform and the application of modern inputs, such as fertilizers and tractors, increased substantially in farming after the reform. On the whole, a new agrarian system replaced the traditional landlord–peasant system which was believed to have hampered agricultural and rural development. Nevertheless, the agricultural growth rate did not rise markedly and there was hardly any improvement in the standard of living of the majority of the peasantry.

There is no doubt that many political, social, and economic factors contributed to bring about the prevailing situation in rural Iran. There are also many valuable studies which discuss various aspects of the Iranian land reform. However, most of the literature, especially the books, focuses on socio-political aspects of the land reform programme. The major purpose of the present study is to provide a detailed economic analysis of the reform and its economic consequences. The emphasis, however, will be on the effects of the reform upon the levels of living of the rural poor.

The pattern of the study

The study is divided into eight chapters. Chapter 1 is the introduction, which also contains some general information on Iranian agriculture. Chapters 2 and 3 are devoted to the analysis of traditional tenancy systems in general and the Iranian pre-reform agrarian system in particular. This will pave the way for understanding why agrarian changes are believed to be necessary in such systems and how land reform, which is a widely recommended means for such changes, can affect the prevailing conditions in these systems. The latter issue is discussed at length in Chapter 4 and will be followed by a study of the Iranian land reform in Chapters 5 to 8. Let us look at each chapter more closely.

Since the traditional agrarian system in Iran was a tenancy system, Chapter 2 begins the study with a general discussion of agricultural backwardness in traditional tenancy systems. We examine the major features of such systems which are often alleged to hinder agricultural progress and cause rural poverty. The emphasis of the discussion, however, will be on the theory of sharecropping because (a) sharecropping was the dominant form of land tenure in Iranian agriculture before the reform; and (b) the system was widely believed to have certain shortcomings that were detrimental to agricultural development.

In the light of the theoretical analysis of Chapter 2, Chapter 3 explores the traditional agrarian system in Iran. We want to see whether the system retarded agricultural development and contributed to the impoverishment and indebtedness of most of the rural people. The arguments for changing traditional agrarian systems are mainly based on the presence of such problems in these systems.

Chapter 4 will turn to the theory of land reform. Land reform is often suggested as a means to bring about desirable agrarian changes in many developing countries. The major aim of this chapter is to explain in detail how land reform may, or may not, succeed in improving the lot of the peasantry and stimulating agricultural progress. We argue that the ultimate results of a land redistribution programme depend highly on whether the programme alters a traditional tenancy system into a 'bimodal' or a 'unimodal' system. The former comprises principally very small and very large farmers while the latter consists of small family farms. We maintain that it is the unimodal agrarian system which can raise the levels of living of the majority of the rural people.

Chapter 5 introduces the basic features of the Iranian land reform programme. We will examine the laws and regulations of

the reform to see whether they aimed at a unimodal or a bimodal agrarian system. This will be followed by Chapter 6 which studies the actual impact of the reform programme on the agrarian system in Iran. To begin with, Chapter 6 traces the implementation of the land reform laws in order to estimate the amount of land that was redistributed and the proportion of the rural population who received land. This chapter will then try to identify the major characteristics of the post-reform Iranian agrarian system.

Having followed the formation of a new agrarian system in Iran as a result of the land reform, we will study the economic consequences of the change in the agrarian system in Chapter 7. We intend to examine post-reform Iranian agriculture to find out, for instance, whether the majority of farmers could modernize their methods of cultivation and improve their production after the land reform. The importance of these changes for the well-being of the majority of the rural population will be discussed in Chapter 8. This chapter concludes the study by examining the living conditions of the rural people in the 1970s to see whether or not they experienced any marked improvement after the land reform.

Difficulties concerning the statistics

It is appropriate here perhaps to examine some of the shortcomings concerning the data although we will refer to such problems repeatedly throughout the study. To begin with, we should point out that our empirical analyses are principally based on official statistics.[1] These statistics, which have been released by various ministries and governmental organizations, are notoriously unreliable and often inconsistent. For instance, there are sometimes considerable discrepancies between the two main sources of agricultural data, namely the Ministry of Agriculture and the Statistical Centre of Iran. As a matter of fact, the former frequently reported higher production figures for almost all types of crops (see Table 7.11, p. 127). On the whole, the agricultural data of the Statistical Centre are often believed to be more reliable.[2]

One major source of the unreliability of the official statistics might have been the officials' anxiety to give a favourable impression of the prevailing situation. There is also sporadic evidence which confirms such a belief. For example, when the consumption inquiries of the Central Bank indicated the deterioration of food consumption between 1959 and 1966, the Central Bank maintained that such a result 'did not appear to be plausible'. Therefore, the higher consumption figures of 1959 were brought down to the level of the 1966 figures (see Bank Markazi Iran 1969: 106–7 and Brun

and Dumont 1978: 16). It is then reasonable to believe that the official statistics generally exaggerate the successes and underestimate the failures.

Lack of statistics is another problem. For example, there is little information available about the amount of land which was distributed among the peasants under the land reform programme. In such cases, one has to rely on the estimates made by others or try to make one's own estimates on reasonable grounds.

We treat, therefore, the official statistics critically and attempt to give a more realistic picture by using other available information. We are also cautious about our conclusions and compare them with the findings of other Iranian rural studies, especially the field research if available. Consequently, we believe that the major conclusions of the study broadly reflect the reality.

Iranian agriculture: a general view

Like most of the developing countries, Iran was predominantly an agricultural country in the early 1960s. The agricultural sector, which consists of farming, animal husbandry, forestry and fishing, employed more than 50 per cent of the labour force and produced about one-third of GNP at that time (Statistical Centre of Iran 1966: 88; Bank Markazi Iran 1974: 32–3). On the whole, over 65 per cent of the population lived in the rural areas and must have relied mainly on agriculture for their living.

However, the significance of agriculture has been decreasing in recent decades. In 1976, only 34 per cent of labour force was engaged in agriculture and produced less than 10 per cent of GNP (Statistical Centre of Iran 1984: 65; Bank Markazi Iran, *Annual Report and Balance Sheet* 1977: 10; for further information about changes in the Iranian economy in this period see MacLachlan 1977). The rural population also decreased to 53 per cent of the total by 1976 (Statistical Centre of Iran 1984: 42).

Rural people principally live and work in about 60,000 villages.[3] These villages differ from each other in many respects. Some of them have several hundred families while some others consist of a few farmsteads with only a handful of families.[4] Some are located in more fertile areas and the villagers are prosperous; others are located in less fertile areas and the villagers have more difficulties in achieving a bare subsistence.

The prosperity of the Iranian villagers is generally determined by the availability of water for farming. Generally speaking, it is due to the shortage of water that only one-tenth of the total area of the country is cultivable (see Table 1.1).

Table 1.1 Land utilization in Iran, 1977 (thousand hectares)

Total area of the country		165,000
Total arable land		15,400
Rain-fed arable land:		
Under cultivation	5,600	
Fallow	4,200	
Irrigated arable land:		
Under cultivation	3,600	
Fallow	2,000	

Source: Statistical Centre of Iran (1984: 335)

The average rainfall for the country is roughly 300 mm per annum, i.e. it is only one-third of the average annual rainfall for the world. As Figure 1.1. shows, there are also great regional variations in average annual precipitation. For example, most of the central, southern and south-eastern regions receive between 100 and 200 mm of rainfall a year while the average precipitation in parts of the Caspian littoral is from 1,500 to 2,000 mm (for further information see Aresvik 1976: 1–9).

If we consider the fact that the rain-fed farming requires at least 250 mm of rain, Figure 1.2 illustrates that such farming is possible mainly in certain areas of Iran. As a result, rain-fed agriculture is generally practised in those provinces which lie in these regions, i.e. Gilan, Mazandaran, Azarbayjan, Kordestan, Kermanshah, and parts of Khorasan, Fars and Khuzestan (see further Aresvik 1976; Taqavi 1983). In the other areas, agriculture is largely dependent on irrigation water. As Table 1.1 indicates, the lands which are cultivated with the help of irrigation water constitute more than one-third of the total arable land.

Irrigation water is provided by underground channels (*qanats*), wells, dams, rivers, and springs. Until recently, *qanats* were by far the most important source of irrigation in Iran. By 1960, more than 75 per cent of all water used for irrigation came from these underground channels (Wulff 1966: 251).

The *qanat* system shows how skilfully Iranian peasants have tried to cope with water shortages. *Qanats* were dug to collect underground water and bring it to the surface. The construction of a *qanat* requires a great deal of experience and skill. The *qanat* builder has to determine first where a *qanat* system could be dug considering the natural conditions and the estimated flow of water. A head well is sunk into the water-bearing layer, above the land to be irrigated. A sloping conduit is then dug, the gradient of which is

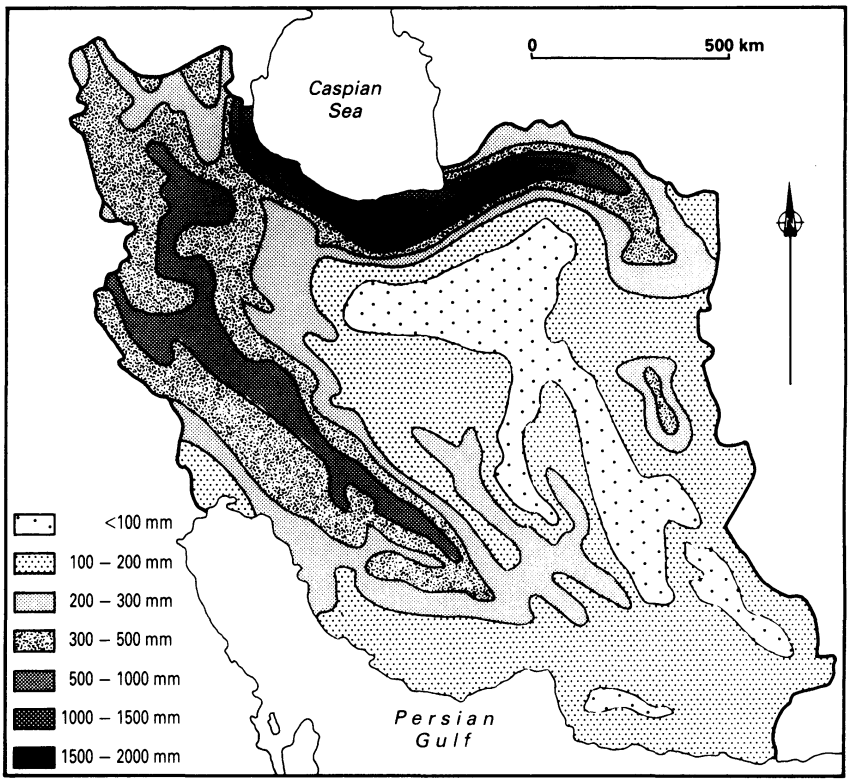


Figure 1.1 Annual precipitation in Iran
Source: Aresvik (1976: 4)

such that the water is able to flow along it under the influence of gravity and so reach the surface. Wells are dug at intervals along the length of the *qanat* to provide fresh air for the diggers and through which the soil can be brought to the surface and emptied. (For more information about *qanat* construction see Wulff 1966: 251–4.) Figure 1.2 illustrates a sectional view of a *qanat*.

The depth of a *qanat*, its length and the amount of water it provides depends on the local conditions. The depth of a head well could be between 20 to 90 m while the length of a *qanat* could be up to 70 km. A *qanat* may irrigate only a few hectares or more than one village (Lambton 1953: 217–20; Khosrovi 1979a: 97). The numbers given for *qanats* vary greatly. It is estimated that they might have numbered about 50,000, with an average length of 5 km; i.e. an underground channel of 250,000 km (Khirabi 1981:

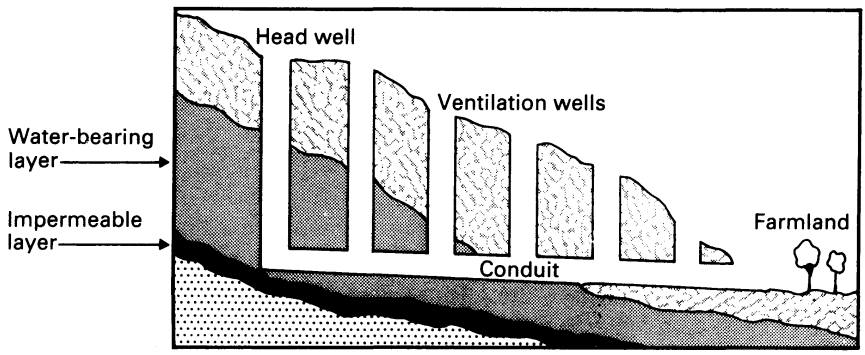


Figure 1.2 Sectional view of a *qanat*
Source: Wulff (1966: 249)

56). The first official statistics in 1954 estimated that there were about 21,000 *qanats*, 8,500 of them out of use (Taqavi 1983: 71). A large number of *qanats* lie in the central, eastern and south-eastern regions, i.e. the less rainy areas (Taqavi 1983: 69).

Springs and wells were also main sources of irrigation water, next in importance to *qanats*. Usually, a channel dug on the soil surface led such water to farms. Spring water was sometimes conducted into the *qanats* rather than being used directly (Adabi 1955: 157). There were also special *qanats* which were dug only to bring the river water to farms. These *qanats* were constructed where the configuration of the surface prevented farmers from using ordinary channels (Wulff 1966: 152).

Since there are high climatic differences from region to region, cultivating periods vary in different regions. For example, wheat is harvested during March and April in the Khuzestan plains while it is harvested from July to September in the mountain regions (Aresvik 1976: 73). During the winter, which is the most rainy period, lands are normally allocated to wheat and barely in most parts of the country. These are the principle staple crops in Iran, except in Gilan and Mazandaran in the Caspian littoral, where rice is the major staple crop. In 1960, more than 76 per cent of the cultivated land was under wheat and barley (Agah Book Collection 1982: 181). Some land is also cultivated during the summer. Summer crops consist mainly of vegetables and fruits such as cucumber and melon in many parts of the country.

Chapter two

Agricultural backwardness in theory

Development studies generally consider traditional agrarian systems as an impediment to agricultural progress and rural development in Third World countries. They argue that traditional institutions in these countries are principally responsible for low productivity, widespread unemployment and underemployment, low investments, and uneven distribution of income in rural areas. The traditional agrarian system in Iran, known as the landlord–peasant system, was often criticized on similar grounds.

The Iranian landlord–peasant system was in fact a traditional tenancy system. That is to say, most of land ownership was concentrated in a few hands but cultivated by tenants in the form of small farms. Moreover, a large majority of the Iranian tenants were sharecroppers which meant that they paid a share of the crop to the landlords for farming the land.

Studying the essence of the above arguments in the context of such systems is important for understanding agricultural and rural development strategies. Thus, the present chapter is devoted to the major problems of traditional agrarian systems in developing countries focusing on the issues which are relevant to the Iranian landlord–peasant system.

We begin the discussion by examining the main features and problems of traditional tenancy systems and we will then turn to the sharecropping debate. Sharecropping was not only a prominent feature of the Iranian agrarian system but also a tenure system which was widely attacked as being defective. There was a general consensus that sharecropping reduced peasants' labour input and deprived them of the incentive to invest in land. In view of the dominance of sharecropping in Iran and its importance in land reform discussion, we will study the theory of sharecropping in some detail.

Traditional tenancy systems: major features and problems

Traditional tenancy systems, or landlord–peasant systems, are the dominant type of agrarian systems in many countries in Asia and the Middle East.[1] Land ownership is concentrated in the hands of a few landlords, but in spite of the high degree of land concentration, land is usually cultivated in the form of smallholdings. In other words, landlords divide their land into small farms and lease them to farmers. The tenants cultivate the farms and pay rent. The rent can be a share of the crop (sharecropping) or a fixed amount per unit of land, in kind or cash.

A great number of landowners live in cities and often have other economic activities there. These absentee landowners administer their landed property through salaried agents. Agriculture as such may be of secondary importance for these landlords. They hold the land not only for income but also for social status and political power. Rent collection may be their only function in agriculture. Sometimes they may also provide consumption credits, construct and repair irrigation systems, etc. (Warriner 1969: 46).

A high degree of land concentration is a key issue in most criticisms of traditional tenancy systems. First and foremost, land concentration results in a highly skewed distribution of income in the agricultural sector. This in turn may adversely affect agricultural progress when it is combined with absenteeism. In fact, absentee landlordism is often singled out as a factor which deprives agriculture of its potential investment resources.

Large landowners normally transfer most of their agriculture income to the cities to finance their consumption and other activities there. In fact, high propensities to consume (an alleged feature of large landlords), in combination with the higher profitability of their urban financial and commercial activities, limit the investment capacity of the rent collected in agriculture.[2]

The tenants would probably not invest in land either. In the first place, a large majority of them are very poor and have nothing to invest. Second, they may have little incentive to invest in the tenanted farms since they lack security of tenancy. As a matter of fact, the tenants can be evicted at will.

It is reasonable to believe that insecurity of tenancy impedes investment in land by tenants. Even a quick-yielding investment, such as fertilizers, needs more than one crop year to bear its full fruits but tenants do not usually know whether their tenancy will last long enough to reap the returns of such investments. In fact, there is always the risk of being evicted the following year without any compensation for investment costs. A tenant is hardly ready to

invest in such insecure circumstances. Moreover, an improvement or innovation adopted by a tenant may promise a higher return and increase the value of his holding which might result in the transfer of the lease to another tenant willing to pay higher rent (Lundahl 1979: 603–5). Such a risk will naturally reinforce the tenant's hesitation to invest in his holding. There is perhaps an additional risk of eviction in the case of sharecropping: a landowner will gladly demand his share of an increased output which results from a tenant's improvement practices, but it seems doubtful that he would extend the contract with an unsuccessful innovative tenant.[3]

The above discussion indicates that part of the problem in the tenancy system stems from imperfectly defined property rights which do not guarantee that an individual will reap the benefits of his activities.

The risk of innovation is another factor which can deprive poor tenants of investment. Using new methods and seeds involves many more risks than the old ones since they are new and there is no previous experience upon which to predict the outcome. In addition, new seeds, fertilizers, and farming systems may all increase the variability of the harvest, i.e. increase the fluctuations around the average. On the other hand, the traditional methods do not promise high production but present low variance. The old technologies used over the generations have proved to guarantee their survival. Poor peasants are very likely to prefer, and get stuck with, traditional methods since these methods provide a better security for their survival.[4]

In sum, both landlords and tenants are reluctant to invest in agriculture in a traditional landlord–tenant system. This will hinder any technological change and innovation in agriculture and limits agricultural growth considerably. As a consequence, there might be further impoverishment of rural people as their number increases in the course of time.

Apart from the above criticisms that are directed toward the tenancy system as a whole, sharecropping, as a special kind of tenancy, has been widely perceived to have its own defects. The argument that sharecropping induces sharecroppers to undersupply variable inputs, discourages them to invest in land, and reduces output is perhaps as old as economics itself.[5] In fact, demands for the reform of sharecropping systems have been largely based on the premise that sharecropping is inefficient and causes underutilization of land and labour. This is, however, a controversial issue today, and in view of the importance of sharecropping in pre-reform Iran we will examine it in more detail.

Sharecropping and efficiency

The conventional wisdom, notably Marshall, argues that a sharecropper has an incentive to apply variable inputs less intensively than an owner-operator or a fixed-rent tenant (Marshall 1920: 534–7). Sharing violates the equality of marginal products of the same input in alternative uses and leads to a Pareto-inefficient allocation of resources. Let us look at this argument in detail.

Suppose that a crop with output Q is produced with land of area H and a labour force of L; i.e. $Q = F(H,L)$. Two farmers, an owner-operator and a sharecropper, cultivate holdings of the same size and fertility in a perfectly competitive and certain world. Further, assume that the sharecropper pays a fixed proportion, r, of gross output to the landlord and that r is traditionally determined.

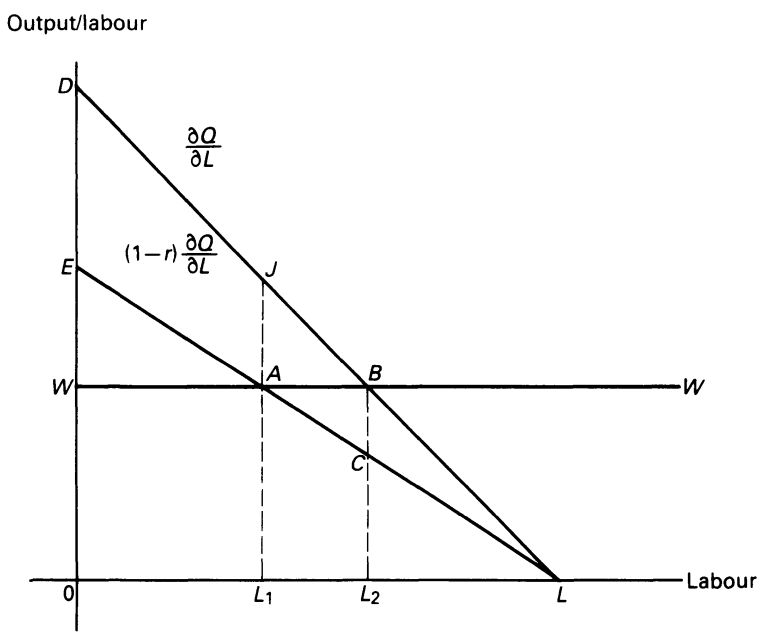


Figure 2.1 Determination of labour intensity in sharecropped and non-sharecropped lands: the traditional approach

In Figure 2.1 $\partial Q/\partial L$ represents the marginal productivity of labour in both farms. This is equal to the marginal income of the owner-operator, but the sharecropper receives only a percentage of this amount. Consequently, his marginal revenue is $(1-r)\partial Q/\partial L$. In terms of Figure 2.1, the sharecropper pays $DE/D0 = r$ of the

crop to the landlord and keeps $E0/D0 = (1-r)$ of it for himself. The marginal cost of cultivation for both farmers is equal and constant at the market wage rate, W. To maximize revenue, each farmer equates his marginal revenue and cost which means that the owner-operator, or a fixed-tenant, uses L_2 of labour and produces $0DBL_2$ while the sharecropper employs only L_1 of labour and produces $0DJL_1$.[6] The marginal product of labour for sharecroppers is greater than his marginal cost by AJ, and the situation is not Pareto-optimal. The area AJB represents an economic loss which is imposed on the society by this tenure system.

The above approach leads to a surprising conclusion: in equilibrium, the marginal product of sharecropping land is equal to zero if the sharecropper is free to decide not only his labour input but also his land input. In fact, if the tenant is allowed, he determines the amount of his labour and the land he rents so that his income, y^t, is maximized. With the above mentioned assumptions, it means that we have:

$$\underset{(L,H)}{\text{Max}}\ y^t = (1-r)\mathrm{F}(H,L) - WL \tag{1}$$

and the necessary conditions for maximization are

$$(1-r)\mathrm{F}_L = W \tag{2}$$

$$\mathrm{F}_H = 0 \tag{3}$$

Equation (3) implies that tenants tend to rent more and more land until the value of the marginal product of land becomes zero.[7] Before this point, they can combine additional land with the same amount of labour and increase the output; they receive $(1-r)$ of the extra output without any extra cost for themselves.

Gale Johnson who was the first to note this implication of the traditional analysis, cast doubt upon its validity (Johnson 1950). It is hardly believable, he argues, that landlords allow tenants to cultivate the land so extensively. He suggests that landlords can force their tenants to cultivate the land intensively by undertaking one, or a mixture, of the following practices: (1) to specify in detail what the tenant is required to do; (2) to share in the costs in the same proportion as output is shared; and (3) to grant only short-term leases and check periodically the tenants' performance (Johnson 1950: 118).[8] However reluctant to accept the implication of inefficiency argument, Johnson does not provide an alternative explanation of sharecropping.[9]

Cheung was the first who formulated a simple general equilibrium

model and arrived at the conclusion that resource allocation under sharecropping was efficient (Cheung 1968, 1969). Cheung's model is based on the crucial assumption that every landlord stipulates the tenant's labour intensity in the contract and is able to enforce such a provision. Under such circumstances, the landlord determines the sharecropper's labour input and the rental share so that he can maximize his income and the tenant receives a share equal to his alternative wage earning. The landlord's maximum income will accordingly be equal to what he could gain if he either cultivated the land himself or rented it at a fixed rate. Cheung concludes that the landlord will stipulate the minimum labour input of the sharecropper at L_2, in terms of Figure 2.1, and thus the inefficiency noted in the traditional argument is eliminated.[10]

The landlord's behaviour in Cheung's model has raised many criticisms. Bardhan and Srinivasan criticize Cheung because he allows only the landlord to maximize his income and, moreover, to determine the rental share as a monopolist.[11] Another problem with Cheung's model is the way he treats the sharecropper. The sharecropper is required to deliver a certain amount of labour for a wage-equivalent income and thus the difference between a sharecropper and a labourer remains only in name. Last but not least, the enforcibility of the minimum labour intensity provision of a contract was questioned.[12]

However, after Cheung's work, the number of theoretical contributions to sharecropping theory has increased greatly. These works generally follow one of two lines: the inefficiency line or the efficiency line.

The first school follows the traditional argument, in the form of different models, and argues that sharecropping is inefficient. An outstanding example is Bardhan and Srinivasan's work which extends the traditional partial equilibrium analysis to a fairly complete general equilibrium model.[13]

In this model, the tenant himself decides about the labour intensity of the sharecropped land and, like the landlord, maximizes his total utility. The demand for and the supply of the leased land are derived from utility maximization by the tenant and the landlord, respectively. For an exogenously given real wage rate, the supply and the demand determine both the rental share and the size of the leased land.

Bell and Zusman (1976) choose another approach which also shows the inefficiency of the sharecropping. First of all, they accept the existence of an excess demand for land in Bardhan and Srinivasan's model which enables the landlords to ration their land.[14] On the other hand, they argue that the difficulty of enforcing

a minimum labour intensity limits the power of landlords. Moreover, they argue, tenants and landless labourers are not alike. The former usually own some land, have draft animals, and possess entrepreneurial skills, all of which are hardly obtainable in the local markets of developing countries. Consequently, both the landlord and the sharecropper face market imperfections and are placed in a bargaining situation (Bell and Zusman 1976: 578–81).

Accordingly, Bell and Zusman adopt a bargaining game approach and formulate a model which determines an equilibrium rental share through a bargaining process. Examining the effects of production technologies with zero, unity, and infinite elasticities of substitution between land and labour on the rental share, they show that, given the wage rate, as the elasticity of substitution increases, the landlord has to provide greater incentives for the tenant to work intensively (Bell and Zusman 1976: 585).

In contrast to the above models, the efficiency school tries to prove that sharecropping does not interfere with efficient resource allocation. Boxley (1971) and Newbery (1974) try to show that Cheung's model is correct. Reid (1976) formulates another sharecropping model in which even the tenant, unlike Cheung's tenant, maximizes his income, and like Cheung arrives at a Pareto-optimal equilibrium. In Reid's model, each agent maximizes his total income with respect to those variables over which he has unilateral control, subject to the special constraint that his endowment is fully employed between its alternative uses (Reid 1976: 550–6).[15] The main problem with the model is the assumption that the tenant's labour intensity on sharecropped land is not decided by the tenant himself but by the market and thus the same problem as Cheung's minimum labour assumption may also be raised here.

Hsiao (1975) presents another efficiency model where the equilibrium is reached through a bargaining process with side payments. He argues that if the labour input increases from L_1 to L_2 (Fig. 2.1), the total gain exceeds the total cost for both the landlord and the tenants.[16] This implies that there is a strong incentive for both parts to negotiate. Hsiao concludes that, if monopoly power and negotiation costs are absent, the landlord will compensate the tenant by ABC and the tenant will work up to L_2; i.e. equilibrium moves from A to B which is a stable and efficient equilibrium (Hsiao 1975: 1026–7).

Finally, Lucas (1979) chooses a new approach to the problem. He points out that there is an unrealistic assumption coupled with the sharecropping discussion, i.e. the assumption that a labourer hired to work on a farm will not shirk. In practice, in order to ensure that workers will work hard enough and not shirk, they

have to be monitored or supervised. This involves an extra cost and therefore wage contracts are not as efficient as is always assumed. Lucas argues that sharecropping is indeed a way to cope with the labourers' natural tendency to shirk because it gives them an incentive to work diligently and unobserved (Lucas 1979: 504).[17] Consequently, two distortions exist in a system with both share and wage contracts: a tax-equivalent share in the former and a monitoring cost in the latter. Lucas shows that the marginal product of labour in both contracts exceeds the wage rate although they are not necessarily equal (Lucas 1979: 510–11). This means that neither system is efficient and therefore a comparison has to be made between two second-best solutions.[18] This is what Caballero shows in another model that contains a similar argument, concluding, among other things, that labour intensity under sharecropping may or may not be larger than under a wage contract (Caballero 1983: 113).

The disagreement between the two schools is reflected in the empirical studies as well. For example, Johnson provides evidence which shows that there was not much difference between rents per acre of cash-leased and sharecropped farms in Iowa in the period from 1925 to 1946 (Johnson 1950: 118).[19] He believes that such evidence substantiates the efficiency thesis. Heady studies a sample of Iowa farms for the production years of 1949 and 1950 and comes to the opposite conclusion: cropshare farms are cultivated less intensively than the cash-rented farms (Heady 1955). Rao criticizes Heady for not paying attention to the size of the land which was cultivated by each group and thus considers this factor in studying the sharecroppers and owner-operators of West Godavari, India, for 1957–9 (Rao 1971: 590–1). He concludes, in general, that small, owner-cultivated farms were cultivated more intensively than corresponding sharecropped farms while the opposite was true for larger farms (Rao 1971: 591–2).

Cheung gives some evidence from China, 1921–5, which indicates that yields per acre were approximately equal for farms with different tenure arrangements (Cheung 1969: 59). Finally, Bell (1977) suggests that the efficiency hypothesis can be tested more accurately by comparing the difference between input and output per acre on owned and sharecropped plots which are cultivated by the same farmer. Such a comparison for a sample of thirty-one owner-tenants for Purnea, India, reveals that owned plots are cultivated more intensively than the sharecropped land; yield per acre was higher on the former (Bell 1977: 329–35).

Sharecropping, investment and innovation

We have already discussed the adverse effects of the insecurity of tenancy and peasant poverty on the incentive to invest and to innovate (see pp. 9–10). As share tenancy is usually insecure and as sharecroppers are usually poor, one would not expect much improvement in an agricultural sector which is dominated by sharecropping. Beside this, share tenancy has been alleged to inhibit sharecroppers from investing because of the nature of the contract. In fact, a sharecropper who pays 50 per cent of his annual crop to his landlord will find a new method or input attractive if, other things being equal, it can increase the marginal production by more than twice the marginal costs. If output rises less than that, the net return will not even cover the sharecropper's costs and thus he avoids innovation altogether.[20]

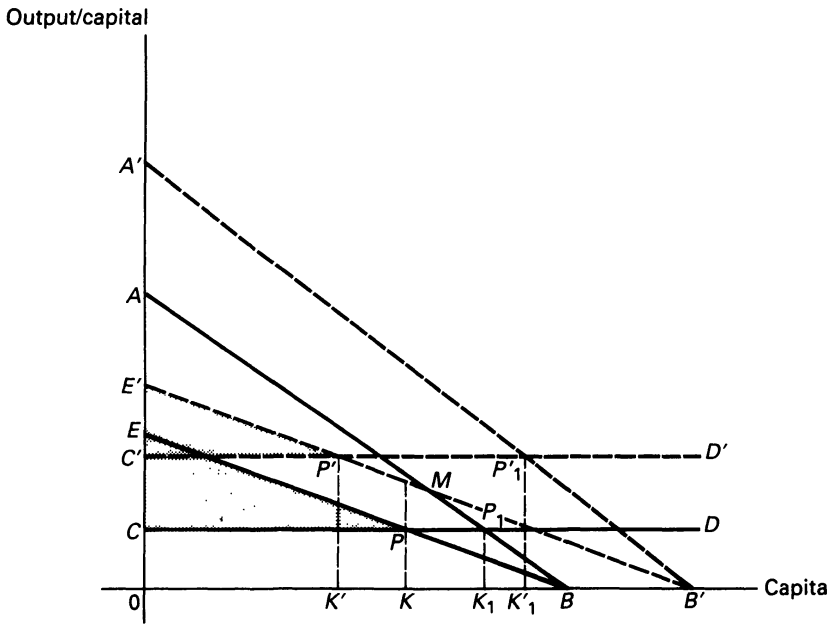


Figure 2.2 Sharecroppers' disincentive for innovation: an illustration
Source: Issawi (1957)

Issawi (1957) illustrates the above argument diagrammatically by using Figure 2.2. In this figure, the horizontal axis represents units of capital and the vertical axis shows output per unit of capital. *AB* depicts the marginal productivity of capital and *EB*, the tenant's share of it. The marginal cost of capital is supposed to be *CD*.

Now, assume that the use of a new kind of equipment raises both the marginal product of capital to $A'B'$ and the marginal cost to $C'D'$, so that $AA' > CC' > EE' = (1-r)AA'$. The tenant's share of the new returns is given by $E'B'$ which would intersect his new marginal cost at P', resulting in input of $0K' < 0K$. The tenant's net income would be $C'E'P'$ which is less than his previous net income, CEP. Since such an investment reduces the tenant's net income, he will be reluctant to invest. On the other hand, the new equipment provides sufficient incentive for an owner-operator, or a fixed-rent tenant to invest. Such an investment increases his net capital return from CAP_1 to $C'A'P_1'$. In other words, although the new investment is profitable under other tenure systems, it is unprofitable for the sharecropper who gives r per cent of the returns to the landlord and pays all the costs. He will be interested in the innovation only if his own share of the increase in unit returns exceeds the increase in unit costs, i.e. $(1-r)AA' > CC'$. In this case, P' will lie to the right of P and the net income of the sharecropper increases.

From Figure 2.2, it can be concluded that sharecropping, by precluding the tenant from innovating, may impose an economic loss on society which is equivalent to the difference between $C'A'P_1'$ and CAP_1. In other words, the net production would have increased by $C'A'P_1' - CAP_1$ if the cultivator had not been a sharecropper and had invested in the new technology (cf. Adams and Rask 1968: 939). This is a loss from the failure to innovate and is different from PMP_1 which results from undersupply of the variable input suggested by the traditional theory. (The area PMP_1 corresponds to the area AJB in Figure 2.1.)

Cheung maintains that if an investment raises the income of the land, it will be made in one of two ways (Cheung 1969: 26). Either the landlord invests in land and charges a higher rental share or he requires the tenants to invest and charges a lower rental share. Since the landowner is the decision maker in Cheung's model, it is he who decides whether or not to innovate in respect to his own interest and irrespective of the impact of innovation on the sharecropper. In fact, in this model innovations and investments leave the tenant's income intact and therefore he may be indifferent to such practices.[21] Moreover, Bardhan and Srinivasan show that when a sharecropper devotes all his labour to sharecropped land, the increase in rental share that follows cost sharing will reduce the tenant's income (Bardhan and Srinivasan 1971: 61–2).[22] Accordingly, such a sharecropper will resist adopting the innovations if they require cost sharing.

If the rental share is given exogenously, tenants will obviously

gain from cost sharing and will be encouraged to use variable inputs more intensively. This is what the traditional analysis has recognized for a long time and has suggested that if the landlord shares costs in the same proportion as output is shared, the sharecropper's disincentive to work intensively and invest in land will be eliminated; under certain circumstances, cost sharing may also become attractive to landlords even with exogenously given rental shares (see Adams and Rask 1968: 937–8; Bell 1976). For example, Bell shows that, other things being equal, if an innovation increases the output elasticity of all inputs combined, at the expense of land and draught inputs, and increases production variations, the landlord will be inclined to encourage his sharecroppers to adopt the innovation through cost sharing and therefore its adoption becomes more probable (Bell 1976: 188).[23]

A few remarks have to be made here. We saw earlier that one of the ways which could make sharecropping efficient was to stipulate the minimum labour input, and other non-land inputs, by the landlord and we also discussed the enforceability problems of such provisions (see pp. 12–13). Suppose that the landlord has learned to estimate the tenant's labour input at the harvest. Let us assume further that, other things being equal, technical progress leads to higher output than is available from existing technique while it allows the substitution of land for labour and makes the estimation of labour input very difficult. In such a case, the tenant may be tempted not to apply labour efficiently. Then, the potential extra profit may not be realized and the landlord will resist innovation (see Newbery 1975a: 273). This will especially inhibit innovations that rely on the landlord's support for their adoption. An improvement might also make sharecroppers prosperous and strengthen their power relative to the landlord. Such changes can destroy the political and economic control of the landowner over his tenants and make them more independent. This may also increase his resistance to the new technology even though on exclusively economic grounds the new technology is profitable for him (see Bhaduri 1973: 136). Last but not least, the introduction of innovations which demand cost sharing requires the landlord to play a more active role in farming than is traditionally the case. On the one hand, they have to invest in land, and, on the other hand, they have to supervise their tenants more closely. None of this is likely to be welcomed by traditional landowners.

In the absence of a satisfactory theory of sharecropping, however, it is difficult to decide whether this system is efficient or not, or whether it really inhibits investment in land. For example, if it is efficient, landlords and tenants must be indifferent to alternative

tenure systems. Then why do they prefer one to another? And, if sharecropping is not efficient, why has its use become so widespread? This is especially important as one frequently observes the coexistence of different tenure systems in one and the same region.

If we do not believe that people are irrational, the choice of contracts needs explanation. Naturally, they do not prefer one system to another exclusively on economic grounds 'but an economist persists in believing that the "relations of production" ultimately reflect some of the basic economic "forces of production" ' (Bardhan and Srinivasan 1971: 48). Attempts to explain the economic reasons for the choice of tenures and the coexistence of different tenure systems reflect such a belief. In the following sections, we examine the theories which try to explain why landlords and tenants prefer sharecropping to other tenure systems.

Sharecropping as risk sharing

A high degree of uncertainty is a distinct feature of production in agriculture. Yields vary greatly from year to year mainly because of uncontrollable variations in the state of nature. Uncertainty is thus an exceptionally important factor for decision making in farming. Such a factor will certainly play a significant role in the choice of tenure, considering the fact that the distribution of risk between contracting parties differs widely between different tenure systems. An owner-operator bears all of the risks of output variations himself while labourers will almost certainly receive their wages. On the contrary, all risks under the fixed-rent tenancy are left to the tenant, the landlord gets his rent and bears no risk. Sharecropping is between these two extremes, risks are shared by both parties; each agent bears a share of risk in proportion to his share of output (see further Griffin 1974: 22–6 and Ho 1976: 88–9).

The risk-sharing capacity of sharecropping has been broadly accepted as an attractive aspect of this tenure system. Using Figure 2.3, Ho shows that the tenant's attitude toward risk affects his choice of contract (Ho 1976). He assumes that each tenant's attitude toward risk is closely dependent upon the position of his income relative to (i) the minimum physiological subsistence level (y_{mp}), (ii) the minimum level required for a standard of living socially and culturally accepted (y_{ms}), and (iii) his desired income (y_D).

The most plausible relationship between these three levels of income is: $y_{mp} \leq y_{ms} \leq y_D$. While y_{mp} can be considered constant, y_{ms}, which is influenced by the average standard of living and other

social factors, may change over time. As the economy develops and the standard of living rises, the difference between y_{mp} and y_{ms} increases. It is reasonable to assume that each peasant wishes to reach y_D, but primarily tries to remain above y_{mp} and y_{ms}. The closer he is to y_{ms} or y_{mp}, the greater is the importance of remaining above it. (This is in fact the 'survival algorithm' which was discussed on p. 10.)

Assume further that the tenant knows the mean, μ_Q, and the standard deviation σ_Q, of the output from his past experiences. Consider also that y_f is the tenant's income under a fixed-rent contract with R as rent per unit of land, H. Alternatively, if he sharecrops the land, his income will be y_s. Figure 2.3 shows then how the tenant chooses his tenure system. This figure corresponds to a poor peasant economy where peasants' expected income is normally low and a large variation in output might lead to starvation. y_{mp} and y_{ms} are close to each other and the tenant's average income under both systems is near y_{ms}.

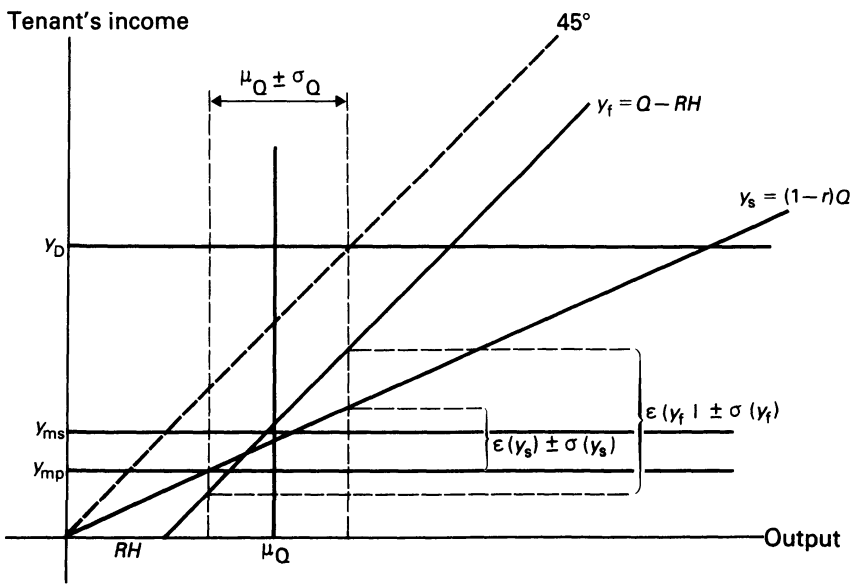


Figure 2.3 Choice of tenure system by a poor peasant

In such a case, a peasant would obviously have a strong aversion to risk and would stick to a system which promises less income variability. As can be seen in Figure 2.3, sharecropping is such a system, and thus the peasant prefers a share contract to a fixed-rent contract even though the latter provides higher expected income.

On the other hand, fixed-rent tenancy would be preferred in a rich economy, or by rich peasants.[24]

Cheung develops a theory of the choice of contract where he combines the risks and transaction costs of different systems (Cheung 1969: 62–72). He argues that any contract involves negotiation and enforcement costs. The contracting parties have to agree about the terms of contract and then check that they are respected. The provisions of a share contract, such as the rental share, the minimum labour input, the farm size, etc., have to be mutually determined and must be controlled thoroughly. In contrast, a fixed-rent tenant or an owner-operator alone decides the amount of the other party's resources he wants to employ at given market prices. This argument implies that the transaction costs of a share contract are higher than those of a fixed-rent or a wage contract. On the other hand, sharecropping offers risk sharing while the other contracts put the risk burden on either party. Cheung concludes that 'the choice of contractural arrangement is made so as to maximize the gain from risk dispersion subject to the constraint of transaction costs' (Cheung 1969: 64). This simply means that a contract is chosen by weighing its transaction costs against its risk, and that, given equal transaction costs, sharecropping is likely to be more widespread in areas of greater uncertainty.

Rao disagrees with the above conclusion (Rao 1971). He believes that greater uncertainty usually reflects a wider scope for entrepreneurship. With a significant scope for decision making, the tenant and the landlord might mutually prefer a fixed-rent contract. This ensures the tenant that he will receive the returns from his entrepreneurial skills and allows the landlord not to become involved in the risks arising from his tenant's decision making. But if the scope for entrepreneurial initiative is limited, the tenant may be anxious to share the uncontrolled risks through a sharecropping contract. The landlord will also accept such a contract provided that the production function does not allow the tenant to restrict the use of non-land inputs (cf. the conclusions of Bell and Zusman (1976: 58) which are explained on pp. 13–14). Accordingly, Rao suggests that a high level of uncertainty brings about fixed-rent tenancy, and sharecropping would be observed more frequently in areas of economic certainty, i.e. where factor substitution and the scope for entrepreneurship are limited (Rao 1971: 593).

Huang (1973) points out that Cheung's theory is not in contradiction with Rao's as long as we distinguish between natural uncertainty and entrepreneurial uncertainty. Natural risks are

beyond the control of the tenant. Sharecropping becomes more attractive where these risks are significant. Where the source of risk is peasant decision making and there is potential profit for entrepreneurship, the fixed-rent contract will be more attractive.

The importance of production risks as the principal explanatory factor in the choice of contract is questioned by Stiglitz (1974). He introduces production risks in a general equilibrium model and shows that the risk sharing opportunities that are provided by a sharecropping system are equivalent to the potential benefits from a combination of pure wage and fixed-rent contracts, provided that transaction costs are zero for all contracts (Stiglitz 1974: 236; see also Newbery 1975b, 1977; Reid 1976; Newbery and Stiglitz 1979). Then, other things being equal, production uncertainty is not sufficient to explain the choice of contractual arrangements on its own.

Additional motives for sharecropping

Reid believes that the production flexibility of sharecropping in the face of unexpected events might also provide a partial explanation of the existence of sharecropping (Reid 1976: 570–4). He argues that a crop year has three stages: planting, cultivating, and harvesting. The loss incurred by a setback at one stage can often be reduced by an appropriate response during the remaining stages. An appropriate response can also increase the gain from an unexpectedly favourable turn of events. For example, unexpectedly favourable weather at the beginning of the growing season increases the anticipated yield of a crop. Then, concentrating efforts more on such a crop than on other relatively less-favoured crops may increase the final gain (Reid 1976: 571).

To change the initial plans and undertake the necessary steps usually involves the rearrangement of the contract. Renegotiation costs and potential disagreements over the distribution of the unexpected loss or gain may prevent the contracting parties from proper actions. Unlike a wage or a fixed-rent contract, a share contract has already settled such issues, and both the tenant and the landlord gain from the timely responses. If this is correct, the average productivities of production factors under sharecropping will be higher than under other tenure forms.

> Consequently, sharecropping's potential for income enhancement and risk reduction – derived from the encouragement a share contract gives the owners of cooperating factors to similarly respond to surprises – rather than its redundant

> capability for risk dispersion emerge as a major impetus to sharecropping.
>
> (Reid 1976: 571–2)

Economies of scale and transaction costs are other factors which may make sharecropping preferable to wage and fixed-rent contracts. It should be borne in mind that in the previous section it was concluded that the risk-sharing opportunities provided by a share contract were identical to those that could be provided by a suitable combination of wage and fixed-rent contracts. Now, if there are economies of scale, it appears to be more profitable to have a piece of land under share tenancy than to divide the land into two parcels under two different contracts that together provide equivalent risk dispersion (Newbery and Stiglitz 1979: 330).

Futhermore, the transaction costs of a single contract for sharecropping is probably lower than the total transaction costs of the combination of wage and fixed-rent contracts which would give similar risk sharing opportunities. Accordingly, Reid contends that share contracts would be used because of lower transaction costs, and not in spite of higher contract costs as assumed by Cheung (Reid 1976: 573).[25]

It has also been argued that a sharecropper has a greater incentive than a wage labourer to follow the landlord's instruction whenever he provides managerial expertise (Ho 1976: 92; Reid 1976: 574). It can even be suggested that the farmers who do not want any advice will rent, those wanting some advice will sharecrop, and those wanting a lot of advice will work for wages (Reid 1977: 405). Accordingly, tenurial arrangements may operate as screening devices and reflect individuals' abilities (Newbery and Stiglitz 1979: 323–8).

Newbery (1975a, 1977) and Jaynes (1982) try to explain the incidence of sharecropping on the basis of the existence of some factor market imperfection. They, respectively, introduce labour market imperfection and capital market imperfection in their models and find that sharecropping becomes more attractive than other tenurial arrangements.[26]

Finally, Mazumdar (1975) brings up another type of labour market imperfection, namely labour market dualism. Mazumdar's model is based on the assumption that the imputed wage rate for a family member working on a family farm is lower than the current wage rate in many peasant economies.[27] This means that the supply curve of family labour lies below that for wage labour in such economies. Considering this fact, Mazumdar shows that sharecropping is more profitable than self-cultivation for landlords. This

can be illustrated with the help of Figure 2.4 which is based on Figure 2.1. *SS* and *S'S'* depict, respectively, the supply curve of labour for the wage labourers and the labour supply curve for family workers.

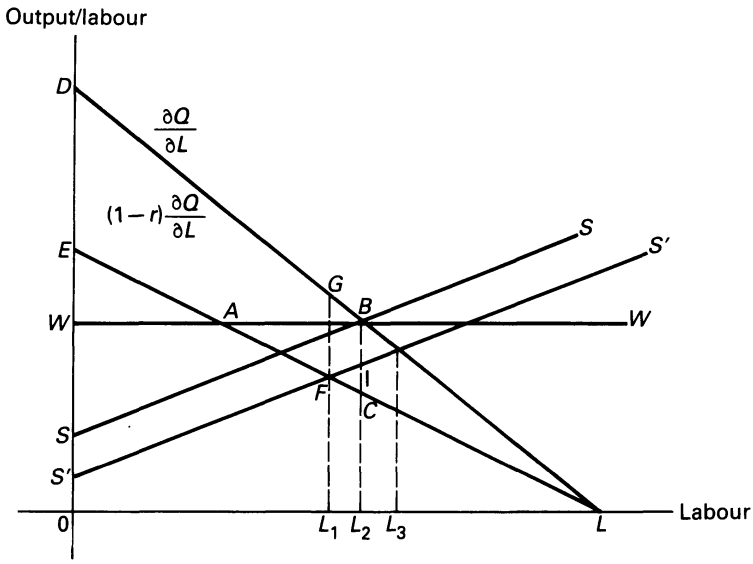


Figure 2.4 The impact of labour market dualism on the choice of tenure system

According to Cheung, the 'net utility' of a worker, i.e. the area *SWB*, is equal to his total satisfaction if he sharecropped the land.[28] But with labour market dualism the supply curve of the sharecropper, who works in the family farm, is *S'S'*. This means that, other things being equal, if the wage worker sharecrops the same farm his net utility increases by the parallelogram *S'SBI*. The landlord can then stipulate a larger labour intensity, L_3, than he would otherwise do. At the same time he can raise the rental share, *r*, so that the net utility of the sharecropper becomes equal to his alternative net utility, i.e. the area *SWB*. Consequently, sharecropping turns out to be more profitable than own-cultivation by hired labour for Cheungian landlords and all tenure systems are not equivalent as Cheung suggests.

After rejecting Cheung's theory, Mazumdar claims that if labour market dualism is introduced into the traditional theory of sharecropping, it can explain why sharecropping may be preferred to own-cultivation with hired labour in spite of its inefficiency

(Mazdumar 1975: 268–71). Let us look at this explanation. We should remember that it is now the sharecropper who determines his labour intensity, as the traditional theory suggests.

When a landlord wants to lease out his land, it is more profitable for him to parcel his land and lease out small holdings. In fact, as the number of tenants rises, the land–labour ratio on the landowner's land decreases and thus the marginal productivity of land in equilibrium increases. The landlord's possibility to reduce farm sizes is, however, limited by the fact that tenants are only ready to sharecrop land if sharecropping provides their alternative net utility, i.e. *SWB* in Figure 2.4. Suppose that *S'EF* is equal to *SWB* and the land is rented free of charge, i.e. $r = 0$. (We suppose for the moment that $EL = \partial Q/\partial L$.) Then tenants are ready to rent any holding which makes the marginal productivity of labour coincide with *EL*. If the tenant is charged for the land, $r > 0$, the share of the tenant from such holdings falls below *EL* and the sharecropper leaves the land. Now, the landlord must offer more land to shift the marginal productivity of labour upwards so that the tenant's share of it remains at *EL*. Consequently, any increase of r has to be accompanied by larger farm sizes. This decreases the total product per unit of land while it increases the landlord's share of it. When these two opposite effects become equal, the landlord's rental income per unit of land is maximized.

Suppose that Figure 2.4 illustrates an equilibrium situation. Then the area *S'EF*, the tenant's satisfaction as a sharecropper, is equal to the area *SWB*, his net utility as a labourer. On the other hand, the area *EDGF* indicates the maximum rental income which the landlord can get under sharecropping. He can alternatively cultivate this parcel of land with hired labour and obtain an income equal to *WDB*. Now, if *EDGF* is greater than *WDB*, he will prefer sharecropping; otherwise he cultivates his land with wage labour. The landlord's choice will depend on the values and elasticities of the functions which were referred to in the analysis.

Concluding remarks

Broadly speaking, the traditional agrarian systems are associated with outmoded methods of production, low land and labour productivity, and rural poverty and inequality. This is especially true for the traditional tenancy systems which cover most of the agricultural land and population in developing countries in Asia and the Middle East. The Iranian agrarian system in the early 1960s was, for example, such a system.

Land concentration and absenteeism constitute the core of most

of the criticisms of the tenancy systems. Land concentration causes a large proportion of agricultural income to accrue to absentee landlords. Most of this income is transferred to the cities to finance their notoriously luxurious standard of living and their non-agricultural activities. Consequently, nothing or very little is left for investments in the rural areas.

On the other hand, poor tenants do not have much to invest either. Insecurity of leases, which is common in landlord–tenant systems, is likely to preclude further tenants from investing in land. There is hardly any farmer who is ready to invest in land when he is not certain to reap the fruits.

The Iranian landlord–peasant system was criticized for similar reasons. Moreover, since most of the Iranian tenants were share-croppers, it was often presumed that the traditional tenancy systems in Iran suffered from further shortcomings attributed to the share tenancy system. In fact, the conventional theory argued that sharecropping caused land and labour underutilization and deprived peasants of the incentive to apply improvement practices. This view of sharecropping has been challenged recently by a new way of thinking which contends that sharecropping is as efficient as other tenurial arrangements.

Although the ongoing debate concerning sharecropping does not solve the efficiency problem, it reveals some interesting aspects of this tenure system. There is also an interesting tendency to focus on such characteristics of sharecropping instead of the efficiency question. These characteristics will in fact contribute to an explanation of why sharecropping has been observed throughout many parts of the world, including Iran.

It has been argued, for example, that a sharecropping contract might be preferred to the other tenure systems because of the production risks, the conditions prevailing in the factor market, supervision costs, etc. For instance, poor tenants may prefer share-cropping to fixed-rent tenancy when there is uncertainty regarding the avoidance of the damaging consequences of harvest failure. A landlord, on the other hand, might be inclined to opt for share tenancy since this is likely to reduce his supervisory costs, and allow him to increase his income in the presence of labour market dualism, etc. Last but not least, the theory does not offer anything that reduces the main defects of the tenancy system such as insecurity of tenancy and concentration of land in a few hands.

In view of the various criticisms of the traditional agrarian systems such as the Iranian landlord–peasant system, it has been suggested that institutional changes are necessary to pave the way for agricultural growth and improving low levels of living in rural

areas of developing countries. Land reform has been proposed as an appropriate means for such changes in, among others, the tenancy system. But before turning to this point, we will examine the pre-reform Iranian agrarian system to see if it really suffered from the problems attached to the traditional tenancy systems.

Chapter three

Pre-reform Iranian village

Having studied the general problems of traditional tenancy systems and discussed the theory of sharecropping in the previous chapter, we now turn to the *arab-raiyati* agrarian system in Iran, which could be considered as a typical tenancy system. We will explore this system to find out whether the Iranian agricultural sector was really faced with sufficient institutional shortcomings to justify an agrarian reform. At the same time, we will recognize the type and extent of the problems that had to be tackled by an agrarian programme that aimed to improve the situation.

On the whole, the pre-reform agrarian system in Iran was characterized by three main features: (i) the concentration of land in the hands of a few absentee landlords; (ii) sharecropping as the major form of land tenure; and (iii) the prevalence of outmoded techniques of production. We will examine these features in some detail, and then focus on rural poverty and inequality, to see how the prevailing agrarian system affected the distribution of income in agriculture and the standard of living of the majority of the rural people in Iran.

Land concentration

On the eve of the land reform programme four different categories of landholdings existed in Iran: private land, state domains, endowed land and crown estates.

Private land

From the Arab Conquest in the seventh century until recently, private land occupied only a small fraction of farm land in Iran.[1] During many centuries the main type of landholding was land assignment. This was called *iqta* or *tuyul*, which was in fact a kind of provisional landownership (Nomani 1980: 201). In this kind of

landholding, land or its revenue was granted to government officials, officers and others by the ruler. These grants were given in return for the provision of troops, or instead of a salary, as a pension, or as a gift.[2] In any form, the assignee could hold the land as long as the ruler wished and it could be withdrawn and reassigned at any moment (Lambton 1969: 21). Consequently, the rise and the fall of dynasties, which was not uncommon, economic pressure, etc., caused frequent changes in the composition of the landholding class.

In the nineteenth century, the dominant form of *tuyul* was a land assignment instead of salary, and there was a general tendency for such land to become hereditary. Once the control of the central government weakened a lot of this land was usurped and converted into private property. After the constitution was passed, privatization of lands continued more rapidly. The abolition of the grants of *tuyuls*, which was one of the first laws passed by the first National Assembly in 1907 followed by further legislation in the 1930s, accelerated and legalized this trend. Accordingly, many *tuyul*-holders who controlled large areas became large landowners (Lambton 1969: 129–93).

In the same period, the sale of crown and state land, *khaliseh*, began and continued for many years.[3] These were mainly the result of conquest, inheritance from former dynasties, and confiscation. Thus their volume varied in different periods (Lambton 1969: 27). *Khaliseh* land was under the control of the government which received the rent directly from the peasants (Nomani 1980: 151–61). In the late nineteenth century this type of land was still a major form of landholding. For the most part, however, it was in a state of decay. As *khaliseh* holdings made very little contribution to the finance of the growing public expenditure, it was decided to sell them (Lambton 1953: 152, 239).[4] Until the late 1950s, most lands of this type were sold according to different laws which permitted their sale.[5] However, the peasants were largely unable to buy this land since they were either too expensive or they were taken over by powerful landlords and officials. Besides the landed families, other people such as merchants, contractors, and government servants of all kinds bought most of the lands offered for sale by the government. Thus the sale of state lands ended up with a further concentration of lands in the hands of landlords and the introduction of new elements into the landowning class. These newcomers were mainly after the social prestige of land ownership, while some were also after profit (for more details see Lambton (1953: 240–55); Homan (1955: 286–97)).

As the result of its historical development, the Iranian large

landholding class was not a homogeneous one. A large landowner could be a royal family member, a senior army officer, a civil servant, a religious leader or a merchant. He might have only one village or possess tens of villages.[6] In any case, large landowners were altogether a small fraction of the population who controlled quite a large proportion of the villages. These villages were mainly the most prosperous ones, which enjoyed a high annual rainfall or sufficient irrigation facilities (Ashraf and Banuazizi 1980: 18).[7]

Unfortunately, exact statistics about landownership do not exist and the available data are usually guesses, sometimes rather confusing. However, they generally indicate a high degree of concentration of landownership. For example, a sample survey of 1,300 villages near Tehran, the capital, in 1949 reveals that the landlords who had more than 100 hectares represented only 0.2 per cent of the total number of families but owned 34 per cent of the land (see Table 3.1). Considering the fact that landownership in this area was less skewed than other parts of the country, this can be said at best to reflect the nationwide situation (Hadary 1951: 185; Lambton 1953: 281). It is then likely that about one per cent of the rural families owned more than 55 per cent of the land.

Table 3.1 Distribution of land ownership in 1,300 villages surveyed in Tehran and Demavand areas, 1949

Area of land owned (hectares)	% of rural families		% of land
0	60		
0–1	25	95	17
1–3	10		
20–100	0.8		22
over 100	0.2		34

Source: Adibi (1955: 149)

There are other estimates which show much higher landownership concentration throughout the country. For example, the Minister of Agriculture was quoted as having said on the eve of land reform that 17,000 villages belonged to the landlords with five or more villages (*Iran Almanac* 1963: 388). It was also estimated that 400 to 500 large landlords owned 57 per cent of all Iranian villages (Keddie 1968: 83).

A large majority of these landlords were absentees. In recent decades there has been a growing tendency among the landowning families to leave the villages and reside in cities and even abroad.

Although some landlords might visit their villages now and then, there were some who never came to their villages.

The property of absentee landlords was supervised by bailiffs and, with growing absenteeism, more and more villages were managed in this way. It was also common for a landlord to lease his land to a third person. The renter could be a non-resident, e.g. a merchant or a contractor, who rented the property for a short period for the profit that he could gain from it. Thus, sometimes even the lessee was an absentee (Lambton 1953: 272).

Not only large landlords but also many medium and small landowners were absentees. In fact, many of them were new landlords who never lived in the villages. Some others were the heirs of big landlords who were gradually reduced to small landlords because of the subdivision of the land according to Islamic law of inheritance. When such landlords lived in villages, they seldom cultivated their land and they preferred to lease it to peasants (Lambton 1953: 282).[8]

Apart from the different types of landlords, there were some farmers who owned part or all the land they cultivated. However, direct peasant proprietorship was not common in Iran. These farmers usually owned tiny plots which could hardly provide a bare subsistence (Nikayin 1976: 75). On the whole, direct peasant proprietorship might have constituted about 5 per cent of the cultivated land most of which was located in less fertile and isolated areas (Lambton 1953: 277; Kazemi and Abrahamian 1978: 270; Azkia 1986: 89). In some areas, where water was more crucial, peasant proprietors owned their holdings but not the irrigation water. This was often owned separately by a third person who leased it to peasants and enjoyed high income and power (Keddie 1972: 400).

State domains

As we saw earlier, most of the state domains, which were an important type of landholding before the beginning of the twentieth century, had been sold by the early 1960s. As a consequence, the state domains, *khaliseh*, covered only about 3 per cent of all villages before the land reform (*Iran Almanac* 1963: 388).[9] Black observed that most of the public domains were located in very dry, barren lands and in arid mountains (Black 1948: 428). This makes it reasonable to believe that the most fertile lands were sold and became private property, and that mainly the poor land remained in the hands of the government. State domains were not always controlled directly by the government. They were often rented by

other persons, e.g. tribal chiefs in Khuzestan, who later let them to the farmers and collected dues as landlords (Research Group 1965b: 161).

Endowed land

Endowed lands, *vaqfs*, were another major type of landholding in Iran. These lands were endowed in perpetuity for some special purposes, either charitable or personal. The benefits of the former were used for the support of religious classes, the maintenance of mosques or shrines, the feeding of the poor, etc. Personal *vaqfs* were virtually private properties because those who endowed lands used the revenues for special personal purposes. According to Islamic law, a *vaqf* cannot be dissolved and it must have an administrator who receives 10 per cent of the benefits for his work. Thus, a person could be the administrator of his endowed land and while using it for his personal purposes, secure it from usurpation.[10] Personal *vaqfs* were also constituted to prevent the subdivision of land because of the Islamic law of inheritance and to escape paying taxes since endowed lands were sometimes exempted from taxation (Nomani 1980: 189).

Vaqf lands were found all over the country, usually mixed with other types of landholdings. They could also appear as the only type of landholding in a given area. For example, it is estimated that about 466 villages belonged to the shrine of Imsam Reza. Most of them were concentrated in Khorasan province (Lambton 1953: 234–5). There is considerable disagreement between the various sources about the amount of *vaqf* lands. For example, one official source estimated that as much as 15 per cent of all cultivated land was *vaqf* while another official source maintains that only 5 per cent of lands were *vaqfs* before the land reform (Dehbod 1963: 62; Momeni 1980: 29).

Vaqf lands were seldom cultivated by their administrators. They were usually leased out to the farmers, or to a third person. Sometimes the administrator rented the land himself. In such a case he paid a fixed amount to the foundation and kept the remainder of the benefit for himself (Lambton 1953: 234).

Crown estates

During the reign of Reza Shah, between 1925 and 1941, an important class of crown lands, known as *amlak*, were the personal property of the ruler. They were separately administered from crown lands proper, which were known as *khaliseh*. These lands

were mainly the result of confiscation and usurpation although some of them were apparently 'bought' or were received as 'gifts', or were exchanged for less valuable lands. When Reza Shah was deposed in 1941, he had acquired enormous landed properties and was the largest landowner in Iran. These lands were among the most fertile lands of Iran, e.g. most of Mazandaran province became crown lands during this period (Black 1948: 428; Lambton 1953: 256–7).

After 1941 some of these lands were returned to their former owners. In 1951, Mohammed Reza Shah ordered the redistribution of crown lands among the peasants. The distribution went on for almost ten years and part of the crown lands were sold, partially to the peasants, but mainly to the wealthy landlords and even members of the royal family (Keddie 1972: 399). However, crown villages constituted approximately 2 to 4 per cent of all villages around 1960 (see, for example, *Iran Almanac* 1963: 388; Ashraf and Banuazizi 1980: 17).

To sum up, before the land reform, more than 80 per cent of the cultivated land was concentrated in the form of large private landownership, state domains, endowed lands or crown estates. The majority of these lands were let to others for cultivation, mainly under the supervision of a third party. More than 10 per cent of the land also belonged to medium and small landlords who were either absentees or non-cultivating residents. Consequently, a large majority of the rural people were left with no land or not enough land necessary to provide their bare subsistence.

Sharecropping

In spite of the high degree of land concentration, large-scale farming was not common in Iran. Instead, like many other countries in Asia and the Middle East, most of the landowners divided their lands into parcels and leased them to the farmers.[11] The farmers paid a share of the crop, or a fixed rent to the landowner. However, sharecropping was far more common than fixed-rent tenancy. As Table 3.2 shows, about 62 per cent of the total cultivated land was sharecropped while only about 9 per cent of land was leased on a fixed-rent basis.

Although sharecropping was the dominant tenure system in the country as a whole, there were regional variations. For example, we mentioned earlier that peasant proprietors usually dominated in less fertile regions. On the other hand, fixed-rent tenancy was more common in the very fertile region along the Caspian Sea (see Khosrovi 1981: 122, 156; Keddie 1972: 400).

Table 3.2 Numbers and areas of Iranian holdings by type of tenure[a], 1960

Type of tenure	*Number of holdings (thousands)*	*% of all holdings*	*Area (thousand hectares)*	*% of total area*[b]
Sharecropping	919	49	7,021	62
Owner-operated[c]	692	37	3,353	30
Fixed-rent tenancy	266	14	982	9
Total	1,877	100	11,356	100

Source: FAO, *Report on the 1960 World Census of Agriculture* (1966)

Notes: a The original table contains 204,000 mixed tenures that occupied 1,315,000 hectares. Here, the share of each mixed-tenure category is distributed among the main tenure systems in proportion to their shares of the sub-total.
b Sometimes rounding of percentages causes slight discrepancies here and in the following tables throughout this study.
c It includes both the family holdings which were cultivated by the owners and large estates which were operated by wage labour.

Risk sharing features of sharecropping might have contributed to the prevalence of sharecropping in most parts of Iran.[12] As a matter of fact, Iranian

> agriculture is subject to interruption by the capricious climate. Drought, due to insufficient spring or winter rain, causing partial or total crop failures, and floods after sudden storms, which destroy irrigation channels and *qanats*, are of common occurrence. Earthquakes are another contributory factor, causing local and temporary dislocation. Ravages by pests and locusts not infrequently cause heavy losses. High winds and violent hailstorms are other detrimental factors.
>
> (Lambton 1969: 5)

On the average, one crop in five is likely to be poor or a complete failure; however, two successive bad years or good years are not unusual (Vreeland 1957: 177–8; Land 1965/6: 115).

The dominance of fixed-rent tenancies in the Caspian region may also indicate the importance of production risks in the choice of tenure. It is reasonable to believe that the peasants in these areas, which were generally more prosperous, were more inclined to take risks to achieve higher incomes (see Chapter 2, pp. 19–22, particularly Figure 2.3). However, in the absence of any empirical analysis of sharecropping in Iran, little can be said about the importance of different factors influencing the choice of tenure system in Iranian agriculture.

Cultivation rights and security of tenure

Sharecropping was practised in many different ways all over the country. Customs governed the contracts and there were many regional and local differences.[13] However, they had much in common and their socio-economic consequences were more or less identical.

Generally, only 60 per cent of all rural families received land from landlords for cultivation, mainly on a sharecropping basis. The rest of the rural families were shopkeepers, village artisans, labourers, etc.; these people, who did not have cultivation rights, were called *khwushnishins*. Those who received land and had cultivation rights were called *nasaqdars*. In many places, the cultivation rights of village land were held jointly by *nasaqdars* and usually none of them were entitled to any specific holding in the village. In practice, cultivation rights were hereditary, but they were seldom alienable. Traditionally one of the able sons, usually the eldest, inherited this right. In this way, the subdivision of rights, and thus holdings, which might have resulted following the Islamic law of inheritance was avoided (Khosrovi 1979a: 21).[14] The heirs normally cultivated the land under the same tenurial arrangements as before and there was apparently no relationship between the form of land tenure and individual agricultural skills in Iran. In fact, such a relationship would have appeared if tenurial systems acted as screening devices (see Chapter 2, pp. 22–5).

It is worth noting that the *nasaqdars* did not form a caste. Any villager who knew how to farm and was permanently settled in the village could become a *nasaqdar* with the landlord's consent (Saedloo 1979: 219, 223). This was only possible if the landlord brought new lands under cultivation or if cultivation rights, for some reason, were available for redistribution. Neither of these cases were very common.

Although every sharecropper received a share of the village land, the shares were not necessarily equal. Each sharecropper had a traditionally specified share of the land. The village land was usually measured in terms of plough-lands, or *jufts*.[15] A plough-land was an amount of land which could be cultivated with a yoke of oxen in a farming year.[16] Obviously, the unit of measure varied from place to place and it could be an area of approximately 4 to 6 hectares (Khosrovi 1979a: 16). A sharecropper might have one or several *jufts*, but in many cases he had only part of a *juft*.

The holdings of the sharecroppers were sometimes determined periodically by drawing lots. Consequently, after a year, or some other specified period, the village land was redistributed among the

peasants. In each land allotment, each farmer received a holding which consisted of both good and bad lands. Accordingly, village holdings comprised scattered plots that sometimes were very far from each other. Table 3.3 shows the dispersion of average holdings in some provinces.

Table 3.3 Average number of plots in average holdings in five provinces, *c.* 1966

Province	*Average size of holding (hectares)*	*Average number of plots*
Gilan	1.8	3
Khuzestan	6.5	6
Mazandaran	4.4	3.4
Sistan and Baluchestan	2.7	4.1
West Azarbayjan	8.7	4.9

Source: Research Group (1965b, 1967, 1968)

Sharecropping contracts were almost verbal and everything depended on the landlord's goodwill, but there were also some traditional checks upon the landlord's arbitrariness. For example, peasants could get a kind of occupancy right if they grew trees, or any other long-lived crop. Periodical land distribution and landlords themselves tended to discourage peasants from planting trees and gaining occupancy rights which could, to some extent, limit the vast power of the landlords. In some places, the peasants owned their houses and their gardens and if the landlord wished to evict them he had to provide compensation; in some areas they had to be compensated for the labour which they put into the land they had cultivated. On the whole, tradition, and in some places the shortage of manpower, limited the extent of peasant eviction in Iran (Lambton 1953: 302; Hadary 1951: 188; Saedloo 1979: 783).

All in all, the allotment system had some egalitarian features. Each peasant had an equal opportunity to get one of the more favourable holdings. Everybody also had access to a share of lands of different qualities. Nevertheless, the system should perhaps be primarily viewed as a means of guaranteeing a bare subsistence and of preventing them from gaining occupancy rights.

Despite all the above facts, the landlords could exercise their power in different ways if they wished to do so. They could transfer a peasant to another area, assign him an inferior holding, and, finally, permanently abrogate his cultivation right to village land (Keddie 1972: 400). Peasants well understood that their means of living were in the hands of the landlords and insecurity was part of their everyday life.[17]

Group farming

The traditional production factors in Iranian villages were land, water, seed, labour and draught animals. In addition to land and water, which a farmer received from his landlord, he was often obliged to hire and rent other factors, too. Some sharecroppers, for example, did not have oxen which were predominantly the main power source and had to rent a team of oxen from an oxen rentier or from a fellow farmer who had one. It was also common for the landlord to provide seed for cultivation. Usually, a sharecropper could not cultivate his holding by himself and had to hire extra hands, especially during the busy periods.

In many parts of the country, for example Qazvin, Kermanshah and Shiraz, sharecroppers did not cultivate their land individually. Instead, they formed a number of groups in each village and each group cultivated the shares of its members from the village land as one big holding. These groups were called *buneh*. Each *buneh* was a co-operative of four to twelve sharecroppers with two to four plough-lands which lasted one farming year. There was a special division of labour among members which mainly reflected their skills and experiences (Safinezhad 1972: 2; Saedloo 1979: 218).

One of the members of each *buneh* was in charge of the activities of the *buneh*. This person, *sarbuneh*, was chosen by the landlord or his bailiff and could keep his position as long as the landlord was satisfied. *Sarbunehs* were normally the most skilled and experienced peasants and kept their position for their lifetime. At the beginning of each agricultural year the *sarbunehs* of a village held a meeting and estimated the amount of land which could be cultivated that year, mainly with respect to the expected water capacity. They decided on the distribution of land among the groups and the water rotation for that year, both of which might be specified by lot. The holding of each group also consisted of scattered plots of varying fertility.

Sarbunehs enjoyed special socio-economic position in the village. They were in the landlords' favour and were more prosperous than the rest of the peasants. Often a *sarbuneh* not only received a larger share of the income of the *buneh* but also had additional incomes such as a yearly bonus from the landlords. As the head of the group, the *sarbuneh* had also some power in the decision making, especially in choosing the members (see further Khosrovi 1979b: 78–92; Hooglund 1982: 24–8; and Safinezhad 1966: 116–17).

Irrigation activities, seed broadcasting, and supervising and directing various activities of *buneh* composed the main part of a *sarbuneh*'s duties. They sometimes chose one or several assistants,

according to the size of the *buneh*, to help them with their work. The assistants, *pabunehs*, might have other duties such as weeding the land, as well. At the bottom of each *buneh* were the cultivators, *barzigars*. They were responsible for almost all the hard work of the *bunehs*: ploughing, cleansing water channels, reaping, threshing, etc.[18]

Group farming was in fact a reasonable response to land dispersion and, especially, widespread water shortage in most parts of Iran (Safinezhad 1972: 2; Saedloo 1979: 213–14). The fact that the *buneh* system was more common in dry areas of the country which relied principally on *qanats* for irrigation may reflect the importance of water in the development of group farming. Teamwork would facilitate the construction and maintenance of these underground water channels (Ajami 1975: 141).[19] Moreover, group farming brought about a partial land consolidation and thus made the use of land and water more efficient.

Large landowners could probably run their large landed properties more easily and efficiently through group farming and with the help of *sarbunehs* who had every reason to please their landlords (Ajami 1975: 141; Hooglund 1982: 27). Therefore, they might also have encouraged group farming as a device that was likely to cope with the disincentive problem of sharecropping at a low supervision cost.[20]

Division of the crop

Theoretically, the annual produce would be divided into five equal parts and each part accrued to anyone who provided one of the five traditional production factors: land, water, seed, labour and oxen. In such a case, any sharecropper could capture a share of up to 80 per cent of an unirrigated crop. On the other hand, if the sharecropper could provide only labour and the crop was irrigated, he would receive only 20 per cent of it. In reality, practices differed widely all over the country and the crop was rarely divided into five equal shares.[21] In practice, a sharecropper's portion ranged from extremes of 12.5 to 90 per cent of the crop (Khatibi 1972: 62).

Although the rental shares were apparently determined by custom, it seems that they might have been influenced by the regional values of production factors, the importance of risk which cultivation imposed on each factor owner, the type of crops, and even the bargaining power of each party. Generally speaking, it is likely that the major differences in sharing the crop stemmed from the provision of seed, the irrigation source if the crop was irrigated, and the type of crop (cf. Momeni 1980: 53–7).

The seed supplier, whether the landlord or the peasant, usually received a share that was higher than 20 per cent and, often, up to one third of the crop. (Here, we are mainly concerned with wheat and barley which were the main staple crops.) This might partly reflect the high cost and risk that the seed owner had to accept. The storage of seed was costly and risky. In past centuries, the villagers were under the permanent threat of nomads, who could plunder them at any time, rival villages or rival groups in the same village, tax collectors, or even the army itself. These, naturally, left their imprint on prevailing custom and this influence has persisted until modern times.[22] Moreover, the seed suppliers risked losing all their seed in the case of a crop failure.

The landlord's share as the supplier of land and water varied from place to place, mainly because of the type of farming. In dry farming, the landlord received not more than 20 per cent of the harvest if he did not provide other factors. On the whole, the crop division was more favourable to the peasants in dry farming. For irrigated farming, the landlord's share increased with the cost of providing water. For example, when a *qanat* or well was used for irrigation, a larger share of the crop accrued to the landlord than in the case of the spring- or river-irrigated crops. Moreover, a greater proportion of the harvest was given to landlords when the *qanats* were deeper and longer, and thus more costly.[23]

Irrigation, manuring, weeding, etc. were different for different crops and thus the sharing of the crop was influenced by the type of the crop. For example, the cultivation of rice needs a great deal of water and when the peasant was responsible for all the other expenses the landlord usually captured up to 50 per cent of the crop (Research Group 1965a: 120).[24] On the other hand, for most of the summer crops and vegetables, manuring and weeding were necessary and peasants usually bore these extra costs. This changed the division of such crops in favour of the peasants.[25]

In some areas, the peasants had to pay certain dues to the landlords in addition to the stipulated shares. Furthermore, the peasants had to pay a share of the crop, sometimes from their own share, to the village officials and craftsmen (Lambton 1969: 26).[26]

Where group farming was common, the crop was first divided between the landlord and the group as a unit. A further division between the members of the group occurred after certain dues, which were the responsibility of the group were deducted. The group's share was subdivided between the members, sometimes equally and sometimes unequally. In a skewed distribution, a *sarbuneh* received higher shares than the others, and even his assistants might get more than the cultivators.[27]

Techniques of production and investment in agriculture

At the beginning of the second half of the twentieth century, Iranian farmers largely used the same agricultural methods as their ancestors had been using for many centuries.[28] Animals and manpower were almost always the main power in the agricultural sector; in 1962 there were about 6,000 tractors in the whole country (*Iran Almanac* 1963: 263).[29] Table 3.4 shows that these were not evenly distributed, and there were places where the peasants relied solely on manpower.[30]

Table 3.4 Kind of power used in agriculture in selected provinces (percentages), *c.* 1960

Province	*Machinery*	*Animals*	*Machinery and animals*	*Manpower alone*
Gilan	0.5	88	2	9.5
East Azarbayjan	3.5	8.5	1	87
Sistan and Baluchestan	2	61	2	35
Mazandaran	14	66	17.5	2.5

Source: Research Group (1968, 1970a)

While animals might be used for ploughing and threshing, all other tasks were performed by hand. The soil was usually ploughed with a wooden plough drawn by an animal. In some places, cultivation was carried out by spade. Digging by spade, which was more common in orchards and vegetable cultivation, required a great deal of manpower. However, it dug the soil much deeper. After ploughing, the ground was broken up by either special animal-drawn ploughs or mallets. The farmers usually sowed seed by broadcasting. Reaping was carried out with a sickle; and the grain was carried to the threshing place by animals or simply by the farmer on his shoulder. Threshing was done by a wooden threshing-machine which was also drawn by draught animals, a job which continued days and nights with men and animals working in shifts. The threshed grain was winnowed in the wind using a special form, and finally the grain was cleaned by sifting it with a sieve.[31]

Cultivation with such simple tools demanded long hours of exhausting, hard work. In order to grow one hectare of wheat and barley a farmer had to work 780 hours, of which 190 hours was with a draught animal, to say nothing of one hectare of cotton which required 2,300 hours of manpower, and 80 hours of animal power (Adibi 1955: 133). The peasants had usually no choice but to employ outside hands, especially at harvesting time.

The discussion of traditional tenancy systems in Chapter 2 showed that these systems were subject to certain factors, such as landlord absenteeism and insecurity of tenancy, which could be detrimental to agricultural progress. It seems likely that the prevalence of outmoded techniques of production and lack of improvement practices in pre-reform Iranian agriculture were in general the outcome of such factors.

The landlords who could invest showed little interest in agricultural investment. As a whole, they are often alleged to have had high propensities to consume. Moreover, when they wanted to invest, they had more profitable alternatives in the cities (Momeni 1980: 27). They usually let their bailiffs run the village affairs or leased their villages to rentiers. The former were notorious for feathering their own nests and the latter tended to squeeze out of the land what they could and then abandon it (Lambton 1953: 271–2).

It is safe to claim that *qanats* were the only investment in traditional Iranian agriculture. However, the available evidence indicates that even this kind of investment was on the decrease as more and more landlords became absentees in recent decades. Since the cost of construction and maintenance of *qanats* was heavy, it was in fact only landowners who afforded such work.[32] As a result,

> not infrequently the *qanats* fell into disrepair when the head of the family died and heirs living in the cities or perhaps Europe, demanding even more income, could not or would not use a sufficient part of the returns to maintain the *qanats* or construct new ones if the old ones could not be repaired.
>
> (Black 1948: 432)

Poverty, periodical reallotment, and lack of security might have precluded peasants from making any investment. Even simple improvement practices such as manuring were rarely carried out. About 83 per cent of the land under cultivation of annual crops received no fertilizer of any kind (Mahdavy 1965: 140). Animal dung was used for fuel rather than for manuring. However, orchards and vegetables which were near towns and villages were manured regularly (Wulff 1966: 269–70).[33] This was mainly due to the easily obtainable human sewerage as well as the fact that such peasants enjoyed more security and were even more prosperous. The peasants did all in their power, with the traditional means at their disposal and based on the experience of many centuries, to cope with agricultural pests and diseases. When they knew no way, or wished to make their methods more effective, they appealed to superstition.[34]

Even though the agricultural methods were primitive, they were excellently suited to the prevailing natural and economic conditions of the village. Despite their general resemblance, there was a large variety of different implements which were made for varying conditions (see Wulff 1966: 260–77 and Lambton 1953: 359–78). Hardary writes, for example, that the type of draught animals in use would not be physically able to draw ploughs of heavier design than the primitive wooden instrument (Hardary 1951: 184). It would have been necessary to improve the breed of animals and feeding practices before it became practicable to introduce more advanced equipment. Furthermore, the shallow, narrow, irregular furrows made by these ploughs prevented wind erosion in the treeless plains which made up a large proportion of the country's arable land. If modern ploughs were introduced without the concurrent addition of tree belts or other methods of soil protection, the result would be decreased rather than increased production. Although the practices which Iranian farmers employed to conserve water were limited in scope and had not progressed much in many centuries, they were in line with modern theories of conservation. Such practices included terracing, the use of irrigation *qanats* to prevent evaporation and conserve water, and the protection of certain trees.[35]

Agricultural production and rural income

Insufficient rainfall, poor soil, lack of investment and improvement practices would naturally result in low production and productivity in Iranian agriculture. The first agricultural census shows that outputs per hectare for wheat and barley, i.e. the major staple crops, were 790 and 770 kg, respectively, in 1960; the yields of wheat and barley in the US were, respectively, 1,433 and 1,509 kg per hectare (FAO 1966; see also Statistical Centre of Iran 1966: 127). There are also reports which give a bleaker picture. For example, a list of yield-to-seed ratios, i.e. the weight of harvested crop to the weight of sown seeds, shows that in nineteen cases average yields for wheat from forty-nine different areas were as low as 5–10 to 1 and only three were up to 50 to 1 on irrigated land in the early 1950s, while comparable figures in the US were 75–150 to 1 (Lambton 1953: 364–5. Platt 1970: 22–3).

The available information also indicates that agricultural production grew very slowly before the land reform. A study which has estimated agricultural production in the period 1935 to 1961 concludes that Iranian agricultural production increased so slowly during that period that per capita production declined (Richards

1975: 6). Similarly, a sample study shows that between 1926 and 1961, crop production increased by about 29 per cent while yields per hectare decreased by 22 per cent (Research Group 1962: 137–54).[36]

On the whole, Iranian agricultural production was rather low especially if we consider the number of people who relied on agriculture for their livelihood. In the early 1960s, more than 65 per cent of the total population lived in villages but only about 30 per cent of the gross national income came from agriculture (see Chapter 1, pp. 4–7). A good deal of this produce was paid to the landlords. A share of the remainder went to those who rendered different services to peasants. Sometimes the peasants had to pay artisans who provided services for cultivation entirely from their own shares. Labourers were also paid mainly by the peasants. They had to pay their loans and the interest after harvesting, as well. Finally, what was left was a small proportion of an already meagre crop. Platt quotes from a minister of agriculture asserting that the peasants retained only 25 per cent of what they produced (Platt 1970: 24).

Three different income groups could be distinguished among the peasants. First, there were prosperous peasants who owned all their production factors except land and enjoyed a relatively comfortable life. This usually meant that they could live at subsistence level without borrowing, and could sometimes even invest in an ox or a luxury. They comprised about 5 per cent of all peasants. Second, there were middle-class peasants who constituted between 30 to 40 per cent of all peasants and earned their living without working for others. However, they sometimes had to borrow to meet their needs. The last group, who were in the majority, consisted of poor peasants. These peasants could hardly survive without additional income. Most of them had to work as casual labourers, mainly in towns, during slack seasons and thus had to compete with agricultural labourers (Momeni 1980: 70–5).

Many peasants kept flocks and wove rugs and carpets to supplement their income. Flocks were kept in almost all villages although their size varied greatly because of climatic differences and the availability of pasture. On the border of the central desert, peasants only kept a small number of animals while in parts of the west they depended equally on flocks and agriculture for their livelihood. The milk products often formed an important part of peasant food, and if there was something left, it could be sold. Goat's hair was used in some areas to weave cloth for tents and to make rope, and sheep's wool was used in carpet weaving. Furthermore, their dung was used as fuel or as manure for cultivation (Lambton 1953: 350). The role of flocks and poultry as savings might also have been

important. In the case of some disaster, it was possible to sell them and have some cash in hand for immediate needs.

Another source of supplementary income in some parts of the country was rug and carpet weaving. Rugs and carpets were made mainly by female members of the family. In some areas, the main part of the income of the family came from this activity.[37]

Even part of these non-farming products was transferred to landlords in the form of different dues. In many areas, the peasants had to pay a certain amount of clarified butter, or cash, per plough land or per share of water; in some other places, they had to pay a pasture fee to the landlord (Lambton 1953: 333, 350–8).[38] A sheep or a lamb, some hens, roosters, eggs, or several kilos of milk products, etc., formed other kinds of dues paid to landlords in different parts of the country for varying pretexts. A gift for permission to marry, a present at the New Year, presents on other occasions, etc., are other examples of additional dues.[39] In general, it seems that there was a direct relation between the amount of additional dues and the share of the peasants from the crops: the larger the shares they received, the heavier the dues they had to pay under other names (Lambton 1953: 308). Thus, one may conclude that the peasants were generally left with as much of their whole income as would keep them on the border of subsistence (cf. Momeni 1980: 48).

We have to remember that so far we have discussed the lot of the villagers who had cultivation rights, i.e. *nasaqdars*, and ignored that of agricultural labourers.[40] These people comprised about one-third of the rural population and were normally poorer than peasants. Their life seems to have been very hard. In fact, those labourers who managed to work every day during the busy season were lucky 'in the sense that they had fewer days of hunger with which to contend' (Hooglund 1982: 33).

Peasant poverty, indebtedness and the adverse effects of markets

Reports on the state of villagers and rural studies before the land reform generally agree that most peasants were unable to make ends meet despite all their efforts (for some examples, see Momeni 1980: 60–7). In many cases, all the peasants' crops were exhausted long before the new harvest. However, there were great differences between rural incomes from region to region. For example, a sample survey in a number of villages in different parts of the country found that the annual income of peasant families in the south-east varied between $8 and $80 in 1954. On the other hand, the average family income in the prosperous northern provinces

of Gilan, Mazandaran and Azarbayjan was $516 (Keddie 1968: 71–2).[41]

However, most of the villagers suffered from hunger and malnutrition. Indeed, the whole rural life was stricken with poverty. Except in rice-growing areas, the most common diet of rural people was bread. Some did not always have enough bread for the entire year. Locusts and clover were the main food of the poorer peasants in a few areas and the rural people in some regions were reduced to eating herbage in bad years (see Keddie 1968: 75 and Momeni 1980: 61–2). Vegetables, root crops, or some fruit might be eaten with bread; meat was used very little.[42] Houses usually consisted of one or more rooms made of mud-bricks. The rooms had no windows and houses might even lack a latrine. Peasant possessions were generally very limited, hardly enough to satisfy their very primary needs.

Many peasants were forced to borrow for their survival. If they managed to maintain a bare subsistence without borrowing, they might be forced to borrow for other expenses such as tools, seeds, and draught animals. In the absence of institutional credits, the peasants had to appeal to landlords, shopkeepers, and moneylenders for loans.[43] The coming harvest was the security for such loans. In most cases no interest was specified but it was hidden in the terms of the contracts. For example, advances and repayments could be calculated at current prices which involved real rent. This was due to the fact that advances were given in the winter when the prices were very high and were repaid at harvest time when prices were at their lowest. Loans were also raised on the condition that the coming harvest would be sold to the lender at stipulated prices. This practice was known as pre-harvest sale and will be discussed below. Irrespective of the type of loan agreement, the interest rates would be very high. Landlords normally provided advances in easier terms, but they financed only a small portion of all the loans.[44]

Reports from several provinces reveal that in 1960 normally 50 per cent of the peasants were indebted and paid at least a 50 per cent interest rate to non-governmental loan sources (see, for example, Research Group 1965a; 1970b). These sources, which supplied 90 per cent of the credit for agriculture, frequently charged interest rates up to 150 per cent (Platt 1970: 30). Borrowing with such high interest rates aggravated the economic situation of the borrowers and, considering their hand-to-mouth economy, perpetuated indebtedness. Besides, the interest rarely remained in villages which implied a further decrease in the investment capacity of rural areas.

Not only the credit market but also the product market was detrimental to the peasants. Clearly enough, the peasants had no, or very little, surplus to sell and still less income to buy anything. However, product market conditions affected them to a certain extent. Their advances and repayments were assessed at current market prices, they had to sell part of their crops to repay their loans or to purchase other necessities, and frequently planted fruit and vegetables for market. In all such cases, they had to deal with monopsonists, either the landlords themselves or, usually, others such as village shopkeepers and city dealers. For a variety of reasons, the villagers were isolated from the outside world and the village market was the only available one.[45] Lack of adequate roads was a major problem in this respect. Many villagers had no roads at all and where there was a rural road it might be closed for parts of the year. The peasant was either unable to reach the outside markets, or was only able to do so at considerable cost. Hence he was forced to accept any price which was offered by the local buyers. On the other hand, when the peasants wanted to buy something, they had to buy from the only village shop, or from peddlers who enjoyed monopoly power. Of course the effects of this monopoly were on the whole insignificant since the purchases of the peasants were normally limited to a few articles such as tea, sugar and tobacco.

Pre-harvest sale of crops was another practice which indicated peasant poverty and a weak position in the markets. The peasants were often forced to sell their crops before harvest, and sometimes very long before it. This was in fact an indirect way of charging interest for the loans and to circumvent the Islamic prohibition against interest (Khatibi 1972: 62). In such contracts, loans were given provided that the whole crop, or a certain quantity of it, would be delivered to the lender at a stipulated price. The price was determined in different ways. It could be specified on the basis of past experience, or at the low level prevailing at harvest time. Sometimes it might be the current price itself. However, prices were mainly set by lenders who even manipulated harvest-time prices in accordance with their own interest (Lodi 1965: 2).

Pre-harvest sale might also have been a way of coping with production risk.[46] Poverty and lack of security measures made peasants very sensitive to crop failure. If this problem arose, there was little that could be done about it. Risk aversion was probably one main reason for pre-harvest selling of fruits and vegetables for two reasons. First, those farmers who grew wheat in addition to fruits and vegetables had recently harvested their staple crop when they agreed to the pre-harvest sale of the latter, i.e. they were not in

any urgent need to borrow. Second, in the case of the pre-harvest sale of grain the peasant was supposed to deliver a certain amount of the crop to the buyer (lender). In the event of a poor harvest, he would have to recover this shortfall. In the case of fruits and vegetables any loss was born by the buyer himself (Lodi 1965: 3). Lodi believes that between 20 to 70 per cent of the total production of crops was sold at prices which were 20 to 40 per cent lower than those obtainable in post-harvest sales (Lodi 1965: 1).

In sum, all such factors not only contributed to the persistence of rural poverty in Iran but also to the exacerbation of the situation. There is evidence which indicates that real peasant income in most areas declined between the late 1920s and the early 1960s (Keddie 1968: 70; see also Lambton 1953: 143–5; Keddie 1972: 365–6).

Concluding remarks

The high degree of land concentration was one of the main features of the Iranian landlord–peasant system. The landowners usually leased out their lands in smallholdings, mainly under sharecropping contracts. The landlords and beneficiaries of non-private land appropriated part of the agricultural income and transferred most of it to the cities since most of them were absentees and lived in urban areas. They did not invest in agriculture, with the exception of the construction and maintenance of irrigation systems in certain parts of the country. Even this kind of investment was on the decrease in recent decades.

Generally speaking, the peasants' share of agricultural income was only sufficient to keep them on the edge of subsistence. They had little to invest. In fact, many of them had to borrow to survive. The lenders, who were usually local moneylenders or shopkeepers, charged high interest rates. They also transferred their profits to the cities.

Poverty, however, was not the only obstacle to peasant investment; insecurity of tenancy and periodical land reallotment deprived the sharecroppers of incentives to invest in land. On the other hand, lack of data does not allow any comparison between investment in land under different tenurial arrangements. It is therefore difficult to say whether sharecropping itself was detrimental to agricultural investment or not.

Poverty and indebtedness seems to have been the lot of the majority of the Iranian villagers. Half of all peasant families, or one-third of the rural population, could hardly survive without additional income, for example, from casual work in the cities during the slack seasons. The life of families of agricultural

labourers who comprised another third of the rural people was even worse. They had no access to land and relied on the limited work opportunities in agriculture and casual jobs in the cities. The standard of living of the majority of the remaining one-third was perhaps only slightly better.

We can say that the problems of the Iranian landlord–peasant system were in general similar to those of the tenancy systems that were discussed in the previous chapter. As we saw there, most of such problems are generally assumed to result from the shortcomings of the traditional agrarian systems themselves. Accordingly, many observers argue that changing such systems is a prerequisite for an improvement in rural life and for agricultural progress. It is to this issue, i.e. agrarian change, that we turn in the next chapter.

Chapter four

Theoretical analysis of land reform

The argument that agricultural backwardness and rural poverty are principally a consequence of the institutional characteristics of traditional agrarian systems leads many observers to conclude that changing such systems is a necessary step for both agricultural and rural development. In view of the importance of agriculture in developing countries and its role in economic development, some observers have gone on to argue that agrarian change is in fact a pre-condition for economic development in most Third World countries.

Land reform (agrarian reform) is widely recognized as a promising means of producing desirable agrarian change. It is enthusiastically supported by masses of peasants who believe that land reform will improve their living standards. This belief is shared by many economists who argue that land reform can reduce inequality and poverty in rural areas and induce agricultural growth. The aim of the present chapter is to examine this argument with respect to traditional tenancy systems. We will try to lay down an analytical framework within which we can explain the impact of land reform on agricultural development in general and rural poverty and inequality in particular.

Despite the widespread support for land reform, or perhaps because of it, there is no general agreement about the meaning and content of reform. Thus, we will begin this chapter with a discussion of the definition of land reform. This is followed by a brief examination of the political aspects of the reforms. Finally, we turn to the economic analysis of land reform, which comprises most of the chapter. We attempt to see how land reform might change traditional tenancy systems and will analyse the economic consequences of such a change.

Definition of land reform

Land reform is generally conceived to be a programme that is designed to cope with inequality and poverty in the rural areas. This has primarily been understood to be equivalent to land redistribution. However, it has been argued that there are other programmes which can improve the living conditions of peasants and which accordingly should be included in the definition of land reform. Tenancy reforms, which aim at making the terms of tenancy better for the peasants, are the most prominent examples among the alternative programmes.

There has also been a tendency to widen the concept of land reform mainly on developmental grounds. This has originated from a belief that traditional agrarian systems hinder economic development. Accordingly land reform is assumed to be a means of removing the defects of such systems. For example, almost any programme that leads to a change which improves the manner in which land is held or used is sometimes described as land reform (Barlowe 1953: 174). In a similar vein, the United Nations' various reports on progress in land reform provide a still broader definition that covers even those measures that are usually assumed to be complementary to land reform. That is to say, land reform 'also includes the establishment or strengthening of essential governmental cooperative or commercial agencies or services related to agricultural credit, supply, marketing, extension, and research' (UN 1963: iv).

Accordingly, the term 'land reform' has taken on many different meanings. It may mean a simple rent reduction policy, a radical land redistribution, or a comprehensive rural development plan. This lack of a generally accepted definition of land reform has made almost every author who deals with the subject provide his or her conception of land reform (for some examples see Bergmann 1977; Flores 1970; Lipton 1974; Raup 1968). In general, the equity-oriented definitions emphasize changes in land distribution in favour of the peasants while the development-oriented definitions turn out to be the wider ones.

If rural poverty and inequality are supposed to be the main concern of a land reform programme, these wider definitions may cause confusion. They usually include many different programmes which may not have any meaningful effect on rural life.[1] When land is the main source of income with no or little alternative available, restructuring the distribution of land is likely to be the most effective way to improve the lot of the peasantry in many countries.[2] This is not, of course, a sufficient step. Peasants need

credit, extension services, etc., and they have to receive services to raise their level of living. These measures, however, cannot basically change income distribution in agriculture by themselves. Therefore, they must be applied to support land reform, not to replace it (Flores 1970: 897).

In the present study, land reform is defined as the takeover of land or landownership from large landowners, the land being subsequently placed at the disposal of small farmers and landless labourers. This definition is close to the common usage of land reform and the generally accepted sense of the term (see Warriner 1969: xiv–xv; Lipton 1974: 269–74). Moreover, it addresses itself more directly to rural poverty and inequality which is the major objective of our study. Other measures that might theoretically improve the living standards of peasants fall outside our definition. Such measures as tenancy reforms, land reclamation and settlement projects, and progressive land tax have largely shown themselves to be of little help to the rural poor. They might be considered as 'substitute policies' when land reform is not feasible.[3]

Land reform usually comprises two phases: (1) expropriation, and (2) the establishment of a new structure that reduces poverty by spreading the benefits of land more equally than before (see further Warriner 1969: 17–22). The new agrarian system may be one in which peasants own their land and cultivate it personally or co-operatively, or one in which land is nationalized and where large-scale farming such as co-operatives and collective farms generally substitute peasant farms. If land reform aims at developing the former system, it is distributive, and when its object is the latter, it is collective. In this study, we consider only distributive land reform.

Political dimension of land reform

Land is the most important source of income and wealth in traditional agrarian systems. In view of the dominant role of the agricultural sector in the socio-economic life of many developing countries, the importance of land normally extends much further than agriculture; land often means political power and social status in the society as a whole. Therefore, land reform which involves the restructuring of patterns of wealth, income, social status and prestige is predominantly political (Raup 1968: 303). It is also politicians who generally initiate the land reform debate.

The politicians who advocate land reform claim that their major motive is the search for social, political and economic equality in society. Their real motives, however, are likely to be more complex

than that and reflect the interests of various classes and groups in the society. As a matter of fact, different politicians strive for different types of changes in a traditional agrarian system and the associated power structure.

Political motives can be classified into three broad groups in regard to their ultimate purposes: liberal, populist and socialist. The liberal motive is represented by the urban elites and even some landowners who desire a modern agrarian system which is freed from the numerous traditions that limit their economic activities. Moreover, the urban elites want to be able to reduce the power of landlords and thereby achieve a more equitable distribution of political power. On the whole, these liberal forces do not seek a radical land reform that would completely erode the power of the landlords and limit the size of land property. They use land reform, above all, as a measure by which pressure can be applied to traditional landlords in order to weaken their resistance to changes in the status quo.

On the other hand, populists and socialists support widespread and fundamental land reforms. They want to eliminate entirely the power of large landlords and redistribute social status and political power. Agricultural labourers, peasants, urban middle classes and workers constitute the major social forces that support such land reform programmes.

In spite of their common interest in the abolition of large landlordism, populists and socialists generally aim at different objectives. The populists, for example, want a transformation from the traditional agrarian system tc a peasant system, i.e. a family-farm agrarian system. They believe that such a system would act as a bulwark of political democracy and stability. Moreover, they consider family farms to be more efficient.

Socialist objectives, on the other hand, aim at restructuring society as a whole. They consider agrarian changes both as an important part of overall social change and as a means to gain power and consolidate it in order to carry out other changes. Furthermore, they prefer large-scale farming such as co-operatives and collectives. The main justification for the latter is that large farms are considered to be more efficient than small farms. In addition, socialists sometimes argue that co-operation is morally superior to the individual profit motive.[4]

In addition to internal forces, external factors can influence the process of land reform in a given country. Foreign powers may, for example, exert a marked impact on a land reform programme by strengthening or weakening certain forces. They may also attempt to initiate, speed up, slow down, stop, or even reverse a land

reform process. This can be done through direct intervention, providing provisional aid and assistance, or via the propagation of ideologies.[5] The example of agrarian reforms and their results in other countries can also to some extent influence land reform activities, especially in neighbouring countries (Bergmann 1977: 21).

To sum up, the land reform question involves different interest groups in society and provokes different reactions. Landlords as a whole oppose land reform activities which endanger the source of their wealth and income. They do their best to thwart land reform programmes. At the other extreme, there are forces which aim at far-reaching changes through land reform and therefore advocate radical reforms. It is eventually the balance of power between different forces in a society which determines whether or not land reform will be introduced, and if it does, how far it will go in redistributing land, and what changes it will ultimately bring about.

Economics of land reform

As we saw in Chapter 2, the traditional tenancy system has been criticized for adversely affecting agricultural investment and innovation, and for the prevalence of rural poverty and inequality in many developing countries. The concentration of land in the hands of a few absentee landlords seems to be a central element in most of the criticisms of tenancy systems. Therefore, land reform is widely considered to be an effective measure which can promote rural equality and prosperity and foster agricultural progress.

Broadly speaking, the explanation is rather straightforward. First of all, land reform redistributes the main source of income in agriculture and thereby brings about a more equal distribution of income. Moreover, redistribution of income presumably occurs in favour of the peasants who live in rural areas. This in turn means that a larger share of agricultural income remains in rural areas and does not go directly to the cities. Lastly, land reform gives peasants secure titles to land which would in turn stimulate investment. It is not then difficult to see that the interaction of these factors might increase investment and innovation in agriculture and raise the living standard of the peasants.

The traditional approach to sharecropping would certainly suggest that there would be a further improvement in agricultural employment, investment, and output as a result of land reform since it would eliminate the disincentive problem of the system (see Chapter 2, pp. 11–19). On the other hand, the modern line, which contends that sharecropping is efficient, would deny such consequences.[6]

Given the lack of a generally accepted theory of sharecropping, it seems rather difficult to deduce the economic outcome from land reform in such a system. We should, however, consider two important points. First, the efficiency line has devoted much effort to show that sharecropping is as efficient as fixed-rent tenancy and owner-cultivation with hired labour. There have also been some attempts to show that sharecropping has certain advantages over these two systems. But, as we will discuss later in this chapter, these advantages are not always valid when comparing sharecropping with direct peasant proprietorship. One can then conclude that transferring the ownership of the land to sharecroppers is likely to leave the levels of production, investment and employment intact according to the efficiency line.

Second, the sharecropping literature has generally concentrated on the production effects of the system and has overlooked the distributional effects. The fact that sharecropping is often associated with skewed income distribution and poverty is rarely brought up. This is probably presumed to be self-evident: unequal distribution of landownership will result in an unequal distribution of income and has nothing to do with sharecropping. As a consequence, there is no reason to believe that the sharecropping theory opposes a more equal distribution of landownership.

On the whole, the prevailing impression among many observers is that land reform can generally cope with the problems associated with traditional tenancy systems. Concerning sharecropping, one can conclude that the efficiency line views land reform as a purely redistributive measure with no efficiency consequences while the traditional line considers that it is also likely to promote better utilization of production factors.

We maintain that it is the post-reform agrarian system which determines the ultimate outcome of land reform. To express it more precisely, if landowners are not powerful enough to prevent the transformation of the traditional agrarian system, a new agrarian system will emerge out of the reform. This new system may be either 'unimodal' or 'bimodal'.[7] Our major hypothesis is that the economic consequences of land reform are closely bound up with which of these systems prevails after the reform. We begin the discussion with the bimodal system.

The bimodal agrarian system

Land reform laws usually specify the maximum amount of land which a landowner can retain for himself and allow the government to expropriate the lands in excess of that ceiling. Moreover, in

many countries certain categories of land are exempted from expropriations such as land belonging to religious institutions and mechanized farms. These, sometimes deliberate, loopholes in the law provide good opportunities for landlords to retain most of their lands. Since landlords are normally free to choose the land that they want to hold, the retained land usually turns out to be among the best.

At the same time, landlords may increasingly evict their tenants and form modern large- or medium-scale holdings. Exemptions of mechanized farms from expropriation, the prospect of increased profitability from new agricultural technologies, attempts to compensate the loss of income from the expropriated land by making the best possible use of the remaining land, the need to create an argument against continued land reform, are among the factors that may contribute to persuade the landowning class to substitute owner-cultivation for the tenancy tenure system (see Alexander 1974; Barraclough 1970; de Janvry 1981; Warriner 1969). Although the landowning families, i.e. former landlords or their younger family members, are mainly the managers of the new, modern farms, some merchants, army officers, local and foreign agribusiness companies, etc., may succeed in taking over large and medium holdings. On the whole, while the traditional system changes radically, land concentration undergoes only minor changes and large landowners continue to be a very influential class both politically and economically.

Alongside the farms of the large landowners, there will be another sector comprising small farms which belong to small peasant proprietors. They largely consist of the ex-sharecroppers and the pre-reform small peasants. They will only own a small fraction of the arable land. This means that the reform will create two main groups of farmers: large farmers who own and control large holdings, or some large and some medium holdings, and small farmers who own small plots of land.[8] This is what we call a 'bimodal system'.

The bimodal system, factor-price distortions, and misallocation of agricultural resources[9] Unequal access to land, which is a basic feature of the bimodal system, implies different land prices for the two groups of farmers. Land is relatively abundant for the large farmers and scarce for the small ones. It is then plausible to argue that land is valued relatively lower than its social opportunity cost for the first group while the second group has to pay higher prices for it. Two reasons can be mentioned in this context. First the large farmers set a monopolistic price for land because they are by far the

main source of land supply to the local land market.[10] Moreover, small farmers need land much more than large landlords and this makes their demand less elastic. It is then possible for the large landowners to practise price discrimination, imposing a higher price on small farmers than (implicitly) on themselves (Lundahl 1984: 525).[11]

In the capital market, landowners are able to obtain loans very easily and at lower rates because their wealth and income minimizes the risk of lending to them. Moreover, the socio-political status of the large landowners tends to direct government credit policy to the satisfaction of their interests (see, for example, Barraclough 1970: 910; de Janvry 1981: 156–7). They have their own savings as well. Small peasants, on the other hand, can hardly obtain credit from institutional resources and usually only have access to the local credit market with notoriously high interest rates. Interest rates are high because the local credit sources such as moneylenders and shopkeepers have monopoly power in the market and because the risk of default is high. Large landowners thus pay a price for capital which is lower than the social opportunity cost while the mass of the peasantry pay a price which is certainly higher than the social opportunity cost of capital.

Finally, control of land and credit associated with limited alternative opportunities for off-farm employment open to labour may enable landlords to exercise control over labour as well. In other words, their monopoly power in the local land market provides them with monopsony power in local labour markets. As monopsonists, they limit their labour demand and push down the number of hired labourers and the wage rate below competitive levels. Nevertheless, the large farmers who rely mostly on hired labour may have to pay higher wages than the implicit wage rate for the small farmers who rely mainly on the family members' labour. In some cases, for example, large landowners are legally required to pay a fixed minimum wage rate. On the other hand, family members may prefer to work on the family farms at (implicitly) lower wage rates rather than moving away for a variety of reasons.[12] We may conclude that the wage is lower than its social opportunity cost for both groups, although it is likely that large farmers pay higher rates than the small ones.

The characteristics of factor markets in a bimodal system can thus be summarized as follows:

$$r_b < r < r_s$$

$$i_b < i < i_s$$

$$w_s \leq w_b < w$$

where r, i and w represent social opportunity costs of land, capital and labour respectively. The subscripts b and s are used to show, respectively, the price of similar factors for large and small farmers.

These factor price distortions cause an inefficient resource allocation in agriculture. Large farmers make extensive use of land and capital, which are comparatively scarce in developing countries, but sparse use of labour, which is relatively abundant. The small farmers, on the other hand, are forced to use too much labour for their available land and capital. In other words, large farmers tend to use techniques of production that are too land- and capital-intensive relative to the social optimum while on the other hand, small farmers use too labour-intensive methods. This is an important conclusion. Let us look at this in some detail.

Biased technical progress under the bimodal system A farmer can produce a crop in many different ways. He will naturally choose a method which minimizes his costs in terms of the prevailing relative factor prices. Other things being equal, all producers of a certain crop use one and the same technique in a competitive world because they face similar prices. On the other hand, the farmers choose different methods when they face different factor prices as in the case of a bimodal system.

Suppose that qq, in Figure 4.1, represents an isoquant for cultivating a crop with different combinations of labour (L) and material inputs (M). (For simplicity we assume that there is continuous substitutability between labour and material inputs.) Further, assume that bb_1 represents the relative factor prices for large farmers whose labour cost is relatively more expensive while ss_1 stands for the relative factor prices of small farmers, who pay relatively more for their material inputs. We can easily see that rich farmers choose technique α which employs relatively more capital and less labour than technique β that is chosen by small peasants.

Given factor price differentials, the two groups of farmers may even show different attitudes toward the new methods of cultivation. In other words, an innovation may only be attractive to one group of farmers. This can be explained with the help of Figure 4.1. We have specified five zones in this figure and an innovation may occur in each of these five zones. The adoption of any innovation depends on where it occurs.

Innovations that fall in zone 1 are 'technologically irrelevant' since they use more inputs per unit of output than existing methods. Those in zone 3, the shaded areas, are 'economically irrelevant' given the current prices. The first group of innovations

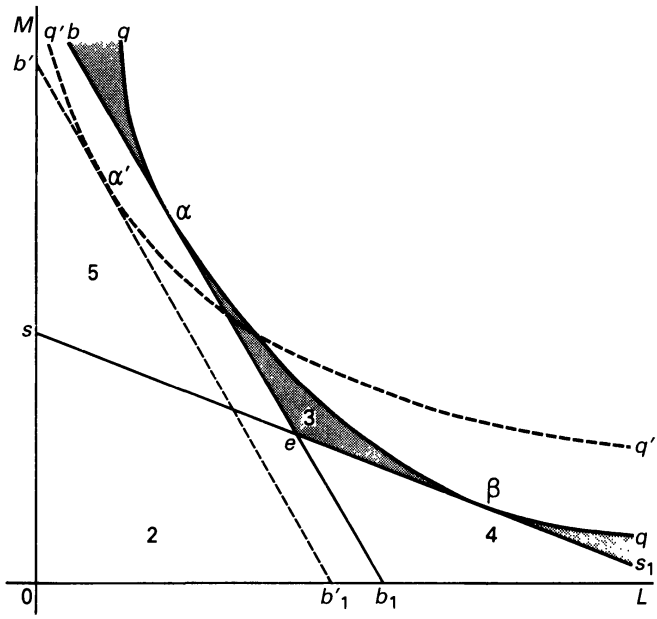


Figure 4.1 Choice of techniques of cultivation

would never be adopted irrespective of the factor prices but the second group are not adopted at present factor prices but may be adopted at certain factor prices. On the other hand, innovations that occur in zone 2, the $0seb_1$ area, will be adopted by everyone.

Any innovation that falls in zone 5, the *seb* area, will be attractive for those who have bb_1 relative factor prices (the landlords), but such an innovation is economically irrelevant for those with ss_1 relative factor prices (the peasants). These innovations could be called 'landord-biased' since only large farmers benefit from them. For similar reasons, new techniques that occur in zone 4, the b_1es_1 area, will only be adopted by small peasants and are therefore 'peasant-biased'. For example, if a new technology is represented by isoquant $q'q'$ it is 'landlord-biased' and will be adopted by large farmers (techniques α') while small farmers will continue to use technique β.

In recent decades, many innovations were only attractive to large farmers since they have been capital-intensive in nature, and thus landlord-biased.[13] This may be explained on two grounds. First, innovations are usually imported from developed countries where

labour has not been an abundant factor. Second, research, extension services, and investments in agriculture sponsored by governments, have themselves been landlord-biased. This can partly be explained by the fact that many governments preferred landlord-biased innovations because they have primarily been concerned with raising (explorable) output. Since large farmers control most of the land, 'a landlord-biased innovation could be expected to spread over most of the cultivated surface and have a large impact on production' (Griffin 1974: 51). Moreover, it has been argued in certain quarters that agricultural improvement programmes should be directed towards the large farmers who are more innovative than the small, poor farmers.[14] Finally, where large landowners have remained a powerful class in society they will inevitably influence the government to adopt landlord-biased policies.[15]

The bimodal system, rural poverty, and agricultural growth The mechanism of the bimodal system will gradually create two quite different sections in the rural economy. At one pole, there would be large farmers who use modern techniques and who produce mainly for the market. The other pole would comprise small farmers who are stuck with traditional methods of cultivation and whose production is mainly subsistence-oriented.

Large farmers are able to follow a 'crash' modernization strategy which is largely based on imported equipment. These imports are usually supported by subsidizing interest rates and undervalued foreign exchange. A key aspect of this modernization process is mechanization, i.e. the use of increasingly agricultural machinery. These are capital-intensive innovations and are welcomed by large farmers to reduce their labour requirement (see Griffin 1974: 67–73). The tendency towards mechanization in large landowners' farms are reinforced by two factors: (1) most agricultural machinery and associated equipment can be readily transferred and their introduction is encouraged by the large international corporations that manufacture them via the provision of suitable service facilities; and (2) the local climate of opinion is often favourable to machinery such as tractors since they are considered to be a symbol of modernity (Johnston and Kilby 1975: 152).

Large farmers produce for the market and can more easily move towards crops that are more profitable. If the profitable crops are cash crops, the small peasants can hardly switch to them because they often do not have access to markets and lack the knowledge and experience of growing cash crops. They cannot afford to buy this know-how either. Moreover, they are likely to be reluctant to run the two additional risks which production of cash crops will

impose on them: first the price of the crops may decline or the marketing system may become disrupted and, second, the price of inputs which are necessary for growing the crops may rise or the supply of the inputs may break down (Griffin 1974: 63). Thus, while the prospects of large farmers for increasing their income through cultivating cash crops will probably grow, small farmers will be bound to their subsistence crops.

On the other hand, staple food crops might turn out to be more profitable because it is likely that population growth, growing urbanization, and rising income cause excess demand for food. This would tend to raise the price of food crops and thereby generate increased income for the small farmers whose potentially marketable output is food crops.[16] But there are reasons to believe that price rises may increase income inequality and impoverishment in rural areas. First of all, an increase in the price of food negatively affects landless labourers and deficit farmers who have to buy food considering the fact that real wages will in all probability decline. Second, large farmers who are the main surplus producers will benefit from price rises and the relative stagnation of money wage rates. The marketable surplus of food will then increase by squeezing the consumption of the rural poor and the distribution of rural incomes worsens (Ghose and Griffin 1980: 560–1).

The polarization of the agricultural sector into two sub-sectors is then the most plausible outcome of the bimodal system. Large commercial farms constitute one sub-sector which can enjoy imperfections in the factor market, receive various forms of support from the government, and use new techniques of production. Another sub-sector comprises small peasants and may be adversely affected by all the above factors. Large farmers are able to increase their incomes by both reducing costs and increasing their relatively high price output. Small farmers can hardly increase their incomes and may in fact even experience a fall in their incomes.

Landless labourers and the tenants who are evicted from tenanted land will probably face even greater difficulties. Considering the limited labour demand in rural areas, which might even shrink as large farmers increasingly adopt labour-saving methods, migration will be the only alternative open to these people.

On the whole, there are many factors in the bimodal system which operate against the rural poor. They will even counteract the effects of a secure title to land and a partial redistribution of income which would occur after the redistribution of some land to the peasants. It is not therefore unreasonable to believe that rural poverty and inequality will tend to increase after a land reform that produces a

bimodal system. However, the few peasants who can, for one reason or another, advance into the medium-size farmer class or who succeed in getting permanent employment on large commercial farms are likely to experience an increase in their levels of living.

Agricultural production will probably rise, as landlords become actively involved in agriculture and as they, and the newly arrived commercial farmers, invest in land and use modern techniques. However, this rise is not socially efficient and can hardly promote economic development. The rich farmers in fact operate in a system which is subject to factor-price distortions and thus fail to use resources in the most efficient combinations (see pp. 55–7).

Modernization of farming on large farms will also occur in competition with the other sectors of the economy which require foreign exchange to develop. Moreover, if the large farmers specialize in producing cash crops, which are often for export, a large part of the scarce resources in agriculture, land and capital, will be shifted away from the increasingly required food crops. This will reinforce the need to import food crops which puts further pressure on the foreign exchange resources of the country. Finally, the amount of foreign exchange which is spent to cover equipment and food crop imports could even exceed the amount of foreign exchange that large farms earn by exporting cash crops.

The unimodal agrarian system

The post-reform agrarian system would be unimodal if the reform succeeds in transferring all of the land of large landowners to tenants and landless labourers. Landlords might be allowed to retain some land but this would be limited so that the size of their holdings would not exceed that of the average peasants' holdings. A land reform of this kind may lead to an immediate reduction in rural inequality. In other words, agricultural income would be redistributed in favour of the beneficiaries of land reform who are mainly the ex-tenants. They would experience a rise in their standard of living since they retain all of their crops for themselves. They may pay annual instalments for the land they receive but compensation will presumably not be substantial given the favourable conditions for peasants which have been strong enough to promote a radical land reform.

The growing capacity of a system of small-scale family farms might be questioned because of a possible trade-off between equity and capital formation. This is based on the premise that large, rich farmers save a greater proportion of their incomes and therefore can afford to invest in agriculture. In contrast to such an argument

some empirical studies show that there is not a meaningful difference between the marginal propensities to save of small and large farmers and, thus, there is no reason to believe that rural saving and investment will be reduced because of such a difference (see Adams 1973: 134).

We even have good grounds for believing that investment tends to increase in a unimodal system. First, land reform stimulates investment and improvement activities since it transfers land ownership to the tenants and secures their rights on their holdings. The farmers are now guaranteed to reap the benefits of their work. This will provide an incentive to increased effort which can lead to higher output and investment.[17]

> The prospect of long and secure tenure may thus create a condition in which maximum incentive is given for investment of family labor time in productive undertakings. Much agricultural capital formation can be explained in this fashion. Livestock care, repair and maintenance of structure, improvement of water supplies, drainage, soil improvement practices, and a variety of similar tasks are often accomplished in agriculture in what might otherwise be leisure time.
>
> (Raup 1968: 274)

Second, even if land redistribution reduces the proportion of income saved, it can increase the volume of savings as the result of increases in the total amount of agricultural output after the reform (Lipton 1974: 291).

The strength of a unimodal system is its capacity to employ the relatively abundant factor, labour, in order to utilize more intensively the relatively more scarce factors, land and capital. In other words, the peasantry generally employs method β in terms of Figure 4.1. It is also 'peasant-biased' innovations which become attractive in this system. The fact that 'output per unit of land is inversely related to farm size' reflects the potential capability of a system of small, family-farms to increase land productivity and output (see, for example, Dorner 1972: 120; Cornia 1985). Redistribution of land among villagers in more equal but still viable holdings can therefore extend the benefits of the reform to some landless families, who do not have access to land, at the same time as it may raise output.[18] 'Viability', of course, limits the number of beneficiaries in densely populated countries. However, a land reform can hardly contribute to the creation of an innovative or happy peasantry if the holdings are too small to provide a decent minimum level of income for the beneficiaries (Lipton 1974: 283).

As peasant incomes rise and their fear of starvation decreases, their capacity to take certain risks and innovate may increase. This can lead to a progressive modernization which implies widespread adoption of a sequence of innovations that is mainly based on simple, cheap implements that can be produced at home. Moreover, there are many agricultural innovations such as hybrid seeds, fertilizer, pesticides, etc. that are technologically scale neutral. Lack of complementary measures like irrigation and credit most often prevent small farmers from applying these inputs in a bimodal system. It is likely that the institutions and government policies which encourage and support the introduction of such inputs, and would probably serve the large land owners in a bimodal system, become peasant-oriented in the unimodal system. Whenever necessary, even indivisible inputs such as irrigation systems and tractors could be made divisible by special arrangements (see Johnston and Kilby 1975: 148–53). All these would facilitate the modernization process in agriculture.

In a unimodal system, all peasants who are confronted with similar climatic and soil conditions are supposed to have approximately equal access to land, water, labour, technical knowledge, etc. If an innovation proves to be profitable to one of these farmers, it is profitable to all the other farmers and it will be adopted rapidly and widely (Griffin 1974: 52). These innovations will be labour-absorbing since their profitability is determined by labour-abundant smallholders.[19] Once the decision to hire tractor services is made by these farmers, there is a good chance that the input will also be a complement to the labour of family members and, in consequence, it will not displace labour (Johnston and Kilby 1975: 149).[20] It is thus reasonable to assume that modernization in this system will increase labour demand which may also improve the lot of some of the landless families.

A small, family-farm system can also further economic development since the total output of the sector is likely to rise. It is important to note that agricultural growth in this system is largely financed by the direct diversion of the small farmers' inactivity, and this hardly affects the nation's resources for financing growth elsewhere (Lipton 1974: 289).

We should observe that the equalization and improvement effects of a unimodal system would be limited under certain conditions. First, although the extent of inequality can be reduced significantly in the rural areas, it does not disappear altogether. Peasants' different initial resources, ability and desire will naturally influence the speed and efficiency with which they can seize new opportunities. Moreover, technological progress will also

create differences among peasants since it has an uneven impact on different regions and types of farming. But these differences are not as significant as those in a bimodal system (Johnston and Kilby 1975: 148). In other words, the peasantry will not be polarized between large wealthy, and small destitute farmers.

Second, lack of enough cultivable land normally limits the benefits of land reform to most of the landless labourers who are among the poorest of all villagers. It is true that some of them may benefit from the reform by either receiving land or finding employment as the result of increased labour demand in agriculture. But it is also true that even a highly egalitarian reform which does not intend to create non-viable holdings will only cover a small fraction of landless families in many countries whose agriculture is dominated by a tenancy system (see Lipton 1974: 285–6). Furthermore, subdivision of the larger farms implies reduced levels of employment for landless labourers while the increased labour demand after the reform may fail to provide these people with sufficient employment opportunities. Most of the increased labour demand may be met by increased labour supply of family members.[21] Therefore, alternative employment opportunities in rural areas, e.g. through supporting rural small-scale industries, seems necessary to improve the living standards of these people. The promising opportunity of urban migration would naturally help to alleviate the rural unemployment problem.

The unimodal system and the importance of auxiliary services

Apart from the complementary programme for helping the non-beneficiaries of a land reform, there are also some follow-up services which seem necessary for those who receive land. Over and above land, the new owners need other things: they need credit and advisory services, especially in respect of the utilization of the new inputs. They will also need the help of research and development institutes for the creation of new agricultural inputs and for the testing and adoption of these and imported inputs into the local environment. Moreover, they need marketing and input delivery services, etc. The provision of such services for the peasants will naturally resolve some of their difficulties after the reform and speed up the improvement process which is likely to occur in the unimodal system.

This is not, of course, to suggest that the failure or success of the reform will closely depend on a complicated package of such follow-up programmes. The unimodal system will to some degree guarantee the success of the reform in reducing rural poverty and inequality. First and foremost, this depends on the fact that in our

model the unimodal system replaces a traditional tenancy system. In other words, the land reform beneficiary is normally an ex-tenant or an ex-sharecropper who

> already knows about farming; he can now innovate and devote more time to his land, so he needs some extension, but he probably knows how to get it and whether it is worth the effort. Similarly, when his marketing needs change as he moves into surplus, he has some relevant knowledge.
>
> (Lipton 1974: 296)[22]

A problem might arise here with respect to sharecroppers. The study of sharecropping theory in Chapter 2 has shown that sharecropping might offer certain benefits to the contracting parties that other tenurial arrangements might not (see pp. 19–25). It is then likely that sharecroppers lose these 'privileges' when they become peasant proprietors. This may cause some difficulties for them.

As was mentioned earlier, however, most sharecropping models compare this type of tenure system with fixed-rent tenancy and owner-cultivation with hired labour. It is principally within such a context that the advantages of sharecropping appear. However, many of these advantages lose their importance when sharecropping is compared with direct peasant proprietorship. For example, as a peasant proprietor, the land reform beneficiary relies on his own (family) labour which does not as such require supervision.[23] Moreover, he, himself, is the manager and he would certainly make appropriate responses to unpleasant or pleasant surprises in order to reduce the damage from the former or to increase the gain from the latter. In other words, supervision and production flexibility advantages of sharecropping are of little relevance with regard to peasant proprietorship.[24]

On the other hand, the new owner may face certain problems since he will be deprived of sharing risks through sharecropping and will not have access to, for example, the managerial skills provided by landlords. In view of the increasing absenteeism within traditional tenancy systems, the latter has probably lost most of its significance. It is also reasonable to believe that the risk problem will be less severe after the land reform since the ex-sharecroppers will retain almost all their crops for themselves; as they become more prosperous, they will be less vulnerable to risks. Nevertheless, some insurance against a harvest failure in the years immediately after the reform might be necessary to reduce the extent of the problems that would arise in such a case.

On the whole, one cannot ignore the importance of such programmes. First, the improvement process may be very slow if it is left to itself. The fact that many beneficiaries are indebted and have to repay their pre-reform debts will reduce their investment and innovation capacity over a period of time. Second, lack of alternative ancillary services may give the former rural elite the chance to provide certain services and retain its dominance in another form (Lipton 1974: 296–7). Provision of different services will then accelerate progress in improving the lot of the peasantry and will reduce the risk of the survival of the monopoly powers.

However, the mass peasantry in a unimodal system can probably make their voice better heard and may thus influence the state to provide the necessary services. This will be especially true if the political climate favourable to a radical land reform also persists after the reform.

Concluding remarks

Land reform is usually conceived as an egalitarian measure whose main object is to help the rural poor. The definition at the beginning of this chapter reflects such a belief. The effectiveness of a land reform programme depends, above all, on the political climate of the country. While many different socio-political forces might support land reform, all of them do not advocate radical land reform. Moreover, it is very difficult to overcome the resistance of the landowning class who are usually very powerful in developing countries.

Land concentration in a few hands is the main feature of the traditional agrarian systems in Asia and Latin America. Since land is the major source of income in agriculture, the principal impact of such systems is a highly uneven income distribution in the agricultural sector. A land reform programme is then capable of redistributing income and reducing rural inequality. Moreover, the theoretical analyses indicate that the traditional agrarian systems are likely to be responsible for low output and productivity, rural unemployment and underemployment, and insufficient incentives and/or ability to invest. Consequently, a land reform which changes the traditional agrarian system might also contribute to agricultural development.

Although the micro-analysis concludes that redistribution of land, or landownership, might improve income distribution and promote growth, the actual outcome will largely depend on the kind of agrarian system which develops out of the reform.

A bimodal system emerges when land reform expropriates and

redistributes land on a limited scale but succeeds in compelling and encouraging the large landowners to turn away from their traditional behaviour and become modern agriculturalists. The agrarian system will then be characterized by two different types of farms. On the one hand, there will be small farmers who own small plots of land and, on the other hand, there will be large farmers who own large amounts of land and manage them in the form of large estates or even medium-sized farms. Large farmers enjoy monopoly power in local markets and can influence the government to direct, for instance, agricultural policies towards their interests. They will use modern technology and inputs and can increase their total output and profit. The small farmers, on the other hand, because of their poverty and the prevalence of landlord-biased innovations and government policies, will be unable to change their primitive methods of cultivation. In the final analysis, the new system might contribute to agricultural growth, but it will also increase inequality and poverty in the rural areas. At the same time, production will not be socially efficient since large farmers produce at distorted factor prices.

If the political climate permits a radical land redistribution, the post-reform agrarian system will consist of small family farms. As the peasants own their farms and know that they will retain the results of their effort, they have sufficient incentive to work hard and cultivate their farms as intensively as possible. These farmers have abundant family labour at their disposal and will use it to increase agricultural output and capital. The peasants will then be able to innovate, improve the production methods and inputs. As labour is relatively cheap for these farmers, they will tend to adopt labour-intensive innovations. This will be a general tendency among all peasants since every peasant is faced with more or less similar factor prices and produces in a competitive market. That is to say, a unimodal system might not only benefit the majority of peasants and create growth but it also guarantees the social efficiency of agricultural production. Nevertheless, follow-up services and rural work programmes for surplus labour seem to be necessary for rapid improvements in the levels of living of the rural poor.

Chapter five

An analysis of the Iranian land reform programme

The early appeals for land reform in Iran can be traced back to the early twentieth century. Some minor steps were also taken in this direction by the mid-1950s. For example, as was mentioned in Chapter 3, a royal decree was issued in 1951 for the sale of the crown estates to the peasants. Such attempts had, however, little effect on the Iranian landlord–peasant system and the pressures for agrarian change, especially through a land reform programme, were increasing toward the end of the 1950s.

Consequently, a land reform bill was drafted by the government and an amended version of it was passed by both houses of parliament in 1960. The law was badly drafted and contained many pro-landowner provisions that made it virtually ineffective. A more radical land reform law was then introduced two years later, on 9 January 1962. Although this law was technically an amendation of the land reform law of 1960, it was virtually a new law (Lambton 1969: 63). The land reform law of 1962 was almost immediately put into operation and started the first stage of the Iranian land reform programme. The second and third stages of land reform followed subsequently in 1965 and 1969, respectively, with the implementation of two new amendments to the law.

In this chapter, we want to introduce the basic features of the Iranian land reform programme and examine its contents in order to find out if it could benefit the peasants. As we discussed earlier, this depends largely on the post-reform agrarian system. Therefore, we will try to trace out the land reform laws to see whether they directed the Iranian traditional system toward a pro-peasant, unimodal or a landlord-biased, bimodal agrarian system.

The present chapter, however, is only concerned with the implications of the land reform laws and other agricultural laws for the development of the new agrarian system. The implementation of the law and the actual consequences of the land reform programme will be discussed in the next chapters. But before analysing the land

reform laws, we will look very briefly at the different forces which worked to promote the land reform programme.

A review of the background of the land reform

In the late 1950s, Iran faced a series of social, economic, and political crises. Certain fundamental changes in society seemed unavoidable if the stability of the country was to be maintained. Due to the importance of agriculture in the life of the society, the agrarian question was inevitably a key ingredient of any reform programme.

Between 1955 and 1960, the Iranian economy enjoyed an annual growth rate of about 7 to 8 per cent (Bharier 1971: 45; see also Mahdavy 1965: 135). This growth, which was initiated by increased oil revenues and foreign loans and grants, was accompanied by high public and private expenditures. Imports increased enormously. Most of these imports were consumer goods which might be considered as luxuries or unnecessary at that time. Moreover, they were sometimes in excess of local demands.[1] Investments in trade, construction, and industry increased. Investments were also made in agriculture but these were mainly directed towards large dam projects with doubtful economic returns (Momeni 1980: 163–4; Bharier 1971: 92). But, as inflation began to increase and foreign exchange reserves were depleted, economic problems started to appear. In 1960, the government had to carry out an economic stabilization programme. The volume of internal borrowing, imports and government expenditure was reduced drastically and this resulted in an almost complete stagnation of the country.[2] Meanwhile, signs of discontent were appearing in the urban areas.

As a matter of fact, economic problems in the cities were accompanied by social and political crises in the early 1960s. The increase in urban unemployment, the exacerbation of the living conditions of the urban poor, the widening of the gap between the poor and the rich, and the growth of corruption contributed to create unrest in the urban areas. The established political authority in the country was challenged by both the urban elites, who had gained in strength during the boom years, and the suppressed opposition, which got the chance to operate actively again. Demonstrations and meetings were held in many cities, especially in the capital, although they were treated violently by the government forces (Mahdavy 1965: 134–6; Ashraf and Banuazizi 1980: 25; Momeni 1980: 177–90; see also Zonis 1971: 71–3). On the whole, pressures for changes in the society, not least for changes in the agrarian system, were mounting.

The traditional landlord–peasant system had in fact been criticized from different directions for a long time. It was generally believed that this system was feudal, and that it prevented agricultural development and caused rural poverty. Although the reformation of the agrarian system was demanded by various groups and parties, they often suggested different solutions. For example, some aimed at the abolition of 'feudalism' and redistribution of the landlords' land among the peasants while others believed that the major problem was the landlord–peasant relationship and not the concentration of landownership. Therefore, they argued, it was the latter which had to be changed so that the large landowners became modern agriculturists who cultivated their land with the help of agricultural workers.

There were also certain large landowners who, for a variety of reasons, had been advocating various measures that could change the prevailing situation in the rural areas. Some, for instance, wanted to sell their landed properties if the government could guarantee that they would receive the proceeds from the sale of their land in order to invest it in other sectors. Others wanted to recapture their lands for mechanized self-cultivation, while a third group was anxious about peasant revolution.[3]

The need for change in the conditions of the rural areas was also apparent to the government. Consequently, certain attempts were made to improve conditions for the tenants of the state and crown lands who were now given the opportunity to purchase these lands. However, these attempts to modify the situation progressed very slowly. Moreover, most of the land offered for sale went to people other than the local peasants (see Chapter 3, p. 29 and pp. 32–3).

In contrast to these land sale policies which were supposed to benefit the peasants of certain areas, the government issued two decrees in the early 1950s which concerned all sharecroppers throughout the country. These decrees were in line with tenancy reforms and did not promise meaningful changes in the prevailing system. According to the first decree, 20 per cent of the landowner's share was to be deducted, 10 per cent of which had to be paid to the sharecroppers and the remaining 10 per cent was to be allocated to local development projects. The second decree abolished all the dues that the peasant used to pay to the landlord over and above his share of the crop and limited landlords to using peasant labour or their agricultural implements without payment (for more details, see Lambton 1969: 37–41). However, this government did not last long enough to implement these decrees throughout the country and the subsequent government abandoned both of them.[4]

During the 1950s, landlords generally reduced their investments in irrigation systems because other sectors became much more attractive. For the same reason, traditional moneylenders limited their loans to the peasants. Consequently, both the area under cultivation and production decreased during this period, especially, that of staple crops (see Momeni 1980: 181–2). Living standards in the rural areas deteriorated since the rural share of agricultural production failed to keep pace with the growing population. Although some of the villages that were near big towns and main roads experienced improvement in their economic conditions, poverty remained the lot of the majority (Lambton 1971: 15).

The tendency to blame the landlords for many of the evils in society in earlier years gave place to a campaign of vilification in 1959 and 1960 (Lambton 1969: 55). At the same time, Iran was exposed to strong external pressures for the introduction of a land reform programme. First, the revolution in Iraq in July 1958 had significant repercussions on Iranian political life. On the one hand, Iraq began a land reform programme, and, on the other hand, the revolutionary government put propaganda pressure on Iran (Ashraf and Banuazizi 1980: 26; Momeni 1980: 143). Second, the situation was exacerbated by the numerous radio programmes which launched bitter attacks on the regime.[5] This anti-regime propaganda was especially important because the radio had spread to the most distant villages by 1959 (Lambton 1971: 15). Third, the United States, as the major foreign supporter of the Iranian regime, was urging Iran to carry out a land reform programme.[6] Finally, when most of the countries in the region had had a type of land reform, the absence of any such measures in Iran seemed too conspicuous (Mahdavy 1965: 137).

The government drafted a land reform bill in December 1959 to meet at least part of these pressures. In early 1960, an amended version of the bill was passed by the landlord-dominated parliament.[7] The law, which was already biased towards landlords, contained many loopholes and provisions in their favour.[8] In principle, the law laid down that each landlord could retain only 400 hectares of irrigated land or 800 hectares of unirrigated land. He was permitted to transfer 600 hectares of irrigated and 1,200 hectares of unirrigated land to his heirs. Gardens, woodlands, and land which during the previous two years had been cultivated by mechanized means and agricultural labour were among those categories of land which were exempted from the law. Furthermore, the landlords were able to sell land in excess of the maximum to the peasants, or the government, or retain it and pay an annual tax. In brief, 'the legislators of the Land Distribution Law had thought of every

possible way of evading its enforcement and had adopted them' (Land 1965b: 92).[9]

Such an abortive law could hardly bring about any change in a society which was badly in need of change. During August 1960 and May 1961 social unrest led to the fall of two cabinets and the cancellation of the parliamentary election. A new cabinet took office on 6 May 1961, and the newly elected parliament was once again dissolved. The new government declared its intention to struggle against corruption, carry out financial stabilization measures, and introduce a land reform programme.

The new land reform law was declared on 9 January 1962. This law came to be known as 'the law of the first stage of land reform' and was followed by two other laws in the 1960s. These laws, which were in fact amendments to the land reform law, established the framework for the second and the third stage of land reform programme. We will review the main features of the three stages of land reform separately in the following sections.

The first stage of land reform

On 9 January 1962, the land reform law was passed by the cabinet in the form of a decree.[10] There was nothing in the text of the law about the aims of the land reform. However, different government authorities stated that land reform aimed at, among other things, increasing agricultural production, raising the income of the rural people, improving the distribution of agricultural income, and setting free the peasants from the control of the landlords (Research Group 1970a: 51; see also Momeni 1980: 204–5, 217–19; 225–6).

The main features of the land reform law were as follows. (1) Each landlord could retain only one six-*dang* village of his landed property and any land in excess of this maximum was to be sold to government at a given price. (2) The lands which the government bought had to be sold to the peasants who cultivated them on the basis of the existing field layout of the village (*nasaq*). (3) Such peasants could only receive land if they became members of a co-operative society. (4) Fruit orchards, tea plantations, woodlands, 'mechanized lands', and endowed lands (*vaqfs*) were exempted from the law.[11] A Land Reform Organization was set up to carry out the reform under the supervision of a Land Reform Council.[12]

The law declared that all the members of a family that consisted of the wife and dependent children would be calculated as one person. Therefore, they could altogether retain only one whole village. This was confirmed later by the Land Reform Council

which decided that the wife and the dependent children could not hold one village each in addition to the head of the family.[13] This decision was reversed during the following eight months and both the wife and each dependent child of a landlord became entitled to hold one whole village each.[14] The new decisions did not affect the lands already distributed to the peasants but they applied to the lands that had been purchased but not distributed (Lambton 1969: 68; Research Group 1970a: 55).

According to the law, a village was the place of residence and work of a number of families who were engaged in farming, where the income of the majority of them came from agriculture, and which was locally recognized as a village. Since the definition was vague and caused difficulties, the Land Reform Council explained more precisely the definition of a village (Lambton 1969: 65–6; Research Group 1970a: 57–8). The Council also tried to limit the choices open to landlords by stating that land reform officials should recommend landowners to sell to the government the villages that had grown into municipalities.[15] However, this decision was modified later and allowed landlords to retain even such villages.[16]

As was mentioned in Chapter 3, there were also large landowners who owned parts of different villages. Considering this fact, the land reform law allowed the landlords to retain six *dangs* of different villages instead of one whole village. Many landlords then planned to redistribute their retained land among several villages in order to maintain their influence in several villages. This was first prohibited by the Land Reform Council but subsequently modified to allow landlords to divide their villages subject to concession made by the regional land reform organizations.[17]

The price of the land which was subject to distribution was to be assessed on the basis of the annual land tax paid by its owner before 1962 multiplied by a special index for each region. This index was determined by the Ministry of Agriculture and was based on, among other things, the distance of the village from cities, the type of crops grown, and the method of division of the crop.[18] The price of the land was paid by the government in ten annual instalments. This was later increased to fifteen.

The government had to sell the purchased lands immediately to the peasants who were cultivating them. These peasants were to pay the government-purchased price plus 10 per cent in fifteen annual instalments.[19] If the land was irrigated, it was to be transferred to the peasants with its traditional water rights.

Finally, the law provided some measures of protection for the peasants of those villages which were either retained by landlords

or were not affected by the law. Accordingly, (1) the periodic redistribution of holdings was prohibited (see Chapter 3, pp. 35–6); (2) the peasants were given security of tenure on the land which they were cultivating on the date when the law was passed; and (3) the peasant's share was to be increased by 5 per cent on irrigated land and 10 per cent on unirrigated land and this was to be deducted from the net share of the landowner. The landowner could, however, recover the land for self-cultivation if the peasant agreed to sell his cultivation rights (see further Lambton 1969: 81–4).[20]

It should be mentioned that the law did not define 'mechanized lands'. It only stipulated that the land which was farmed by mechanized means with wage labour on the date when the law was passed would be excluded from the law.

The land reform law of 1962 was immediately put into operation in Maragheh in East Azarbayjan.[21] In fact, lack of trained personnel prevented the Land Reform Organization from carrying out the reform simultaneously throughout the whole country and thus the land reform law had to be executed district by district (Hooglund 1982: 55).[22]

The second stage of land reform

About one year after the implementation of the land reform law of 1962, the government revealed its intention to extend the land reform legislation to the retained villages and the villages owned by small landowners (Lambton 1969: 103–4). The new stage was supposed to mitigate the dissatisfaction of the peasants unaffected by the law and to settle the disputes between the landlords and the peasants.[23]

The new law, which governed the second stage, is known as the Additional Articles. Contrary to the original law, the Additional Articles were put into operation after a long delay. They were first issued on 17 January 1963 but the regulations for their execution were passed by the parliament on 25 July 1964.[24] The implementation of the Additional Articles began on 22 February 1965.

According to the Additional Articles and the regulation for their implementation, the owner of the land which was not affected by the original law had to choose one of five options. He could (1) rent his land to his sharecropper for thirty years; (2) sell the land to the occupying peasants; (3) divide the land between himself and the sharecropper; (4) form an agricultural unit with the peasants; or (5) buy the cultivation rights of the peasants.

Mechanized farms remained exempted from the provision of the Additional Articles. The law defined 'mechanized land' as land

which was, at least, ploughed by mechanical means and was cultivated by wage labourers. The charitable *vaqfs* were not exempted at this stage but treated in a special way. They had to be leased to the occupying peasants for ninety-nine years. Personal *vaqfs*, on the other hand, were to be sold to the occupying peasants or rented to them if they did not want to buy the land. Orchards and different kinds of plantations also became the subject of the land reform provided that the occupying peasants had certain traditional rights on them. Such properties were to be sold to the peasants, or bought from them by the landlord, or shared between them in accordance with their respective traditional shares (Lambton 1969: 211–12).

The first method allowed the landlord to lease his land to the occupying peasants for a period of thirty years. The rent was fixed in cash and was based on the average net income of the landowner in the preceding three years. The rent could be revised every five years. If the tenant did not pay the rent within three months of its falling due, the landlord could, with the knowledge of the Land Reform Organization, evict the tenant and lease the land to someone else.

The second option was sale of the land to the occupying peasants. If this method was chosen by the landlord and the peasant, they had to mutually agree on the price and terms of the payment. The government would provide the peasants with a loan of up to one-third of the price of the land at 3 per cent interest. They would start to repay this loan from the sixth year in ten annual instalments.

If the landlord and the peasant agreed on the third method, they had to divide the land between themselves in the same proportion as they traditionally shared the crop. The price of the land was assessed at the highest local coefficient and the peasant had to pay only two-fifths of this price to the landlord.[25] The peasant received a loan to pay one-third of the price and the remainder would be paid in ten annual instalments. The rules for the loans which the peasant received were similar to those under the second option.[26]

The fourth option could only be chosen if the majority of the peasants and landowner(s) in a village agreed on forming an agricultural unit in the village. The unit was to be run by a committee which consisted of three persons: the representative of the landowner(s), the representative of the peasants, and a third person chosen by both the landowner(s) and the peasants. The members of the agricultural unit were to share the proceeds in accordance with local custom (for further information, see Lambton 1969: 204–6).

The last option was in principle one of the provisions of the original law. Accordingly, the landlord could buy the cultivation right of a peasant and farm the land himself if the peasant agreed to

the sale. However, the landlord could only buy the peasant out if the land held by the landowner did not exceed a specified maximum limit. This maximum limit was determined by the legislation for different regions and varied between 20 and 150 hectares (Lambton 1969: 206–7).

According to the provisions of the law, landlords could choose the first alternative regardless of peasant wishes. But they could only choose other options if they were able to reach agreement with the peasants. The land reform officials were not supposed to influence the choice of the two parties, but when they agreed on a method, the officials had to be satisfied that the decision was made freely (Lambton 1969: 207–8).

The third stage of land reform

The third stage of land reform began with the implementation of the Law for the Distribution and Sale of Rented Farms in the spring of 1969.[27] It has been argued that this stage was introduced because, (1) the tenants were deeply dissatisfied and there were growing tensions between them and landlords; and (2) some policy makers believed that tenancy inhibited modernization and production growth (see Denman 1973: 143–4; Momeni 1980: 239–40; and Hooglund 1982: 67–9).

This third stage stipulated that the landlords who had rented their lands or formed agricultural units with the peasants were required either to sell their lands to the tenants or to divide the lands between themselves and the tenants. In the latter case, the landowner and the tenant had to divide the land between themselves in the same proportion as they traditionally shared the crop.

When the landlord wanted to sell the land, the price was fixed at ten times the annual rent for cash. The tenant might pay the purchase price in twelve equal annual instalments; in this case the price was twelve times the rent. The cash price of the divided land was equivalent to two-fifths of ten years of rent, minus 15 per cent.[28] The landowners were reimbursed by the government if the peasants defaulted and they could also receive the value of each instalment from the government three months before their due date (Momeni 1980: 240; Hooglund 1982: 69).

Charitable *vaqfs* were exempted from the provisions of this law. Theoretically, *vaqfs* cannot be dissolved, and therefore, they could not be sold to the occupying tenants. But, the civil code permitted the sale of *vaqfs* if the value was enhanced by sale and was then used for the purchase of another property.[29] Accordingly, a law was passed in April 1971 which permitted the sale of these lands to

the tenants in accordance with this article of the civil code (Momeni 1980: 241). It was declared that 22 September 1971 was the end of 'one decade of land reform programme'. The lands which were not settled by that date were considered to be sold to the tenants and the rent they paid would be regarded as instalments of the purchase price (King 1977: 409).

Analysis of the land reform legislation, and the government's agrarian attitude and policies

Having reveiwed the Iranian land reforms laws, we now want to examine them in order to see whether they could benefit the rural poor. We have already concluded, in the previous chapter, that a land reform would help the majority of the rural people, and would also contribute to economic development, if it could substitute a unimodal agrarian system for the traditional system. Thus, our main object in this section is to find out whether the laws served the formation of a unimodal system or whether they were 'bimodal-oriented'.

In spite of its shortcomings, the original law of 1962 created favourable conditions for a rapid redistribution of a large part of the land of the large landowners among the peasants. First, specifying one village as the upper limit of landownership and redistributing the land among the occupying peasants enabled the rapid implementation of its provisions (Lambton 1969: 64; Warriner 1969: 115). The landlords who would be expropriated and the peasants who would receive the land could be easily specified under the prevailing agrarian system. If the law limited landownership and distributed the land among the peasants in a way which required a cadastral survey and land measurement, it would certainly become a dead letter.[30] Second, we saw earlier that there was a high degree of land concentration in Iran before the reform, and that mechanization was not common (see Chapter 3, pp. 30–1, 40). Accordingly, a large proportion of the agricultural land would have been subject to redistribution. There is a likelihood that more than 30 per cent of the villages were to be redistributed to the peasants during the first stage (cf. Momeni 1980: 248–9). Third, the purchase price might be favourable for the peasants if we consider the fact that in practice landlords paid very low taxes (see Momeni 1980: 230; Hooglund 1982: 53–4).

On the other hand, the first stage was to give land principally to those who cultivated it and the distribution was to occur on the basis of the prevailing village lay-out. In the first place, this would imply that the landless labourers, i.e. about one-third of the rural

population, might not receive any land in practice. Furthermore, cultivation rights were not distributed evenly and giving lands to peasants on the basis of these rights would perpetuate the existing inequalities in the villages.

There were also other factors which could limit land redistribution. For example, the land reform, by allowing the landlord to choose which village he would retain, was likely to deprive the peasants of the best lands. This was a serious problem especially when the landlords decided to choose their six-*dangs* from different villages because (i) the landlord was likely to choose the best land in the village; and (ii) he would retain his influence in a large area. Moreover, imprecise definition of a 'six-*dang* village' or lack of a definition of 'mechanized lands' could be used by landlords to evade the law. For example, they could claim that several villages were only one single village and or they could employ a tractor in a village and exclude it from the reform under the pretext of 'mechanized' farming.

As was pointed out earlier, however, the first stage contained pro-peasant elements. Moreover, the anti-landlord political climate which prevailed in the early period of the land reform strengthened the pro-peasant authorities in the government and enabled them to seek ways to limit the landlords' manoeuvres. The landlords were prohibited from dividing their whole villages, the provision of the law which allowed only the head of a landlord family to retain one whole village was confirmed, the definition of the village became more precise, and there were also some attempts to specify and limit the meaning of the 'mechanized lands'.[31]

Although large areas may have remained in the hands of large landowners who owned one or part of one village, we cannot conclude that the agrarian system after the first stage would be bimodal. A bimodal system consists of large and small farmers who compete for scarce resources on unequal terms (see Chapter 4, pp. 54–7). To create such a system in Iran, the landlords had to be able to take over their lands which had been rented out in order to form large and medium-size farms under their own management. The original law generally did not permit such a transformation. According to the law, the retained lands were to remain at the disposal of the occupying cultivators; therefore, they could not become large farms.

A bimodal system was likely to emerge, however, if the landlords could illegally evict the peasants and begin to cultivate their land with hired labour. In fact, lack of peasant participation in the enforcement of the legislation and shortage of trained officials were two factors that could result in slow implementation and

evasion of the law. Peasant participation would ensure the effective implementation of the law since peasants knew the local conditions better, they could cope with evasive practices of landlords and reduce the work of the officials.[32]

The necessity of executing the law in one district at a time afforded landowners the opportunity to recover from the initial shock of the reform (Hooglund 1982: 56).[33] The announcement of the draft legislation for the second stage in January 1963 indicated that the opponents of the land reform had finally managed to bring the reform under control. The Additional Articles did not force the landlords to sell their lands but gave them several alternatives and, in practice, they were able to choose whichever alternative they liked best. This seems to have been a pro-landlord move. The subsequent events showed that this was in fact the turning point in the direction of the land reform programme, i.e. a shift from the peasants to the landlords. From then on, two bimodal-biased tendencies emerged in government policies. First, there was a quick move to limit the distribution of land among peasants (cf. Katouzian 1974: 228). Second, landowners, and others, were urged to establish large-scale, modern farms. Both trends are reflected in the changes in some of the regulations of the first stage, the introduction of the second and third stage, and other government agricultural policies.

If we consider the decisions of the Land Reform Council after 1963, we can see that the earlier pro-peasant decisions were replaced by pro-landowner ones. Accordingly, the wife and each of the children of landowners became entitled to one whole village. In other words, a typical family of five could then retain five villages instead of the single village which had previously formed the upper limit. Furthermore, the landlords were permitted to choose parts of different villages as well as villages which had grown and were virtually small towns.

The second stage of the land reform might seem to be in line with a tenancy reform. In reality, the Additional Articles could hardly improve the position of the peasantry. The law in general favoured the landowners at the expense of the peasants. Moreover, there was a long delay between the announcement of the bill, its passing by parliament, and its implementation. The landowners, who had experienced the first stage, were given sufficient time to enable them to take measures to avoid its effects. They could sell or transfer their lands to others, or they could 'mechanize' their lands by ploughing them with a tractor. None of these measures were illegal according to the Additional Articles (Momeni 1980: 236–7).[34] We now turn to the legislation related to the second stage and examine its provisions more closely.

The first method, i.e. tenancy, had certain disadvantages for the peasants. First, the rent, which was the average income of three years, meant that the tenants would pay the same rent as they paid earlier while they were now responsible for all the costs. This was especially important since the landlords' share was normally higher where they provided other services. In such cases, tenants must put up with both higher rents and higher costs. Second, the rent could be revised every five years which would allow the landlord to capture part of a possible production increase of the farms without contributing to the costs or taking any risks. Third, peasants could survive a crop failure under the sharecropping system since they were not required to pay any rent and, moreover, they often received advances from the landlords. In the new system, they both lost this occasional help and had to pay the rent under all circumstances (Momeni 1980: 253). Finally, a bad year might even lead to a peasant's eviction since the law permitted a landowner to abrogate the lease if the tenant fell more than three months behind on payment of the rent (Hooglund 1982: 67). Thus, even security of tenancy, which was apparently the only advantage of the new system, was not seriously guaranteed.

Even selling the land to peasants, the second option, was less favourable to peasants than the first stage. The price of the land and the arrangement to pay the greater part of it were to be determined mutually between landlords and peasants. Given the bargaining power of the two sides, the peasants who received land in the second stage were likely to pay higher prices than those who had received land in the first stage.

The third method, which allowed division of land between landlords and peasants, might benefit the peasants if they had received a large share of the crop. However, the landlord would continue to exert his influence over the village since his share of the whole village could provide him with a large farm. The landlord would, in all probability, also occupy the best lands too. The situation of the peasants in the villages where the sharecropping agreement had been unfavourable to them was likely to be aggravated immediately if this option was chosen. These peasants had to eke out their livelihood on very small, inferior farms. At the same time, they had to pay the instalments for the purchase of the land.

The last two options were also in favour of the landlords. If an agricultural unit was formed in a village, the sharecroppers would farm the land and share the crop with the landlord as before. Thus the agricultural unit was simply a device for preserving the traditional distribution of income from land (Warriner 1969: 124; see also Lambton 1969: 205; Momeni 1980: 253). Permitting landlords

to buy peasants out was clearly in contradiction of the main aim of a land reform programme, i.e. land redistribution among the peasants. Interestingly enough, this option also existed in the original law but landlords had not been allowed to use it (see p. 74).

The legislation concerning the third stage of land reform was in fact a shortened version of the previous law. The landlords were to choose either option 2 or option 3 of the five possible methods which were open to them according to the law of the second stage. Considering the later step in selling the *vaqf* lands to the peasants, absentee landlordism must have gone after the reform. Those who owned land were large and small farmers who were to engage actively in cultivation.

The above achievement, by itself, might not be promising from the perspective of the peasants. On the contrary, peasants would be worse off if they had many large farmers beside them. Indeed, this was likely to be the ultimate outcome of the whole Iranian land reform programme. The spirit of the original law was pro-peasant and it initiated a unimodal trend. However, the first stage was not an end in itself. The land reform left the traditional system unaffected in many places.[35] The future of these lands would have a decisive role in the development of the coming system. If such lands were also distributed among peasants, the new system would probably be unimodal. On the other hand, the new system was likely to be unimodal if the landowners could take over all or most of these lands. The subsequent steps of the Iranian reform programme were in line with the latter. Moreover, they even weakened the initial policy by allowing the landowners to retain more lands. The process of moving away from a peasant-oriented system was accelerated further through the changes in the government's attitude toward peasants as well as its agricultural policies. These latter conclusions remain to be explained in detail in the following sections.

The government attitude toward land reform

Broadly speaking, governments usually have two important tasks in a land reform programme. First, they draft land reform legislation and the regulations for their implementation and, second, they are responsible for the implementation of the law.[36] The attitudes of governments towards land reform will thus play a decisive role in the ultimate outcome of the reforms. This is especially true when a government, such as the Iranian one, undertakes full responsibility for formulating and implementing the land reform and denies peasant participation.

As was discussed earlier, the government which took office in 1961 was determined to carry out cetain socio-economic reforms (see p. 72). Land reform was a key ingredient in this programme. Ali Amini, the Prime Minister, considered land reform to be one of its major tasks (Momeni 1980: 204–5). He appointed Hassan Arsanjani as the Minister of Agriculture who was to be in charge of implementation of the land reform. Arsanjani had advocated land reform since 1945 and was known to have radical views on the subject (Lambton 1969: 61–3; Zonis 1971: 54; Hooglund 1982: 42–3, 50–3). Not all the members of the cabinet shared Arsanjani's view and some of them were even against them (Land 1965/6: 97–8, 101). The year 1961, however, provided an opportunity for the advocates of more radical land reform and thus the land reform law of 1962 could be introduced.

When the implementation of the first stage began, the propaganda campaign against the landlords was intensified (Land 1965/6: 101–3; Momeni 1980: 270–5). The government sided with the peasantry and the prevailing impression was that the reform would be extended so that all occupying peasants would receive the land that they were cultivating (see Momeni 1980: 270–5). The principle that land should belong to the tiller was never explicitly stated in the law, but it was announced. Posters with this slogan were put up in the land reform officers in the capital and the provinces (Warriner 1969: 120; Lambton 1969: 100).

The implementation of the law was the responsibility of the Land Reform Organization which was set up under the land reform law. This was a new organization. In a country such as Iran, where many landlords had high government positions and enjoyed great influence in existing institutions, this was an advantage for the reform.[37] Many of the officials in this organization were young and were dedicated to their work. In the early stages they generally enjoyed the confidence of the peasants while they were not welcomed by the landlords (Lambton 1969: 99, 122–3).

Shortly after the beginning of the land reform programme, the cabinet resigned. The new government was less prepared to support the programme in its original form.[38] Arsanjani, however, continued in office. While the land reform was progressing, its opponents could organize and put strong pressure on the government. They mainly used the growing tensions between landlords and tenants in villages not affected by the reform as evidence that land reform was not planned properly (Hooglund 1982: 56). The government showed increasing concern about stabilization in the countryside. The second stage in fact aimed at this goal. Soon after the proclamation of a bill for the second stage, the government was

reconstructed and new ministers, who were probably opposed to the land reform, were appointed (see Zonis 1971: 59; Momeni 1980: 284). A high-level economic conference that was subsequently arranged also advocated law and order in the villages as a means for encouraging investment in agriculture (Momeni 1980: 285). The replacement of Arsanjani by an army officer might be considered as a sign that the government was no longer supporting a 'pro-peasant' land reform.[39]

The same change of attitude had also been underway among the land reform officials. The number of these officials had been increasing considerably and there were many officials who were temporarily transferred from other departments of the Ministry of Agriculture (Lambton 1969: 122).[40] The sense of dedication among the officials decreased as their number grew (Lambton 1969: 123). In part, this might have been the result of the extension of landlord influence over this organization in the course of time as well as the arrival of some pro-landlord sympathies from the transferred officials. There were other reasons as well. Security forces, which together with all other military and civilian officials were supposed to help the agricultural ministry to execute the land reform, began to disrupt the work of land reform officials (see Land 1965/6: 107; Zonis 1971: 59). Moreover, the new orders and instructions, which contradicted the previous pro-peasantry ones, disappointed many officials who were working enthusiastically. They were either dismissed or resigned (Momeni 1980: 286).

These changes would naturally slow down and limit land redistribution in those districts where the first stage still remained to be implemented. Moreover, they would further the landlords' benefits from the pro-landlord laws of the subsequent stages.

Agricultural policies and co-operatives

As the government lost interest in land reform as such, it concerned itself more and more with production growth in the agricultural sector. It was believed that modernization, especially mechanization, was necessary to increase agricultural production. This was accompanied by the growing influence of the school which maintained that small peasants lacked capital and experience to increase production. They suggested that the government had to encourage and support men with capital to invest in mechanized farming at the expense of the peasantry (Lambton 1969: 346).[41] In the course of time, this became the core of Iranian agricultural policies. Many agricultural policies such as 'The Agricultural Development Fund', 'Farm Corporations', 'Protection and Utilization of Forests and

Natural Pastures', and 'Establishment of Companies for Development of Lands Downstream of Dams' which were introduced in the subsequent years all urged the creation of large-scale commercial farms (cf. Momeni 1980: 244; Katouzian 1983: 319).

The Agricultural Development Fund was established in 1968 to encourage private investment in barren and dead land. The fund was mainly intended to support the large-scale projects (Denman 1973: 264).[42] The law of 'Protection and Utilization of Forests and Natural Pastures' would in general have similar consequences.[43] This law permitted the government to rent or sell forests and natural pastures in the north of the country to individual persons or companies that wanted to turn them into (large) agricultural enterprises. Some of the machinery they would import was exempted from import taxes (Momeni 1980: 245).

New legislation covering the 'Formation of Farm Corporations' was passed in 1968. Some of the main objectives of the corporations, according to this law, were: (i) to increase the income of the farmers, (ii) to create large-scale mechanized farms, and (iii) to prevent the fragmentation of holdings (see further Freivalds 1972: 185; Richards 1975: 10). Farm corporations were to be established by pooling all the lands of several villages. All farmers in these villages were eligible to be the members of the corporations. This meant that the membership was open to landowners, small peasants, and tenants. A member gave up his land rights permanently and received shares in the corporation mainly in proportion to these rights. He might or might not be employed in the corporations. Employment depended on the need of the corporation.[44]

The government selected the suitable regions for farm corporations and asked the would-be members if they agreed to the setting up of a corporation. The farm would be established when the majority accepted the proposal.[45] Once the formation of corporation was approved, the government would nominate its manager. The government tended to provide generous facilities for the corporation.

A shareholder could sell his shares to other shareholders. Considering the fact that a shareholder was not guaranteed employment in the corporation, small shareholders without employment were likely to sell their shares since the dividend might not meet their ends. Moreover, the general policy of the corporations was to encourage the small shareholders to sell their shares to others or to the corporation (Momeni 1980: 246; Brun and Dumont 1978: 18; Mohammadi 1973: 36). It was then possible that a peasant who had recently received a plot of land of his own sold it back to his former landlord.

The law governing 'Establishment of Companies for Development of Lands Downstream of Dams', passed in 1968, allowed the government to establish very large holdings on lands irrigated by dams. Each holding, known as an 'agribusiness', was to be between 5,000 and 20,000 hectares. Both foreign and domestic investors could receive lands to set up 'agribusinesses'. The government was to provide every incentive to encourage investors, e.g. tax relief, long-term loans, and cost sharing of major construction projects (see Hooglund 1982: 84; Azkia 1986: 225). If there were peasants on the lands involved, the government would buy these peasants out. Considering the huge size of agribusinesses, their development was likely to dispossess thousands of peasants. Thus, the spirit of 'corporations' and 'agribusinesses' was virtually the negation of the original purpose of the land reform programme (Hooglund 1982: 84).

While the government increasingly supported large-scale farming, it provided only limited help for small peasants. This help was supposed to reach the peasants through rural co-operative societies. The stated objectives of these societies were, among others, to provide credit, to buy the agricultural produce of its members; to help the members with storage, transport, and sale of their produce and to provide its members with agricultural inputs and machinery such as fertilizers, pesticides, and tractors (Lambton 1969: 293–4; Momeni 1980: 242–3).

In practice, such multi-purpose societies were few and most of the societies only provided credit for their members. This was of course an important function and would have helped the peasants considerably if it had been worked properly.[46] But the government, which was providing the resources for loans, tended to restrict the total amount of credit to the societies (Hooglund 1982: 109). Moreover, some of the terms of the loans limited further the flow of available credit. For example, new loans were paid to the members of a society only if all the members repaid their previous loans. If one of them defaulted, none of them would receive a new loan (Hooglund 1982: 109–10; Lambton 1969: 314).

The practice of providing only short-term loans to members was also detrimental to the peasants. The members could receive loans for a maximum term of twelve months. This would naturally discourage the peasants from using the loans for productive investment which yielded over a longer period of time. Such investments meant that the peasant had to repay the whole loan from his coming harvest. This was difficult and therefore many peasants preferred to use the co-operative loans for annually recurring agricultural expenses (Hooglund 1982: 110).

On the whole, the general policy of the government tended to limit the scope for co-operative activities. For example, in the first stage of land reform, membership of rural co-operatives was a precondition for receiving land. The policy was to set up separate societies for two to four villages. Each society was to cover between 50 to 300 members (see Hooglund 1982: 106; Lambton 1969: 304). In 1972, the government decided to consolidate the societies so that each society would have more than 1,000 members and would be located in the largest village. This meant that some villages lost their functioning societies (Hooglund 1982: 106). Considering the dispersion of villages and communication problems, many members might in this way have been deprived in practice from the services which they received from their societies.

Concluding remarks

The Iranian land reform programme began in 1962 under both internal and external pressures. This programme passed through three different stages during the 1960s. During the first stage, the reform was pro-peasant although it suffered from certain shortcomings. The reform aimed at limiting large landownership to one village and redistributing the villages in excess of this upper limit among the occupying peasants as quickly as possible. The relatively radical spirit of the law was generally supported by the government. Regulations for the execution of the law and interpretation of its provision were directed towards the peasants and most of the land reform officials did their job enthusiastically. There was considerable hope that a subsequent stage would impose greater restrictions on the landlords and provide more land for the sharecroppers.

However, the landlords were quickly able to regain sufficient strength to control the reform. The introduction of the second stage of land reform, which was the result of a new political situation, reversed the initial peasant-biased policy. In theory, the second stage extended the 'reform' to the retained lands and the lands which were unaffected by the first stage because their owners had one or less than one village. But, in reality, the new law largely secured the property rights of the owners of these lands and encouraged them to abandon the sharecropping system. They were even permitted to buy their tenants out. At the same time, the legislation for the first stage was gradually adjusted so that the upper limit, i.e. one village, was increased. The landowners could now retain more land. These changes, which indicated the pro-landlord attitude of the decision-makers, were accompanied by the weakening

of the pro-peasant sympathies among land reform officials. The result of this could increase the benefits of the 'reform' for the landlord beyond the terms of the law.

Meanwhile, the encouragement of the establishment of large-scale modern farms became an important element in the government's general agricultural policy. This was partly due to the belief that only such farms were able to increase agricultural production. On the other hand, small peasants were faced by a growing coolness on the part of the government.

On the whole, many factors seem to have been at work to create a bimodal agrarian system out of the Iranian land reform. Landlords were likely to take over a large, and more fertile, part of their land and cultivate it with the help of agricultural labourers. Some sharecroppers would receive lands of varying, but generally small, size. Some others might be evicted legally, while many more illegally because the prevailing circumstances permitted landlords to engage in such practices. Finally, the third stage of land reform was introduced to force those landlords who still hesitated to give up their old habits either to farm part of their lands themselves or to sell their lands. Legally there was no place for any kind of tenancy.

Chapter six

The agrarian system after the land reform

In Chapter 4, we argued that a distributive land reform may transform a traditional agrarian system into either a 'unimodal' or a 'bimodal' agrarian system. We saw that the economic consequences of these two systems were very different. For example, while the majority of farmers are likely to improve their levels of living in a unimodal system, they would not only fail to do so in a bimodal system but may also experience a fall in their standard of living. This implies that the outcome of a land reform will depend on the kind of agrarian system which develops as a result of the reform. Therefore, we examined the Iranian land reform programme in the previous chapter to see whether it aimed at creating a bimodal or a unimodal agrarian system.

The analysis of the Iranian land reform programme showed that it allowed a rapid redistribution of a relatively large proportion of land in its early stage. This situation, however, lasted only about one year, i.e. until the introduction of the second stage of the land reform programme. From then on, there was a general tendency in the new laws and regulations to limit redistribution and, at the same time, to convince the landlords that they had to give up their traditional form of land use, such as sharecropping. Consequently, it is reasonable to believe that the reform, reinforced by numerous government policies, adopted a bimodal direction. In other words, the post-reform agrarian system in Iran would consist mainly of very small and very large farmers.

In the present chapter, we will seek to present empirical evidence to back up the above conclusions. We will first explore the implementation process of the land reform programme in its three stages especially to find how much land was redistributed among the peasants. As a matter of fact, very few statistics are available with regard to the redistributed land and, therefore, we try to arrive at some reasonable estimates.

On the basis of the above background, we will single out the

major characteristics of the post-reform agrarian system in the rest of the chapter. This will pave the way toward an examination of the economic consequences of the reform and its ultimate impact on rural poverty and inequality in the next two chapters.

The implementation of the Iranian land reform

The land reform law of 1962 was put into operation in Maragheh in Eastern Azarbayjan only one week after it was approved by the cabinet.[1] The Government's ambition was to complete the implementation of this law by September 1963. But this plan was not realized and the purchase and distribution of the land which was subject to this law continued even after the introduction and execution of the other stages. These other two stages of land reform were implemented between 1965 and the early 1970s. The whole land reform programme was officially terminated on 22 September 1971. However, the implementation of the reform was not really completed by the date and continued even after 1971.

The amount of land transferred to the peasants during this reform period had obviously a decisive role in forming the new Iranian agrarian system. Since each stage of the land reform had a certain effect on the distribution of land, we will examine the outcome of each stage separately. In this way, we can follow the development of the new system and we will be able to see how powerful large owners managed to save their landed properties and change the direction of the land reform programme.

The first stage of land reform

On the eve of 16 January 1962, when notification was given to the landowners of Maragheh to declare their land, the Minister of Agriculture stated that Maragheh was chosen because its landlords were of the worst type in the country (Land 1965/6: 100). He also said: 'If we can prove that land reform will work in Maragheh, then we have proved that it will work anywhere else' (Land 1965/6: 100). But there might be more plausible explanations. The peasants in this area were economically better off than in many parts of the country. They were good farmers who could immediately benefit from the land reform and exemplify its success. Moreover, the landowners did not have much political power and their position in relation to their peasants had already been weakened. They were likely not to oppose the land reform as hard as those landowners who still enjoyed considerable political power (Lambton 1969: 87–8).[2] On the whole, Maragheh was a 'pilot project' for the

government in every respect. The difficulties of enforcing the reform were discovered here. Suitable solutions were then found and passed as legal decrees (Salour 1963: 55–7).

Those authorities who supported the land reform were anxious to begin land redistribution as soon as possible. Thus, when one landowner volunteered to sell her land to the government before the legal deadline, the transfer deed was signed at midnight on 1 March 1962 (Momeni 1980: 271; Land 1965/6: 103).[3] On 12 March, the Minister of Agriculture stated that 99 per cent of the Maragheh landowners had submitted their declarations and the first series of title deeds was handed over to the peasants the next day (Land 1965/6: 103). The implementation of the law was quickly extended to other areas. By the end of 1963 the first stage of land reform had already begun in about twenty-one regions (see Table 6.2, p. 91).

Table 6.1 Progress made in obtaining villages subject to redistribution in the first stage, 1962–75

Period	*Number of villages purchased*	*% of total villages purchased*
Jan. 1962–Sept. 1963	8,042	48
Nov. 1963–March 1971	8,109	48
Apr. 1971–Dec. 1975	666	4
Total	16,817	100

Sources: Salour (1963: 58); Bank Markazi Iran, *Annual Report and Balance Sheet* (1965–72); Denman (1978: 266)

The Land Reform Organization made good progress in implementing the land reform up to September 1963. As we saw in the previous chapter, the landlords had succeeded in controlling the reform by this time. Consequently, further implementation of the law was considerably slowed down. For example, in the first half of 1963 the number of villages which were purchased was 4,337 while in the second half of the year only 665 villages were purchased. Table 6.1 shows that the total number of villages which were obtained for redistribution in the first eighteen months of the reform were approximately equal to those which were purchased during the next nine years.[4] This table also reveals that the implementation of the first stage was not completed even in 1971, i.e. when the land reform was officially announced to be finished. It should be noted that there was always an interval between the purchase and redistribution of the eligible villages. Consequently, the first stage was still not completed in 1975; by 1976 some fifty-two purchased villages remained outstanding (Denman 1978: 266).

Table 6.2 Villages purchased in different regions in the first stage up to September 1963

Region	*Total villages*	*Villages purchased*	*Percentage purchase in the region*
1. East Azarbayjan	3,960	1,184	30
1. West Azarbayjan	2,710	285	11
2. Khorasan[a]	6,716	1,133	17
3. Gilan and Shahsavar	2,284	305	14
3. Mazandaran	2,331	261	11
4. Hamadan	665	196	30
4. Kermanshah	2,765	631	23
4. Kordestan	1,681	325	20
4. Salas Shahrestan	599	129	22
4. Zanjan	971	399	41
5. Khuzestan	1,825	459	25
6. Fars	2,779	989	36
7. Arak	479	291	61
7. Central Regions	2,311	337	15
7. Esfahan	1,709	284	17
7. Qazvin	820	159	20
7. Tehran	4,052	240	6
8. Oman Coast	754	16	2
8. Persian Gulf Coast	451	58	13
9. Kerman	5,457	359	7
10. Sistan and Baluchestan	1,283	2	1
10. Yazd	2,000		—
Total	48,592	8,042	16.5

Sources: Research Group (1964); Lambton (1953: 270–1)

Note: a In fact only northern Khorasan belongs to this group while southern Khorasan comes between groups 7 and 8

Some have argued that the first stage was carried out more quickly in regions where landowners were 'larger' since there were fewer landowners to deal with a greater number of villages could be handled in each transaction (Denman 1973: 116–17). Table 6.2 shows that such a statement is incorrect. In this table the names of different regions are listed in the order of their degree of land concentration: number one represents the highest. Those villages which

have been given the same number had approximately the same proportion of large landed properties.[5]

We can see that the best results were not achieved in the regions with the highest land concentration, i.e. regions 1, 2, and 3, but in the regions which had a lower land concentration, i.e. Arak (region 7), Zanjan (region 4), and Fars (region 6). In fact, except for East Azarbayjan, only a low percentage of the villages with the 'largest' landowners were acquired. The record of East Azarbayjan was not very high either if we consider the fact that it was here that the law was first implemented and continued to be applied during the most intense phase of the anti-landlord campaign. On the whole, it can be concluded that in the initial and most pro-peasant period of land reform, good progress was only made in the area with a 'medium' degree of land concentration and land fertility.

We should remember that the figures in Table 6.2 refer to the total number of whole and partially purchased villages which were bought in each region. It is estimated that only 37 per cent of these villages, namely about 3,000 villages, were entire, or six-*dang*, villages (Momeni 1980: 309). However, we do not know the proportion of six-*dang* villages in each region. This might limit the above conclusion since a larger proportion of the villages bought in the areas with 'larger' landlords might have been whole villages. We can, however, use Table 6.3 as a good approximation of the whole villages that were purchased in each region. Comparisons of those regions which appear in Tables 6.2 and 6.3 reveal that the proportion of the entire villages was generally not higher in the areas with higher land concentration.

It is interesting to note that between September 1963 and the end of 1975 only about 1,000 whole villages were bought while the number of villages partially purchased increased by about 6,000. In other words, the whole villages acquired for redistribution after September 1963 increased by 33 per cent while the partial villages increased by 120 per cent. This huge increase in the number of villages of less than six-*dang* is not surprising since in March 1963, the landlords were allowed to retain six *dangs* of different villages instead of one entire six-*dang* village (see Chapter 5, p. 73). In all probability, this encouraged many landlords to choose parts of several villages in order to retain the most fertile parts of their landed property as well as to maintain their influence in all these villages.[6]

There is plenty of evidence to indicate that landlords successfully used a variety of subterfuges to retain a large proportion of their land. One method was to transfer the land to relatives or others and to antedate the documents. As early as 20 January 1962 a daily

Table 6.3 Completion of the first stage as of December 1975

Region	*Villages purchased*		*Families who received land*
	6 dang	*less than 6* dang	
Bandar Abbas	3	13	219
Bushehr	49	132	6,111
Chahar Mahal Bakhtiari	6	102	5,978
East Azarbayjan	510	1,317	123,138
Esfahan	60	338	22,192
Fars	247	1,405	63,307
Gilan	57	690	41,213
Gorgan and Gonbad	2	21	5,990
Hamadan	189	648	80,906
Ilam	210	89	9,547
Kerman	97	454	6,219
Kermanshah	487	1,215	59,013
Khorasan	222	1,069	26,343
Khuzestan	256	452	22,464
Kohkilouyeh and Bouyer Ahmadi	335	43	12,413
Kordestan	166	511	32,145
Lorestan	99	449	16,602
Mazandaran	174	537	69,767
Semnan	10	59	698
Sistan and Baluchestan	1	3	45
Tehran	221	617	50,504
West Azarbayjan	422	435	34,681
Yazd	—	15	82
Zanjan	169	429	20,141
Total	3,992	11,043	709,718

Source: Denman (1978: 268–9)

newspaper, *Kayhan*, wrote that some of the landlords who had over forty to fifty villages had recently transferred part of them to their heirs or sold land to others and consequently reduced their landed properties to less than six villages (cited in Momeni 1980: 288; see also Hooglund 1980: 8). At the beginning of the first stage wives were not entitled to retain any land. Consequently, there were landlords who nominally divorced their wives and transferred part of their land to them (Mohammadi 1973: 32). Another common method used to evade the first stage was to declare falsely that the land had been mechanized and, therefore, exempted from the land reform.[7] In order to justify such a claim, some bought tractors for their farms or, as long as mechanized land was not defined, simply forced the sharecroppers to sign a wage list in order to show that they were labourers (Momeni 1980: 288–9; Platt 1970: 97).[8] After

the introduction of the draft of the Additional Articles and its simple definition of mechanized farms, there were landlords who ploughed once with a tractor, which could even be borrowed, and claimed that the cultivation was mechanized.[9] There were also landlords who took the best village away from the peasants and began to cultivate it with tractors.[10]

The exemption of orchards was another loophole which allowed landlords to evade the law. The landlords developed their orchards by appropriating more village land and water. The considerable increase in orchard products between 1960 and 1968 may partly reflect such an abuse of the law. The output of pistachio rose by 600 per cent, apples and pears by 600 per cent, and citrus fruits by 300 per cent (OIPFG 1973: 111; see also Aresvik 1976: 77). There were also landlords who got round the reform by including several villages in one registration unit and retaining them as their 'chosen village'. For instance, a big landlord in Fars declared ten villages as one six-*dang* village and kept all of them (OIPFG 1973: 43–4).[11]

The activities of landlords to retain a large proportion of their land was not limited to the above methods. They tried whatever ways the situation permitted. They bribed the reform officials, forced peasants to leave by bulldozing their houses, deceived peasants to sell their cultivation rights, etc.[12] It seems that many of the landlords' illegal practices were identified but little was done to protect the interests of the peasants, especially after 1963 (see Momeni 1980: 293–5; Lambton 1969: 257–62; Platt 1970: 97; OIPFG 1973: 116–20). Unfortunately, there is no data which can reveal the extent of the land that was retained in this way. However, the estimate given below for the total amount of land redistributed during the first stage suggests that the impact of such actions was far from being limited.

It is almost impossible to determine the amount of land which was distributed among the peasants and all we can find out is the number of the redistributed villages. First of all, Table 6.3 shows that 11,043 out of the total of 15,035 purchased villages, i.e. more than 70 per cent of them, were less than six-*dangs*. A less than six-*dang* village might be anything from a small fraction of the village to almost the whole village. It is therefore not clear how many entire villages were redistributed. Like others, we arbitrarily assume that each partial village is equivalent to half a village on average (see Research Group 1964: 140; Momeni 1980: 266; Katouzian 1974: 229). Accordingly, partially redistributed villages amount to 5,522 entire villages which implies that a total of 9,514 six-*dang* villages were redistributed among the peasants. In other words, about 16 per cent of all Iranian villages were purchased and

resold to the peasants under the first stage of the land reform.[13] It is also interesting to note that about 58 per cent of all these villages had been acquired by September 1963 which reflects, more realistically than Table 6.1, the speed of land reform in its earlier period.[14]

According to Table 6.3, a little less than 710,000 families received land in the first stage. This figure is certainly inflated by the practice of many-fold counting. A fieldwork study found out that when only a fraction of a village was sold by one landlord, the entire households of the village were registered as the beneficiaries of the reform and when another landlord sold another fraction of the village, once again, all households were counted to have benefited from the reform (Keddie 1968: 83). In view of this fact, we tried to obtain a more accurate figure by using the average number of families who lived in a village in each region. We ignored the fact that sometimes a large number of the villages were *khwushnishins*, who were practically excluded from land reform, and assumed that in each entire village all families received land.[15] Such a generous estimation shows that about 570,000 families, or 19 per cent of all rural families, might have received some land in the first stage.[16]

The size and quality of the land that the peasant families received varied greatly. This is not surprising since, among other things, the size of the villages differed, their soil quality varied, and the families' traditional shares of the village land were usually not equal. Some maldistribution also emerged as the result of loopholes and illegal practices. It seems that in the early period of land reform, the acquired land was redistributed more equally than the later period. In some cases the enthusiastic officials even gave land to the agricultural labourers (Hooglund 1982: 88–9; Momeni 1980: 316–17; Jones 1967: 114; Lambton 1969: 98).

The second stage of land reform

In contrast to the first stage, the second stage was put into operation long after the introduction of the first draft of its law. When the second stage began after a two-year delay, those landlords who wished to evade the provision of the law had already done so. The implementation of the second stage started in each region when the first stage had been completed and reform officials became available (Lambton 1969: 218).

Although the implementation of the second stage involved considerable office work, most of the work was completed in the first year (see Table 6.4). This, above all, may reflect the pro-landlord

Table 6.4 Implementation of the second stage, February 1965–February 1966

	Villages	*Hamlets*
Land subject to the second stage	52,533	15,166
Land in which the second stage completed	45,513	13,013
Completed cases as % of total	83	86

Source: Lambton (1969: 219)

nature of the second stage. It is likely that the landlords, in general, co-operated with the officials in implementing the law rather than resisting it. In practice, they chose any of the five alternatives which suited them best: they (1) rented their land to the peasants for thirty years; (2) sold the land to the peasants: (3) divided the land between themselves and the peasants; (4) formed agricultural units with the peasants, which implied the persistence of the traditional relationships; or (5) bought out the peasants.[17] The change in the attitude of the government toward the land reform in favour of the landlords certainly contributed to the rapid implementation of the second stage (see Chapter 5, pp. 81–3). The land reform officials now tended to stay in their offices in towns where much of the settlement work was now carried out, sometimes in the absence of the peasants (Lambton 1969: 217; Hooglund 1982: 65). It was also common to ask a team, which was sent to the villages without sufficient preparation, to settle a specific number of villages within one day. They then travelled to the villages which had received no advance notification, rounded up as many peasants as could be found, read the available options and insisted upon immediate decisions (Hooglund 1982: 65).[18]

The statistics concerning the progress of the second stage are puzzling and inconsistent. For example, the statistics for 1966 indicate that 15,024 landlords had bought the traditional right from their peasants. Two years later, the total number of such landlords during all three years was reported to be 7,852! (see Bank Markazi Iran, *Annual Report and Balance Sheet* 1966: 135; Bank Markazi Iran, *Bulletin* 1968). A similar reduction can also be observed in the total number of peasants who had sold their cultivation rights by 1966 and 1972 (see Denman 1973: tables 10 and 10a). The number of landlords and peasants affected by the second stage is another source of confusion in the statistics.[19]

Table 6.5 Peasants affected by different options of the second stage up to October 1972

Region	Tenancy (1)[a]		Sale (2)		Division (3)		Corporation (4)		Peasant right sale (5)	
	No.	% of HT[b]	No.	% of HT	No.	% of HT	No.	% of HT	No.	% of HT
Bandar Abbas	299	55	—	—	200	37	—	—	45	8
Bushehr	6,933	74	2,425	26	—	—	—	—	—	—
Chahar Mahal Bakhtiari	22,496	95	128	<1[c]	994	4	—	—	22	<1
East Azarbayjan	179,814	89	13,449	6	7,559	4	1,975	1	46	<1
Esfahan	51,343	68	4,252	6	18,532	25	613	1	464	1
Fars	77,825	44	3,422	2	77,121	44	15,988	9	1,188	1
Gilan	188,536	96	6,313	3	—	—	—	—	564	<1
Gorgan and Gonbad	43,135	90	4,761	10	—	—	—	—	—	—
Hamadan	52,197	72	2,300	3	15,797	22	1,668	2	98	<1
Ilam	24,128	100	—	—	—	—	—	—	—	—
Kerman	3,702	16	22	<1	1,865	8	5,778	25	11,687	51
Kermanshah	48,692	81	392	1	10,679	18	53	<1	185	<1
Khorasan	17,428	23	1,238	2	2,983	4	46,974	62	6,779	9
Khuzestan	79,262	96	24	<1	2,908	4	11	<1	—	—
Kohkilouyeh and Bouyer Ahmadi	32,497	100	—	—	127	<1	—	—	—	—
Kordestan	56,691	93	2,113	3	2,126	3	—	—	97	<1
Lorestan	49,660	87	1,050	2	2,363	4	3,850	7	104	<1
Mazandaran	132,604	99	327	<1	330	<1	—	—	164	<1
Semnan	2,294	94	102	4	—	—	—	—	33	1
Sistan and Baluchestan	254	17	213	14	483	32	22	1	525	35
Tehran	85,203	84	6,101	6	9,136	9	—	—	1,241	1
West Azarbayjan	48,007	79	5,963	10	2,963	5	3,891	6	159	<1
Yazd	18,614	100	—	—	—	—	—	—	—	—
Zanjan	24,948	83	2,569	9	115	<1	2,444	8	—	—
Vertical total	1,246,652	—	57,164	—	156,279	—	83,267	—	23,401	—
% of grand total	80		4		10		5			

Sources: Denman (1973: 338–9); Lambton (1969: 221)

Notes: a Numbers in the parentheses refer to the different options of the second stage. b Horizontal total. c The percentage is less than one.

However, we can arrive at a few general conclusions irrespective of the available data that is examined. We have chosen Table 6.5 which shows the number of peasants affected as late as 1972. The last column of the table is corrected with the help of a similar table for 1966. We compared the total number of peasants who sold their cultivation rights and inserted the higher figure in Table 6.5. The rationale is simple: the number of such peasants could only increase or remain constant after 1966.

Table 6.5 shows that alternative (1), i.e. renting the land to the peasant, was predominantly preferred by the landlords in all regions except in Kerman (16 per cent), Sistan and Baluchistan (17 per cent), Khorasan (23 per cent) and Fars (44 per cent). For the country as a whole, about 80 per cent of the cases were settled by the tenancy arrangement. Although the landlords could maintain the traditional sharecropping system in the form of owner–peasant joint stock corporations, only a few chose to do so, except in Khorasan. These facts suggest that the tenancy option was highly in favour of the landlords and it was even more attractive than the prevailing system. The campaign conducted against the old agrarian system, and sharecropping, as one of its main elements at the beginning of the land reform, might have acclerated the switch over to the fixed-rent tenancy.

Division of the land between peasants and landlords which comprised only 10 per cent of all cases in the whole country, was the second popular choice among landlords (see Table 6.5). This option was widely chosen in Fars where 44 per cent of the peasants received a share of the land. Fars alone was responsible for about half of all cases which were settled by division in the whole country. The main reason for this was probably that the landlords generally received a larger share than the peasants in this region; in many cases, their share was three-quarters or four-fifths of the crop (Lambton 1969: 242). Where land was settled by division, landlords retained the most fertile part of it and sometimes occupied more land than they were actually entitled to since there was no cadastral survey and many land reform officials worked in favour of the landowners. A large number of the peasants who received land by division ended up with a plot of land that was too small to provide them with a living and were therefore dissatisfied (see Lambton 1969: 242–9; OIPFG 1973: 56–9).

A small number of the peasants, 4 per cent, received land from their landlords who selected option (2). On the other hand, about one per cent of the peasants lost their land as the landlords managed to 'persuade' them to sell their cultivation rights. The available evidence indicates that not all such peasants sold their

rights of their own free will. Some were forced by the landlords and some others were deceived by their landlords and signed the documents of sale without knowing what they were doing (see Lambton 1969: 253–5; OIPFG 1973: 60–3; Azkia 1986: 115–16). However, the great majority of the peasants who sold their rights, i.e. 79 per cent, were from Kerman and Khorasan.

According to the official statistics (see Table 6.5) 57,164 peasants bought the land that they had been cultivating and 156,279 peasants received only part of the land they used to cultivate. Thus, about 213,500 peasant families received some land in the second stage. Meanwhile, 23,401 peasants 'sold' their traditional rights and became landless. The number of peasants who became landless may certainly have been higher since there were landlords who deprived the peasants of their traditional rights by declaring their land mechanized or by other methods.[20] We should remember that a) more than 70 per cent of the peasants who received land obtained only small plots of land through the division option, and b) in addition to the likely exaggeration in the number of beneficiaries, which was revealed earlier in the case of the first stage, some of the beneficiaries of the second stage had already received land in the first stage and, therefore, the above number also contains some double-counting in this respect.[21]

We have no information about the amount of land which was distributed in the second stage. If we assume that a) the peasants who received land by division, obtained on average one-third of the land that they had been cultivating, and b) every fifty families who received land, represent one six-*dang* village, we find that the distributed land in the second stage was approximately equivalent to 2,185 entire villages. The first assumption is based on the fact, mentioned above, that land division occurred mainly in places where crop division was generally in favour of landowners, e.g. in Fars. The second assumption relies on the official data which show that the average population of the Iranian villages is fifty households. We have thus overlooked the fact that we could assume that seventy-five peasants represent one village according to the inflated official statistics (see p. 95).

In sum, about 4 per cent of all Iranian villages were distributed as a result of the second stage of land reform among 7 per cent of the rural families. This slow-down in the land redistribution programme underlines the pro-landlord nature of the second stage, especially in comparison with the initial period of the first stage.

The third stage of land reform

The second stage of land reform turned 1,246,652 sharecroppers into fixed-rent tenants while the situation of 83,267 peasants who were in the owner–peasant joint stock corporations remained unchanged (see Table 6.5). Moreover, 172,103 peasants who sharecropped the endowed land (*vaqf*) received 99-year tenancy contracts according to the provisions of the law of the second stage (see Hooglund 1982: 65, Table 3). The third stage of land reform which was introduced in January 1969 should be seen in relation to the deep sense of dissatisfaction of these 1.5 million peasants (see Hooglund 1982: 67–9 and Momeni 1980: 239–41). Accordingly, all the rented land, except *vaqf* land, as well as the land of owner–peasant corporations were either to be sold to the peasants or to be divided between the two parties in proportion to their traditional share of the crop.

Although the law was passed at the beginning of 1969, its implementation progressed very slowly. If we assume that all eligible landlords were those given for 1973 in Table 6.6, i.e. 342,000, only one-third of them had declared their intention to sell or divide their estates by May 1971 (see Denman 1973: 149, Table 12). Table 6.6 is convincing evidence that the total number of landlords who actually sold or divided their land must have been much lower. We can also see that as late as March 1972, i.e. six months after the reform had officially been terminated, only 45 per cent of all the eligible landlords had transferred their land, or part of it, to the peasants.

Table 6.6 Implementation of the third stage of land reform

	March 1972	*March 1973*
Sale:		
(i) Landlords who agreed to sell	281,470	305,000
(ii) Landlords who actually sold	144,219	229,000
(iii) Beneficiaries	564,203	842,000
Division:		
(i) Landlords who agreed to divide	41,221	37,000
(ii) Landlords who actually divided	10,129	24,000
(iii) Beneficiaries	32,031	75,000

Source: Bank Markazi Iran, *Annual Report and Balance Sheet* (1972–3)

Several reasons have been given for the sluggish progress of the final stage of land reform. Hooglund argues that the government did little to implement the third stage. The government avoided taking part in the process of transferring ownership or educating the peasants about the law and, instead, tried to persuade the owners to sell voluntarily to the peasants (Hooglund 1982: 69–71). Denman, on the other hand, maintains that landowners and their tenants preferred different methods of settlement and this resulted in a delay in the implementation of the third stage (Denman 1973: 145–6). This argument is based on the result of an inquiry about the preferences of landlords which was carried out at the beginning of the third stage. It showed that 88 per cent of the 222,360 who answered the inquiry preferred land division (Denman 1973: 145–6). However, such differences could not have had considerable effects on the progress of the third stage, according to the law, it was the landlords who were able to select the option they wanted. Moreover, the pro-landlord attitude of the government would have enabled the landlords to conclude 'agreements' with their tenants at their own convenience (see Chapter 5, pp. 81–3; also Hooglund 1982: 70).

However, we can see from Table 6.6 that there might have been a sudden increase in land reform activities after 1972. Other official statistics indicate that the implementation of the third stage must have been completed by 1974. Accordingly, all the small landowners covered by the law either sold their lands to 1,154,578 tenants or chose the division option and transferred part of their land to 110,347 farmers (Statistical Centre of Iran 1976: 212–13). This means that 1,264,925 peasants might have received land under the third stage.

There are grounds for questioning the reliability of these data. First, most eligible landowners seem to have preferred division to sale (see OIPFG 1973: 68 and Momeni 1982: 258).[22] The official statistics show the opposite. The figures for the last year might have also been inflated to take account of the wishes of the authorities who wanted to bring the land reform to a close.

An estimate must now be provided for the number of villages that were distributed during this stage. The estimate, however, cannot be based on 1,264,925 peasants who received land under the third stage since many of them were not first-time beneficiaries. Salmanzadeh and Jones (1979: 121) have shown that about 35 per cent of all beneficiaries of the third stage in 169 villages had already received land in the first stage. It is then likely that 842,283 farmers – two-thirds of the total – were first-time beneficiaries of this stage (cf. Majd 1987). Provided that, once more, fifty peasants represent

one entire village, and ignoring that in the division cases tenants received only part of the rented land, the redistributed land might have been equivalent to about 16,866 villages.[23]

This figure is a remarkable result because it constitutes about 28 per cent of all Iranian villages and is higher than the result of the first stage. However, a comparison between the number of affected landowners and peasants in Table 6.6 indicates that these villages were in general the property of small landowners. Moreover, a large part of the redistributed lands were probably in less fertile regions, since the landed properties of most small landowners were located in such regions.[24]

Apart from the lands redistributed under different stages of the reform, some land was also distributed among peasants as a result of the law of transfer of ownership of endowed lands (*vaqfs*) to tenants, which was passed in 1971. We will assume that all 172,103 tenants of these lands received the ownership of the tenanted farms (see p. 100).[25]

The distribution of land ownership and land tenure after the reform

Table 6.7 summarizes our discussion about the results of the land redistribution in Iran after 1962. In respect of the number of beneficiaries, the results of the reform appear to be very impressive: 92 per cent of eligible families received land after 1962. Moreover, redistribution of state domains and crown estates which had begun before 1962 gave another 157,000 peasants the ownership of 2,676 whole villages (4 per cent of all villages).[26] This implies that by the end of the land reform *all* sharecroppers and fixed-rent tenants must have become peasant proprietors.

In reality, however, progress was certainly not as impressive as on paper. We have already shown that the official figures are highly exaggerated. Furthermore, some cultivators 'sold' their cultivation rights during the second stage and some others also lost their rights in one way or another. Finally, there were tenants who continued to cultivate the landlords' land in the same manner as they had done before the reform.[27] All in all, we can conclude that land reform succeeded in giving land to a large majority of the eligible peasants while most of the remainder lost their rights and joined the landless labourers.

In sum, our relatively generous estimates indicate that about 60 per cent of all villages were redistributed among peasants. This, however, in no way means that 60 per cent of the agricultural land was transferred to the beneficiaries. Village lands varied greatly in

Table 6.7 Results of the Iranian land reform, 1962–75

	Beneficiaries		*Redistributed land (village equivalent)*	
	No. of families	*% of eligible families*	*No.*	*% of all villages*
First stage	709,718	34	9,514	16
Second stage	213,443	10	2,185	4
Third stage	843,283	40	16,866	28
Vaqf land	172,103	8	3,442	6
Total	1,938,547	92	32,007	53

Note: a This is assumed to be equal to the number of families with cultivation rights, i.e. approximately 70 per cent of rural families who were potentially eligible to receive land.

size and soil quality. The villages of the large landlords tended to be located in more fertile regions and they were usually the larger ones (see Chapter 3, pp. 30–1).[28] It was also the large landowners who might have retained a large proportion of their landed properties. Considering the fact that the large landowners were principally affected by the first stage, Table 6.7 suggests that most of the redistributed land might have belonged to small landowners (cf. Momeni 1980: 266 and Hooglund 1982: 79). It is thus plausible to believe that the 40 per cent of villages which remained in the hands of landowners were among the more fertile and larger ones. This implies that landowners, especially the large ones, might have owned at least half of the agricultural land after the reform.

It is not then surprising that a large majority of peasants received only small plots of land which in many cases were probably less than the holdings they used to cultivate. In some cases, for example, when part of a village was redistributed in the first stage, all the peasants who had cultivation rights on the entire land of the village shared that part and, therefore, ended up with smaller farms than before (OIPFG 1973: 45). This was also true for the land divided between peasants and landlords under the second and third stages. We can imagine the size of this group of peasants if we look at the numbers of such peasants in Tables 6.3, 6.5, and 6.6. At the same time, a minority received proportionately larger and more fertile farms. These people were generally former village-headmen and *buneh* leaders (Hooglund 1982: 89; Azkia 1986: 126–7; Lambton 1969: 177). Consequently, while many peasants received less than 3 hectares, few received as much as 50 hectares or even more. Different studies suggest that the holdings of the majority of the land reform beneficiaries might have been between 0.5 and 10 hectares.[29]

Most of the estimates on landownership after the reform are based on statistics which show peasant holdings, i.e. data similar to those in Table 6.8. For instance a study assumes, on the basis of personal observation, that all holdings over 20 hectares, which constituted about 50 per cent of the crop land, belonged to approximately 200,000 absentee landowners (Hooglund 1982: 78–9).[30] In view of the fact that some landowners had divided their land into holdings of different sizes and nominally transferred them to their relatives and friends, there must have been owners who possessed several holdings (see Bavandi 1982: 90; Hooglund 1982: 91, Table 5; Khosrovi 1981: 30). This implies that landownership was more skewed than can be inferred from such data. Some information which has come out recently reveals that there were landowners in Gorgan and Khorasan who had more than 10,000 hectares each. The names of nine landlords who owned between 2,000 and 9,000 hectares have also been published (*Ettela'at* 1980, cited in Azkia 1980: 64; *Ettela'at* 1986: 11).

The large, and medium, landowners who were comparatively very small in number still owned almost half of the agricultural land while the other half belonged to about 2 million farmers. A large majority of these farmers became owners as a result of the redistribution of part of the land of large landlords and a major part of the land of small landowners.

Some of the landowners continued to lease out their land, or part of it, for a fixed rent or a share of the crop even after the reform. However, the post-reform tenancy and sharecropping systems were not always similar to those of the pre-reform era. The tenants and sharecroppers did not have any special rights to the land. They cultivated the land for a specific period of time, e.g. for six months, and the contract might or might not be extended when it expired (see Momeni 1980: 342; Antoun 1981: 226). Moreover, not only the landlords but also small peasants might lease out their lands. Such small peasants earned their living as workers or construction labourers in the cities and therefore had to let their lands. The tenants could be well-to-do and competent farmers who came from other regions, especially from Yazd on the borders of the central desert. They cultivated summer crops and vegetables and might leave the village after the harvest (Azkia 1980: 113). Different types of sharecropping also developed after the reform.

Pump cultivation was one type of share tenancy which became common after the reform. In this sytem, tenants dug a well in non-irrigated, and sometimes uncultivated, land and installed water pumps on them. They used modern techniques to cultivate the land and paid one-tenth of the crop to the landowner (Ashraf and Banuazizi 1980: 38–9; Azkia 1980: 116–17).[31]

Official statistics show, for example, that 2 per cent of all peasants were tenants who cultivated about 3 per cent of the farmland around 1975 (Khosrovi 1981: 27–30). If sharecropping, tenant–owner cultivation, and other types of tenancies are taken into account, it could be said that the tenanted land might have accounted for about 10 per cent of the cultivated land (Bavandi 1982: 88; see also Khosrovi 1982: 25–43).

The above statistics indicate that the tenants were distributed among all different sizes of holdings but they were concentrated in the holdings which were less than five hectares (Khosrovi 1981: 30). This means that most tenants were small farmers and might have experienced little change after the reform. Case studies show that this was true for a lot of peasants who still sharecropped under the traditional conditions (Momeni 1980: 338–41).

In spite of the presence of the traditional and new types of tenancy after the reform, the significance of tenancy has considerably diminished. On the other hand, peasant proprietorship and large, modern estates which were uncommon before the reform became the predominant type of tenure as the result of the land reform.

Although most peasants owned the lands they cultivated, a large majority of them had to eke out their livelihood from a small plot of land. At the same time, most former landlords stopped renting their lands, which was common before the reform. The general tendency among them was to form large, and/or medium farms, under their direct or indirect management. These farms, although small in number, covered most of the cultivated land. This is clearly illustrated in Table 6.8. We should remember that the classification of all farms in different regions of Iran on the basis of the same size criteria may involve some overgeneralization. There is in fact a great deal of regional variation. For example, in Gilan, where agricultural conditions are very favourable and farms are normally small, farms of 20 to 30 hectares could be considered as large farms. On the other hand, in Sistan and Baluchestan with detrimental conditions even 10–20 hectare farms display the characteristics of small farms (cf. Khosrovi 1981: 11–12). However, such differences are within certain limits and do not materially affect the general picture which is provided by Table 6.8.

We should note that the farm categories shown in Table 6.8 do not necessarily correspond to different farmer groups, i.e. small, medium and large farmers. These were very likely large farmers who owned, for example, several medium-sized farms. In order to better understand the post-reform agrarian system and developments we should examine large and small farmers in more detail.

Table 6.8 Distribution of land holdings by size, 1974

Size of holding (hectares)	*Number*	*% of total number*	*Area (hectares)*	*% of total area*
Small				
<2	1,056,467	43	703,562	4
2– 5	541,592	22	1,732,868	11
5–10	427,934	17	2,953,447	18
Medium				
10–20	304,000	12	4,487,000	27
20–30	97,000	4	1,904,000	12
Large				
30–50	27,074	1	1,109,741	7
>50	25,822	1	3,526,603	21
Total	2,479,989	100	16,417,221	100

Source: Khosrovi (1981: 57; 1982: 13–14)

Large farmers

According to our classification, farms of 30 hectares or more should be considered as large farms. But there is not a general agreement about this. For example, some assume that farms of 50 or more hectares were large farms while others maintain that farms which were as large as 100 hectares or more were truly large-scale holdings (for instance, see Khosrovi 1981: 41–5; Azkia 1986: 120–1; Hooglund 1982: 82–3; and Aresvik 1976: 100–1). We have included 30–50 hectare farms among large farms for several reasons. First of all, we are interested in the land concentration phenomenon which causes market imperfections and this small group also possessed a proportionately high percentage of the land. Second, farms between 10 and 50 hectares are usually assumed to be medium farms whose owners are rich by Iranian standards (see, among others, Khosrovi 1982: 41–4 and Azkia 1986: 126–7). Accordingly, farms of 30–50 hectares land are generally considered to be upper-medium and their owners are among the richest peasants who enjoy more power than the majority. These farmers are then very close to those commonly called large owners and our assumption does not differ significantly from many others. Finally, there is evidence which shows that from the point of view of land and water, these farms were in a radically better position than those which were less than 30 hectares (Khosrovi 1982: 42).[32]

On the whole, large farms consisted of three types of farms: individual large farms, agribusinesses, and farm corporations. The first

kind of large farms were the dominant form of large-scale holdings and covered more than 85 per cent of all the land which belonged to such farms. It is interesting to note that large farms of 100 hectares and above comprised only 18 per cent of these farms but occupied more than 50 per cent of their land (see Khosrovi 1981: 57, Table 9). Although many large farmers were former large and medium landowners, other social groups could also be found among them; for example, merchants, military officers, former village headmen, etc. A large number of large holdings were concentrated in Gorgan, Khorasan and Khuzestan which are among the very fertile regions (Hooglund 1982: 83; Azkia 1986: 121). In fact, even other large farms were concentrated mainly in the more fertile regions of the country.[33] We should remember that some of the large farmers owned several holdings which were either large or small, or both. This implies that large farmers must have controlled much more land than large farms cover, i.e. 28 per cent as shown in Table 6.8.

Agribusinesses emerged as a result of the law of 1968 which encouraged the establishment of large and modern estates on the land irrigated by dams (see Chapter 5, pp. 83–5). By 1978, thirty-seven agribusinesses had been established on 238,000 hectares of irrigated land (Azkia 1980: 128). Fifteen of these farms were very large estates, i.e. between 5,000 to 25,000 hectares, and controlled by multinational companies, while the remaining twenty-two agribusinesses were smaller, i.e. 1,000–5,000 hectares, and belonged to Iranian shareholders (Azkia 1980: 128).[34] In some cases, the government forced the peasants to sell their land to make way for agribusinesses. For example, it is estimated that about 55,000 peasants must have 'sold' their land to the government which leased it for thirty years to eleven large agribusinesses. It is therefore plausible to believe that at least 75,000 peasants were displaced for the establishment of the thirty-seven agribusinesses (Azkia 1980: 129; Hooglund 1982: 85; Momeni 1980: 358).

The third type of large holdings were farm corporations. They were created in accordance with the 1967 law which governed their establishment. Only four months after the law was passed, the first corporation was established and the number of such farms increased to ninety-four by 1978. These corporations were spread over the whole country but they were mostly established in the more fertile regions and on lands downstream of dams. They tended to be close to towns which means that they usually enjoyed good communication facilities as well. Each farm corporation, on the average comprised between eight and ten villages and all ninety-four corporations encompassed 850 villages which covered about 400,000 hectares. These farm corporations had more than 35,000

shareholders and covered about 300,000 villagers (Denman 1973: 213–15; Momeni 1980: 359–63; Hooglund 1982: 86–7; Azkia 1986: 235).[35]

It might be argued that the farm corporations differed from other large farms in the sense that they were controlled by a large number of shareholders rather than a few people. Moreover, the shareholders consisted of both large landowners and peasants who had one vote each irrespective of the number of shares they hold.[36] In practice, however, it was principally the large shareholders who occupied the key positions in farm corporations, such as on the board of directors (see Azkia 1986: 237 and Momeni 1980: 366). Moreover, the managing directors and their expert assistants, who were nominated civil servants, conducted the corporations like large landowners. They seldom consulted even the board of directors (Momeni 1980: 366–7; Group of Experts 1980: 41–4). It is therefore not unreasonable to assume that the corporations were also large farms under the control of a handful of people.

Small farmers

On average, an Iranian farmer has to have about 7 hectares to be able to support his family (see Hooglund 1982: 77 and Momeni 1980: 70–1). This might be enough to explain why farmers with less than 10 hectares are widely assumed to be small farmers. Most of these farmers have received land under different stages of the land reform. Ignoring the fact that some of the small farms belonged to others and were rented by their cultivators, we can assume that the number of small farmers was approximately equivalent to that of the small farms shown in Table 6.8.

As Table 6.8 illustrates, the land distribution was very unequal among the small farmers. More than 50 per cent of all the small peasants possessed holdings which were less than 2 hectares, i.e. their farms were much smaller than the size which was necessary for basic subsistence (see Table 6.8). Even those peasants of this group whose lands were in Gilan where there was abundant rainfall could not rely solely on their farms for their living and had to work for others as well (see Khosrovi 1981: 141–3). It was in fact quite common that at least one of the family members of the small peasants worked for medium and large farmers or as casual labourers in the cities (see Khosrovi 1979: 81–3; Azkia 1986: 124; and Lambton 1969: 146).

A large majority of the peasants who owned between 2 and 5 hectares were normally in the same situation as the owners of the mini-holdings. Only a small number of such peasants who

cultivated rice, vegetables and summer crops were better off than the others (Khosrovi 1981: 34). In other words, from Table 6.8 we can say that 65 per cent of all farmers in Iran, or about 80 per cent of the small peasants, had to work more or less outside their farms in order to earn their living.

The smallness of farms was not the only problem which faced the small farmers. Irrigation difficulties and land fragmentation were two other factors which affected them adversely.

We know from Chapter 3 that before the reform peasant holdings consisted of several separate plots of varying quality. Since the reform transferred to the peasants the land that they were cultivating before the reform, their farms remained fragmented even after the reform. To the extent that the peasants received all the land of their village and its water resources, the reform did not aggravate the old water and land fragmentation problems. In some cases, such beneficiaries could even improve their irrigation systems immediately after the reform (Lambton 1969: 288).

The situation was, however, quite different in partially redistributed villages. We observed in the previous sections that the landlords retained the best lands and sold the poorest lands to the peasants. The small peasants had to reallocate their share of land which was again divided into small plots. Consequently, the lands of these peasants also consisted of scattered plots while their holdings diminished in size. Furthermore, the landowners controlled water resources in such villages and allocated a much larger proportion of the village water to their own large holdings. If we recall that most of the villages were only partially redistributed, it is reasonable to assume that irrigation and land fragmentation problems became more serious for a large majority of the peasants (see Khosrovi 1979b: 165–9; Hooglund 1982: 90–1).[37]

Concluding remarks

The Iranian land reform succeeded in redistributing, at best, about 50 per cent of the agricultural land among the farmers. A large amount of the redistributed land belonged to the small landowners. Large landed properties were in fact affected mainly by the first stage of the land reform, especially in the early phase of the reform when the law was enforced rather vigorously. However, even then, the large landowners evaded the law and took full advantage of its loopholes. They managed to retain a great part of their land. This was the best land with easy access to water sources in the majority of the affected villages. On the whole, a small group of people still controlled a large proportion of the land and water sources.

Large owners generally took over their land rather than leasing it to others, which was a common practice in the pre-reform era. There was an increasing tendency among this group to form large and/or medium-sized farms and cultivate them under their direct or indirect management. This was the main reason for the sharp increase in the number of large farms and the land they occupied. Part of the increase in the number and size of the large farms might have resulted from bringing new land under cultivation. At the same time, the number of very small farms grew rather quickly as well. However, the increase of the landholdings of such farms fell far behind the increase in their number. The transfer to the peasants of only part of their rights to land might have been the principal reason for this process.

Table 6.9 summarizes the changes which took place in the structure of the Iranian landholding system after the reform. It illustrates that although the total number of farms and the area under cultivation increased rather quickly after the land reform, it was the large farms which increased most of all. In fact, the largest farms tended to become both more numerous and larger. On the other hand, we can see from Table 6.9 that the number of small farms increased relative to their area. This implies that the small farms also grew in number but they generally became smaller.

Table 6.9 Growth of holdings and their area, 1960–74 (percentage)

Size (hectares)	*Number*	*Area*
<2	41	23
2– 5	14	12
5–10	26	22
10–20	36	49
20–50	60	36
>50	106	127
Total	32	45

Sources: FAO (1966: 96–7); Table 6.8

We can reasonably argue that the traditional agrarian system moved toward a bimodal system as the result of the reform. However, did the difference between the size of the land which was owned and cultivated by the large landowners and small peasants really spread to credit and labour markets? Was there any difference between the two types of farms with respect to their methods of cultivation, the crops they grew, etc.? We will try to answer these and other related questions in the next chapter.

Chapter seven

Economic consequences of the land reform

We saw in the previous chapter that the concentration of land in a few hands persisted in post-reform Iran. However, in contrast to pre-reform Iran, the large landowners tended to become 'large farmers'. That is to say, they increasingly formed large and medium-sized farms under their own management. A large majority of former tenants constituted the numerous group of 'small farmers' as they received some land under different stages of the land reform. In other words, it appears that the traditional tenancy system became transformed into a bimodal system after the reform.

The analysis of Chapter 4 concluded that large farmers in such a system not only have more land but also have easier access to capital. Small farmers, on the other hand, have relatively more labour at their disposal. Hence the conditions prevailing in the factor markets are not the same for all farmers. This in turn causes inefficient allocation of resources. Moreover, the two groups of farmers who face different factor prices would exhibit different patterns of economic behaviour. While large farmers will tend to use modern technology and inputs and will increasingly cultivate cash crops, peasants will be confined to traditional methods of production and the cultivation of subsistence crops. Government policies in a bimodal system are normally in favour of large farmers which will strengthen the above tendencies.

In the present chapter, we want to examine the Iranian bimodal system in order to find out if large and small farmers really faced different conditions in factor markets other than land, and if there were growing contrasts between the two groups of farmers as the theory suggests. We will also study the outcome of the government's policies for each group. Finally, we will try to see how the post-reform changes affected agricultural growth in Iran. But, in view of the importance of irrigation in Iranian agriculture, we should perhaps begin with an examination of the impact of the land

reform on the accessibility of irrigation water to different groups of farmers before turning to the other issues.

The land reform and irrigation sources

Access to water sources became more difficult for the peasants after the land reform for several reasons. First, the law allowed the landowners to choose the land they wanted to retain and they chose the best land with good water resources. This was especially common where only part of the village land was transferred to the peasants. In these villages, peasants received plots of land which were often far from the water supply. If such villages depended on *qanats* or wells for irrigation they generally remained in the hands of the landowners although, according to the law, the irrigated land had to be transferred to the peasants inclusive of its water rights (see Azkia 1980: 112; Hooglund 1982: 91–2; for some examples see OIPFG 1974).

Apart from retaining the title to source of water supply, landlords exercised other practices which limited irrigation opportunities for the peasants. In many cases, the landlords reallocated the irrigation water so that their retained land received more water than they were legally entitled to. One way, for instance, was to change the rotation of water. Let us assume that the peasant farms had the right to water once in eight days. The rotation could then be increased so that the peasants would only get water once in, say, fifteen days (Lambton 1969: 284–7; see also OIPG 1974).

Landowners were sometimes alleged to demand high prices for irrigation water. There were cases where many peasants felt that they were unable to pay water charges and began to rely exclusively on the uncertainties of rain, or to leave increasingly more of their land fallow every year, or to rent out part of it to the larger owners (Hooglund 1982: 92–3).

The second reason that facilitated irrigation for the larger farmers, and which probably adversely affected the peasants, was the development of power-operated wells. Large farmers were able to invest in power-operated, deep and semi-deep wells, something which peasants normally could not afford. This type of investment which began before the reform increased markedly afterwards (see, for example, Momeni 1980: 326). It is believed that the sinking of deep wells was in part responsible for the reduction of water to some *qanats* which irrigated the peasants' farms in several regions (Lambton 1969: 289–90; Momeni 1980: 399; OIPFG 1974: 2).

Finally, the government's irrigation policies and projects were

much more beneficial to large farmers than to the peasants; they were sometimes prohibitive for the latter. For example, the Ministry of Water and Power abolished customary water rights which existed in parts of Azarbayjan and demanded water dues from the farmers. The peasants complained that water was sold to the highest bidder or that they had to pay high water dues although they grew only grain crops which used little water in comparison to cash crops (Lambton 1969: 278–9).

The government correctly considered the scarcity of water to be a major problem for Iranian agriculture and gave priority to irrigation projects in its agricultural investments. For instance, irrigation projects absorbed about 45 per cent of total government expenditure in agriculture between 1963 and 1972 (Bank Markazi Iran, *Annual Report and Balance Sheet* 1965–1975). However, most of the government outlay for irrigation was allocated to the construction of big dams and water networks which yielded little benefit to the peasants. As we saw in the previous chapter, the land under the dams was allocated to agribusinesses and farm corporations.[1] These dams might reduce the flow of water in the streams that were used by the peasants or even dry them up, something which happened in Sistan (Momeni 1980: 397–8). Where peasants had access to the water of these dams, they had to pay more than agribusiness for the water: each cubic metre of the irrigation water cost the peasants 0.15 rials while the agribusiness companies paid 0.12 rials for it (Azkia 1980: 127; Dumont 1978: 17).

In sum, sources of water supply were under the control of the larger farmers, which normally meant that they had easier access to water than peasants. The government's irrigation policies often reinforced the prevailing situation. We should note that the unequal access to water implies that if land is measured in efficiency units, the extent of land concentration will in fact be greater than shown by the crude measures of land area such as numbers of hectares.[2]

The rural capital market

The borrowing conditions in the rural capital market were not the same for the two groups of farmers. First and foremost, large farmers were among the richest in the society and in all probability had large savings. They could draw from them whenever they needed money. Most of the peasants, on the other hand, were in debt when they received their land. While it is impossible to prove the former assertion, our discussion in Chapter 3 substantiates the latter (see pp. 44–7). Moreover, the peasants were

principally land reform beneficiaries and had to pay, in fifteen years, the instalments of the land they received.

For their financial needs, farmers could borrow from commercial banks, the government or the informal credit market. The Iranian commercial banks provided about 20 per cent of the agricultural credit between 1963 and 1973 (see Table 7.1). The difficulties involved in obtaining loans from commercial banks, such as the necessity of providing security or finding a guarantor, deprived the peasants of borrowing from these sources (Khosrovi 1979b: 160). It was then principally large farmers, processors, merchants and exporters who received most of the agricultural loans from these banks (Aresvik 1976: 171).

As Table 7.1 shows, more than 60 per cent of the institutional credits of the agricultural sector consisted of state loans. These loans were given to farmers through co-operative societies, the Agricultural Bank (the Agricultural Co-operative Bank of Iran), the Agricultural Development Bank of Iran, and some special government agencies.[3]

Table 7.1 Relative importance of various sources of agricultural credit, 1963–72 (estimated average)

Source of credit	*Amount (billion rials)*	*%*
Agricultural Co-operative Bank	9.0	14
Agricultural Development Bank	0.4	1
Commercial banks	13.2	20
Government agencies	5.1	8
Rural co-operatives	6	9
Non-institutional sources	32.3	49
Total	66	100

Source: Aresvik (1976: 169)

Rural co-operative societies were in practice the most accessible source of institutional credit available to peasants. These co-operatives were supposed to cover all the land reform beneficiaries. According to the provisions of the first stage, the eligible peasants received land if they were members of rural co-operative societies (see Chapter 5, pp. 85–6). The number of co-operatives, however, increased slowly. There had been 8,361 co-operatives serving about 30,000 villages by the end of 1972 (Statistical Centre of Iran 1976: 215). Some of these co-operatives existed only on paper and about half of the villages can be said not to have had co-operative societies. In 1972, it was decided to consolidate the societies and

consequently their number decreased to 2,717 in 1973 (Statistical Centre of Iran 1976: 215). As a result, the members of 6,000 dissolved co-operatives had to rely on the bigger co-operatives which were located in larger villages. In some cases, for example, peasants in twenty different villages were members of one and the same society (Hooglund 1982: 106). Considering the dispersion of villages in Iran and the communication difficulties, this might in practice have deprived some members of the benefits of the societies. This may partly explain the bitterness and objections of the members of smaller co-operatives toward such a plan from its early beginning (see Denman 1973: 202).

The major activity of the co-operatives was to provide loans to their members. The loans were principally financed by the Agricultural Bank and a small share of them was provided out of the co-operatives' own capital and reserves. But these sources were not enough to meet peasants' credit needs and therefore co-operative loans comprised on average only 9 per cent of agricultural credit between 1963 and 1972 (See Table 7.1). The loans which were granted by the societies were principally short-term. For instance, in the period 1965 to 1972 more than 90 per cent of the loans granted by the societies were for 6 to 12 months (Azakia 1980: 98; see also Hooglund 1982: 110).

The amounts of most of the loans were also low. While the average expenditure for producing 1 hectare of dry-farmed winter wheat was about 6,000 rials, the average co-operative loans hardly exceeded this amount during the period 1963 to 1971. The amount of loans increased rapidly after 1971 but, considering the peasants' needs and inflation, such increases were insignificant. For example, loans per head in 1972–3 show an increase of 25 per cent over 1970–1 while the inflation rate for the same period was about 20 per cent (see Table 7.3 and Bank Markazi Iran, *Bulletin* 1975).[4] Moreover, co-operatives were not even able to provide such limited loans to all their members: between 1963 and 1973, only 55 per cent of the total members received co-operative loans (Statistical Centre of Iran 1975: 183, 211; 1976: 184, 215). This figure is certainly an overestimate since it was quite common to grant a 'new' loan to the peasant who could not repay his previous loan and take the new loan as payment for the old one (see Katouzian 1974: 233).

There is also evidence which indicates that the richer members of the societies had easier access to co-operative loans than the poorer ones. First, the amount of the loan was related to the number of shares owned by each member. Each member could borrow up to ten times the value of his shares, but in no case more than 20,000 rials (Hooglund 1982: 109; Azukia 1980: 280). Second, it was the

manager of each society who in practice decided about the loans. In most of the co-operatives the more prosperous farmers occupied such positions, especially after the consolidation of the co-operatives, and the abuse of power, for example to ensure more credit for themselves, their friends, etc., was a widespread practice among them (see Azkia 1980: 280–1; Hooglund 1982: 106; Lambton 1969: 337–8).

Small farmers could also borrow money from the Agricultural Bank. However, about 65 per cent of the loans from this source were granted to members of co-operatives through the societies and only 5 per cent were directly lent to the members as individuals. Around 25 per cent of loans were given to non-members and 5 per cent to supervised credit projects (Aresvik 1976: 171–2).

Table 7.2 illustrates some basic features of the loans that were granted by the Agricultural Bank between 1963 and 1973. It can be seen that most of the loans of the bank were small and short-term. Table 7.2 also shows that the percentage of loans which were below 10,000 rials decreased while that between 10,000 and 50,000 increased. In other words, the size of the loans tended to increase during this period. However, the facts that the Agricultural Bank mainly paid loans to peasants via the co-operative societies and that the average loan of the co-operatives was close to 10,000 rials may imply that the smaller loans were much more frequent in the group of 10,000–50,000 rials. In accordance with co-operative loans, a large majority of small loans were in all probability short-term.

Moreover, Table 7.2 indicates that 93 to 98 per cent of the loans of the Agricultural Bank amounted to less than 50,000 rials. But the available information indicates that the loans which were more than 50,000, although small in numbers, absorbed a large proportion of the credits. Between 1965 and 1969, for example, about 36 per cent of the total credits of the bank was allocated to such loans (Statistical Centre of Iran 1976: 183; Bank Markazi Iran, *Annual Report and Balance Sheet* 1970: 126). Considering the percentages of short-term loans (Table 7.2) and the fact that a great majority of them were in fact co-operative loans it is reasonable to conclude that most of the large loans might have been medium- and long-term.

On the whole, we can divide credits which were provided by the Agricultural Bank into two groups: (a) small, short-term loan which were mostly given to peasants through co-operatives and (b) large, medium- and long-term loans which were given direct farmers. Small farmers were in practice deprived of the latter lo First and foremost, the granting and the size of loans was cl related to the collateral offered by the borrowers (Aresvik 1976:

172). Small farmers could hardly provide such a security, especially for larger loans. Moreover, the branches of the bank were in large towns. Regarding detailed security requirements, paperwork and delays, it sometimes took a week or more to process a loan. During this period, the peasants had to stay in towns, or if possible, travel there several times. In each case, the costs of getting loans to peasants were certainly prohibitive (see Aresvik 1976: 174; Katouzian 1974: 234; Momeni 1980: 405).

Table 7.2 Credit allocated by the Agricultural Bank, 1963–73

	1963–7	*1968–9*	*1970–1*	*1972–3*
Total number (thousands)	2,048	675	445	609
Total amount (million rials)	23,394	10,705	18,491	34,374
% of loans < 10,000 rials	76[a]	69	62	34
% of loans 10,000–50,000 rials	20[a]	29	32	59
% of total amount of short-term loans[c]	49	63	83[b]	75[b]

Sources: Statistical Centre of Iran (1975, 1976) and Bank Markazi Iran, *Annual Report and Balance Sheet*, (1971, 1973, 1974)

Notes: a For 1966–7
b The figures stand for only one year in each group, i.e. 1971 and 1973
c Up to two years

Large long-term loans were also provided by the Agricultural Development Bank. This bank was founded to finance the development of large farms. The bank normally lent to farmers with holdings in excess of 500 hectares. However, farmers with farms as 'small' as 100 hectares also had access to its loans. The minimum loan was one million rials and the bank could lend up to 80 per cent of investment outlays. Its only borrowers were naturally large farms such as farm corporations and agribusinesses (Aresvik 1976: 172; Katouzian 1983: 333).

In general, it can be said that peasants only had access to institutional credits only through the co-operatives while the large farmers were able to borrow both from the Agricultural Bank and from the Agricultural Development Bank.[5] Table 7.3 shows that the average loan varied greatly from source to source. The co-operatives gave the smallest loans per capita. It should be noted that if we could eliminate the negative effect of the small co-operative loans which are included in the loans of the Agricultural Bank, the difference in the average loan available from these sources would become much larger.

Unequal access to government credits was in fact the natural outcome of the government's general agricultural policy. The

analysis of agricultural policies in Chapter 5 showed that they generally supported large farmers. It was also as a result of such a policy that the government granted large sums of money to farm corporations or provided agribusinesses access to road and drainage facilities, etc. free of charge (see Denman 1978: 287; Azkia 1980: 126–7; Bagley 1976: 30). The privileges of farm corporations are clearly evident in Table 7.4. It illustrates that each peasant who borrowed through co-operatives in the period 1968 to 1971 received, on average, loans which were only 11 to 18 per cent of the loans and grants per shareholder received by corporations in the same period. Moreover, more than 70 per cent of the sums which corporations received were in fact grants. It should be also noted that corporation loans received were in fact grants. It should be also noted that corporation loans were medium- and long-term with an effective interest rate of one per cent. The interest rate of co-operative loans was 6 per cent.[6]

Table 7.3 Average state loans per recipient by source of credit, 1968–74 (thousand rials)

Source	*1968–9*	*1970–1*	*1972–3*	*1974*
Cooperatives[a]	6,800	7,400	9,300	14,500
Agricultural Bank	15,200	41,600	56,400	n.a.
Agricultural Development Bank[b]	18,600	20,500	8,900	29,500

Sources: Bank Markazi Iran, *Annual Report and Balance Sheet* (1971–5); Statistical Centre of Iran (1976)

Notes: a A small part of co-operative loans was financed by their own capital and reserves
b The figures are loan and government participation per project

Table 7.4 State loans and grants to co-operative societies and farm corporations, 1968–71 (rials per head)[a]

	1968	*1969*	*1970*	*1971*
1. Co-operative: loan	6,821	6,816	6,992	7,776
2. Corporation: loan and grant	60,238	48,630	39,876	51,674
3. Grants as % of (2)	76	77	70	76
4. (1) as % of (2)	11	14	18	15

Source: Statistical Centre of Iran (1976: 184, 216)

Notes: a For co-operatives 'per head' stands for 'per borrower' and for corporations it stands for 'per shareholder'

Since the organized credit market did not meet the credit need of most small farmers, they were forced to borrow in the informal credit market. One decade after the reform began, about half of the

total agricultural credit was still obtained on the latter market, i.e. from moneylenders, shopkeepers, relatives, etc. (see Table 7.1). In fact, these were the only sources of credit which were available to many peasants since they did not have co-operative societies. But even the peasants who had access to co-operative loans often had to obtain non-institutional credit to supplement co-operative loans. In many cases, loans which were obtained from moneylenders, shopkeepers, relatives, etc., constituted between 40 and 80 per cent of the peasants' total loans (see Momeni 1980: 406–7; Khosrovi 1979b: 160–1; Denman 1973: 201). These loans generally carried interest rates which amounted to between 20 and 50 per cent (see, among others, Momeni 1980: 407–8; Aresvik 1976: 169).[7]

In sum, the peasants' loans consisted of low, short-term institutional credits and loans financed by moneylenders at notoriously high interest rates. Such loans could hardly be used for productive investments and as a result they were principally spent on current agricultural expenditure and consumption (see Khosrovi 1979b: 157; Momeni 1980: 391–4; Hooglund 1982: 96, 110). On the other hand, as would be expected in a bimodal agrarian system, the Iranian large farmers generally had easier access to long- and medium-term institutional credits.

The rural labour market

We saw that large farmers controlled land and water in the Iranian rural areas and had also easier access to capital. Theoretically, this situation may give large landowners some power even in the village labour market since there are normally few employment opportunities near by other than the farms of large farmers. One can hardly expect small farms to provide any significant employment for outsiders. These farms often have difficulty in providing sufficient work even for family members. There are reasons to believe that large farmers could have gained some power in labour markets in the Iranian villages after the reform.

Table 7.5 shows, for example, that about 70 per cent of farms under 10 hectares were cultivated exclusively by family members and 26 per cent of such land employed few additional labourers. In fact, numerous small farmers could not rely entirely on their farms for their living and, therefore, they had to seek jobs on other farms (see Chapter 6, pp. 108–9; Hooglund 1982: 97). Moreover, some peasants lost their rights to land in one way or another in different stages of the land reform and became landless labourers. Consequently, competition for the available jobs in each village, or in close neighbouring villages, could have increased, which in turn

Table 7.5 Percentage share of wage labour in total labour employed by farms of different size-groups, 1974

Farm size (hectares)	*Only family labour*	*Mainly family labour*	*Mainly wage labour*
< 10	70	26	4
10–50	44	48	7
> 50	15	39	47
All farms	66	29	5

Source: Khosrovi (1981: 53)

would have increased the influence of large farmers in their local labour market.

However, the increasing non-agricultural employment opportunities operated against such a likely outcome. The oil boom of the early 1970s created an enormous demand for labour, especially for unskilled labour. Demand increases occurred mainly in the urban areas which led to a widening gap between rural and urban wages (Kazemi 1980: 42; see also Khosrovi 1979: 112–13). This, and the gloomy employment outlook in the rural areas, might have been the main reason behind the mass rural–urban migration of the 1970s. It has been estimated that about 2 million villagers, or more than 70 per cent of the population increase in villages, migrated to cities in the period 1966 to 1976 (Hooglund 1982: 118–19; see also Kazemi 1980: 14). The existence of a great outlet for surplus labour must have helped to reduce rural unemployment to the low level of about 3 per cent in 1976.[8]

Although the availability of non-agricultural employment was beneficial to the rural surplus labour, it might have been detrimental to the agricultural sector in general and poor peasants in particular. The loss of too many labourers was in fact a resource movement out of agriculture which might have affected agricultural production adversely (see pp. 132–4). The migration of the labour force was perhaps more troublesome for small farmers since an increasing number of family members left the villages and they thus lost their principal, and relatively cheap, production factor. In consequence, they had to turn to wage labour. It seems that many of them could not afford this and chose instead to cultivate less land than before (cf. Hooglund 1982: 120; see also Haririan 1976: 341–2; Saeedi 1975: 652). On the other hand large farmers, who might have faced labour shortages and increasing wages, began to employ 'imported' Afghani labourers who demanded lower wages than the prevailing ones (Hooglund 1982: 87, 116, 120; see also Mahdavi 1982: 61).

Modernization of farming techniques after the reform

Some of the landowners had begun large-scale farming before the reform. They tended to utilize new technology on their farms and especially invested in tractors. Such activities were interrupted when the land reform and anti-feudal campaign started in 1962. For example, a government agency imported agricultural machinery and sold it to agricultural companies and landowners before the reform, but since the demand for agricultural machinery decreased after the reform, the agency curtailed its activities in 1963 (Bank Markazi Iran, *Annual Report and Balance Sheet* 1965: 131; see also Momeni 1980: 323–4). However, landlords generally began to invest in modern technology after they succeeded in controlling the reform and became more involved in agricultural activities. Table 7.6, which shows the sales of tractors and combines in the years immediately before and after the reform, reflects this trend very clearly.

Table 7.6 Sales of tractors and combines, 1961–5

	1961	*1962*	*1963*	*1964*	*1965*
Tractors	1,223	645	133	714	6,300
Combines	168	195	6	36	91

Sources: Statistical Centre of Iran (1971: 44); Katouzian (1978: 365)

Table 7.6 also shows that there was a tremendous increase in the sale of tractors in 1965. This might be partly attributed to the implementation of the second stage of the land reform. The land which was worked by labourers and ploughed by mechanized means was exempted from land reform during this stage.[9] However, the modernization process continued after the reform and the application of tractors, different kinds of fertilizers and pesticides, etc., became more common in agriculture in Iran. Table 7.7 illustrates the increase in the use of three major technical inputs after the reform. We can see that the amount of fertilizers had increased by more than 4,000 times, and tractors by about eight times up to 1977.

Notwithstanding large increases in the application of modern inputs, the employment of these inputs was far from being comprehensive enough to cover all 2.5 million Iranian farmers. The available statistics show, for example, that there were wide differences in the use of machinery in various parts of the country. In some regions such as Sistan and Baluchestan, Yazd and Ilam, the amount of machinery in use was negligible in 1972 (Statistical Centre of Iran 1976: 199). At the same time, the bulk of all types of machinery was concentrated in several regions (see Table 7.8). Among these regions

Mazandaran, Khorasan and Markazi provinces had comparatively more large farms (Hooglund 1982: 83; Azkia 1986: 218; Khosrovi 1981: 91). Moreover, a large number of farm corporations were located in Fars (see Azkia 1986: 244).[10] It is then reasonable to believe that it was mainly large farmers who employed machinery on their farms (cf. Agah Book Collection 1982: 165–6; Khosrovi 1982: 47; Momeni 1980: 326–7).

Table 7.7 The application of machinery and fertilizer to agriculture, selected years

Years	*No. of combines*	*No. of tractors*	*Fertilizers (tons)*
1962	900	6,000	18,000
1967	n.a.	17,500	49,000
1971	1,800	23,000	32,800
1977	2,500	53,000	772,000

Sources: Statistical Centre of Iran (1976, 1984); *Iran Almanac* (1963: 263); FAO, *Production Yearbook* (1966, 1971)

Table 7.8 Distribution of machinery by selected provinces, 1972 (percentage)

Province	*Tractors*	*Combines*	*Other machinery*
Fars	6	8	5
Khorasan	13	2	10
Kordestan	7	11	<1[a]
Markazi	10	5	2
Mazandaran	18	53	50
West Azarbayjan	10	5	<1
Rest of the country	36	16	33

Source: Statistical Centre of Iran (1976: 199)

Note: a Less than 1 per cent

The application of modern technology was comparatively widespread in Mazandaran. According to Table 7.8, Mazandaran was the most mechanized region of the country in 1972. Large-scale, modern farming had begun in this region, which mostly belonged to the royal family, before the land reform (see Chapter 3, pp. 32–3, 40). In spite of the land reform and the redistribution of crown lands, large farms prevailed in this region after the reform. Moreover, many of these farms belonged to influential figures such as members of the royal family, high-rank army officers, lawyers, etc. (see Nikayin 1976: 78–9; Momeni 1980: 300–1). One of the shah's brothers had a farm in this region which was as large as

3,000 hectares and was cultivated with twenty tractors, nine combines, thirteen tillers, etc. (Mohammadi 1973: 39–40).

Large farmers were generally inclined to use labour-saving methods. As a result, mechanization of major operations such as ploughing and harvesting became very common on large farms (see Momeni 1980: 325–6; Aresvik 1976: 164; Hooglund 1982: 97; Antoun 1976: 8). Unfortunately, we have little information about individual large farms but there is more information about agribusinesses and farm corporations which shows that they followed a labour-saving, capital-intensive policy. It was because of this policy that the use of tractors and combines in farm corporations increased to 100 per cent and 95 per cent, respectively, of the cultivated area in 1975 (Ashraf and Banuazizi 1980: 41). Sometimes the houses in the villages covered by corporations were bulldozed to make sure that the surplus labour would go elsewhere. New accommodation was subsequently built only for the retained workers and their families (Platt 1970: 81). However, if the labourers had not been driven out of the villages by force, the decreasing employment opportunities on corporations would have forced them to migrate (Azkia 1980: 221, 239; see also Aresvik 1976: 123).

From the outset, the emphasis of agribusinesses was on capital-intensive methods. The agribusinesses which were established below the Dez Dam in Khuzestan provide a good example. Here only 6,000 labourers were employed on 68,000 hectares, i.e. one worker for every 11 hectares of irrigated land; peasant holdings employed one person for every 2 hectares (Dumont 1978: 17). The tendency to use capital-intensive methods on larger farms was accompanied by relatively extensive use of land. We can see from Table 7.9 that the amount of land which was left fallow increased as farms became larger. It should be borne in mind that land, first and foremost, was left fallow because of water shortage. However, the land of large farmers was generally very fertile with easier access to water and it could have been cultivated more intensively than the small farms. The extensive land use and the increasing tendency of large farmers to employ capital-intensive methods lend support in part to our earlier conclusion that land and capital would be cheaper for large farmers than the peasants in a bimodal system (see Chapter 4, pp. 54–9).

In addition to its credit policy, the government supported the development of capital-intensive methods in different ways. For example, all the machinery which agribusinesses needed to import for their activities was exempted from import taxes, provided that such machinery was not produced in the country (Azkia 1980: 126–7). A government agency was also responsible for importing tractors and selling them to farmers on instalment.

Table 7.9 Distribution of fallow land by farm size, 1974 (hectares)

Farm size	*Total area*	*Fallow land*	*Fallow land as % of total area*
< 10	5,389,877	1,589,154	29
10–30	6,391,526	2,413,456	38
> 30	4,635,818	1,952,737	42
All farms	16,417,221	5,955,347	36

Source: Khosrovi (1981: 57; 1982: 68)

Some attempts were made by the government to introduce improved seeds in agriculture but they had little effect. Accordingly, high-yielding varieties of wheat and improved cottonseed, vegetable seeds, fruit shoots, etc., were distributed among farmers. Considering the fact that wheat had been the staple crop in Iran, the spread of high-yielding varieties of wheat would have largely benefited small farmers. However, the programme which aimed at extending the use of high-yielding varieties of wheat suffered from serious shortcomings, and as late as 1976 the improved varieties were cultivated on only about 350,000 hectares of irrigated land. Inadequate supply of improved seeds and fertilizers, and insufficient work, etc., were responsible for the limited effects of the programme (Aresvik 1976: 136–43).

It is necessary to point out that the modernization process was not strictly confined to large farms. There were also many farmers with medium-sized holdings and some prosperous small farmers who began to use new methods of production. For example, the upper strata of small farmers in Gilan and Mazandaran, who were considerably better off than most of their counterparts in the rest of the country, increasingly used two-wheel tractors on their farms (Khosrovi 1981: 144–5; Aresvik 1976: 160–3). Moreover, a recent statistical study shows that small farmers in fact used more fertilizers per hectare on all their land than the large farmers in 1974 (Agah Book Collection 1982: 185). The data for the consumption of fertilizers in 1971 shows, however, that around 74 per cent of all fertilizers were only used in five provinces (Statistical Centre of Iran 1976: 185).[11] Among these Esfahan, Gilan and Mazandaran are known to have comparatively prosperous peasants which suggests that it might again have been the richer peasants who mostly used fertilizers.

On the whole, there is some evidence which suggests that the large farmers generally tended to modernize their farming after the land reform. There were also small peasants who, sometimes jointly, invested in agricultural machinery, began to use fertilizers,

improved seeds, etc., and who hired tractors for ploughing (see, for example, Momeni 1980: 322–7; Craig 1978: 148; Lambton 1969: 224–8). However, the great majority of peasants experienced little change in their farming techniques and relied predominantly on the traditional methods of production in the late 1970s.

The new agrarian system and agricultural production

We saw in Chapter 3 that most of the cropland was under foodgrains, i.e. wheat, barley and rice, before the reform. Cash crops such as cotton, sugar beets, fruits, and oil seeds occupied only a small proportion of all the cultivated land.[12] Parallel to other changes in the agrarian system after the land reform, this crop pattern also began to change. Accordingly, more and more land was allocated to cash crops and the comparative share of land under staple food products declined. Table 7.10 summarizes this trend.

The table shows that although the land under cultivation increased for both staple food crops and cash crops, during the period 1960 to 1975, the latter increased much more markedly than

Table 7.10 Land under staple and cash crops, 1960–75 (thousand hectares)

Type of crop	*1960*	*1975*	*% change*
Foodgrains[a]	5,579	7,778	39
Irrigated	(2,080)	(2,318)	(11)
Unirrigated	(3,499)	(5,463)	(56)
Cash crops[b]	1,126	2,670	131
Irrigated	(619)	(1,786)[c]	(189)
Unirrigated	(507)	(884)	(74)
Others	98	168	71
Irrigated	(78)	(137)[c]	(76)
Unirrigated	(20)	(31)	(55)
Total	6,803	10,616	56
Foodgrains as % of total	82	73	−11

Source: Agah Book Collection (1982: 181)

Notes: a Mainly wheat, barley, and rice
b Cotton, sugar beets, vegetables, fruits, etc.
c The distinction between irrigated and non-irrigated land for fruit and non-fruit for 1975 was not available; consequently 1976 data were used

the former. Consequently, the share of food grain land decreased from 82 per cent of the cultivated land to 73 per cent in 1975 (see Table 7.10). Its share dropped further to about 70 per cent in 1978 (Statistical Centre of Iran 1984: 335). Table 7.10 also shows that the share of cash crops was partly increased at the expense of staple crops (cf. Agah Book Collection 1982: 163). Moreover, it clearly illustrates that cash crop cultivation was developed predominantly on irrigated land while the increase in foodgrain land mainly consisted of unirrigated, and probably less fertile land (see also Aresvik 1976: 49).

As was pointed out in Chapter 1, there are frequent discrepancies between the crop production data which are given out by different official sources (see p. 3). As a result, different sources may give different pictures of the changes in crop production after the reform. This is clearly reflected in Table 7.11 which illustrates production data for the staple and cash crops by the two major sources of agricultural statistics, namely the Statistical Centre of Iran and the Ministry of Agriculture.

The two sources indicate a similar trend in crop production up to 1973. They show that although the production of all crops increased, cash crop production, as a whole, grew more rapidly (see Table 7.11). On the other hand, they produce a rather different picture for the period after 1973. The data of the Statistical Centre show that the same trend persisted even after 1973 while the very high figures of the Ministry of Agriculture for staple crops obscure this trend in its statistics. The fact that imports of foodgrains increased from 904,000 tons in 1973 to about 1,804,000 tons in 1974 and remained high in the following years seems to suggest that (i) the production of staple food products could not have been as favourable as the Ministry of Agriculture reported, and (ii) the data of the Statistical Centre of Iran, as was mentioned earlier, are comparatively more reliable (see Bank Markazi Iran, *Annual Report and Balance Sheet* 1976: 79).

Table 7.12, which is based on the data provided by the Ministry of Agriculture, provides a detailed illustration of the growth of agricultural products for a period when the production statistics of the two sources are largely in agreement with each other. We can see from Table 7.12 that all agricultural products increased in this period but cash crops increased more steadily. The growth rates of wheat and barley, which were comparatively high between 1963 and 1967, fell close to zero in the following period. On the other hand, most cash crops increased at higher rates in the second period.

There are several reasons to believe that large farmers might have

Table 7.11 Production of staple and cash crops by source of information, 1968–77 (thousand tons)

Source of information and year	*Staple crops*[a]	*Cash crops*[b]
A. Statistical Centre		
1968	5,995	4,538
1971	5,340	4,988
1972	6,633	5,136
1973	6,641	6,003
1974	4,463	5,766
1975	6,408	6,298
1977	5,779	5,416[c]
B. Ministry of Agriculture		
1968	5,990	4,576[d]
1971	5,650	5,304
1972	6,755	5,400
1973	6,857	6,177
1974	6,876	6,418
1975	8,330	6,580
1977	8,130	5,961[e]

Sources: Agah Book Collection (1982: 182–3); Bank Markazi Iran, *Annual Report and Balance Sheet* (1972–7)

Notes: a Wheat, barley and rice
b Cotton, sugar beet, tobacco, pulses, sugar cane and oil seeds
c For pulses the data of the Ministry of Agriculture were used
d For tobacco the data of the Statistical Centre were used
e For sugar cane the data of the Statistical Centre were used

been mainly responsible for the increase in the land under cash crops and in cash production. First, the process of bimodal-oriented changes which started in 1963 must have progressed greatly by 1968. This process was very likely reinforced by the landlord-biased agricultural policies and the third stage of land reform. They were mainly initiated in 1968 (see Chapter 5, pp. 83–5). Consequently, large farmers were certainly more active in farming after 1968 and the coincidence of the increase of growth rates of cash crops with this increased activity may indicate that there was probably a positive relationship between the two trends.

Second, since orchards were exempted from the land reform, many landlords extended their former orchards or planted new ones to retain their land (see Chapter 6, p. 94). Moreover, fruits and nuts were much more profitable than staple crops and this must also have encouraged farmers to develop orchards.[13] However, it would take time before orchards came to bear fruit and most peasants could not afford it. Thus it was in all probability large owners and more prosperous farmers who planted most of the fruit-bearing trees after the land reform. As a result, the orchard

land increased by about 60 per cent and the number of trees by 50 per cent between 1960 and 1972 (Statistical Centre of Iran: 1976: 193, 207).

Table 7.12 Annual growth rates of crop production, selected periods (percentage)

	1963–7	*1968–72*
Grains		
Wheat	7.1	0.5
Barley	6.2	−0.7
Rice	6.5	4.6
Pulses	9	4.7
Vegetables		
Potatoes	3.1	−1
Onions	3.9	7.7
Tomatoes	3.8	3.1[a]
Melons	3.6	5.9
Industrial crops		
Cotton	4.5	7.8
Sugar beet	27	7.7
Oil seed	8.5	24.8
Fruit[a]		
Citrus	4.4	6.2
Apples	4	21.1
Grapes	0.2	6.8

Source: Aresvik (1976: 251)

Note: a 1972 data were not available and thus the growth rates for the second period refer to 1968–71

Third, cash crops needed comparatively more water and were therefore mainly cultivated on irrigated land (see Table 7.1). We saw earlier that peasants generally had difficult access to water and they could have contributed only marginally to the increase of cash crops. Finally, farmers often had to learn the cultivation of 'new' crops. Moreover, cultivation of cash crops requires more cash and involves difficulties such as buying new inputs and selling the crop. We have already seen that a large majority of peasants had considerable difficulty in rasing loans or obtaining new inputs such as fertilizers and consequently were not able to cultivate cash crops easily. There were even cases where the peasants reduced their cash crop cultivation or stopped it altogether after the reform since they lost their landlords' financial support (see Momeni 1980: 332).

The available data also reveal that cash crop production was in general concentrated in certain regions. Table 7.13 shows, for example, that production of several major cash crops, including

rice, was highly concentrated in only four regions. The comparatively higher proportion of large farms in Mazandaran and Khorasan and the existence of the largest number of farm corporations in Fars could have contributed to such a concentration (see Azkia 1986: 244).[14] On the other hand, a large majority of rice-growers, who were predominantly small farmers, were Gilani and Mazandarani peasants (see Kamali 1982). They were normally more prosperous than their counterparts in the rest of the country. Perhaps most of the small cotton cultivators were also Mazandarani peasants. They were encouraged to shift to cotton by the success of large farmers who had initiated its cultivation (see Okazaki 1968: 28–35). Large farmers in fact diffused the know-how of cotton cultivation in the region and demonstrated its profitability which together reduced the risk of shifting to the 'new' crop for the peasants. Finally, we can see from Table 7.13 that the concentration of production of cotton, sugar beet and rice in the above-mentioned provinces increased further between 1973 and 1978.

All in all, it is plausible to conclude that it was especially the large and more prosperous farmers who were in a position to produce cash crops. In fact, they increasingly did so. Many studies emphasize that those farmers allocated more and more land and capital to cash crops such as cotton, sugar beet and oil seeds (see Momeni 1980: 333; Hooglund 1982: 112; Haririan 1976: 340).

Table 7.13 Production shares of selected provinces in the total production of principal cash crops, 1973 and 1978

	1973			*1978*		
Province	*Cotton*	*Sugar beets*	*Rice*	*Cotton*	*Sugar beets*	*Rice*
Fars	10	12	4	4	20	10
Gilan	—	—	44	—	—	50
Khorasan	11	45	—	13	46	—
Mazandaran	65	—	41	76	—	34
Total	86	57	89	93	66	94

Source: Statistical Centre of Iran (1984: 281–4)

Staple food crops and agricultural growth

In spite of the gradual shift from staple crops to cash crops, the former continued to be predominant in Iran. Wheat, barley and rice comprised 51, 15, and 3 per cent, respectively, of cultivated land as late as 1978 (Statistical Centre of Iran 1984: 335). Most of this land, except rice land, was rain-fed and more than 80 per cent of it

belonged to peasants. In consequence, changes in the production and productivity of staple food products had decisive effects on both agricultural growth and peasant living standards.

Scarcity of credit and lack of peasant-biased innovations, which are typical in a bimodal agrarian system, hindered peasants from improving the production of staple crops. Large farms also failed to invest in these crops. The main investment of many large farmers in staple crops seems to have involved the mechanization of wheat and barley production and the expansion of their production in unirrigated land (see Aresvik 1976: 53). Consequently, rates of land productivity for wheat and barley were stagnant and remained low after the land reform. The yield per hectare for rice improved to some degree during this period (cf. Kamali 1982: 142–3; Katouzian 1983: 327–8). On the other hand, the yield of cash crops, e.g. sugar beet and cotton, increased rather well in the post-reform period (cf. Aresvik 1976: 49–50). Table 7.14 summarizes the changes in yields of five major crops after the reform. We have to note that the FAO figures for staple crops are usually higher than those of the Statistical Centre of Iran, especially in more recent years. For example, according to the data of the Statistical Centre of Iran, the yields (kilograms per hectare) of the above crops in 1977–8 were as follows: wheat 807, barley 791, rice 2,457, sugar beet 26,442 and cotton 1,778 (Statistical Centre of Iran 1984: 278, 284, 335).

Table 7.14 Yield of major crops in Iran, 1961–78 (kilograms per hectare)

Crop	*1961–5*	*1966–70*	*1971–5*	*1977–8*	*Average of near east region 1977–8*	*Average of developing countries 1977–8*
Wheat	802	896	884	1,011	1,408	1,347
Barley	800	891	753	841	1,204	1,218
Rice	2,914	3,355	3,455	3,315	3,866	2,484
Cotton	939	1,198	1,650	1,598	1,762	914
Sugar beets	19,202	21,082	24,243	25,314	31,088	30,938

Source: FAO, *Production Yearbook* (1971–9)

The government's price policy was an important factor which discouraged farmers from investing in staple crops and improving their production. This was a particular deterrent in the case of wheat which was by far the most important crop. The price policy for wheat, which started in the early 1960s, aimed to stabilize the wholesale price of wheat in order to prevent, above all, increases in the price of bread in the cities and at the same time guarantee a

minimum price to the producers. However, this policy suffered from certain shortcomings and failed to benefit wheat producers. (The following three paragraphs are largely based on Aresvik 1976: 143–9).

First, the minimum price of wheat which was set by the government was rather low in comparison to its production costs. Between 1960 and 1974, the minimum prices of wheat remained almost unchanged while all other prices increased. In 1974, the minimum price was raised by 40 per cent but this price and the following revisions in the minimum price were not realistic either. In 1978, for example, the floor price of wheat was 12 rials per kilo but the production costs to the farmer averaged between 14 and 17 rials per kilo (Hooglund 1982: 114).

Second, the minimum price used to be announced every year after the wheat harvest was partly over. This made it ineffective as a floor price for the producers and could naturally not affect their decision to improve wheat production. In fact, by such a late date, most farmers, especially the small farmers, had already entered into contracts with private wholesalers at prices which were usually lower than the minimum price subsequently announced (Hooglund 1982: 113). In good crop years, wheat was often sold at up to 40 per cent less than the official prices.

Third, the Cereals Organization, which was responsible for carrying out the government's wheat policy, was not ready to buy all the wheat offered at the official minimum price. The purchases of this organization comprised only a small proportion of total wheat production and never reached even 6 per cent of total production by 1974.[15] Consequently, even after a considerable increase in the official floor price in 1974 and the government's decision to expand its purchase, most of the peasants did not regard sales to the Cereals Organization as a realistic alternative (see Aresvik 1976: 145; Hooglund 1982: 113). As late as 1975, only about 8 per cent of total wheat production was sold to the Cereals Organization. Pre-harvest buyers, dealers, shopkeepers, etc., continued to be the major, and often the only, alternatives which were available for peasants wishing to sell their products.

On the whole, this policy did not benefit wheat growers and failed to stimulate them to invest in yield-improving practices. It also did not succeed in pegging the wholesale prices of wheat which fluctuated widely during this period (Aresvik 1976: 145–6). In order to control the price of bread in the urban areas and fill the gap between the growing demand and sluggish supply of wheat, the government increasingly imported wheat. The amount of imported wheat increased from 127,000 tons in 1962 to 1,440,000 tons in

1975 (Bank Markazi Iran, *Annual Report and Balance Sheet* 1966, 1976; the figure for 1962 includes both wheat and flour imports). It is worth noting that the total cost of each ton of imported wheat exceeded the price of each ton of wheat which the government was ready to pay to the Iranian farmers. For example, the imported wheat cost the government 13,400 rials per ton in 1975 while it paid only 10,000 rials for domestic wheat (Azkia 1986: 165–6; see also Aresvik 1976: 149). The government subsequently sold the wheat to millers at 10,000 rials per ton. In other words, the imported wheat was subsidized by 3,400 rials per ton (see Azkia 1986: 165–6; Aresvik 1976: 149).

Broadly speaking, the stagnation of staple food production adversely affected the agricultural growth rates in the post-reform years. According to the official figures, agriculture grew at the annual rate of 4.5 per cent during the period 1962–7, 2.6 per cent during the period 1967–72, and 4.9 per cent during the period 1972–7 (Katouzian 1983: 320–1). Considering the population growth rate of 3 per cent per annum, per capita agricultural production might have increased slowly after the reform. This is clearly reflected in the FAO statistics which are illustrated in Table 7.15. Such growth rates, however, overestimate the real growth rates. These figures are in all probability based on the statistics of the Ministry of Agriculture. Many researchers estimate that the real agricultural growth rates might have been 2 to 3 per cent per annum after the reform (see among others, Katouzian 1983: 320–1; Dumont 1978: 8–9; Halliday 1979: 126–7).

Table 7.15 Production indices of agriculture, 1967–77 (1969–71 = 100)

Year	*Total agriculture*	*Per capita agriculture*
1967	87	95
1969	99	102
1971	100	97
1973	117	107
1975	128	110
1977	132	107

Source: FAO, *Production Yearbook* (1979: 78, 82)

In sum, the growth of most of the agricultural products, especially staple crops, failed to meet domestic needs. The government then imported an increasing amount of agricultural products in order to satisfy domestic demand, particularly after sharp increases of oil revenues in 1973. As a result, the gap between the value of exports and imports of agricultural products widened

rapidly after the reform and the gain to the country from exporting some cash crops fell far below its costs of importing food products (see Table 7.16).

Table 7.16 Value of agricultural exports and imports, 1961–73 (million rials)

	1961–3	*1966–8*	*1971–3*
Exports	6,167	8,521	15,762
Imports	6,353	9,140	25,069
Balance	−186	−619	−9,307

Source: Aresvik (1976: 224)

We must point out that our theoretical analysis concluded that there could be rapid agricultural growth even in a bimodal agrarian system (see Chapter 4, pp. 59–61). This implies that the sluggish agricultural growth in Iran can hardly be explained by the post-reform agrarian system alone. There must be some other factors behind the unsatisfactory performance of the agricultural sector as well. The oil-boom of the early 1970s was perhaps the most important contributing factor.

The 'Dutch disease' theory provides a theoretical ground to believe that the boom in one sector of an economy, such as the oil industry, might cause a decline in another sectors, e.g. agriculture. In its simplest version, the theory argues that a booming sector would influence a non-booming sector through a resource movement effect and a spending effect (see, among others, Corden and Neary 1982; Benjamin *et al.* 1989; Scherr 1989). The former draws resources out of the non-booming sector while the latter turns relative prices against it. We saw earlier that a mass rural–urban migration followed the 1970s' oil-boom in Iran which presumably caused serious problems for the peasantry. This resource movement out of agriculture must have had some negative impact on the performance of the sector.[16]

We maintain, however, that the 'Dutch disease' may damage an agricultural sector more seriously in a bimodal agrarian system than in a unimodal system. The intersectoral mobility of production factors is likely to be greater in a bimodal system and this makes the agricultural sector more vulnerable to the pressures of a booming sector. To put it more precisely, landless labourers and small peasants who can hardly make a living in agriculture would very easily respond to the employment opportunities elsewhere. On the contrary, peasants with viable farms would likely need a much stronger incentive to leave their villages. Moreover, it is relatively

easy for large landowners, who have very often had other non-agricultural activities, to move their capital to more beneficial sectors. Small farmers have usually little capital and lack the experience to do so.

Concluding remarks

Generally speaking, the differences between large and small Iranian farmers were more noticeable in relation to the size of their land and their access to irrigation water in the initial period of the post-reform era. But, as expected in a bimodal agrarian system, the differences tended to spread to every aspect of their economic lives. To begin with the large farmers could more easily receive large, long-term credits especially from the government institutions, sometimes at very low rates of interest. On the other hand, the peasants had access generally to small, short-term loans which were extended to them through the rural co-operative societies. Not infrequently, they could not even receive such loans or the loans fell short of their immediate needs. Consequently, they had to borrow money from moneylenders at very high interest rates.

There were also growing differences between the production methods used by the two groups. The large farmers were generally inclined to modernize their farming, mainly through capital-intensive methods, while the majority of the peasants continued to rely on traditional methods of production, which meant they continued to depend on the normally abundant peasant labour force. However, the oil-boom and the following mass rural–urban migration deprived many peasants of their relatively cheap labour. In some cases, this led to a reduction of the cultivated area of peasants' farms.

Finally, large farmers tended increasingly to substitute cultivation of more profitable cash crops for that of the staple crops. it was also the increase in cash-crop production that was responsible for much of the agricultural growth, which at most seems to have matched population growth after the reform. This implies that the production of staple food crops lagged behind the rapidly increasing food demand which resulted from population growth and income rises. The country was then forced to import an increasing amount of food products each year. In general, the government's agricultural policies and the oil boom tended to reinforce the above process.

Chapter eight

Rural poverty and inequality after the reform

The previous chapter showed that the development of a bimodal agrarian system after the land reform in Iran was accompanied by certain imperfections in the rural factor markets and resulted in growing differences between large and small farmers with respect to their farming methods and the kinds of crops that they grew. In addition to land, large farmers enjoyed easier access to water, employed modern, capital-intensive techniques, and increasingly cultivated cash crops. On the other hand, most of the peasants continued to use the traditional methods of cultivation and predominantly grew wheat and barley.

We did not, however, discuss the effects of the changes that followed the land reform on rural poverty and inequality. According to the theoretical analysis in Chapter 4, if a land reform leads to a bimodal agrarian system, it can hardly improve the lot of the majority of the rural people and reduce rural inequalities. The aim of our concluding chapter is to see if this was true in the Iranian rural areas after the reform. In other words, we want to find out whether the rural poor in Iran experienced any meaningful improvement in their living conditions after the reform.

Prevalence of subsistence farming

We saw earlier that the post-reform conditions in the Iranian rural economy generally preclude the majority of peasants from investing in land and improving their traditional methods of cultivation. As a result, many peasants could hardly have experienced substantial increases in their production after the reform. Considering the fact that the peasants predominantly produce food crops and usually sold only their surpluses, Table 8.1 suggests that a large proportion of the peasants lack any surplus over and above their own consumption. This is especially true for small farmers because that part of their crops which they sold was not always in excess of

their own consumption: they sold it because they needed to buy other necessities such as sugar and tea (Khosrovi 1981: 27). However, we can see from Table 8.1 that there were substantial regional variations. In more prosperous regions, e.g. Gilan, peasants were more market-oriented particularly because of favourable natural conditions and/or better access to the urban market. On the contrary, almost all peasants were subsistence farmers in less fertile regions such as Sistan and Baluchestan where a negligible percentage of farmers produced mainly for the market.

Table 8.1 Distribution of farmers by share of crop sold and by farm size, 1974 (percentage)

Farm size	*Not sold*	*Partially sold*[a]	*Mainly sold*[b]
< 10	53	27	21
10–50	48	28	23
< 50	2	1	97
All holdings:			
1. Iran	51	27	22
2. Gilan	26	25	49
3. Markazi	34	34	32
4. Sistan and Baluchestan	86	10	4

Source: Khosrovi (1981)

Notes: a Less than 50 per cent of the produce was sold
b At least 50 per cent was sold

As we can see, the information provided by Table 8.1 concerns only the annual crops, especially wheat and barley. Nevertheless, considering other agricultural products, such as fruits and poultry, does not change the general picture. According to the available data, a comparatively larger proportion of orchard products was sold by the peasants while the share of their sale of livestock production was lower than those of the annual crops (Khosrovi 1981: 21, 52). A higher rate of commercialization of permanent crops, especially among small farmers, might reflect the fact that some small farms consisted entirely of orchards and were market-oriented in nature. If we could exclude the marketed products of this group of farmers, we would very likely find out that the rest of the orchard products were principally consumed by the peasants themselves. On the whole, it is plausible to conclude that self-consumption was still the prime motive of most peasants in post-reform agriculture (cf. Khosrovi 1979: 98–102; Hooglund 1982: 93–4).

We should note, however, that self-consumption does not necessarily mean self-sufficiency. A rough estimate, for example,

indicates that peasants with less than 2 hectares (i.e. 43 per cent of the Iranian farmers) produced only 28 per cent of their total wheat consumption in 1978 (Khosrovi 1982: 30). Despite the shortcomings of such estimates, it is reasonable to believe that the production of many peasants, especially the smaller ones, was less than their consumption.[1]

One reason for the lack of improvement in the production of most small peasants and the persistence of the subsistence economy in most of the Iranian villages must have been the small size of the peasants' farms (see Hooglund 1982: 93–4). In fact, given the prevailing farming practices, more than 70 per cent of the farmers received farms which were smaller than the size which is generally considered to be essential for a subsistence standard of living, i.e. 7 hectares (see Chapter 6, pp. 108–9). This was, however, a less important factor compared with the post-reform agrarian system. If the peasants had not been deprived of improving their production methods by the bimodal mechanism, they might have largely overcome the size problem and increased their production. This fact is clearly reflected in a study carried out by the World Bank in 1975 (see Price 1975, cited in Aresvik 1976: 102–4).

Table 8.2 Approximate size of farms required from $500 income per capita (hectares)

Type of farm	*Common practices in 1975*	*Improved practices*
Dryland cereal farm	70	40
Partly irrigated field-crop farm[a]	9	5
Intensively irrigated field-crop farm[b]	5	3.3
Rice farm	2.5	1.5
Orchard	0.6	0.35

Source: Aresvik (1976: 104)

Notes: a 60% wheat, 30% sugar beets or similar crops and 10% fallow
b 90% sugar beets, vegetables, and similar crops, and 10% fallow

According to the above study, a peasant family would attain an income of $500 per member in 1980 from a considerably smaller farm if the peasant was able to modernize his farm. For example, we can see from Table 8.2 that applying improved practices would almost halve the size a partially irrigated field-crop farm required to meet the target income. Table 8.2 also reveals that the income of a great number of small farmers must have been considerably lower than the target income in 1975: their farms were comparatively too small with respect to the fact that they suffered from insufficient

access to water, applied the traditional methods and grew mainly wheat and barley. This issue, however, will be discussed in more detail in the next section.

Persistence of low levels of living in rural areas

The lack of sufficient data limits our study of the developments of the rural levels of living after the reform. The available information, however, tends to suggest that rural poverty persisted in the Iranian rural areas and the gap between rich and poor villagers might have widened.

For example, Table 8.3 which is based on a World Bank study, shows that one decade after the reform the average income of a large majority of the rural population was very low. The income of about half of the rural people even fell short of the per capita consumption in rural areas in the same year, which was about $102 (see Parvin and Zamani 1979: 47; we assumed that 73 rials = $1). This means that these people, who were principally landless labourers and the very small farmers, could not even afford the average rural consumption, which was lower than the nation-wide consumption level (see Table 8.4).

Table 8.3 Per capita rural income, 1972 (US dollars)

Size of farm (hectares)	*Mean income*	*% of rural population*
<3[a]	70	47
3–10	131	33
11–50[b]	302	19
>50	1,000	1

Source: Halliday (1979: 132)

Notes: a Includes landless labourers
b Includes animal herders

Table 8.4 which shows per capita consumption expenditures can be used to throw some light on changes in levels of living in rural areas. Accordingly, rural consumption increased very slowly between 1962 and 1972 and rose sharply after the oil-boom in 1973. Considering the fact that official statistics were notoriously inclined to report good results, even Table 8.4 suggests that there had been little change in rural standards of living up to 1973. All in all, the official figures show that rural consumption increased by about 40 per cent between 1962 and 1975. There are some facts, however, which indicate that a large number of rural families might have hardly enjoyed any improvement in their levels of consumption.

Table 8.4 Per capita consumption expenditures of the urban and rural households, selected years (1959 prices)

Year	*Rural*	*Urban*	*Rural as % of urban*
1959	7,012	14,923	47
1962	7,638	16,502	46
1965	8,375	16,277	51
1968	8,590	22,027	39
1971	8,036	25,866	31
1973	8,351	31,843	26
1976	10,706	40,789	26

Source: Parvin and Zamani (1974: 47)

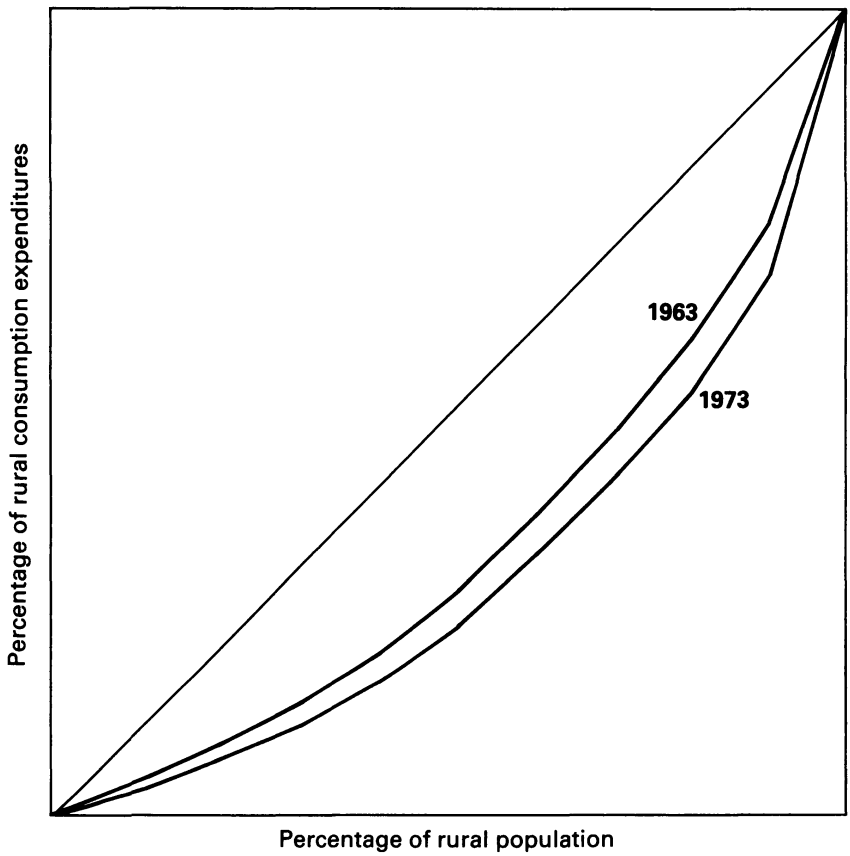
First, Iran is characterized by numerous regional variations and, therefore, the peasants of a region, or even a village, might have considerably improved their income and consumption while some other peasants in their neighbourhood might have experienced the opposite. For example, a comparison between incomes of different regions in 1977 reveals that the peasants in West Azarbayjan earned, on average, more than 2.4 times the average income of the peasants in Kohkilouyeh and Bouyer Ahmadi (Statistical Centre of Iran 1981: 156). Similarly, monthly expenditures of a rural household in different regions varied between 11,652 and 25,392 rials in the same year (Statistical Centre of Iran 1981: 128–35; see also Asayesh 1974: 116).

Second, as Figure 8.1 demonstrates, consumption inequality increased after the reform. Thus, some peasants must have experienced a larger consumption increase than others. In fact, it has been estimated that the consumption level of the bottom 40 per cent of the rural population deteriorated between 1963 and 1973: the monthly consumption expenditures of these people declined from 2,500 rials to about 2,400 rials or less in this period (Parvin and Zamani 1979: 49).[2]

We should point out that income inequalities in the rural areas might have been more pronounced than consumption inequalities because the rich often have some savings while the poor might frequently face income deficits and thus dissave (cf. Parvin and Zamani 1979: 47–8; Katouzian 1983: 344). We may then conclude that alongside the consumption inequalities, the distribution of income also became more skewed. This is of course not unexpected in a bimodal agrarian system.

The mass rural–urban migration must have alleviated rural poverty to some extent and improved the lot of those who migrated. Nevertheless, this movement could not have led to any significant

Figure 8.1 Lorenz curve for the rural consumption of expenditure



Source: Parvin and Zamani (1979: 48)

improvement in the situation of the landless labourers who remained in the villages. The small farmers could hardly afford to employ wage labour and the large farmers increasingly mechanized their farms. The availability of cheap Afghani workers did not make things easier for the rural labourers. We should remember that migration might have affected many peasants negatively since it drained them not only of surplus family labour but also of essential workers.

Irrespective of small differences, however, the rural poor generally consisted of both landless labourers and small peasants. Considering our earlier conclusion that the production of a large number of small farmers was insufficient for their own

consumption, one may believe that the income of many of the rural poor was not enough to reach a subsistence level. As a matter of fact, the statistics for rural incomes, which are only available for the recent years, reveal that total rural incomes fell below total rural consumption expenditure. For instance, total rural incomes in 1976, 1977 and 1978 covered only 79, 80 and 82 per cent, respectively, of total consumption expenditure (Statistical Centre of Iran 1984: 753–4). The small farmers and landless labourers must have been principally responsible for the deficits.

In many cases, field studies substantiate the above conclusions. They report that there was often a wide income gap between the villagers and that a large number of them suffered from income deficits (many examples can be found in Momeni 1980: 383–94; see further Rafipoor 1985: 206–7; Antoun 1976: 4; Mahdavi 1982: 62). In extreme cases, as many as 90 per cent of the villagers claimed that their income did not meet their living expenses (see Momeni 1980: 389–90). This may partly explain why most of the loans which were received by the peasants were spent on consumption purposes (see Chapter 7, pp. 113–19). Moreover, population growth, higher rises in the prices of the goods bought by the peasants than in the prices of their products, repayment of former debts, and lack of any significant improvement in their production might have contributed to increase income deficits and indebtedness in the rural areas.[3]

It seems that the conditions of most of the shareholders of farm corporations were not much better. These corporations, as we saw earlier, enjoyed generous government support and there are studies which show that the average income of their shareholders grew considerably after the formation of the corporations (see, for example, Denman 1978: 288–92; Azkia 1986: 246–7; Stickley and Najafi 1971: 19). But most of these studies also point out that the number of shares which each shareholder held varied greatly. Sometimes the difference between the number of shares of different shareholders of a farm corporation was about 200 shares (Khosrovi 1979b: 177; see further Momeni 1980: 368–70; Azkia 1986: 236; Mohammadi 1973: 36–7; Stickley and Najafi 1971: 19). Thus, there were plenty of shareholders whose incomes, which mainly consisted of their dividends, were hardly sufficient for life at subsistence level. They had to seek additional incomes to support themselves and thus often migrated to cities (Momeni 1980: 369–70; Azkia 1986: 242–51). This was probably one of the principal reasons that contributed to the dissolution of 88 out of the total 93 corporations and their subdivision among their shareholders, i.e. mainly the former peasants, after the 1979 Iranian Revolution.[4]

There is no question that some rural people improved their welfare after the reform. Many new, modern houses were built in rural areas, the number of the villagers who owned sewing-machines, bicycles, refrigerators, radios, etc. was on the increase, and so on. But these improvements were limited and failed to bring about any remarkable change in the living conditions of the majority of the rural people. Furthermore, these changes tended to exacerbate rural inequality.

The official statistics show that as late as 1977, about 58 per cent of the villagers had radios, 10 per cent had bicycles, 7 per cent had refrigerators, and only 3 per cent had television sets (Statistical Centre of Iran 1981: 120). It is evident that, with the exception of radios, only a minority of the rural people could have possessed such living facilities. This is clearly reflected in a sample study of the villages of a relatively less prosperous region (Yazd): 30 per cent of the sample peasants had no valuable belongings, including agricultural implements, and the total value of all belongings of another 33 per cent was less than 10,000 rials ($141) per household in 1980 (Rafipoor 1985: 216).

Table 8.5 provides further insights into the rural poverty. It illustrates, in general, that more than one decade after the reform a large proportion of rural families still lived in one or two rooms without electricity or piped water. Table 8.5 also indicates that there were large differences in housing conditions between the rural and urban areas. The fact that the majority of villagers who migrated to the cities lived in shanty towns may lead one to presume that most of the ill-conditioned houses in urban areas belonged to the rural migrants as well (see Kazemi 1980: 46–53).

Last but not least, the household expenditure surveys for rural areas in the period 1976 to 1978 reveal that around 50 per cent of

Table 8.5 Housing in rural and urban areas, 1976

	Rural areas		*Urban areas*	
Description of houses	*No. (thousands)*	*%*	*No. (thousands)*	*%*
Made of traditional materials[a]	2,381	81	676	28
Without electricity	2,512	86	234	10
Without piped water	2,297	78	255	11
With one room	568	19	167	7
With two rooms	1,029	35	557	23

Source: Statistical Centre of Iran (1984)

Note: a This includes all houses which are made of sun-dried brick and wooden beam, sun-dried brick and mud, wood, straw, etc. as well as tents

the consumption expenditure was devoted to foodstuffs. This percentage was probably higher for the poorer villages. Nevertheless, there is reason to believe that many of them were undernourished. A study from 1972 shows that between 30 to 40 per cent of the rural population suffered from malnutrition: 700,000 of these people, or 4 per cent of the villagers, were severely undernourished in the sense that they received 75 per cent or less of a certain minimum calorie requirement (Katouzian 1983: 346–8).[5]

Conclusion

The Iranian agrarian system in the early 1960s displayed the major features of traditional tenancy systems that have often been criticized for the prevalence of agricultural backwardness and rural poverty in many developing countries. Sharecropping was the predominant form of tenure in the Iranian peasant–landlord system. Lack of sufficient information, however, precludes us from drawing any conclusions regarding the efficiency implications of sharecropping. We can only say that the high degree of uncertainty, which was a salient feature of Iranian farming, might have been a principal factor for the widespread occurrence of sharecropping in Iran.

Broadly speaking, the concentration of land in the hands of a few absentee landlords has been the key element in most criticisms of traditional tenancy systems. Therefore, the redistribution of land among the tenants has usually been considered to be a measure that would promote agricultural and rural development. We have argued, however, that this depends on whether a bimodal or a unimodal agrarian system emerges out of a land reform programme. The economic consequences of the two systems are quite different.

A unimodal agrarian sytem may follow a radical land redistribution programme and consists of family farms. Our analysis shows that this system would bring about an immediate income redistribution in favour of the majority of farmers and generate agricultural development. The essence of the argument is rather simple: the farmers have every incentive to use their resources, especially the abundant family labour, to increase their output and capital. Moreover, the important point is that the prevailing circumstances in a unimodal system as well as government policies would presumably encourage the peasants.

On the other hand, a land reform can redistribute only a limited amount of land but succeed in convincing the large landlords that they have to give up their traditional behaviour and become modern agriculturists. This leads to a bimodal agrarian system in which

very small peasants coexist with large landowners. We maintain that the large landowners would enjoy both monopoly power in the market and government support and may increasingly use modern, imported technology. As a result, they would tend to become richer. On the contrary, the small peasants would fail to improve their situation because the imperfect markets and government policies often operate against them. They may eventually experience a fall in their standard of living.

The land reform experience in Iran exemplifies the latter case. As a consequence of the land reform programme which started in 1962 and of generally bimodal-oriented government policies, a bimodal system replaced the Iranian landlord–peasant system. The large farmers in post-reform agriculture owned a relatively large proportion of the arable land, had easier access to water and capital, and enjoyed government support. They were therefore encouraged to invest in capital-intensive methods and cultivate an increasing amount of cash crops. The growing of cash crops turned out to be more beneficial especially as a consequence of the government policy to control the price of wheat, which is the major staple crop in Iran.

In contrast to the large farmers, the Iranian small farmers generally owned small patches of land, often faced difficulties with respect to the irrigation water, and could hardly obtain sufficient credits. Moreover, no serious attempt was made to introduce follow-up programmes after the reform and the small farmers were practically deprived of government support. The small farmers had by and large little chance to improve their farming and their standards of living under the conditions which prevailed in Iranian agriculture after the land reform. They continued to cultivate the staple crops, largely by traditional methods. They produced these crops chiefly for their own consumption; nevertheless, the production of many of the peasants was insufficient to support their families.

There were also large numbers of landless labourers who were among the poor in the villages. They generally received no land under the land reform and their number in fact increased as some tenants were legally or illegally deprived of their rights to receive land. Decreasing agricultural employment opportunities as a result of the tendency towards mechanization among the large farmers in addition to poverty and small size of land on the part of peasants put further pressure on the rural labourers.

Sharp increases in oil income and the resulting urban boom might have come as a blessing for many poor people in the rural areas. The following mass migration to cities might have raised the

well-being of those who left and reduced somewhat the pressures on those who did not. Nevertheless, only a small group of rural people experienced increasing improvements in their welfare and poverty remained the lot of the majority.

Glossary

arbab	landlord; master
barzigar	cultivator; the lowest member of a *buneh*
buneh	a team of peasants who grouped together to cultivate a specified amount of land
dang	one-sixth part of a village
iqta	land assignment; *tuyul*
juft	plough-land; an amount of land which is approximately between 4 to 6 hectares
khaliseh	state land
khwushnishin	a villager who did not have cultivation rights before the land reform; a landless villager
nasaq	a cultivation right of the village land before the land reform; the lay-out of a village lands
nasaqdar	a villager who had cultivation rights
pabuneh	an assistant of a *sarbuneh*
qanat	underground water channel
raiyat	a peasant who cultivated the land of a landlord in accordance with the local customs
sarbuneh	the head of a *buneh*
vaqf	land endowed for certain purposes, usually religious or charitable

Notes

Chapter 1 Introduction

1 In some cases, we had no access to the original statistics and had to rely on second-hand sources. If it is not stated to the contrary, even these statistics should be regarded as official.
2 For example, Research Group (1962), which studies the production of wheat and barley in Zanjan region, concludes that the statistics of the Ministry of Agriculture were extremely exaggerated for this region and prefers those of the Statistical Centre.
3 The reported number of villages varies greatly. At the beginning of the land reform in 1962 it was reported to be about 49,000 villages. Later reports give much higher figures. According to the 1973 agricultural census, there were 59,189 villages and 23,122 independent and attached farms. About 99 per cent of the rural population lived in the villages. See Ashraf (1979). See also Momeni (1980: 22–4) for a discussion about the number of villages.
4 For instance, most villages in the south-east and east of the country consist of two or three peasant families. On the other hand, in Azarbayjan, villages are much larger and often have several thousand people (Lambton 1969: 8–9).

Chapter 2 Agricultural backwardness in theory

1 It should be considered that since the end of the Second World War a great number of agrarian policies, especially land reform, have been adopted in many developing countries. As a consequence, agrarian systems in some of these countries might have changed partially or completely over the past four decades. See, for example, Dovring (1970), Warriner (1969) and King (1977).
2 There is little empirical evidence for notoriously high consumption propensities of landlords. For an attempt to provide such evidence for Latin America, see Stenberg (1970).
3 Nevertheless, sharing the crop might be seen as a risk-sharing device which can provide an incentive for a sharecropper to innovate. For further discussion, see pp. 19–22.

4 This is what Lipton (1968) calls the 'survival algorithm'. That is to say, poor farmers cannot afford a harvest failure, which means starvation or loss of land, and therefore focus on survival rather than profit under conditions of uncertainty. In other words, they avert from the risks which jeopardize their 'certain' subsistence harvest and thereby dispense with the expected risk premium. For further discussion see *ibid.* and Weeks (1970).

5 Johnson (1950) provides a brief summary of this traditional view about sharecropping and an expansion of this summary can be found in Cheung (1969).

6 For a fixed-rent tenant the rent is a fixed cost and his marginal cost will be equal to that of an owner-operator and his labour inputs for income maximization becomes equal, too.

7 Note that equation (2) is the core of inefficiency argument, i.e. tenants apply labour until their share of marginal products equals the wage rates.

8 Johnson emphasizes the last practice. Griffin (1976: 122), too, believes that insecurity of tenancy is in fact a means by which landlords are able to get rid of bad tenants who are not efficient.

9 In fact, Johnson accepts that theoretically sharecropping is less productive than fixed-rent tenancy although the two may be identical in practice (Cheung 1969: 49–51).

10 Cheung's landlord allows his tenant to receive only an income equal to his alternative wage earning. To maximize his income, the landlord divides his total land, s, equally among m tenants so that $H = s/m$ and his total rent is $y_1 = mrQ$. He then maximizes his income by choosing m, r, and L, subject to the constraint that each tenant receives $y_t = WL$. That is to say:

$$\underset{(m,r,L)}{\text{Max}} \quad y_1 = mr\text{F}(H,L), \text{ subject to } WL = (1-r)\text{F}(H,L) \qquad \text{(I)}$$

Considering the fact that $H = s/m$, we get the following necessary conditions:

$$\text{F}_H = r\frac{Q}{H} \qquad \text{(II)}$$

$$\text{F}_L = W \qquad \text{(III)}$$

$$r = \frac{Q - WL}{Q} \qquad \text{(IV)}$$

We can see that equation (III), in contrast to equation (2), implies that sharecroppers apply labour as much as they otherwise would do.

11 Bardhan and Srinivasan (1971: 52) argue that an atomistic landlord in a competitive market takes r, which is determined in the market, as given. See also Koo (1973: 572).

12 Sen (1975: 67) also refers to this problem. He argues that it is one thing to stipulate the number of men who have to work in a piece of land, but it is quite another to specify the amount of labour hours to be applied, and still more complex to have a contract on the exact intensity of effort. See also Stiglitz (1974: 242–3) for further discussion of this problem.

13 Bardhan and Srinivasan (1971) assume that the tenant devotes part of his labour to the sharecropped land and sells part of it in the labour market at a given wage rate. The landlord also rents part of his land to sharecroppers and retains part which he cultivates with hired labour. The utility of each agent is then a function of his total income and his leisure.

14 This is a point which was initially brought up by Newbery (1974) who argues that Bardhan and Srinivasan's model shows that, at any rental share, peasants want to rent more and more land until its marginal utility falls to zero. This makes the demand for land exceed its supply.

15 In this system, there are three kinds of contracts, namely sharecropping, fixed-rent tenancy, and own cultivation with hired labour. Then the above assumption means that the tenant has to devote all his labour to work on the land he sharecrops and rents under fixed-rent tenancy, and what remains of his total labour has to be sold in the market. The landowner, in his turn, has to divide all his land between sharecropped land, fixed-rented land, and own-cultivation.

16 The total gain is L_1JBL_2 which is greater than the landlord's total cost L_1ACL_2. It is also greater than the wage cost to the tenant, L_1ABL_2.

17 On incentive creating features of sharecropping, see further Stiglitz (1974: 242–51) and Reid (1977: 405).

18 Accordingly, Lucas compares the two systems with Cobb-Douglas production functions, separable utility functions, and a linear monitoring cost function. He finds out that, 'the share-cum-wage economy is . . . technically efficient in the sense of equating marginal products of factors in both uses' (Lucas 1979: 519).

19 At the same time, Johnson admits that the data were very rough and the lands rented under the two lease systems were not proved to be comparable (Johnson 1950: 118).

20 Provided that the sharecropper alone stands for all the costs, as is normal in many traditional rural economies. Adam Smith (1937: 367) puts the above argument in this way: 'It could never, however, be the interest [of the sharecroppers] to lay out in the further improvement of the land any part of the little stock which they might have saved from their own share of produce, because the lord, who laid out nothing, was to get one-half of whatever it produced.'

21 We should recall that Cheung's sharecropper (see p. 13) is satisfied with receiving his alternative wage earning and is passive in decision making. The landlord, however, makes sure that the tenant receives his wage-equivalent share anyhow. Such a tenant would hardly use and take care of the new investments properly without some extra supervision costs.

22 Bardhan and Srinivasan conclude further that such a tenant would prefer the landlord not to share costs if he had the choice (Bardhan and Srinivasan 1971: 61–2).
23 In fact, Bell (1976: 183, 188) concludes that a landlord is more likely to prefer a cost-sharing: (i) the higher the rental share, (ii) the larger the returns to scale in all inputs combined, (iii) the more risk averse his tenant relative to himself, and (iv) the greater the production risks.
24 In a rich economy, y_{ms} is likely to be substantially above y_{mp}, reflecting mainly an average Q which is much greater than y_{mp}. As the risk of falling below y_{mp} or y_{ms} becomes very small, a tenant would try mainly to attain his desired income level, y_D. He is ready to bear more risk to increase his chance of achieving this income level. A fixed-rent contract provides such an opportunity and is thus more attractive.
25 There are also several models which try to explain the occurrence of sharecropping mainly on the basis of transaction costs. See, for example, Agrawal (1984) and Pertev (1986) who focus on supervision costs.
26 For example, Newbery (1977) argues that production decisions in agriculture are affected by various unforeseen events during a crop year. This will affect the demand for labour over the year in a random way and thus the labour market cannot guarantee full employment at a constant, or even predictable, wage rate. It is then plausible to assume that wage rates are risky and equal to $w\phi(s)$ because the state of the world, s, is not known. He then introduces this factor first in a system where only fixed-rent tenancy is available and, then, in a system where both fixed-rent tenancy and sharecropping are available. He finds out that only the latter system is efficient.
27 Family labour cannot, or will not, always move away from the family farms for a variety of reasons: they may simply prefer to work on the family farm, they may be forbidden by custom to accept outside work, as is the case for women and children in some places, they may be ready to work part-time but restrained by the problem of location, etc. See further Mazumdar (1975: 261–4), Bardhan (1973: 1379–81), and Sen (1975: 53–5).
28 In other words, the area WEA = the area ABC. We should remember that the landlords in Cheung model stipulate a minimum labour input at L_2. See p. 13.

Chapter 3 Pre-reform Iranian village

1 This was an important land category before the seventh century but, for a variety of reasons, such as the increase of land assignment, commendation and usurpation, its volume decreased sharply. For more details see Nomani (1980: 174–87).
2 In fact, it was not until the eleventh century that land assignment became the dominant type of landholding under the name of *iqta*.

Later, the term was gradually replaced by *tuyul*. See Lambton (1953: 28) and Nomani (1980: 201–28).

3 Theoretically, *khaliseh* (crown land) and *divani* (state domains) were different. The income of the former was supposed to finance the expenses of the court while that of the latter was supposed to be spent by government for the whole society (Nomani 1980: 151–74). In practice, they were not really distinguishable. In modern usage *khaliseh* is state domains and *amlak*, the crown estates. See pp. 32–3.

4 This process began when 'The need for more money to pay for extravagances of the court, to provide for administration of the country and the upkeep of the army, became ever more pressing in the latter part of the Qajar period' (Lambton 1953: 151).

5 Since the 1930s several laws passed which dealt with these lands: in 1937 for the sale of *khaliseh* around Tehran, and Sistan, in 1947 for *khaliseh* in Khuzestan, and in 1955 for the sale of state domains in general. See Lambton (1953: 240–55), Homan (1955: 286–97), and Denman (1978: 260–2).

6 Land property was expressed in terms of villages and parts of villages in Iran. Each village consisted of six equal parts, regardless of its size, and each part was called a *dang*. There were further subdivisions, as well.

7 Lambton (1953: 269–71) reports the results of an unpublished study which showed that the areas with higher average annual rainfall coincided with a higher degree of land concentration.

8 It is worth noting that there were other landlords who owned shares in different villages and were called small-owners, but were virtually large landowners if one considers the sum of their shares. In official statistics such large landowners, small-owners, and peasant proprietors are all called *khordeh malekin*, and therefore it is difficult to estimate the amount of land that smallholders owned. However, *khordeh malekin*, i.e. those with less than one *dang* in any one village, owned 42 per cent of villages (Kazemi and Abrahamian 1978: 269).

9 There are, however, great variations in the estimates of state domains. See, for example, Ashraf and Banuazizi (1980: 18) and Azkia (1986: 86).

10 Although theoretically a *vaqf* cannot be dissolved and thus usurped, there have been numerous cases of usurpation of such lands throughout history. Consequently, to endow lands as personal *vaqfs* was not always a safe protection against usurpation. See Nomani (1980: 189–91) and Lambton (1953: 131–2) for some examples.

11 It should be emphasized that all four land categories described in the foregoing section were used in rather similar ways. In state and endowed lands the renters or administrators acted as the landlords and followed the same fashion.

12 For discussion about sharecropping as a risk-sharing device, see Chapter 2, pp. 19–22).

13 A section of the Civil Code deals with sharecropping. However, the law approaches such contracts very generally and it mainly legalizes the system rather than trying to unify practices. An English translation of this section can be found in Lambton (1953: 402–4).

14 Saedloo (1979: 781–98) who discusses the inheritance issue in the villages in detail concludes that other heirs virtually benefited from the right as long as they lived in the family. If anyone of them left, he would receive a share of annual product even though it was sometimes very small.
15 There were also other common terms for plough-lands. The same is true for most of the terms used in this study. The terms used here should only be considered as the more common ones in the rural study literature.
16 In less fertile regions, and especially in peasant proprietor areas, and where ploughing was done by spade, the village holdings were reckoned by reference to a share of water. In such cases, the village water was divided in a varying number of shares of equal duration in time which could be used for irrigation (Lambton 1953: 7).
17 A headman is quoted to have said: 'Every year we must go to the landlord at harvest time and give him gifts for the opportunity of cultivating his land for next year. . . . Owners throw out gavband [group-farming] members at their slightest whim. If one of us forgets to say "Good morning" to the landlord he would throw us off the land' (Keddie 1972: 382). See also Azkia (1980: 32).
18 For a detailed list of the various tasks of the members of each *buneh*, see Safinezhad (1972: 46–9). Lambton (1953: 299–302) provides some examples of such tasks in several areas. Generally, a hierarchy existed within each *buneh*, *sarbunehs* at the top and *barzigars* on the bottom, and the hard tasks were pushed downward. See Safinezhad (1972: 30).
19 For more information about *qanats* see Chapter. 1, pp. 5–6.
20 For a discussion about sharecropping and the incentive question, see Chapter 2, pp. 11–15.
21 For example, one can see an extensive list of different practices in Lambton (1953: 308–29). See also Research Group (1965a: 116–20).
22 The importance of such problems are studied, among others, by Lambton (1953: 157–60, 283–94), Abrahamian (1975: 140–7), and Nomani (1980: 451–71).
23 For example in Semirom and Daran where spring water is used, a farmer who supplied the oxen, labour, manure and seed got about four-fifths of the wheat while a similar farmer received between a third to a half of the crop in Yazd, where the water is obtained from long deep *qanats* (Research Group 1965: 119). Reviewing the division of the crop in the province of Kerman, Lambton (1953: 312) concludes that the share of the peasant was considerably less than in most areas which was partly due to the heavy maintenance expenses of the very lengthy *qanats*.
24 It should be noted that rice was usually cultivated on a fixed-rent basis, in cash or in kind. See Lambton (1953: 321–2).
25 Soodagar (1978: 134) believes that two other factors strengthened the position of such sharecroppers. These crops required more skilled labour, and they were usually grown around towns by more socially conscious farmers.
26 The extra dues and their effects on peasant income is discussed later, see p. 44.

27 Safinezhad (1972: 30) shows that in some places the annual income of a *sarbuneh* was about 5 to 12 per cent, and that of a *pabuneh* was about 2 to 3 per cent higher than a *barzigar's* income. In some cases the share of the group subdivided equally but the group cultivated a small piece of land whose income accrued exclusively to the *sarbuneh*, or *barzigars* worked one day a year free for him (Lambton 1953: 299; Khosrovi 1979b: 79–80).

28 There are striking similarities between the cultivation methods that historians, travellers, etc., described centuries ago with what were in use in 1950s. For example, see Nomani (1980: 293–325).

29 These were what a government company bought and sold to farmers between 1952 and 1962 but we don't know how many of them were in use by 1962. Some tractors had been imported before but only 240 of them were in service by 1955 (Adibi 1955: 132).

30 It can be seen from Table 3.4 that machinery was concentrated in Mazandaran, where commercial farming was growing. In fact, a sluggish movement towards mechanization was under way before the land reform. This was much more rapid in the 1950s when machines were bought to cultivate the idle lands and pastures. This occurred particularly in the highly fertile plain of Gorgan in Mazandaran province. Most of the lands in this province were crown lands and it was the royal family members, and some other upper class entrepreneurs from capital who were developing mechanized farms here. See further Okazaki (1968) and Nikayin (1976: 78–80).

31 For detailed description of various tools used for cultivation see Wulff (1966: 260–77).

32 It is estimated that a medium length *qanat* of about 9.5 km required $13,500–34,000 for construction and 0.5 per cent of that for maintenance (Wulff 1966: 254). At the same time, different estimates show that a peasant's family annual income, on average, might have been between $100 and $250 (Platt 1970: 24–5).

33 Around Esfahan, towers were made to collect pigeon droppings which were used as manure for fruit trees. For further details see Wulff (1965: 269–70).

34 Many examples of traditional practices and beliefs can be found in an old agricultural handbook reproduced by Yavari (1981: 95–167). See also Bolukbashi (1981).

35 For more examples about advantages of some traditional practices, see Siyahpoosh (1981).

36 On the other hand, Bharier (1971: 133–4) shows that several major crops increased fairly well between 1925 and 1959, and the increase in the production of wheat, for instance, might have matched the population increase. He concludes, however, that 'most of the production increases were the result of extending the area of cultivation rather than of using more intensive methods'.

37 This was mainly true for villages on the borders of the central desert. For example, one study (OIPFG 1974) shows that some of the villages of Kerman, a province in the neighbourhood of the central desert,

depended almost entirely upon rug and carpet weaving. See also Khosrovi (1979b: 133).

38 Remember that a plough-land, or *juft*, was between 4 to 6 hectares. See p. 35 for further information.

39 For further information about different dues, see Lambton (1953: 333–5) and Momeni (1980: 48–9).

40 In fact, agricultural labourers comprised more than 80 per cent of the *khwushnishins* who were the villagers without cultivation rights. Shopkeepers, money-lenders, village artisans, bathkeepers, etc., constituted the rest of the *khwushnishin* population. It is believed that around 40 per cent of Iran's rural population were *khwushnishins*. However, the proportion of *khwushnishins* to all villagers varied greatly from village to village. In general, a greater percentage of the population were *khwushnishins* in the larger villages and a lesser percentage in the smaller ones. For further information see Hooglund (1973 and 1982: 28–34).

41 The author, however, argues that these figures might have been unusually low (Keddie 1968: 71–2).

42 Lambton (1953: 389) describes the common diets of a peasant very accurately: 'The main feature of his diet apart from bread is *abgusht*, or soup, made chiefly by boiling a little meat with split peas and occasionally with some other pulses. In summer, he supplements his diet in the fruit growing areas with fruit and vegetables such as cucumber. If he owns flocks he also eats a small quantity of cheese and curds.'

43 There were some governmental institutions which provided agricultural credit such as the Agricultural Bank of Iran and the Rural Fund. The limited loans of these institutions were mainly received by landlords (Lambton 1953: 380; Platt 1970: 30).

44 For example, the reports of Research Group (1965a, 1970b) show that between 2 and 23 per cent of loans were received from landlords.

45 Kazemi and Abrahamian (1978: 275–80) argue that, apart from the communication problems, lack of enough surplus to sell and thus of income to consume, the need for irrigation, linguistic, religious and ethnic differences, periodic attacks by nomadic tribesmen, and the landlords' efforts to perpetuate the insularity of the villages encouraged the communal solidarity within the villages and discouraged the villagers from integrating into the market economy.

46 Lodi (1965: 1) believes that by preharvest selling the peasants avoided the risk of a price fall at harvest time, as well.

Chapter 4 Theoretical analysis of land reform

1 For instance, land registration might be described as land reform since it is useful to know how land is owned, it paves the way for subsequent reforms, etc. But, as far as rural poverty and inequality is concerned, it is absurd to call such measures land reform when they hardly affect the

way of life in rural areas. See further Warriner (1969: xiv–xv) and Lipton (1974: 269–74).

2 It should be noted that in arid regions, water is the determining factor for cultivation. For example, the previous chapter indicated that in many parts of Iran land without water is worthless. It is therefore better to measure land 'in "efficiency units" (indicating its capacity to yield net-value-added per acre in conjunction with optimal hirings of non-land inputs) rather than in crude acres' (Lipton 1974: 272).

3 As is discussed below, land reform involves important economic, social and political changes and requires a favourable socio-political climate. In the absence of favourable conditions, 'substitute measures' which have considerably less impact on the status quo become attractive as alternatives. See Lipton (1974: 274–81) and Flores (1970: 896–8) for discussion about the reliability of such measures as alternatives in practice.

4 For further discussion about the politics of land reform see, among others, Warriner (1969: Ch. 1 and 3); Layman and French (1970); and Huntington (1970).

5 For different examples, see Bergmann (1977: 19–21), Huntington (1970: 385–7); Gittinger (1961); and Ashraf and Banuazizi (1980: 27).

6 Cheung (1969: 100–17) argues that even certain kinds of government policies cause displacement of the already efficiently allocated resources and will lead to inefficiency.

7 Since the two systems correspond to what Johnston and Kilby (1975: 127–81) describe as the outcomes of 'unimodal' and 'bimodal' strategies of agrarian change, we also call the two systems 'unimodal' and 'bimodal'. Cf. Dovring (1970: 2–4).

8 As Griffin (1974: 22) writes, 'there is a spectrum along which landownership or various sizes of farms will be distributed. . . . Nevertheless, it simplifies thought to imagine for the moment that farms are owned either by big landlords or by small peasants.'

9 The following two sections are largely based on Griffin (1974: 17–61).

10 There might be more than one large landowner in each local market who may then behave as collusive oligopolists (Lundahl 1984: 524).

11 For more details about price determination under these circumstances, see Lundahl (1984: Appendix).

12 For more information see Chapter 2, pp. 23–4, where labour-market dualism is explained.

13 It seems that many of the high-yielding seeds are of this nature because their applications are connected with other inputs which require capital. In fact, such innovations are packages of complementary changes. See Griffin (1974: 51–61), and Lundahl (1979: 582–5).

14 As was explained in Chapter 2, p. 10, poorer peasants may be more risk averse simply because they are poor and vulnerable to output variations. This may lead some policy makers to conclude that poorer peasants are 'tradition bound', 'conservative', and 'change resistant' (Weeks 1970: 32).

15 De Janvry (1981: 169–72) provides an explanation for the bias of government policies based on the political power of large landlords.
16 Ghose (1983: 4) argues that growing landlessness and declining levels of rural employment and rural wages in developing countries may reduce the effective demand for food in rural areas. This can reduce, or even outweigh, the demand-pull which tends to raise food prices.
17 This is largely in line with the traditional view of sharecropping. As a sharecropper becomes an owner-operator, he will have no reason to work less intensively than L_2 in terms of Figure 2.1. Then, not only total agricultural output but also employment increases.
18 Berry's and Cline's estimate of the potential output gains from such a land reform in six developing countries shows that even in Asian countries, where farms are predominantly small, total production can rise. Such an increase is estimated to be, for example, 19 per cent in India and 10 per cent in the Philippines (Berry and Cline 1979: 132–3).
19 Multiple cropping is a prominent example. Dorner (1972: 105) writes that the change from single to double cropping of rice raised the demand for labour by 80 per cent in Chekiang Province in China; the change from double to triple cropping increased labour requirements up to two or three times that of double cropping in Taiwan.
20 In cases where timely ploughing by tractors makes double cropping possible, or where tractors enable farmers to plough hard soil which otherwise would not be used, tractors are complementary with labour (Griffin 1974: 53). See also Misawa (1974: 322–4).
21 It is even possible that, under certain circumstances, land reform affects the non-land recipient adversely. In the final analysis, the post-reform situation of these people will largely depend on the changes in the demand for hired labour in the agricultural sector. See further Berry (1971).
22 As Lipton (1974: 296) points out, the situation will be different in Latin America where most landless labourers, without knowledge and experience in farming, might get land. The need for some follow-up is acute in such a case.
23 Even when it becomes necessary to hire a labourer now and then, he would in most circumstances need very little supervision, as his work is controlled by family members who are already in the workplace (Pertev 1986: 33).
24 For further information about these features of sharecropping see Chapter 2, pp. 22–3.

Chapter 5 An analysis of the Iranian land reform programme

1 For example, it is estimated that the televisions and cars which were stored in Iran in 1960 could meet the local demand of these goods for four years (Momeni 1980: 177).
2 The programme was suggested by the IMF and the World Bank as a

precondition for helping the country. See Mahdavy (1965: 135–6) and Bharier (1971: 95).

3 For a very good and well-documented survey of different views with respect to the agrarian question in Iran see Momeni (1980: 81–118, 206–11). See also Hooglund (1982: 36–48).

4 The above decrees were issued by the government of Dr Musaddiq in October 1952. His government, which was supported by a vast majority of people, was overthrown in a coup in August 1953. For further information about this period, see Cottam (1964).

5 The Soviet Union and her allies were broadcasting 132 hours of criticism per month toward Iran and there were three clandestine radio stations sending out anti-regime programmes (Ashraf and Banuazizi 1980: 26). See also Lambton (1969: 61).

6 In 1959, the then US Senator Hubert Humphrey said: 'I think there are revolutionary forces at work in Iran, and I think those forces are at work in the tribes and villages' (Ashraf and Banuazizi 1980: 27). In fact, Americans were convinced that land reform was necessary in Iran and insisted that Iranians should introduce a reform programme (Ashraf and Banuazizi 1980: 27; Fishburne 1964: 162–3, 264–74).

7 It is estimated that 50 to 60 per cent of the members of the parliaments were from the landowning class (Ajami 1973: 122).

8 See Momeni (1980: 144–52) for some information on the proposed bill and the passed law. See also Lambton (1969: 56–8) and Khatibi (1972: 63–4).

9 Quoted from Arsanjani, the Minister of Agriculture, in a press conference, where he gave information about a new land reform law in 1962. See p. 72.

10 It was normally both houses of parliaments, the *Majlis* and the Senate, which were to pass laws but, as we mentioned above, they had already been dissolved. For an English translation of the land reform law of January 1962 see *Iran Almanac* (1962).

11 Only charitable *vaqfs* were exempted while personal *vaqfs* were treated as other private lands whose beneficiaries were able to receive the revenue of a maximum of one village. For information about *vaqfs*, see Chapter 3, p. 32.

12 The Land Reform Council was under the chairmanship of the Minister of Agriculture and consisted of four high officials of the Ministry of Agriculture and the director of the Land Reform Organization. The interpretation of the reform law and drawing up the regulations were among the duties of this council. See Research Group (1970a: 49, 52).

13 The 36th session of the Land Reform Council, 18 December 1962.

14 The decisions were taken in the 43rd (7 June 1963) and 47th (25 August 1963) sessions of the Land Reform Council. In fact, the 43rd session decided that 'women landowners, in the same way as men and irrespective of their position in the household, are entitled to upper legal limit of land' (Research Group 1970a: 54–5). This practically implied that the wife (and perhaps the wives?) were entitled to retain one whole village.

15 This decision (on 28 January 1963) was in line with land reform officials who had not allowed landlords to choose such villages. See Research Group (1970a: 58).
16 The session of the Council on 24 July 1963 (Lambton 1969: 67).
17 The first decision was made at the 25th session in early months and the second decision was made at the 39th session on 18 March 1963.
18 The regional indices were between 100 and 180. For a detailed list of indices, see Denman (1973: 340–1).
19 According to the law, if a beneficiary fails to pay three of his annual instalments, his land could be confiscated by the government and sold to another. The law does not state anything about the repayment of the already paid instalments (Research Group 1970a: 67).
20 According to Lambton (1969: 206) oral instructions were given to the land reform officials not to allow the sale of the cultivation rights.
21 Notification was given to the landlords in Maragheh to declare their holding on 16 January. See further Land (1965: 100) and Lambton (1969: 87).
22 The Land Reform Organization began its activity with only twenty officials in 1962 (Lambton 1969: 99).
23 Such peasants felt that it was unfair to exclude them from the benefits of the reform. Disputes arose since many of these peasants refused to pay the landlords' share of the crops. See Lambton (1969: 104, 192) and Momeni (1980: 234–5).
24 The first draft of these regulations was issued on 7 February 1963 but underwent substantial amendments before it was passed by the *Majlis* and the Senate. See further Lambton (1969: 104–8, 194–215). For an English translation of the Additional Articles and the original draft of the regulations for their execution see also *Iran Almanac* (1963).
25 It was not clear whether the peasant had to pay two-fifths of the value of the total land before division or two-fifths of the value of the land which the peasant received; the text could be read either way (Lambton 1969: 203).
26 In both cases, the Agricultural Credit and Rural Development Bank paid one-third of the price directly to the landlord on behalf of the peasant (Momeni 1980: 237). For more details, see Lambton (1969: 201–4).
27 The draft articles for the third stage were submitted to the parliament in September 1968 and was finally approved on 13 January 1969. See Momeni (1980: 239), and Hooglund (1982: 69–70). An English translation of this law could be found in *Iran Almanac* (1969).
28 Instead of paying in cash, the peasant could pay the price of land in twelve annual instalments. In such a case, the price would be two-fifths of twelve annual rents minus 15 per cent.
29 This was called 'exchange for the better'. See Lambton (1969: 106).
30 It was estimated that a cadastral survey of all villages would have taken thirty years and would have cost about 64 million dollars (Momeni 1980: 151–2).
31 In some cases, for example, the land reform officials in practice found a

quantitative formula for mechanization. They gave farm improvement practices and machinery specific points; for example, a tractor got 10 points, discs 10, a combine 15, use of chemical fertilizers 5. A landlord who scored 51 points or more was judged to have 'mechanized' his land (Denman 1973: 104).

32 Jacoby (1971: 199) believes that one of the main reasons for the success of the Japanese land reform was the active role of the peasants in its implementation.

33 Of course, the abortive law of 1960 served as a warning to landlords and they had two years to sell their lands, transfer their villages to their wives or children, etc., and some exercised such practices (Keddie 1968: 82; Momeni 1980: 153–6). But there were many who did not react even to the news of a forthcoming land reform law, because they hardly believed that such laws could be put into operation. They supposed that the measure would be forgotten like so many other measures before it (Lambton 1969: 63).

34 According to the regulations of the second stage the landlords were allowed to take any action with respect to their land until the date that the execution of the law would begin, provided that they held the land legally. This was in contradiction to the provisions of the first stage which prohibited any transaction from the date of the passing of the law. See Momeni (1980: 237).

35 The unaffected lands were the retained lands, gardens and fruit orchards, *vaqf* lands, and lands of 'petty landowners'.

36 If the laws and regulations are supposed to be passed by a parliament or parliaments, they will be put into operation after they are passed. The parliament(s) can then influence the land reform as well.

37 Zonis (1971: 158–9) provides interesting information about land-holdings of the political elite in Iran.

38 In July 1962, Alam succeeded Amini as prime minister.

39 Arsanjani was succeeded on 12 March 1963 by General Riahi, whose deputy minister was a colonel. Arsanjani left the impression that he might have favoured a more radical second stage (Mahadvy 1965: 139). See also Momeni (1980: 283–4).

40 By September 1963 the number of land reform officials had increased to 1,184 Lambton (1969: 99).

41 Examples of such a point of view could be found in Saedloo (1972) and Doroudian (1976).

42 It should be noted that according to the land reform law of 1962 barren lands of the villages were to be acquired by the government for redistribution. In 1963, the Land Reform Council decided to postpone the sale of these lands. In 1965, the government offered these lands to interested developers at a price fixed by the government. Finally, the government offered these lands free of charge and provided financial help for their developers. See further Lambton (1969: 78–80) and Denman (1973: 154–8).

43 The law was passed in 1967.

44 A new type of farm corporation was also established by the government

from the early 1970s. They were called 'production co-operatives' and differed from corporations in two main respects: (a) the farmers retained their title to their lands, and (b) they were guaranteed a position in team works that cultivated the co-operative land. See Azkia (1980: 135–7).

45 The minimum of 51 per cent of votes was required for the establishment of a corporation. Calculating percentages to ascertain the extent of the majority was not done village by village but over the entire selected region. A village might then be included into a farm corporation contrary to the wishes of the majority (Denman 1973: 213).

46 Lambton (1969: 316) writes: 'Almost universally the peasants expected the cooperative societies to provide first loans and secondly water by wells or other means.'

Chapter 6 The agrarian system after the land reform

1 This was announced by Arsanjani, the Minister of Agriculture, in a press conference on 15 January 1962. He also revealed that a small group of experts from his department had been working to prepare the execution of the land reform in this region since October 1961. See also Land (1965/6: 100).

2 The local Democrat government which was in power in Azarbayjan between 1945 and 1946 adopted a land reform measure. Although this government was overthrown, its land reform stopped, and the landlords largely regained their land, the landlord–peasant relationship was considerably weakened. See Lambton (1969: 36–7) and Hooglund (1982: 41–2).

3 According to Article 8 of the land reform law the landlords had to declare their landholdings within about forty days after the first announcement which showed that the government intended to begin the land reform in a given area.

4 The difference is in fact slightly higher than Table 6.1 shows since the number of redistributed villages in the first period, i.e. 8,042, comprise only private villages redistributed, while the figures for the other two periods, i.e. 8,109 and 666, include state and perhaps crown villages as well. The progress made between January 1962 and September 1963 is more comparable with the total number of villages redistributed in the first stage which is given in Table 6.3, because it excludes state villages. See Denman (1978: 267), Salour (1963: 58), and Statistical Centre of Iran (1966, 1971).

5 The order is based on information in Lambton (1953: 270–1) which shows that, in general, there was a positive relationship between large landlordism and land fertility.

6 There are also reports of such an exercise. For example, Lambton (1969: 259) writes that landowners in Kermanshah deliberately selected parts of different villages with the object of retaining their influence over a wider area. She writes further that even landlords in Kordistan

intended to use this method but were discouraged by the land reform officials.

7 For a discussion about the definition of mechanized land in the law, see Chapter 5, pp. 77–8).

8 A study (OIPFG 1974: 42) points out that some women from a village which was exempted because of 'mechanization' believed that mechanization implied that one received a certain amount of wheat and money instead of sharing the crop. See also Lambton (1969: 101).

9 Warriner (1969: 123) writes that this method was widely used in Tehran and to a lesser extent in Esfahan.

10 Lambton (1969: 262), for instance, points to such cases in Arak and Esfahan.

11 See also Lambton (1969: 258) and Momeni (1980: 291) for other examples.

12 A number of such attempts can be found in Momeni (1980: 287–92).

13 We have assumed that the total number of Iranian villages was 60,000. For more information, see Chapter 1, p. 4.

14 About 3,000 out of a total 8,042 villages were entire villages, and the rest were equivalent to 2,520 entire villages. Thus, around 5,520 entire villages had been purchased by this date.

15 We have used the statistics of the public census of 1956 (Ministry of Interior 1961) for this purpose. It is safe to use these statistics since the number of those with cultivation rights, who were eligible to receive land, could only have slightly increased from that date until the start of the reform and because no account is taken of the *khwushnishin* population. See Chapter 3, p. 35.

16 The percentage is based on the number of rural families in 1966, i.e. 3 million families. Case studies and sporadic reports about land redistribution substantiate our conclusion that even 19 per cent is a generous percentage (Land 1965/6: 105; Momeni 1980: 313–14). As a matter of fact, a simple calculation reveals the likely exaggeration of the official statistics as well as our estimate. The official statistics indicate that the Iranian villages have on the average about 250 inhabitants or roughly 50 families, including *khwushnishins*. See, for example, Ministry of Interior (1961: 21) and Khosrovi (1979b: 9). Provided that the total number of six-*dang* villages which redistributed in the first stage is 9,514, Table 6.3 and our estimate indicate that each redistributed village had on average 75 and 60 households respectively, excluding the *khwushnishins*.

17 For further information about these five alternatives, see Chapter 5, pp. 74–6.

18 Both Lambton (1969: 217) and Hooglund (1982: 65–6) point out that there were notable exceptions.

19 See further Momeni (1980: 262–3) and OIPFG (1973: 60–1). In addition to the differences in the official figures, the implicit assumption in the data of the second stage that each landowner had only one estate is in all probability wrong. There might have been landowners who had small estates, for example, in different villages. Cf. Table 72 in

Bank Markazi Iran, *Annual Report and Balance Sheet* (1966: 135) and a similar table in Lambton (1969: 221).

20 Report issued by the land reform organization in 1966 indicates that 828 six-*dang* villages were claimed to be mechanized and another 1,405 villages were declared to be wholly gardens. See Lambton (1969: 221).

21 A study (Salmanzadeh and Jones 1979: 121) shows that in Dezful, in the Khuzestan Province, about 5 per cent of those who received land by division were second-time beneficiaries.

22 A table in Momeni *ibid* shows that a large majority of landlords declared their preference for division. Although the figures are different, this might be the result of the above mentioned inquiry in Denman (1973: 145).

23 We could either argue that similar to the second stage, those landlords who received a larger share of the crop under the sharecropping system chose to divide their land or assume that in the division cases the land was shared equally between landlords and peasants. However, the difference is insignificant.

24 See note 5 above.

25 There is evidence which suggests that *vaqf* land was not always transferred to the peasants but was sometimes usurped by influential civil servants and military officials. See Momeni (1980: 300–1).

26 See Denman (1973: 81) and *Iran Almanac* (1973: 395). We have assumed that all 491 crown villages (out of 1008) which remained outstanding in 1962 were sold to the peasants. For more information about the sale of state and crown land see Chapter 3, pp. 29, 33.

27 The remnants of the old tenure systems are discussed later in this chapter.

28 For instance McLachlan's (1968: 687) estimate shows that large landowners' landed properties covered about 56 per cent of the total agricultural land although they owned less than 35 per cent of all villages.

29 See a list of examples in Platt (1970: 77) which is extracted from Lambton (1969: Chs 6, 7, and 8), Momeni (1980: 314–15) and Azkia (1986: 123–4).

30 For more examples see also Khosrovi (1982), and Azkia (1986: 120–7).

31 See also Momeni (1980: 343) and Azkia (1980: 122) for other forms of sharecropping.

32 In fact Khosrovi (1982: 43) calls the owners of these farms medium landowners rather than rich peasants.

33 Khosrovi (1982: 43) writes that holdings of 30 to 50 hectares were principally concentrated in Azarbayjan, Bushehr, Gilan, Hamadan, Kermanshah, Kordestan, Khuzestan and Zanjan.

34 Some basic information about all fifteen large agribusinesses is provided in Momeni (1980: 354).

35 Hooglund (1982: 86–7) mentions that there were two corporations which consisted of one single village each and, on the other hand, an exceptionally large corporation with forty-five villages.

36 For more information about the main features of the law governing establishment of farm corporation see Chapter 5, p. 84.

37 Difficulties of the small farmers with respect to irrigation will be discussed in more detail in the next chapter.

Chapter 7 Economic consequences of the land reform

1 For example, the available data reveals that 83 and 87 per cent of all the irrigation expenditures in 1968 and 1969, respectively, was spent on such projects. See Bank Markazi Iran, *Annual Report and Balance Sheet* (1969, 1971). Cf. Azkia (1980: 96).
2 It was mentioned in Chapter 4 (see p. 50) that because of the importance of water in agriculture in arid regions, land should be measured in 'efficiency units'. This would reflect the differences in fertility of land caused by the differences in the availability of water.
3 The Agricultural Development Bank of Iran was established in 1968 under the name of the Agricultural Development Fund of Iran. Its name and some of its regulations were changed in 1973.
4 Katouzian (1974: 233) has estimated that even after the peasants received government credits in 1970 and 1971 their income deficit for each year was about 59 and 73 billion rials, respectively.
5 Unfortunately we do not know much about the credits which different government agencies provided. We know that these credits were given for regional development programmes growing special crops such as tea and sugar, or sale of machinery (Aresvik 1976: 172). As we will discuss later in this chapter, it was predominantly large farmers who cultivated such crops and used machinery. Therefore, we can assume that these farmers might have benefited most from such credits as well.
6 Interest rates on state loans to corporations ranged from 2 to 8 per cent but they could pay only 1 per cent interest and reinvest the rest (Katouzian 1983: 332–3).
7 A small number of these loans, which were probably obtained from relatives and friends, were free of charge. See Khosrovi (1979b: 163).
8 This does not include 600,000 workers, i.e. 11 per cent of the rural labour force, who were only seasonally unemployed. See Statistical Centre of Iran (1984: 61, 78).
9 Mechanized land was even exempted from the reform in the first stage. However, there were two differences: (1) the second stage clearly defined mechanized land and (2) most landlords who were subject to the second stage were in practice 'smaller' landlords who might have needed to prove that their land was mechanized while the bigger, more influential landlords might not. For further information about the regulations of the second stage see Chapter 5, pp. 74–5.
10 We should consider further that Kordestan and West Azarbayjan were among the regions which had proportionately a large number of 10–50 hectare farms, i.e. medium-sized farms and the lower strata of large farms; see Khosrovi (1981: 40).
11 These five provinces were Esfahan, Gilan, Khorasan, Mazandaran and

Markazi which consumed, respectively, 16, 11, 9, 21 and 17 per cent of all fertilizers in 1971.

12 Rice was in fact a semi-cash crop. It was the dominant staple food crop in the rice-growing provinces, namely Gilan and Mazandaran. At the same time, it was a comparatively profitable cash crop for many peasants.

13 In 1975 the net income of each hectare of land under pistachio and citrus fruits was, respectively, eight times and twice that of rice (Khosrovi 1979: 105). At the same time, the net income of one hectare of land under rice was four times the net income of one hectare of land under wheat (Azkia 1986: 168). See also Rafipoor (1985: 199–200).

14 Okazaki (1969: 28) writes that numerous large farms appeared in Mazandaran with the main purpose of cotton production.

15 In fact, the wheat purchase of the Cereals Organization in some years was 0.1 per cent, or even less, of total production. See Aresvik (1976: 144, Table 7.6).

16 Hakimian (1988) argues that many of the agricultural output changes in Iran were in fact a result of the rural–urban migration caused by non-agricultural higher wages.

Chapter 8 Rural poverty and inequality after the reform

1 The main problem with the above estimate is that it ignores the fact that some of these small farmers are rice-growers and orchardgrowers: the latter presumably do not produce wheat and the former principally cultivate and consume rice. Although this will change the estimated percentages it does not change the above conclusion.

2 Another study (Katouzian 1983: 343–4) shows that the share of the poorest rural households in total expenditures increased slightly at the expense of the richest between 1967 and 1972, but concludes that the developments since 1972 lead one to believe that the distribution of rural expenditures, and income, might have worsened in subsequent years. See also Pesaran (1976: 280).

3 Azkia (1980: 253) shows that the peasants' debts in two regions trebled between 1969 and 1978. See also Katouzian (1974: 233) and Azkia (1986: 152).

4 For some information about the corporations after the Revolution see Azkia (1986: 252–5) and a Group of Experts (1980: 46–57).

5 However, undernourishment was more widespread in the urban areas while severe undernourishment was more common in the rural areas (Katouzian 1983: 348).

Bibliography

Abrahamian, E. (1975) 'European feudalism and Middle Eastern despotisms', *Science and Society*, vol. 39.

Adams, D.W. (1973) 'The economics of land reform', *Food Research Institute Studies*, vol. 12.

Adams, D.W. and Rask, N. (1968) 'Economics of cost-share leases in less developed countries, *American Journal of Agricultural Economics*, vol. 50.

Abidi, M. (1955) 'Measure Pre-requisite to the Development of Land and Water in Iran', unpublished dissertation, Colombia University.

Agah Book Collection (1982) *Masā'il- arzi va dihgāni-yi Irān va Khāvarmiyāneh* [Agrarian and peasant problems of Iran and the Middle East], Tehran: Agah Publications.

Agrawal, P. (1984) 'Sharecropping, supervision cost, and rural credit markets', mimeographed, Department of Economics, Stanford University.

Ajami, I. (1973) 'Land reform and modernisation of the farming structure in Iran', *Oxford Agrarian Studies*, vol. 2.

——— (1975) *Shishdāngi: Pizhuhishi dar zamineh-yi jami 'ehshināsi-yi rustā'i* [Six *dangs*: an enquiry in the field of rural sociology], 4th edn, n.p.: Tus Publications.

——— (1976) 'Agrarian reform, modernization of peasants and agricultural development in Iran', in J.W. Jacqz (ed.) *Iran: Past, Present and Future*, New York: Aspen Institute for Humanistic Studies.

Alexander, R.J. (1974) *Agrarian Reform in Latin America*, New York: Macmillan.

Antoun, R.T. (1976) 'The gentry of a traditional peasant community undergoing rapid technological change: an Iranian case study', *Iranian Studies*, vol. 9.

——— (1981) 'The complexity of the lower stratum: sharecroppers and wage-labourers in an Iranian village', *Iranian Studies*, vol. 14.

Aresvik, O. (1976) *The Agricultural Development of Iran*, New York: Praeger Publishers.

Asayesh, H. (1974) 'Barrasi-yi tahlili-yi budjeh khānivārhā-yi shahri va rusta'i-yi Azarbayjan-i Shirqi' [Analysis of income and consumption of rural and urban families of East Azarbayjan], *Nashriyeh Dānishkadeh-yi Adabiyāt va 'ulumi-i Insani-yi Tabriz*, vol. 26.

Ashraf, A. (1979) 'Ābādi' [village], *Ayandeh*, vol. 5.
Ashraf, A. and Banuazizi, A. (1980) 'Policies and strategies of land reform in Iran', in Inayatullah (ed.), *Land Reform: Some Asian Experiences* vol. IV, Kuala Lumpur: Asian and Pacific Development Administration Office.
Azkia, M. (1980) 'The effect of rural development programmes on the Iranian peasantry between 1962 and 1978, with special reference to far corporations', Ph.D. Dissertation, Aberdeen University.
——— (1986) *Jāmi 'ehshinasi-yi tusi eh va tuse 'ehnayāftigi-yi rustā'i-yi Iran* [Sociology of development and rural underdevelopment of Iran], Tehran: Ettel'at Publications.
Bagley, R.F.C. (1976) 'A bright future after oil: dams and agro-industry in Khuzistan', *The Middle East Journal*, vol. 30.
Bank Markazi Iran (The Central Bank of Iran) (annually 1965–77) *Annual Report and Balance Sheet*, Tehran.
——— (1968) *Bulletin*, no. 39.
——— (1969) *National Income of Iran: 1962–67*, Tehran: Bank Melli Iran Press.
——— (1974) *National Income of Iran: 1959–72*, Tehran: Bank Melli Iran Press.
——— (1975) *Bulletin*, no. 76.
Bardhan, P.K. (1973) 'Size, productivity, and returns to scale: an analysis of farm-level data in Indian agriculture, *Journal of Political Economy*, vol. 81.
Bardhan, P.K. and Srinivasan, T.N. (1971) 'Cropsharing in agriculture: a theoretical and empirical analysis', *American Economic Review*, vol. 6.
Barlowe, R. (1953) 'Land reform and economic development, *Journal of Farm Economics*, vol. 35.
Barraclough, S. (1970) 'Alternative land tenure systems resulting from agrarian reform in Latin America', *Land Economics*, vol. 45.
Bavandi, B. (1982) 'Masā'il-i ijtima'i va rifāhi dar rustāhā-yi Iran' [Social and welfare problems in the Iranian villages], *Masā'il-i Kishāvarzi-yi Iran*, vol. 2.
Bell, C. (1976) 'Production conditions, innovation and the choice of lease in agriculture', *Sankhy: The Indian Journal of Statistics*, vol. 38.
——— (1977) 'Alternative theories of sharecropping: some tests using evidence from northeast India, *Journal of Development Studies*, vol. 13.
Bell, C. and Zusman, P. (1976) 'A bargaining theoretic approach to crop sharing contracts', *American Economic Review*, vol. 66.
Benjamin, N.C., Devarajan, S., and Weiner, R.J. (1989) 'The "Dutch" disease in a developing country: oil reserves in Cameroon', *Journal of Developing Economics*, vol. 30.
Bergamann, T. (1977) 'Agrarian reforms and their functions in development process', *Land Reform, Land Settlement, and Co-operatives*, no. 2.
Berry, R.A. (1971) 'Land reform and the agricultural income distribution', *Pakistan Development Review*, vol. 21.
Berry, R.A. and Cline, W.R. (1979) *Agrarian Structure and Productivity in Developing Countries*, Baltimore: Johns Hopkins University Press.

Bhaduri, A. (1973) 'A study in agricultural backwardness under semi-feudalism', *Economic Journal*, vol. 83.
Bharier, J. (1971) *Economic Development in Iran 1900–1970*, London: Oxford University Press.
Black, A.G. (1948) 'Iranian agriculture: present and prospective', *Journal of Farm Economics*.
Bolukbashi, A. (1981) 'Zāri', tabi'at va ravabit-i toulidi dar kishāvarzi-yi Iran' [Peasant, nature, production relations in Iranian agriculture], *Masā'il-i Kishāvarzi-yi Iran*, no. 2.
Boxley, R.F. (1971) 'Cost-share leases revisited . . . again', *American Journal of Agricultural Economics*, vol. 53.
Brun, T. and Dumont, R. (1978) 'Iran: imperial pretensions and agricultural dependence', *Merip Reports*, vol. 8.
Caballero, J.M. (1983) 'Sharecropping as an efficient system: further answers to an old puzzle', in T.J. Byres (ed.), *Sharecropping and Sharecroppers*, London: Frank Cass.
Cheung, S.N.S. (1968) 'Private property rights and sharecropping', *Journal of Political Economy*, vol. 76.
——— (1969) *The Theory of Share Tenancy*, Chicago: University of Chicago Press.
Corden, W.M. and Neary, P. (1982) 'Booming sector and de-industrialisation in a small open economy', *Economic Journal*, vol. 92.
Corina, G.A. (1985) 'Farm size, land yields, and the agricultural production functions: an analysis for fifteen developing countries', *World Development*, vol. 13.
Cottam, R.W. (1964) *Nationalism in Iran*, Pittsburgh: University of Pittsburgh Press.
Craig, O. (1978) 'The impact of land reform on an Iranian village', *Middle East Journal*, vol. 33.
Dehbod, A. (1963) 'Land ownership and use in Iran', *CENTO Symposium on Rural Development*, Tehran: CENTO.
De Janvry, A. (1981) *The Agrarian Question and Reformism in Latin America*, Baltimore: Johns Hopkins University Press.
Denman, D.R. (1973) *The King's Vista: A Land Reform which has Changed the Face of Persia*, Berkhamsted: Geographical Publications.
——— (1978) 'Land reforms of Shah and people', in G. Lenczowski (ed.), *Iran Under the Pahlavis*, California: Hoover Institution Press.
Dorner, P. (1972) *Land Reform and Economic Development*, Harmondsworth: Penguin Books.
Doroudian, A. (1976) 'Modernization of rural economy in Iran', in J.W. Jacqz (ed.) *Iran: Past, Present and Future*, New York: Aspen Institute for Humanistic Studies.
Dovring, F. (1970) 'Economic results of land reform', *AID Spring Review of Land Reform*, vol. XI.
Dumont, R. (1978) *Nukāti chand dabāreh-yi vaz 'nyat-i kishāvarzi-yi Iran* [Some remarks on the situation of Iranian agriculture], Frankfurt: Confederation of Iranian Students.
Ettel'at (1980) 20 February.
——— (1986) 26 October.

FAO (annual) *Production Yearbook*, Rome: FAO.
FAO (1966) *Report on the 1960 World Census of Agriculture*, Rome: FAO.
Fishburne, C. (1964) 'The United States policy toward Iran', unpublished Ph.D. dissertation, Florida State University.
Flores, E. (1970) 'Issues of land reform', *Journal of Political Economy*, vol. 78.
Freivalds, J. (1972) 'Farm corporation in Iran: an alternative to traditional agriculture', *Middle East Journal*, vol. 24.
Ghose, A.K. (1983) 'Agrarian reform in developing countries: issues of theory and problems of practice', in A.K. Ghose (ed.) *Agrarian Reform in Contemporary Developing Countries*, London: Croom Helm.
Ghose, A. and Griffin, K. (1980) 'Rural poverty and development alternatives in south and southeast Asia: some policy issues', *Development and Change*, vol. 11.
Gittinger, J.R. (1961) 'United States policy toward agrarian reform in underdeveloped nations', *Land Economics*, vol. 37.
Griffin, K. (1974) *The Political Economy of Agrarian Change*, London: Macmillan.
——— (1976) *Land Concentration and Rural Poverty*, London: Macmillan.
Group of Experts (1980) 'Darbāreh-yi shirkathā-yi sahami-yi zirā'i va sarnivisht-i ān' [On farm corporations and their destiny], *Masā'il-i Kishavarzi-yi Iran*, no. 1.
Hadary, G. (1951) 'The agrarian reform problem in Iran', *The Middle East Journal*, vol. 5.
Hakimian, H. (1988) 'The impact of the 1970s' oil boom on Iranian agriculture', *Journal of Peasant Studies*, vol. 15.
Halliday, F. (1979) *Iran: Dictatorship and Development*, Harmondsworth: Penguin Books.
Haririan, M. (1976) 'Tahavul-i iqtisād-i rustā'i dar Dasht-i Chamchamāl' [Transformation of rural economy in Chamchal Plain], *Majaleh-yi Dānishkadeh-yi Adabiyāt va 'ulum-i Insani-yi Dānishgah-i Ferdousi*, vol. 46.
Heady, E.O. (1955) 'Marginal resource productivity and imputation of share for a sample of rented farms', *Journal of Political Economy*, vol. 63.
Ho, S.P.S. (1976) 'Uncertainty and the choice of tenure arrangements: some hypotheses', *American Journal of Agricultural Economics*, vol. 58.
Homan, A. (1955) *Iqtisād-i kishvārzi* [Agricultural economics], Tehran: Tehran University Press.
Hooglund, E. (1973) 'The Khwushninin population of Iran', *Iranian Studies*, vol. 6.
——— (1982) *Land and Revolution in Iran, 1960–1980*, Austin: University of Texas Press.
Hooglund, M. (1980) 'One village in the revolution', *MERIP Reports*, no. 87.
Hsiao, J.C. (1975) 'The theory of share tenancy revisited', *Journal of Political Economy*, vol. 83.

Huang, Y. (1973) 'Risk, entrepreneurship, and tenancy', *Journal of Political Economy*, vol. 83.

Huntington, S.P. (1970) 'Political dimensions of land reform', *AID Spring Review of Land Reform*, vol. 11.

Iran Almanac and Book of Facts (annual) Tehran: Echo of Iran.

Issawi, C. (1957) 'Farm output under fixed rents and share tenancy', *Land Economics*, vol. 33.

Jacoby, E. (1971) *Man and Land*, London: Andre Deutsch.

Jacqz, J.W. (ed.) (1976) *Iran: Past, Present and Future*, New York: Aspen Institute for Humanistic Studies.

Jaynes, G.D. (1982) 'Production and distribution in agrarian economies', *Oxford Economic Papers*, vol. 34.

Jones, R. (1967) 'The short-run economic impact of land reform on fedual village irrigated agriculture', unpublished dissertation, University of Maryland.

Johnston, B.F. and Kilby, P. (1975) *Agricultural and Structural Transformation*, London: Oxford University Press.

Johnson, D.G. (1950) 'Resource allocation under share contracts', *Journal of Political Economy*, vol. 58.

Kamali, T. (1982) 'Kisht-i birinj dar Iran' [Rice cultivation in Iran], *Masā'il Kishāvarzi-yi Iran*, no. 6.

Katouzian, H. (1983) 'The agrarian question in Iran', in A.K. Ghose (ed.) *Agrarian Reform in Contemporary Developing Countries*, London: Croom Helm.

Katouzian, M.A. (1974) 'Land reform in Iran: a case study of the political economy of social engineering, *Journal of Peasant Studies*, vol. 1.

——— (1978) 'Oil versus agriculture: a case of dual resource depletion in Iran', *Journal of Peasant Studies*, vol. 5.

Kazemi, F. (1980) *Poverty and Revolution in Iran*, New York: New York University Press.

Kazemi, F. and Abrahamian, E. (1978) 'The nonrevolutionary peasantry of modern Iran', *Iranian Studies*, vol. 11.

Keddie, N.R. (1968) 'The Iranian village before and after land reform', *Journal of Contemporary History*, vol. 3.

——— (1972) 'Stratification, social control, and capitalism in Iranian villages: before and after land reform', in R. Antoun and A.I. Harik (eds) *Rural Politics and Social Change in the Middle East*, London: Indian University Press.

Khatibi, N. (1972) 'Land reform in Iran and its role in rural development', *Land Reform, and Land Settlement and Cooperatives*, no. 1.

Khirabi, J. (1981) 'Intikhāb-i munāsib-i qanāt va ya chāh bar cheh mavāzini bāyad ustuvār bāshid?' [What should be the criteria for the proper choice of *qanat* or well?], *Masā'il-i Kishāvarzi-yi Iran*, no. 3.

Khosrovi, K. (1979a) *Jami'eh-yi dehqāni dar Iran* [Peasant society in Iran], Tehran: Payam Publications, 1978; reprint edn 1979.

——— (1979b) *Jami'ehshināsi-yi rustā'i-yi Iran* [Rural sociology of Iran], Tehran, University of Tehran 1972; reprint edn 1979, Tehran: Payam Publications.

——— (1981) *Mas'leh-yi arzi va dihqānañ-i tuhidast dar Iran* [Agrarian question and poor peasants in Iran], Tehran: Bidari Publication.
——— (1982) *Dihqānān-i tavāngir* [Rich peasants], Tehran: Mitra Publications.
King, R. (1977) *Land Reform: A World Survey*, Colorado: Westview Press.
Koo, A.Y.C. (1973) 'Towards a more general model of land tenancy and reform', *Quarterly Journal of Economics*, vol. 87.
Lambton, A.K.S. (1953) *Landlord and Peasant in Persia*, London: Oxford University Press.
——— (1969) *The Persian Land Reform: 1962–1966*, Oxford: Clarendon Press.
——— (1971) 'Land reform and the rural cooperative societies', in E. Yar-Shater (ed.), *Iran Faces Seventies*, New York: Praeger Publishers.
Land, C.T. (1965/6) 'Land reform in Iran', *Persica*, no. 2.
Layman, P.N. and French, J.T. (1970) 'Political results of land reform', *AID Spring Review of Land Reform*, Washingron DC: AID.
Lipton, M. (1968) 'The theory of optimizing peasant', *Journal of Development Studies*, vol. 4.
——— (1974) 'Towards a theory of land reform', in D. Lehmann (ed.), *Agrarian Reform and Agrarian Reformism*, London: Faber & Faber.
Lodi, H.S.K. (1965) 'Preharvest sale of agricultural produce in Iran', *Monthly Bulletin of Agricultural Economics and Statistics*, vol. 14.
Lucas, R.E.B. (1979) 'Sharing, monitoring, and incentives: Marshallian misallocation reassessed', *Journal of Political Economy*, vol. 87.
Lundahl, M. (1979) *Peasants and Poverty: A Study of Haiti*, London: Croom Helm.
——— (1984) 'Jord och fattigdom i Latinamerika' [Land and poverty in Latin America], *Ekonomisk debatt*, no. 8.
McLachlan, K.S. (1968) 'Land reform in Iran', in W.B. Fisher (ed.), *The Cambridge History of Iran* vol. I, Cambridge: Cambridge University Press.
——— (1977) 'The Iranian economy', in H. Amirsadeghi (ed.), *Twentieth-Century Iran*, London: Heinemann.
Mahdavy, H. (1965) 'The coming crisis in Iran', *Foreign Affairs*, vol. 44.
——— (1982) 'Tahavulāt-i sisaleh-yi yik dih dar Dasht-i Qazvin', [The 30-year changes of a village in Qazvin Plain], in Agah Book Collection, 1982.
Majd, M.G. (1987) 'Land reform policies in Iran', *American Journal of Agricultural Economics*, vol. 69.
Marshall, A. (1920) *Principles of Economics*, London: Macmillan.
Mazumdar, D. (1975) 'The theory of sharecropping with labor market dualism', *Economica*, vol. 42.
Ministry of Interior (1961) *National and Province Statistics of the First Census of Iran: November 1956*, Tehran: Ministry of Interior.
Misawa, T. (1974) 'Agricultural development and employment expansion: a case study of Japan', in N. Islam (ed.), *Agricultural Policy in Developing Countries*, London: Macmillan.

Mohammadi, M. (1973) *Darbāreh-yi mas'leh-yi arzi dar Iran va shiveh-yi hal-i dimokrātik-i ān* [On the agrarian question in Iran and its democratic solution], Stassfurt: Tudeh Publishing Center.

Momeni, B. (1980) *Mas'leh-yi arzi va jang-i tabaqāti dar Iran* [The agrarian question and class struggle in Iran], Tehran: Peyvand Publications.

Newbery, D.M.G. (1974) 'Cropshare tenancy in agriculture: comment', *American Economic Review*, vol. 64.

——— (1975a) 'Tenurial obstacles to innovation', *Journal of Development Studies*, vol. 11.

——— (1975b) 'The choice of rental contract in peasant agriculture', in L.G. Reynolds (ed.), *Agriculture in Development Theory*, New Haven: Yale University Press.

——— (1977) 'Risk sharing, sharecropping, and uncertain labor market', *Review of Economic Studies*, vol. 44.

Newbery, D.M.G. and Stiglitz, J.E. (1979) 'Sharecropping, risk sharing and the importance of imperfect information', in Roumasset, Boussard and Singh (eds), *Risk, Uncertainty and Agricultural Development*, New York: Agricultural Development Council.

Nikayin, A. (1976) 'Silsileh-yi Pahlavi, amlāk zirā'iti, dihqānān' [Pahlavi dynasty, agricultural lands, peasants], *Donya*, vol. 3.

Nomani, F. (1980) *Takāmul-i feodālism dar Iran* [Development of feudalism in Iran], Tehran: Kharazmi Publications.

OIPFG (The Organisation of Iranian People's Feddyi Guerrilla) (1973) *Darbāreh-yi islāhāt arzi va natā'ij-i mustaqim-i ān* [On the land reform in Iran and its direct results], Rural Studies Series of OIPFG, no. 1, n.p.: OIPFG Publications.

——— (1974) *Barrasi-yi sākht-i iqtisādi-yi rustāhā-yi Kirman* [A study of the economic structure of the villages of Kerman], Rural Studies Series of OIPFG, no. 4, n.p.: OIPFG Publications.

Okazaki, S. (1968) *The Development of Large-Scale Farming in Iran: The Case of Province of Gorgan*, Tokyo: Institute of Asian Economic Affairs.

Parvin, M. and Zamani, A. (1979) 'Political economy of growth and destruction: a statistical interpretation of the Iranian case', *Iranian Studies*, vol. 12.

Pertev, R. (1986) 'A new model for sharecropping and peasant holdings', *Journal of Peasants Studies*, vol. 14.

Pesaran, M.H. (1976) 'Income distribution and its major determinants in Iran', in J.W. Jacqz (ed.) *Iran: Past, Present and Future*, New York: Aspen Institute for Humanistic Studies.

Platt, K.B. (1970) 'Land reform in Iran', *AID Spring Review of Land Reform*, vol. 11.

Price, O.T.W. (1975) 'Towards a comprehensive Iranian agricultural policy', Agricultural and Rural Development Advisory Mission (ARDAMI) Report no. 1, Tehran: IBRD.

Rafipoor, F. (1985) *Jāmi'eh-yi rustā'i va niyāzhā-yi ān: Pizhuhishi dar 32 rustā-yi barguzideh-yi ustān Yazd* [Rural society and its needs: an analysis of 32 sample villages of Yazd Province], Tehran: Inteshar Corporation.

Rao, C.H.H. (1971) 'Uncertainty, entrepreneurship, and sharecropping in India', *Journal of Political Economy*, vol. 79.
Raup, P. (1968) 'Land reform and agricultural development', in Southworth and Johnston (eds), *Agricultural Development and Economic Growth*, New York: Cornell University Press.
Reid, D.J. (1976) 'Sharecropping and agricultural uncertainty', *Economic Development and Cultural Change*, vol. 24.
——— (1977) 'The theory of share tenancy revisited again', *Journal of Political Economy*, vol. 85.
Research Group (1962) 'Tarh-i tahqiq-i ghalāt dar Iran' [The research project of crop economy in Iran], *Tahqiqāt-e Eqtesadi*, nos 1 and 2.
——— (1964) 'A review of the statistics of the first stage of land reform', *Tahqiqat-e Eqtesadi*, nos 7 and 8.
——— (1965a) 'Economic report on agriculture in the Isfahan and Yazd area', *Tahqiqat-e Eqtesadi*, nos 9 and 10.
——— (1965b) 'Rural economic problems of Khuzistan, *Tahqiqāt-e Eqtesadi*, nos 9 and 10.
——— (1967) 'A study of rural economic problems of Gilan and Mazandarn', *Tahqiqat-e Eqtesadi*, no. 3.
——— (1968) 'A study of rural economic problems of East and West Azarbaijan', *Tahqiqat-e Eqtesadi*, nos 12 and 13.
——— (1970a) 'An analysis of the law governing the first stage of land reform in Iran', *Tahqiqat-e Eqtesadi*, no. 17.
——— (1970b) 'A study of the rural economic problems of Sistan and Baluchestan', *Tahqiqat-e Eqtesadi*, nos 12 and 13.
Richards, H. (1975) 'Land reform and agribusiness in Iran', *Merip Reports*, no. 43.
Saedloo, H. (1972) 'A critique of a policy for agricultural development at the poles of soil and water', *Tahqiqat-e Eqtesadi*, nos 25 and 26.
——— (1979) 'Huvaiyt-i kishāvarzi-yi Iran' [Identity of the Iranian agriculture], *Ayandeh*, vol. 5.
Saeedi, A. (1975) 'Pitzhuhishi dar Kifaiyt-i muhājirthā-yi rustā'i dar Shamāl-i Khurāsān' [An analysis of the quality of rural migration in the northern Khorasan], *Majaleh-yi Dānniskadeh-yi Adbiyat va 'Ulum-i Insani-yi Dānishgah-i Ferdousi*, vol. 44.
Safinezhad, J. (1966) *Tālibābād: Numuneh-yi jāmi'i az barrasi-yi yik dih* [Talebabad: a detailed example of the study of one village], Tehran: University of Tehran.
——— (1972) *Buneh*, Tehran: University of Tehran Press (in Persian).
Salmanzadeh, C. and Jones, G.E. (1979) 'An approach to the micro analysis of the land reform program in southwestern Iran', *Land Economics*, vol. 55.
Salour, A. (1963) 'Land reform activities in Iran', *CENTO Symposium on Rural Development*, Tehran: CENTO.
Scherr, S.J. (1989) 'Agriculture in an export-boom economy: a comparative analysis of policy and performance in Indonesia, Mexico and Nigeria', *World Development*, vol. 17.
Schickele, R. (1941) 'Effects of tenure systems on agricultural efficiency', *Journal of Farm Economics*.

Sen, A.K. (1975) *Employment, Technology and Development*, Oxford: Clarendon Press.

Siyahpoosh, T. (1981) 'Kishāvarzi-yi sunati ra bishināsim', [We should know the traditional agriculture], *Masa'il-i Kishāvarzi-yi Iran*, no. 3.

Smith, A. (1937) [1776] *Wealth of Nations*, New York: Modern Library edition.

Soodagar, M. (1978) *Nizām-i arbāb ra'iyati dar Iran* [The landlord–peasant system in Iran], Tehran: Pazand Publications.

Statistical Centre of Iran (1966) *Statistical Yearbook: 1966*, Tehran: Statistical Centre of Iran.

——— (1971) *Statistical Yearbook: 1968*, Tehran: Statistical Centre of Iran.

——— (1972) *Statistical Yearbook: 1970*, Tehran: Statistical Centre of Iran.

——— (1975) *Statistical Yearbook: 1972*, Tehran: Statistical Centre of Iran.

——— (1976) *Statistical Yearbook: 1973*, Tehran: Statistical Centre of Iran.

——— (1981) *Natā'ij-i āmargiri az budjeh-i Khānivārhā-yi rustā'i sāl-i 1356* [The results of the 1977 census of budget of rural households], Tehran: Statistical Centre of Iran.

——— (1984) *Sālnāmeh-yi amari-yi 1361* [Statistical Yearbook: 1982], Tehran: Statistical Centre of Iran.

Stenberg, M.J. (1970) 'The economic impact of the latifundista', *Land Reform, Land Settlement, and Cooperatives*, no. 2.

Stickley, S. and Najafi, B. (1971) 'The effectiveness of farm corporations in Iran', *Tahqiqat-e Eqtesadi*, nos 25 and 26.

Stiglitz, J.E. (1974) 'Incentives and risk sharing in sharecropping', *Review of Economic Studies*, vol. 41.

Taqavi, N. (1983) 'Āb va ābyari dar rustāhā-yi Iran' [Water and irrigation in the Iranian villages], *Nashryeh-yi Dānnishkadey-yi Ādabiyat va 'Ulum-i Insani*, vol. 31.

UN (1963) *Progress in Land Reform. Third Report*, New York: Department of Economic and Social Affairs.

Vreeland, H.H. (ed.) (1957) *Iran*, New Haven: Human Relations Area Files.

Walinsky, L. (ed.) (1977) *Agrarian Reform as Unfinished Business: the Selected Papers of Wolf Ladejinsky*, New York: Oxford University Press.

Warriner, D. (1969) *Land Reform in Principle and Practice*, Oxford: Clarendon Press.

Weeks, J. (1970) 'Uncertainty, risk and wealth and income distribution in peasant agriculture', *Journal of Development Studies*, vol. 7.

Wulff, H.E. (1966) *The Traditional Crafts of Persia*, Cambridge: MIT Press.

Yavari, A. (1981) *ShināKhti az Kishāvarzi-yi sunati-yi Iran* [Information on Iranian Traditional Agriculture], Tehran: Book Publication and Translation Company.

Zonis, M. (1971) *The Political Elite of Iran*, Princeton: Princeton University Press.

Index

absentee landowners 9, 25–6, 30–1, 41, 104
agrarian reform *see* land reform
agrarian system: *arab-raiyati* system in Iran 28–48; bimodal 54–61, 66–7, 77, 78, 110–11, 143–4; traditional 8; unimodal 61–6, 67, 77, 143
agribusinesses 83, 107, 113, 123; imported machinery exempt from taxes 123
Agricultural Bank (Agricultural Co-operative Bank of Iran) 114–15, 116–17
Agricultural Development Bank of Iran 114, 117
Agricultural Development Fund 84
agricultural growth: as a result of land reform 1; statistics 132–3
agricultural production 42–3, 83
agriculture: in Iran 4–7; techniques of production 40–2; *see also* farming
Amini, Ali 82
amlak 32
Arsanjani, Hassan 82–3

bank loans 114–17
Bank Markazi Iran 3
Bardhan, P.K. 13, 17
barley 7, 125, 129, 130; yield 42
barzigars 38
beet 130
Bell, C. 14, 15, 18
bimodal agrarian system 54–61, 66–7, 77, 78, 110–11, 143–4
Black, A.G. 31
Boxley, R.F. 14
bread price 130, 131
buneh 37–8, 103

Caballero, J.M. 15
cash crops 59–60, 61, 111, 125–9, 130
Central Bank *see* Bank Markazi Iran
Cereals Organization 131
Cheung, S.N.S. 12–13, 15, 17, 21, 24
co-operative societies 85–6, 114–19
co-operatives (*buneh*) 37–8
conservation of water 42
contracts *see* tenancy contracts
corporations, farm 84, 107–8, 113, 123, 141; privileges 118
corruption 69
cotton 129, 130
credit 45, 64, 85; scarcity 130; *see also* loans
crops: division in a shared tenancy 38–9; food 60, 61, 125–6, 129; pre-harvest sale 45, 46–7; staple 7; subsistence 111; summer 7; *see also* cash crops

Denman, D.R. 101

economics of land reform 53–66

economy of Iran (1955–60) 69
entrepreneurship 21–2
export of cash crops 61

farm corporations 84, 107–8, 113, 123, 141; privileges 118
farming: subsistence 135–8; *see also* agriculture
farms: large 106–8, 110, 134; small 108–9, 134
fertilizers 1, 9, 41, 121; consumption 124
fixed-rent tenancy: contracts 22–3; Iran 32–3; option of the second stage of land reform 75, 98; risk 19–21
food 45; crops 60, 61, 125–6, 129; imports 61, 134
foreign exchange 61

government loans 114, 117–18

Hardary, G. 42
Heady, E.O. 15
Ho, S.P.S. 19
Hooglund, E. 101
Hsiao, J.C. 14
Huang, Y. 21
hunger 45

imports: consumer goods 69; food 61, 134; machinery 123; tax 123; wheat 131–2
income: additional 141; rural 43–4, 47, 108–9, 138–43
inequality 135, 142
inflation 69
inheritance under Islamic law 31, 32, 35
innovation: bimodal agrarian system 57–9, 130; labour-intensive 63, 67; lacking in traditional tenancy systems 10; sharecropping 16–19; unimodal agrarian system 63
interest rates 45, 47, 56; loans 118, 134
investment: lacking in traditional tenancy systems 9–10, 26, 41; sharecropping 16–19, 41; unimodal agrarian system 62
iqta 28–9
Iran: agriculture 4–7; economy (1955–60) 69; implementation of land reform 89–102; land reform law 68, 71–7; land reform programme 1, 68–87, 144; revolution (1979) 141
Iraq: revolution (July 1958) 71
irrigation 5, 39; access post land reform 112–13, 134
Islam: inheritance law 31, 32, 35; prohibition against interest 46
Issawi, C. 16

Jaynes, G.D. 23
Johnson, Gale 12, 15
Jones, G.E. 101
jufts 35

Kayhan 93
khaliseh 29, 31–2
khwushnishins 35, 95

labour market dualism 23–4
labourers 44, 48, 119–25; Afghani 120, 140; landless 60, 64, 77–8, 99, 119, 144
land: crown 29, 32–3, 68; distribution to peasants 94–5, 99, 101–3; endowed 32; exempt from land reform expropriations 55, 72, 74–5; option to divide land between peasant and landlord under the second stage of land reform 75, 98–9; private 28–31; state 29, 31–2; unit of measure 35
land assignment 28–9
land concentration 9, 25–6, 66, 106; pre-reform Iran 28–33, 47
land reform 49–67; first stage 72–3, 77–8, 86, 89–95; second stage 74–6, 79–81, 86–7, 95–9; third stage 76–7, 81, 87, 100–2; and economics 53–66;

and politics 51–3; definition 50–1; effect of legislation 77–81; effectiveness 66; government attitude 81–3; implementation in Iran 89–102; Iranian programme 1, 68–7, 144; lands exempted from the law 55, 72, 74–5; laws 68, 71–7; loopholes in the law 92–4, 109; pro-landlord 79–81, 83, 86–7, 95–6, 101; pro-peasant 72–4, 82, 86; socialist objectives 52
Land Reform Organization 72, 74, 75, 82, 90
land utilization 5
landless 60, 64, 77–8, 99, 119, 144
landlord–peasant system 1, 8–10, 68; criticisms 70; Iran 28–48
landlords: dues 44; evasion of land reform law 92–4; opposition to land reform 53; options provided by the second stage of land reform 74–6, 96; share of the crop 39
landowners: absentee 9, 25–6, 30–1, 41, 104; social status 9, 29, 51
landownership: statistics 30
law: Islamic inheritance 31, 32, 35; land reform 68, 71–7
living standards of Iranian peasantry 1, 71
loans 45, 46, 71, 85; availability 114–19, 134; *see also* credit
Lodi, H.S.K. 47
Lucas, R.E.B. 14–15

machinery *see* mechanization
malnutrition 45, 143
manuring 41
Marshall, A. 11
Mazumdar, D. 23–5
mechanization 83, 121–3; bimodal agrarian system 59
mechanized land: definition 74–5, 78
migration 60; rural–urban 120, 133, 139, 141, 144–5
Ministry of Agriculture: statistics 3
modernization 83; bimodal agrarian system 59, 61; farming techniques 121–5; unimodal agrarian system 63

nasaqdars 35, 44
Newbery, D.M.G. 14, 23

orchards, exempt from land reform 94, 127

pabuneh 38
peasant proprietorship 31, 54, 102, 105
peasants: income 43–4, 47, 108–9; poverty 44–5, 47, 71, 138–43; sale of cultivation rights to the landlord 75–6, 98–9
Platt, K.B. 42
politics of land reform 51–3
population growth 132
poverty 44–5, 47, 71, 135, 138–43
pre-harvest sale of crops 45, 46–7
price of land: first stage of land reform 73, 77; second stage of land reform 75, 80; third stage of land reform 76
price policy for wheat 130–1, 144

qanats (underground channels) 5–7, 38, 39, 41, 112

rainfall 5
Rao, C.H.H. 15, 21
Reid, D.J. 14, 22, 23
research and development 64
Reza Shah 32–3
rice 7, 39, 125, 129, 130
risk: pre-harvest sale of crops 46–7; tenancy systems 19–22, 26

Salmanzadeh, C. 101
sarbuneh 37

savings: and investment in agriculture 61–2; flocks and poultry 43
seed supplier's share of the crop 38–9
seeds, improved 124
sharecroppers as land reform beneficiaries 65
sharecropping 8, 11–25; co-operatives 37–8; contracts 22–3, 26, 36; division of the crop 38–9; economic outcome of land reform 53–4; efficiency 11–15, 26, 54; investment and innovation 16–19, 41; Iran 32–42; post-reform 104; preference to other tenure systems 19–25; risk sharing 19–22, 26
social status of landowners 9, 29, 51
Srinivasan, T.N. 13, 17
Statistical Centre of Iran 3, 130
statistics: first stage of land reform 94–5; second stage of land reform 96–9; third stage of land reform 100–1; crop yields 130; landownership 30; source of official 3–4
Stiglitz, J.E. 22
sugar 130
superstition 41

technology 111, 121–2
tenancy contracts: choice 21–2; fixed-rent vs share contract 22–3, 26; *see also under* sharecropping
tenancy systems 8–10; criticisms 53; insecurity 41, 47; post-reform 104; risk 19–22, 26; *see also* fixed-rent tenancy; sharecropping
tractors 1, 40, 121, 123
tuyul 28–9

uncertainty *see* risk
unemployment 69
unimodal agrarian system 61–6, 67, 77, 143
United States 71

vaqfs 32; effect of land reform 75; 76–7
village: definition 73, 78

water: access after land reform 112–13; 134; availability 4; conservation 42
wheat 7, 125, 129, 130; imported 131–2; price policy 130–1, 144; yield 42

Zusman, P. 14

THE SECURITY OF THE PERSIAN GULF

Edited by
Hossein Amirsadeghi

ROUTLEDGE LIBRARY EDITIONS:
IRAN



ROUTLEDGE LIBRARY EDITIONS:
IRAN

THE SECURITY OF THE PERSIAN GULF



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

THE SECURITY OF THE PERSIAN GULF

Edited by
HOSSEIN AMIRSADEGHI

Volume 26

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published in 1981

This edition first published in 2011
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Avenue, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 1981 Hossein Amirsadeghi

Printed and bound in Great Britain

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN 13: 978-0-415-57033-6 (Set)
eISBN 13: 978-0-203-83010-9 (Set)
ISBN 13: 978-0-415-61050-6 (Volume 26)
eISBN 13: 978-0-203-83218-9 (Volume 26)

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original copies may be apparent.

Disclaimer
The publisher has made every effort to trace copyright holders and would welcome correspondence from those they have been unable to trace.

THE SECURITY OF THE PERSIAN GULF

Edited by
Hossein Amirsadeghi



CROOM HELM LONDON

© 1981 Hossein Amirsadeghi
Croom Helm Ltd, 2-10 St John's Road, London SW11

British Library Cataloguing in Publication Data

The Security of the Persian Gulf.
1. Persian Gulf region — Foreign relations
I. Amirsadeghi, Hossein
953'.605 DS326

ISBN 0-7099-0505-X

Reproduced from copy supplied
printed and bound in Great Britain
by Billing and Sons Limited
Guildford, London, Oxford, Worcester

CONTENTS

List of Figures and Tables

Acknowledgements

Editor's Preface

Maps

Introduction

1. The Gulf Region in the Global Setting *John C. Campbell* 1
2. The Strategic Balance and the Control of the Persian Gulf *Geoffrey Kemp* 26
3. The Soviet Union and the Persian Gulf *Shahram Chubin* 43
4. The United States-Iranian Relationship 1948–1978: a Study in Reverse Influence *C.D. Carr* 57
5. Persian Gulf Nuclearisation: Prospects and Implications *Lewis A. Dunn* 85
6. The Iranian Military: Political Symbolism Versus Military Usefulness *Steven L. Canby* 100
7. Arms Transfers, Indigenous Defence Production and Dependency: the Case of Iran *Stephanie G. Neuman* 131
8. Saudi Arabia and Iran: the Twin Pillars in Revolutionary Times *Richard Haass* 151
9. The Persian Gulf in Regional and International Politics: the Arab Side of the Gulf *John Duke Anthony* 170
10. Iraq: Emergent Gulf Power *Edmund Ghareeb* 197
11. The Iranian Revolution: Triumph or Tragedy? *L.P. Elwell-Sutton* 231
12. Revolution and Energy Policy in Iran: International and Domestic Implications *Fereidun Fesharaki* 255

Bibliography 281

Notes on Contributors 288

Index 290



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

FIGURES AND TABLES

Figure

7.1	IIAF Aircraft History	136

Tables

2.1	Gross National Product and Military Expenditure	36
2.2	Military Establishments of the Gulf States	37
2.3	Revenue and Expenditure of the Gulf States	38
2.4	Balance of Trade	39
2.5	Area, Population and Population Density of the Gulf States	41
2.6	Volume of Crude Oil Exported 1972–8	42
5.1	Growing Technical Capabilities	88
7.1	The Escalator of Development	135
7.2	Maintenance Manpower Requirements	137
7.3	Distribution of Maintenance Technical Skill Level	138
12.1	Iran's Oil in International Perspective, 1978	256
12.2	Iran's Oil Production	257
12.3	Industrial Countries' Dependence on Iranian Oil, 1978	258
12.4	Industrial World's Trade with Iran	260
12.5	Calendar of Oil Price Changes for Selected Crudes	261
12.6	Iranian Budget: Sources and Uses of Funds	270



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

ACKNOWLEDGEMENTS

I would like to thank Shahram Chubin and all the other contributors for their worthy efforts in constructing this volume, bearing in mind the time-lag from start to finish, and the unforeseen changes in the political situation brought about by the Iranian Revolution and the Iraq/Iran war. A special word of thanks to my friend John St John for his initial co-operation, and to all those other worthy people upon whom one chances in connection with such projects.

Hossein Amirsadeghi



Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

EDITOR'S PREFACE

Because of my special interest in the politics of my country and the increasingly vital significance of the Persian Gulf in world affairs, and because of the general dearth of material on the Gulf, I set out three years ago to produce a volume of essays not only to help fill this vacuum with a general overview, but also to produce specialist studies that will help shed some light on the intricacies of the politics and society of this most volatile area.

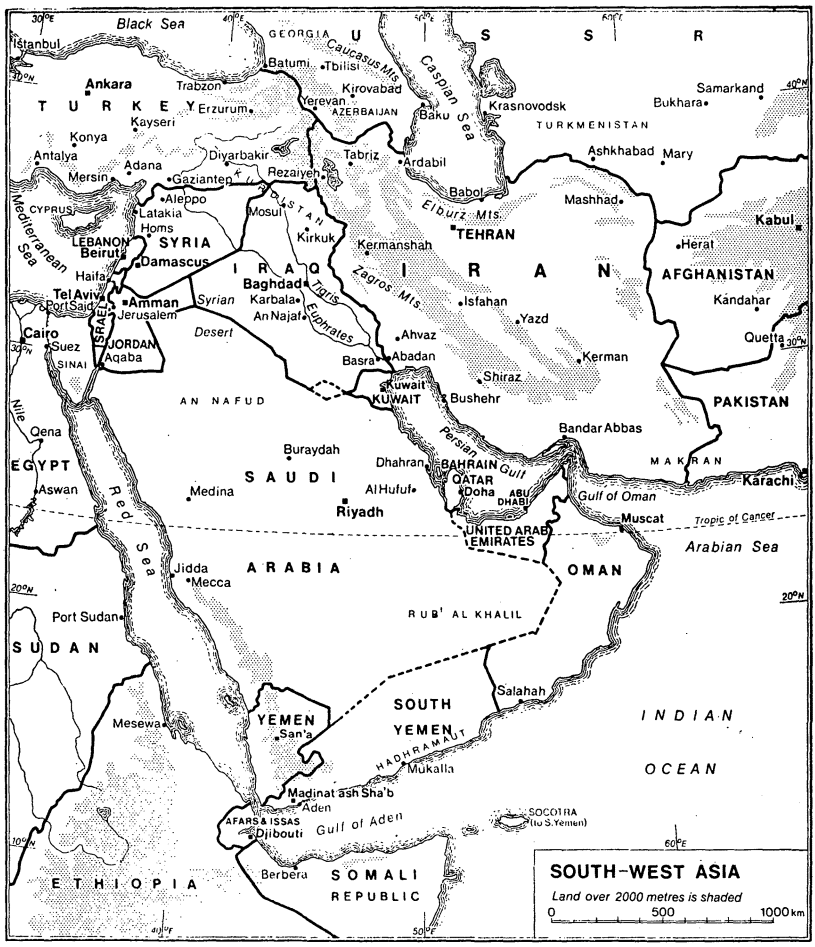
The Persian Gulf, already important because of its vast energy resources, emerged into the limelight of geopolitics at the time of the British Labour government's policy of withdrawal from East of Suez in 1968. Before 1968 it had been recognised that the Gulf lay in the legitimate sphere of influence of Britain, while the United States exerted its influence in the two pivotal littoral states of Iran and Saudi Arabia. The Soviets had been gaining influence in Iraq ever since the overthrow of the monarchy in 1958, and later gained a more secure foothold in South Yemen after the British withdrawal from Aden. The Chinese were also fishing for influence by their support of the Popular Front for the Liberation of the Arabian Gulf, a Marxist guerrilla organisation responsible for the Dhofari rebellion in Oman, and dedicated to the overthrow of the conservative *status quo* of the Gulf countries. But except for some influence in Tehran in the last years of the Shah's regime, the only two remaining arbitrating powers in the Gulf have been the Soviets through their proxies in South Yemen, and to a lesser extent in Iraq and Syria, and the United States, with its increasingly reluctant allies in Riyadh and Muscat.

Since the revolution in Iran and the downfall of the Shah, the whole political balance of the area has been overturned. With their new-found bases on the Iran-Afghan border, the Soviets are only a day's drive from Iran's Gulf ports or the Khuzestan oilfields. With their estimated thirty divisions poised on the Iran-Soviet border, and the continuing turmoil and effective political disintegration of Iran, the Soviet capacity for a direct attack on the Gulf oilfields, and its potential for mischief-making, have increased dramatically. The outcome of the political tug-of-war in Tehran is crucial to the future political map of the Persian Gulf. The collapse of the Khomeini regime will eventually come about; but the question is who or what will replace it. If a moderate regime replaces the present turmoil, the prospects for future stability in the Gulf seem good. If, on the other

hand, extremists gain the upper hand, with no doubt a little help from their Soviet mentors, the prospects for an East-West confrontation over the Gulf's vital energy resources become much more likely.

It is with a view to the above dangers that this volume is directed. The various contributors, each distinguished in his or her own area of knowledge, have been assigned to delve into the manifold intricacies of Persian and Arab politics, society and culture. Although the Iranian Revolution caught everybody by surprise, the chapters have been rewritten with the benefit of hindsight, increasing the scope for analysis. But the reader should bear in mind that the volatile political and socio-religious equation in the Gulf precludes any definitive study of this important subject area in just one volume. With due consideration to the in-built limitations of such a broad-based study, this volume attempts to analyse in one place subjects which have hitherto been approached piecemeal. The success will belong to the contributors, and the limitations are the fault of the editor.

Map1: South-West Asia



Map2: Persian Gulf and Iran



INTRODUCTION

President Eisenhower described the Persian Gulf in 1951 as the most strategically important area in the world. Thirty years and two oil embargoes, a major revolution and a raging war later, there cannot be a more apt description of this most volatile and turbulent area. In terms of sheer riches of petroleum resources, there is no other group of countries that can match it. The eight littoral states of the Gulf hold over 60 per cent of the world's known oil reserves. Their combined petroleum exports in 1978 accounted for 40 per cent of the total world production and their complete shut-off would mean industrial paralysis and devastating economic collapse within a very short time for the industrial nations. The Persian Gulf states' combined imports market represents one of the most lucrative in the world. Vast purchases of armaments, industrial and petroleum plants, manufactured and consumer goods and services are equally vital to the health of world trade. The foreign currency holdings and investments of the Gulf states in Europe and America are becoming staggering. In short, little over 55 million Persians and Arabs hold the free world over the literal barrel.

Little did President Eisenhower imagine how profound his statement at that time would be. Even less could he foresee that thirty years later, and despite America's still awesome power, the value of Washington's friendship would be in such disdain and her ability to influence events in the Gulf so limited. For since the fall of the Shah, with the Carter Administration's muddled abandonment of its longtime Iranian ally, and with its singular ineffectiveness *vis-à-vis* the hostage crisis, all Washington's former friends have scuttled for cover, keeping a safe diplomatic distance, yet demanding increasing security guarantees against encroaching radicalism and the possibilities of direct and indirect aggression by the Soviet Union. Since the Afghan invasion, that threat suddenly seems very real.

In spite of over thirty years of experience in and around the Gulf, the United States seems to have failed in its overall policies. The clear alignments and cold war scenarios began to crumble in the late 1960s, yet Washington could not foresee trends and prepare contingencies to meet inevitable crises with forceful determination. Opportunities were missed and eventually the inability to react with vigour resulted in one of the most devastating post-war reversals for American power and prestige due to the downfall of the Shah. That is not to say that the Iranian revolution was the fault of the United

States, but as the xenophobic hysteria in Iran still points in the direction of the White House, it is only fitting to question why. To have both the remnants of the previous regime and the present Khomeini ensemble blame the US for its demise and troubles does raise disturbing questions.

It is fair to say that Washington was never at home in the 'hotbed' realm of diplomacy and Byzantine intrigue of the Persian Gulf. The British presence up to 1971 ensured a goodly measure of compliance and co-operation, for Whitehall had been at home in the Gulf since the British superseded the Portuguese and the Dutch in the eighteenth century. The Gulf was secured to ensure free passage to India and a British Resident presided over events in the area.

After the discovery of oil at Masjid-e-Suleiman in 1908 in Iran's Khuzistan province, the British looked upon the Gulf as their 'lake' and let it be known to all that it was their 'legitimate' sphere of influence. The Arabian Peninsula was wrested from the Turks, and after the collapse of the Ottoman Empire at the end of the Great War, Mesopotamia was mandated to the British by treaty. Thus, apart from Persia, which was never colonised, the Arab Gulf states were creations of Whitehall and their frontiers drawn up in London. Most became independent states only after the Second World War, and countries like Kuwait have a history of nationhood dating back all of thirty years. The United Arab Emirates was Britain's last imperial hurra, being granted independence in 1971 as Britain finally withdrew from the Gulf.

After their departure, the Shah made it forcefully clear that Iran intended to fill the resultant vacuum and, with the co-operation of the Arab littoral states, to ensure that no foreign power ever again play the dominant role. To thumb his nose at the British, he seized the small islands of the Tumbs and Abu Musa before Whitehall's official date of relinquishment of treaty obligations.

With the advent of the Nixon doctrine of self-help, the Shah soon assumed the mantle of 'policeman of the Gulf' and through Iran's increasing economic, political and military strength — especially after the 1973 OPEC oil price increases — Washington was more than content to let affairs develop and around their 'twin-pillar' strategy of support for Iran and Saudi Arabia. The 1970s saw dramatic economic change in the midst of stability, and the Gulf states managed to sort out their problems with mutual amicability. The threat posed by violent border skirmishing between Iran and Iraq was settled between the Shah and Saddam Hussein at Algiers in 1975, and Iran's increasing military might did not lead her into adventurous or an expansionist foreign policy.

The Shah's forces even put down a long festering Marxist guerrilla war in Oman's Dhofar region, achieving what more limited British

and Jordanian forces had not managed in several years of fighting. The political volatility and unreliability of oil supplies was shown up on the two occasions of the Arab-Israeli wars in 1967 and 1973, at which times the Arab Gulf producers boycotted sales to America and Europe. The Palestinian question continued to radicalise Arab politics, and the occupation of Jerusalem became an emotional issue with the Arabs. Gulf politics became entangled with Pan-Arab interests, and Palestinians living in the Gulf mini-states gave cause for conservative alarm. But the enormous wealth brought about by leap-frogging oil prices allowed a measure of relaxation and brought considerable political and international clout. The concept of buying one's stability and bribing other's friendship became a pronounced and effective lever of diplomacy as utilised by the Gulf Arab states. Meanwhile, overall security of the Gulf was assured by Iran's growing military predominance, albeit reluctantly accepted by the Arab Gulf. Threats from radical Iraq were contained, but an unprecedented arms race ensued in the area.

The Soviet Union was meanwhile making the best out of a bad situation, entrenching its position in South Yemen and trying to retain a base with its increasingly reluctant ally, Iraq. Although the Iraqis had depended entirely on the Soviets for their arms, the French began to sell Mirage jets to both Iraq and Kuwait, and Baghdad became increasingly cool towards and independent of Moscow. Communist manufactures and industry proving shoddy, Baghdad moved closer economically to the West while maintaining an outward radical posture. After the loss of Arab prestige by Sadat due to the Camp David accords, Saddam Hussein moved to fill the leadership role. This was made easier by the collapse of the Shah and the decline of Iran's power in the Gulf. Assuming the mantle of leader of the non-aligned movement after Castro, Saddam Hussein, uneasy because of the overspill possibilities of Iran's Shi'ite revolt (as 60 per cent of Iraqis are Shi'ites while the ruling Takrit cabal are Sunnis) and with increasing signs of internal unrest, stumbled into war with Iran. Having failed to achieve a swift victory, or any victory worth mentioning (at the time of writing), the Gulf became for the first time the scene of military conflict. Whatever the outcome of the Iran/Iraq war, the Gulf has been saddled with a perennial future problem more explosive and destructive than the Palestinian issue could ever be.

With the Arab Gulf states siding — somewhat reluctantly — with Iraq, the old ethno-religious battle lines were redrawn, with Persians and Arabs looking across the Gulf at each other as old enemies.

Iran's much larger population and resource potentials will ensure her long-term supremacy — notwithstanding the short-term revolu-

tionary chaos — in the Gulf and its revenge of the Iraqi invasion will not be long in coming. It will be interesting to see where Tehran eventually turns for succour and arms and how the superpowers will be pulled into this new conflict, even if they do not become entangled this time round. The Soviet's assiduous courting of the unpalatable regime of the Mullas is a sign of what they are willing to swallow in the hope of winning over strategically important Iran, and they have shown that they are quite willing and capable of doing a Somalia-Ethiopia flip between Iran and Iraq given half a chance.

At this point in time, Washington's options *vis-à-vis* Tehran seem rather limited, and, unless there is a quick change in regime in Iran, America will find herself ever more visibly entangled with the fate of Saudi Arabia and the sheikhdoms — a fact which will play into the hands of the radicals and inevitably the Soviet Union. For the danger to stability in the area is not so much from overt outside aggression as from subversion, fifth columns and radical *coups d'états*, as happened in Iran. So until the Iranian débâcle is reversed, Washington will become more bogged down than ever in the Gulf through its forward-basing policies designed to protect the *status quo* in Arabia, risking possible military entanglements. The nature of such contingencies and the opposing forces liable to be sucked in are no doubt the subject of much deliberation in both the Kremlin and the Pentagon. It is precisely such possible entanglements that will lead to superpower confrontation and a very real threat to world peace in the critical first years of the 1980s. With the Soviets bogged down in Afghanistan, their long-term energy resource prospects bleak, increasing cracks within their own colonial empire, a cumbersome and inefficient economy on the way to rapid decline, a prospective unsettling battle for leadership after Brezhnev's death and an unprecedented state of strategic parity with and military superiority over the West, it is not just this writer who is anxious about the future of the thirty Soviet divisions straddled along the length of the thousand-mile-long Iran-Soviet border. Whatever happens in Iran will determine the future shape of politics in the Persian Gulf, and possibly the shape of the political map of the world.

1 January 1981

1 THE GULF REGION IN THE GLOBAL SETTING

John C. Campbell

I

It is no new experience for the nations of the Persian Gulf to find their region an arena for the conflicts of the world's great powers. It is a new experience for the great powers themselves, however, to have to come to terms with their own dependence on the resources, and on the decisions, of those nations which in earlier times had been little more than the objects of their strategy and diplomacy. It is this new factor and its interaction with the old one that provide the new global setting for the Gulf region.

The familiar fact of great-power involvement remains, changed in form but not in essence. The United Kingdom, relieved of past concerns for the security of the routes of empire, is no longer the leading actor. Since the Second World War the United States has assumed the major responsibility for the defence of Western interests and for confronting or containing the southward push of Russian power begun under the tsars and continued under the Soviets. With but two great powers, or superpowers, in the world after the Second World War, ideologically hostile and fearful of each other's intentions, it was inevitable that the areas lying between them should be drawn into the global competition known as the Cold War.

The Gulf and the rest of the Middle East had long been labelled as strategically important by contending great powers for elementary reasons of geography. This region lay athwart major routes from Europe to the Far East and from Asia to Africa. It was on Russia's doorstep, a possible launching point for hostile action against Russia or a prime target of Russian expansion. It contained important resources of oil. Domination or control of particular parts of the Gulf region could give one power a decisive military advantage in case of war. Accordingly, no contending power wanted to risk an unfavourable shift of the balance in time of peace.

These considerations were still present in the period of the Cold War. General Eisenhower was only stating a generally accepted proposition when he said in 1951, 'So far as sheer value of territory is concerned there is no more strategically important area in the world.'[1] But in an age in which any general war would be nuclear and

enormously destructive through direct exchanges between the principal powers, the military value of peripheral areas did not loom so large. The possibility of limited war in the Middle East or the Gulf was not ignored, but neither the Soviet Union nor the United States appeared to accept the proposition that they could fight each other in that area without becoming involved elsewhere. Nevertheless, each felt it must be prepared. American policy in the 1950s rested on the doctrine of massive retaliation and on the building of a barrier of regional organisation and treaty commitments, backed by the ability to deploy military force along the northern tier (i.e. along the southern border of the USSR). Although this latter capability was never established on a credible scale, the general purpose of American policy remained fixed: to deter any Soviet aggression, limited or general, by the totality of the power of the United States.

In addition, the policy of containment was deliberately applied in a global context. Its high ideological content included the conviction that Communist aggression or expansion must not be allowed to succeed anywhere, lest it eventually succeed everywhere. The Gulf area obviously had weight in the global balance. If an appreciable portion of it should be brought under Soviet control, the political effects on the rest of the world might be disastrous for the West. But Washington did not engage in any fine calculations as to what areas of the 'free world' were important enough to defend. Such was the tension of the Cold War, especially after the attack on South Korea in 1950, that any gain for the Soviets (or Chinese) had to be resisted.

The United States was much concerned with what Secretary Dulles called 'indirect aggression', the taking over of a country through subversive agents, guerrilla movements or local Communist parties, all aided and manipulated from outside. It was not easy to resort to military intervention to deal with those contingencies, but it was important to have every feasible means available to prevent or reverse what the Communist powers might achieve in this way. The situation in Greece which gave rise to the Truman Doctrine was one in which an 'armed minority' aided and allegedly directed from outside was trying to overthrow an elected (and pro-Western) government. American military and economic aid to Turkey, Iraq, Iran and Pakistan was aimed at preventing any success which the Communists might gain by any combination of military threat, subversion and political pressure.

For the USSR, of course, the picture was quite different. The Gulf region was close to Soviet territory and far from the United States. Yet it was the latter which was forming blocs, sending arms and military missions, and setting up military bases that could support attacks on the Soviet Union. As a natural reaction the Soviet leaders could be expected to try to break up the American-sponsored bloc

and deny to the United States use of the area for military purposes. To that end they worked to undermine governments which co-operated with the Americans and to encourage those forces which opposed them; that is, they did what the Americans were defining as 'indirect aggression'. That defensive strategy, however, fitted perfectly with an offensive strategy, dating from the Revolution of 1917, of support to local Communist parties and other 'progressive' forces in their efforts to combat Western imperialism and to weaken and replace the existing regimes.

This special pattern of superpower competition, it soon became clear, was less a direct contest of strength than a game in which each had to act through its relations with the local states and in the arena of local politics. Unquestionably the superpowers had a great deal of influence, given the differences in the attributes of power — not only military but others as well — but the local states were not without resources and strengths of their own. They profited from the new atmosphere of the post-war international order, from the existence of the United Nations where they could make their voices heard, and from the anti-colonial revolution then taking place. National sovereignty had acquired a new prestige and could not easily be ignored.

Several other factors, moreover, which marked the historical experience of the Gulf peoples, gave added weight to their ability to cope with outside powers. One was the combination of cultural continuity and tradition, even inertia, that had enabled the peoples of the Ottoman and Persian empires to absorb pressures from outside and survive for centuries despite internal weakness. A second was the fact that British power was for years exerted along the routes important to the Royal Navy and to British commerce in a pattern of treaties and arrangements (unequal ones, to be sure) with local rulers, rather than in the pattern of the conquest of India. The British found no reason to try to dominate the interior of the Arabian Peninsula or to destroy the independence of Persia. A third factor was the room for manoeuvre which the balance between outside powers and the configurations of world politics allowed to smaller states caught in the middle; Iran, for example, was able to play the British against the Russians, and the Germans and later the Americans against both. Finally, with the upheavals of two world wars, the growth of nationalism, the Russian Revolution and its world-wide impact, creation of the mandate system and acceptance of the principle of eventual independence for Iraq and other Arab states, and the increasing influence of the United States as a friend of democracy and self-determination, the basis was laid for the Gulf states to be subjects, not just objects, of the international relations of their region.

The containment policy of the United States, which sprang into being during 1945–6 with diplomatic moves supporting Turkey and Iran against Soviet pressure, and was given solemn official expression in the Truman Doctrine of 1947, was intended not just as an American or Western enterprise but as a collaborative effort with the states of the Middle East and the Gulf. Its purpose, in the genuine belief of the United States, was linked to their vital interest in the maintenance of their integrity and independence. The collaborative aspect of the policy thus rested on the assumption that the governments and peoples of the region saw matters in this light, that they looked upon the Soviet Union and Communism as the main threat to their independence and upon the United States and the West as their natural allies in meeting it.

This assumption was only partially valid. In those countries which bordered on the Soviet Union it was generally accepted. It made possible the formation of the Baghdad Pact, intended as a barrier to Soviet aggression or penetration along the northern tier. It found a mixed and preponderantly negative response in the Arab world, where in the eyes of powerful political forces, in government and outside, the principal threat was not the Soviet Union but Israel, other Arab states, the European colonial powers, or the United States itself. American policy in practice represented an effort to apply a rather simple formula — common defence of the free world against Communism — to a very complex local situation. It involved the United States in the local conflicts of the Middle East and the Gulf and in the internal politics of individual states, for in seeking, finding and aiding friends and allies in the region the Americans became the enemies of all those who for reasons of their own were opposed to those friends and allies. Some of the strongest forces of Arab nationalism, in order to achieve ends of the highest legitimacy to themselves — liberation, independence, unity, recovery of Palestine — turned for help to the only power which could counterbalance American influence, the Soviet Union. It was partly the American policy of containment that opened the doors to a larger role for the Soviet Union in the Middle East. The most dramatic signal of this change in the Eastern Mediterranean was Egypt's turning to the Soviets for arms in 1955, and in the Gulf region was the revolution of 1958 in Iraq.

In the 1950s and 1960s the Soviet Union made extraordinary gains, increasing its political influence and improving its military position. But as one looks back on that period, it is perhaps more extraordinary that the gains were not greater. It could even be said that despite its failures the American policy of containment succeeded, in that no country of the region became a Soviet satellite on the East European model. That result was not due entirely to the

skill of American diplomacy. If anything, Soviet diplomacy was even less effective in handling local leaders and exploiting opportunities. The basic reason was that, in situations where the overwhelming force at the disposal of the outside powers was not a viable instrument, the political gains they made were always subject to reversal in any local state by a change in regime or in policy. Governments which showed their independence by calling in the Soviet Union to frustrate the British or the Americans were not of a mind to give up that independence to their new friends.

The relatively stable situation with respect to security in the area of the Gulf and the Arabian Peninsula, in contrast to the Arab-Israel region, was due to a number of favourable factors. One was the continuance of a British role for more than a decade after the Suez affair of 1956 brought British influence in the Eastern Mediterranean area close to the vanishing point. By retaining membership in the Central Treaty Organization (CENTO), a military presence in the Gulf, treaty relations with Oman and the Trucial States, and a security guarantee to Kuwait, the British were a stabilising force even as they debated at length why they were there and whether they could afford to stay. Their departure from Aden in 1968, to be succeeded there by a radical Marxist regime, was no contribution to the stability of the region, but the coming to full independence of Bahrain and the United Arab Emirates was accomplished with a minimum of difficulty.

The British presence in the Gulf could only be effective, however, because of the steadiness of the two large states on either side of the Gulf, Iran and Saudi Arabia. Thanks to oil wealth, firm monarchical leadership (in Iran after 1953 and in Saudi Arabia after 1964), and their ability to contain the expansive thrust of radical Arab nationalism or Marxism — whether from Cairo, Baghdad, Moscow or elsewhere — they laid a basis for regional security that made it possible for the British to withdraw in 1971 without bringing on a period of unrest and crisis or creating a vacuum to be filled by another outside power.

When the United States, in the Nixon Administration, began to talk about a 'two-pillar' system for the security of the Gulf, with Iran and Saudi Arabia as the pillars, there was an element of wishful thinking in the idea, as Saudi Arabia had but a very limited capability and Iran's willingness to co-operate with Arab states was open to question. Yet the American approach, in stressing the responsibility of local states and refraining from moving in to replace the British, was well attuned to the mood of the Gulf rulers and well calculated to keep the Soviets from moving in. It was based also on firmly established relations with both Iran and Saudi Arabia, including security commitments to both of them and a strong mutual

interest involving oil, trade and development. It was this 'special relationship' with each of the two Gulf states, rather than the fading CENTO, that the United States saw as offering the best chance for maintaining a stable balance in the Gulf region, keeping Soviet influence to a safe level and preserving the vital Western interest in access to Gulf oil. The events of 1973–4 and then those of 1978–9 put those views and policies to a severe test.

II

In the closing months of 1973, as the October War shook not only the Middle East but the entire world, the oil states of the Gulf asserted a new 'oil power' that neither they nor the industrial countries which were the main consumers of their oil were truly aware they possessed. The Arab states, including America's great friend, Saudi Arabia, restricted the production and export of oil, in order to strike at those countries, primarily the United States, that were aiding Israel. The Arab-Israel conflict and relations with the Gulf states, which American diplomacy for years had tried to keep separated, had now merged indissolubly together. Then OPEC, with the full participation and support of Saudi Arabia and Iran, raised the price of oil fourfold. At the time, this reversal of the historic dependence of the underdeveloped on the developed, loudly applauded by the Third World and by the Soviet Union, seemed to put the vital interests of the West in jeopardy, to break its solidarity, and to call into question the very foundations of American policy in the region. It was at this time, when the Europeans and Japanese were bending over backwards to win the favour of the oil producers, that articles appeared in the American press suggesting the use of force to seize Gulf oilfields, American officials spoke of the need to force down the price of oil, and the Secretary of State said that a nation would inevitably use force if faced with 'strangulation'.[2]

The persisting atmosphere of confrontation, however, did not reflect the underlying long-term interests on both sides. The success of American diplomacy in arranging negotiated truce arrangements which Israel agreed with Egypt and with Syria and in mending American relations with the Arab world gave Saudi Arabia the opportunity to end the oil embargo and restore traditional ties of co-operation with the United States. The oil-consuming countries organised to protect themselves against new interruptions of supply, but at the same time they adjusted as best they could to the new OPEC prices. The practical outcome of the crisis was that the oil states of the Gulf had a great deal more money and a great deal more power. What they wanted to do with it, however, was not to bank-

rupt the West but to buy the West's goods, skills, technology and arms, and with them to complete the process of emancipation from outside influence and solve their own national and regional problems to their satisfaction.

The general shape of the security problem in the Gulf did not change. Both Iran and Saudi Arabia began to buy American arms in much larger quantities in order to fulfil their expanded roles and ambitions. All the oil states of the Gulf embarked on vast new plans of development, using their expanded revenues to buy capital goods and technology from America, Europe and Japan. Even Iraq, though closely tied to the USSR in economic and military arrangements, sealed by a formal treaty in 1972, turned to the West as the only place it could get what it needed to keep up the pace of development. The *rapprochement* of Iraq with Iran in 1975 was made possible by the former's success against the Kurdish rebellion and the latter's decision to cease aiding the Kurds, but in a larger sense it was the result of the new power and independence of all the Gulf states. They were acting on their own, not as clients of one or another superpower. Whether Moscow approved or disapproved of Iraq's actions is not clear. For public consumption it voiced its approval, but it was evident that Soviet influence in Baghdad and in the region as a whole had not grown as a result of the oil crisis and the change in bargaining power between the Gulf states and the West.[3] Russia's principal gain was economic: its own oil now brought higher prices on the world market.

A major change took place in 1978–9: the revolution in Iran. It grew from internal conditions, not from external stimuli, and the new regime, virulently anti-American, gave no sign of alignment with the USSR. But the change still represented a shift in the superpower balance, whether seen in geopolitical or psychological terms, because of the obvious loss of prestige and position suffered by the United States. The regime of the Shah, who had been a friend and ally, was gone; so were the ties with the Iranian military, the monitoring stations near the Soviet frontier, and the close involvement with the Shah's development projects. Above all, the psychological effects were deep: on Iran, where the victorious revolutionary forces were taking every opportunity to humiliate the United States, the ally and supporter of the hated Shah, even to the point of attacking its embassy in Tehran and holding captive its diplomatic personnel; and on America, whose people had a sense of disaster, of weakness and failure, calling into question the entire basis of policy in the Gulf region.

Iran's future alignment was not clear, because nothing was clear in the chaotic situation that evolved. The Ayatollah Khomeini and the dominant religious elements in the leadership had no sympathy or

ties with the USSR, but it was not apparent that the same could be said of Marxist and other radical groups which had shared in the revolution (although the Communist Tudeh Party had a very limited following). Dissidence and separatism among non-Iranian peoples in frontier areas, moreover, were open to exploitation by the Soviets. The latter, using due caution and not undertaking any adventures, were giving verbal encouragement to all Iranian moves directed against America and seeking to improve opportunities for exploitation when the time might be ripe, towards the ultimate aim of a 'friendly' (i.e. compliant) Iran.

The Soviet invasion of Afghanistan in December 1979 was not directly related to the events in Iran. Yet they may have contributed to it in two ways: (1) through the astonishing strength of militant Islam which, in the ascendant in Iran, gave impetus to the resistance to the pro-Soviet regime that had been in power in Kabul since April 1978 and could have repercussions in the Soviet Union itself; and (2) through the obvious weakening of the US position in the region, which may have influenced the decision of the Soviets to move in.

The effect, of course, was a shift in the balance, bringing the Soviet ground and air forces closer to the Indian Ocean and to the oilfields and oil transport routes of the Persian Gulf, and increasing Soviet pressure on Iran and Pakistan. Whether the move into Afghanistan was primarily defensive, to save a satellite regime and avert an intolerable blow to Soviet prestige, or to counter manoeuvres by China, or was a step towards further expansion, the rest of the world had to deal with the fact that the military frontier of the USSR had been moved southward and that a former buffer state had been absorbed.

III

The present security problems are of three kinds: (1) those stemming from the global concerns of the superpowers which extend to the Gulf region, and in a sense are imposed on its peoples, just as they extend to other regions of the world; (2) those related to the policies of the regional state towards outside powers, whether based on fear or confidence, economic strength or economic need, ideological affinity or cultural aversion; and (3) those arising from the local and regional aims and policies of the Gulf states within and among themselves.

Let us consider for a moment, perhaps in too mechanistic a way but as a necessary means of capturing their primary concerns, how the problem looks to the two superpowers. The peace of the world, in default of an accepted working system of international law and

security, depends on the balance of military power between them: the global balance in strategic weapons, which provides mutual deterrence on that level, supplemented by a balance at a lower level in regions where both powers have vital interests and where military action could either immediately provoke global nuclear war or set in motion a series of actions and counteractions leading to the same result. These are truisms that have been accepted for so long that people tend to forget them. In the Gulf area the two superpowers have been and are in competition with each other, but they have had a common interest in keeping that competition below the level of military conflict and presumably wish to continue to do so. But neither knows what risks the rival power will take, and neither wishes by displaying weakness to encourage risk-taking by the other.

To preserve the military stand-off, each power has to have the capability of bringing military force to bear in the Gulf region. It is by no means clear on the Western side how much military force, and in what locations, is necessary for this purpose. It may be that in a crisis that threatened to attain global proportions the United States would not wish to commit forces to this region at all. The main purpose is not to be sure of winning a local war but to deter any military move from the other side that would jeopardise a vital American interest. It is a matter of adequately understood intentions and of political will as well as of military forces and dispositions, but the latter constitute an essential element. The reach of Soviet military power has been steadily increasing with the build-up of air and naval power, and Washington can at least imagine a series of Soviet moves that could threaten the vital oil routes from the Gulf through the Strait of Hormuz and across the Indian Ocean.[4]

Provoked by the invasion of Afghanistan, the United States saw the need for creating a stronger capability for military action in the region. It was to consist of greater naval power in the Indian Ocean, the building up of a 'rapid deployment force' that could be moved to where it was needed in case of crisis, and the use of naval and air facilities in nearby states willing to co-operate.[5] This capability was not yet in being at the time President Carter declared, on 23 January 1980, that the United States would resist with all appropriate means, including military force, any Soviet move (presumably in the direction of the oilfields, the Gulf or the Indian Ocean) representing an assault on the vital interests of the United States. The aim of the declaratory policy was deterrence: figuratively drawing a line on the map and letting the Kremlin know that crossing it would mean conflict with the United States, whether or not the latter was ready to contest every metre of Middle East territory.

The President's declaration, incidentally, bore some resemblance to the statements which both the Soviet Union and the United States

had made regarding Iran in the autumn of 1978, as the Shah's rule was disintegrating and neither power was sure what the other might do; so each warned the other of the dangers of military intervention. But so far as the United States was concerned, the Carter statement of January 1980 was a much more measured and portentous announcement of policy.

For the most part the superpowers have created this system of security, based on balance and deterrence, by their own unilateral actions, building up their respective military positions and seeking allies and clients in the Gulf region. Each has posed as defender and protector of the true interests of the local nations against the designs of the other. That was the rationale on the American side, as we have seen, of the Truman Doctrine and all the military dispositions, alliances and aid programmes that followed. On the Soviet side it was the rationale for concentration of forces on the southern frontiers of the USSR, for open threats of force against the 'aggressive' Baghdad Pact and the Central Treaty Organization, for aid to 'national liberation movements' and 'progressive forces' and to states such as Iraq where those forces were in power.

There is another and less dangerous way, however, by which the superpowers might contribute to security in the region, and that is by agreement with one another to control and lower the level of their military confrontation or even to accept the idea of spheres of influence. The 1970s, after all, were the era of *détente* in Europe with a *de facto* mutual acceptance of spheres of influence, creating a situation of greater stability than existed in the time of extreme Cold War. In the Middle East and the Gulf such a stabilisation has not been possible. Neither superpower has recognised the other as entitled to have a preferred position in any Gulf state (even though in fact it might have such a position). The Soviet Union has shown that it is prepared to live for long periods with a situation it regards as basically intolerable. It accepts Turkey's membership in NATO. It conducted normal relations and economic co-operation with the Shah's Iran despite the Shah's close ties with the United States. But the Soviets have not regarded such situations as anything but transitory. While they have shown that they can be extraordinarily patient, and a transitory situation may last a long time, a study of Soviet writing and Soviet policy over the years leaves little doubt that what they seek in the long run in the Gulf region is a position of preponderant influence.[6] Their relative gains in military power, their sense that the 'correlation of forces' is working in their favour, must have strengthened their view that that goal is attainable.

There is no clear line, as there is in Europe, beyond which military ventures by the superpowers are ruled out. The situation in the Middle East and in the Gulf area has no such stability, no equivalent

to the Helsinki Agreement. The United States has always thought of the containment line in the Middle East and the Gulf region as running along the southern boundary of the Soviet Union. Now, with the Soviet occupation of Afghanistan, perhaps the line is displaced southward *de facto*, even though the Soviet Union is being pressed verbally to withdraw its forces. The remainder of the region, in the American definition, is to be kept free, i.e. out of Soviet hands. The Soviets, of course, see that interpretation as an attempt to claim the entire region as a sphere of American influence. That is unacceptable to them and actually is not insisted upon by the United States. The Americans have had to deal at one time or another with the reality of preponderant Soviet influence in Afghanistan (well before 1979), Iraq, Syria, Yemen and the PDRY, but they gave no formal recognition of it, just as Moscow accepted no American preponderance in Saudi Arabia, Kuwait, the United Arab Emirates or Iran. Each superpower hopes and works for changes of regime and of alignment in the states which now lean to the other side or which choose strict non-alignment. The result is continuing tension and insecurity in the superpower balance, even as, on the global level, that balance operates to maintain the peace.

The Gulf states, in any case, do not wish to be labelled as being in any power's sphere of influence, formally or *de facto*. If the superpowers can reduce the risks of war by moderating their competition, well and good, but not if it is done at the expense of the states of the region.

If spheres of influence are unacceptable and the present open competition for the alignment or realignment of local states is uncertain and dangerous, there remain two possibilities of achieving greater stability through agreement of the superpowers. One is the recognition by both powers of the non-alignment of the entire group of Gulf states, thus making it unnecessary for either power to seek allies there as a means of denying the area to the other. That is a serious and promising possibility, to which we shall come back later. The other field for negotiation and agreement by the superpowers is that of the limitation of arms levels, their own and those of the local states.

Limiting arms deliveries is a subject much talked about in the United States. Both the Carter Administration and Congress professed to see the need for checking the flow of arms into the Gulf region as it reached flood proportions in the late 1970s. But they found it difficult to do very much about it, since arms sales are an integral part of the complex military and economic relationships the United States has developed with its friends in the region, and those relationships they do not wish to put in jeopardy. It was an extraneous factor, the revolution in Iran, that put an end to the largest of

the arms sales programmes, but deliveries to Saudi Arabia continued at a high level, and the Yemen Arab Republic and Oman were added to the list of recipients. Against the strong arguments that the provision of arms is necessary both to enhance security and to keep friendly governments friendly, the more general considerations — that arms races increase the risks of war, that the United States should not have the role of the world's leading 'merchant of death', and that the Gulf countries cannot really use the advanced weapons they are buying and should spend their money more wisely — do not seem weighty enough to bring the United States to adopt severe limits on its own arms deliveries or to make an agreement with other suppliers (including the USSR) to impose such limits. Nor have the Soviets shown much interest, although they may well do so in the future; since most of the arms going in are from the United States, it will be to the Soviet Union's advantage to have limits placed on them.

Agreement with the Soviets to limit or reduce the two powers' own military presence in the region is another matter. Here it is or should be possible to agree on limits or reductions that do not change the existing balance of power and deterrence and at the same time reduce the tension and the costs that go with it. The bilateral negotiations on the naval deployments of the two powers in the Indian Ocean are a case in point. UN resolutions for the denuclearisation of the Indian Ocean or for declaring it a zone of peace have no practical effect. Serious negotiations between Moscow and Washington to freeze or eventually reduce their naval presence, such as took place in 1977 and 1978, could conceivably produce results. But those talks virtually broke down even before the Soviet move into Afghanistan, and after that they could hardly be resurrected, as a build-up of naval power in the Indian Ocean was a central part of the American response.[7] Soviet-American *détente*, since the hopeful days of 1972, has demonstrated its fragility and its limitations. It has not yet been extended to the Middle East or to the Gulf region except in the most elementary sense of recognising the common interest in avoiding a nuclear war.

IV

Considering the awesome power arrayed on both sides of the global balance, the states of the Gulf region might appear to be wholly without influence on the maintenance of security in their area. That is not the case. They have to adjust to the facts of powers, yes, but the balance between the superpowers and the recognised rights of sovereignty give them the opportunity to make choices. Each state's

main concern is the preservation of its own independence and security. If it fears attack or pressure from one superpower, it may align itself clearly and openly with the other; it may be more cautious, leaning to one side but avoiding unambiguous alignment with it while seeking reinsurance through agreements and friendly relations with the other; or it may choose a neutral position, depending on the balance, on the absence of a threat to either side, and on an ability to manoeuvre between them. These are not sharply defined alternatives but they give an idea of the range of choice. It can be argued that Iran during the past thirty years has followed one or other of these policies at different stages, and at times has combined elements of all three. The regional states have made their decisions, moreover, not only against the background of relations with the big powers. For them security and national interest is as much involved, perhaps more so, in their relations with each other, and their alignments with outside powers may be useful, or dangerous, to them in that context.

The point is that while the global balance and rivalry of the superpowers may dominate the destiny of the Gulf region, as it does other regions, by virtue of the potential for war and the use of force, neither Soviet expansionism nor Western imperialism dominates the states of that region; and indeed the politics of those states, in the absence of war, has a determining influence on the shape which the superpower competition and balance takes there. The considerations of national security, economic interest or ideology that move Middle East governments and peoples determine whether an outside power secures a military base, a military alliance, a privileged position or a special relationship: or encounters indifference or hostility.

Let us look at some cases. Iraq, in the era of Nuri es-Said, signed the Baghdad Pact and thus aligned itself with the West. Behind that decision were various reasons: the historic ties of the ruling group in Iraq with the United Kingdom, the desire to get arms from the United States; the fact that Iraq did not feel threatened by the West; the fear of radical movements in the Arab world and at home. Nuri saw stability on the international level linked with stability at home. On the Western side the position of the United States and the United Kingdom *vis-á-vis* the USSR was perceptibly strengthened by his decision to join the new regional security pact, and therefore they acquired an interest in his maintaining his position at home. So much were these factors tied together that when revolution came to Iraq in 1958, a combined movement of Arab nationalism and social protest, it brought about almost automatically a *de facto* reversal of alliances, withdrawal from the Baghdad Pact, a turning to the USSR and the Soviet bloc, and a potentially drastic shift in the global balance. From a mood of satisfaction at having contained Soviet

expansion by building a wall of Western-allied states across the northern tier, Washington was thrown into a mood of pessimism and fear that Soviet power now extended through Iraq to the shores of the Persian Gulf and would move on further into the Arab world. Iraq by its own actions determined how the two superpowers looked at their own security problem.

Under a succession of radical regimes since 1958 Iraq has been generally at odds with the Western powers and especially with the United States, because of the Palestine question, disputes with Western oil companies and ideological antagonism towards 'the imperialists'. Iraq opened the door to military and other aid from Moscow, and these ties were measurably strengthened after the Ba'th regime took power in 1968. But the road of collaboration was not always smooth. The ideology of anti-imperialism was a factor of solidarity, although the Iraqi Communist Party was often more an obstacle than a help to good relations. The global interests of a superpower did not automatically coincide with the more parochial interests of a regional state. Iraq was more or less isolated in the Gulf and in inter-Arab councils, where its only friend was Libya. It was never able to get full Soviet support for its Arab policies, particularly its hostility to Syria and its extreme stand on Palestine; nor was Moscow able to count on Iraq's support for all its international policies. The role of patron of the Baghdad regime created obstacles for the Soviet Union in its relations with other Gulf states, but since it had so few friends and allies it did not wish to lose its one foothold in Iraq. Iraq, for its part, still felt the need for the Moscow connection but took care to limit the military aspects (reports of a Soviet naval base at the port of Um Qasr were not substantiated) and to look for other ties, both in the region and further afield, that suited its own rather than Moscow's interests.

Iran differed from Iraq in that it had long experience of invasion and pressure from the north. Exposed also to British incursions from the south, it had developed a policy of neutrality to safeguard its independence. But with the disappearance of the British threat after the Second World War the Soviet threat was all the more ominous, and it was natural for Iran to turn to the distant United States as the guarantor of its security. In this case too the ideological factor influenced foreign policy. The government of Iran feared Russia not only as an imperial power but as the propagator of an alien ideology, Communism, with ties to a revolutionary party in Iran, the Tudeh. In the revolutionary situation that developed from the oil nationalisation crisis of the early 1950s, the Tudeh and the anti-Western nationalist forces of Mohammed Mosaddeq were not able effectively to join forces to overthrow the Shah, and the Soviet Union cautiously kept hands off.[8] Thereafter Iran made an advantageous

deal with the Western oil companies, joined the Baghdad Pact and relied increasingly on the United States.

For twenty-five years the foreign policy of Iran reflected the outward strength and enhanced confidence of the regime. The Shah retained his concern with the threat from the Soviets, but by the 1960s he felt strong enough to give them a pledge that he would not permit any power (i.e. the United States) to have strategic bases on Iran's territory; he was also ready to establish normal relations with them, and to make large-scale economic deals. He retained his reliance on the United States for ultimate security and for arms supply, but he regulated the military relationship in accordance with his concept of Iran's interest. As his military establishment grew, he envisaged for his country a dominant role in the Gulf and even beyond it; he gave substance to his pretensions by sending Iranian armed forces to Oman to help put down the radical revolt in Dhofar and to safeguard navigation through the Strait of Hormuz. Yet it was not military power but the ability to make its own decisions that gave Iran the capacity to affect the superpower balance and the nature of the security system in the Gulf. Although Washington encouraged the idea of Iran as a pillar of security in the Gulf, the Shah did not consider himself a surrogate of the United States. His goal was a strong and independent role for Iran.

The revolution that brought Ayatollah Khomeini to power changed Iran's domestic scene totally, but its effect on foreign policy was less than appeared on the surface. True, the new regime repudiated the Shah and all his works, including the American connection and the effort to make Iran a leading power in the region. But the revolution did not change the fundamental national interest in the country's sovereignty and integrity. It strengthened the drive for an independent policy, not tied to that of any big power, and for a return to historic traditions in dealing with the outside world. America was the declared enemy, but there was no doubt of the fear of and resistance to Russia, especially after the latter's take-over of Afghanistan. The chaotic state of affairs within the country in the first year of the revolutionary regime raised the question whether Iran could have any effective foreign policy at all. Yet it was clear enough that, barring intervention by force from outside, Iran would make its own decisions.

Saudi Arabia has some similarities to Iran in its geopolitical position, but also significant differences. It has never been under threat of attack from the Soviet Union, but its ruling family has had a fixed ideological aversion to Communism and a tendency to equate it with forces of radical change that were a more immediate threat to the existing system. The special relationship with the United States, which began even before the Second World War when King Abdul

Aziz ibn Saud, distrustful of the British, granted an oil concession to an American company, was a natural one for the Saudis in that it brought real economic benefits as well as security. Saudi Arabia was both a rich and a weak state. It needed an external connection to compensate for its weakness in the face of regional threats (for example from Nasser's Egypt in the 1960s or from radical forces elsewhere in the peninsula in the 1970s), and to help it translate its oil wealth into development. Saudi Arabia never joined any American-sponsored regional security system directed against the USSR, but by its congenital anti-Communism and its special relationship with the United States it played its part in shaping the superpower balance in the Gulf region. Even so, while relying ultimately on America for security, the Saudi rulers sought an independent foreign policy and even a guiding role in inter-Arab relations, using their money as an instrument of political influence.

The three large states — Iran, Saudi Arabia and Iraq — largely determined the international relations of the Gulf. The others were too small to have a significant impact, although Kuwait and the UAE had large quantities of oil and therefore plenty of money. In the realm of high international politics they tended to follow rather than to lead. But some had a potential for trouble, either of their own or of others' making: Kuwait because of its wealth, the volatile Palestinian element in its population, and its vulnerability to aggression by Iraq, which coveted its coastline; Bahrain, the UAE and Oman because of the newness of their independence, possible political instability and the danger of being squeezed by Arab-Iranian rivalry; the PDRY because of the threat of revolution it represented and the danger of its being used by others, the USSR in particular, to undermine other governments in the region.

V

As the Gulf states have chosen what kinds of relations and what degrees of association to have with the two superpowers, there has always been in the background an alternative — non-alignment. This can be a very strong trend because it has historical roots and receives nominal adherence even when it does not wholly square with the facts. Saudi Arabia and Oman are close to the United States in security matters; Iraq has a security relationship with the Soviet Union, as has the PDRY. All of them, however, together with the new Iran, the Yemen Arab Republic and the smaller Gulf States, take part in the world-wide non-aligned movement and attended its latest conference at Havana in 1979.[9] Whatever their ties with the great powers, this is where they feel most at home. They are not and

will not for many years be in the 'first world' of the industrial democracies and have no desire to join the 'second world' of the Soviet bloc. China may be a useful counterweight but is too distant, unknown and unknowable to invite close association; and China itself in this context assumes the guise of a Third World country. The major Gulf states are, therefore, in and of the Third World of the non-aligned, and because of their wealth they aspire to a role of leadership in that company.

There is hardly a Gulf State which has not declared its aim of greater independence and freedom from the military presence of outside powers. Most of them are aware, however, that non-alignment can be more effective if it reflects regional and not purely national policy. The Shah of Iran broached the idea of a regional security system in the Gulf to Iraq, Saudi Arabia and others in 1975. The move proved premature. The Arab states, already members of a formal multilateral treaty system of their own, the Arab League, were not ready for it. Past differences and potential Arab-Iranian rivalry in the Gulf, ideological conflict between Iraq's radical Ba'athists and the conservative regimes elsewhere in the region and the struggle then going on in Oman's Dhofar province all combined to make it more difficult.

The mere fact of the discussions, however, bore witness to a common sense of responsibility and a new way of looking at continuing ties of individual states with outside powers. Each would keep its existing ties — Iraq would continue getting arms and other aid from the Soviets, while others got theirs from America — but the concept of regional solidarity put some limits on the ways in which those ties could or would be used, either by regional states or by the outside powers.

For the near future the great question mark is revolutionary Iran. Until it has an effective government in control of the country it can hardly have a regional role, except perhaps in the influence of its Islamic ideology, which tends to be supportive of Afghan resistance to the Soviets but disruptive (because it is Shi'a rather than Sunni) of relations with nearby Arab countries. Nevertheless, the changes in Iran and Afghanistan have stimulated new thinking and new exchanges by Gulf states on the subject of the common security, including consultations between the two leading but ideologically divergent Arab states, Iraq and Saudi Arabia.

The Soviet invasion of Afghanistan, of course, also provoked the United States into a more active effort to organise security in the Gulf region against any further Soviet advance. Yet the degree of military co-operation forthcoming from states of the area was, to say the least, uncertain. The fate of CENTO, which had been dealt a mortal blow by the Iranian Revolution and was formally dissolved

shortly thereafter, was not an encouraging precedent for a new endeavour to create a Western-sponsored regional security system. Iran and Iraq, two key states whose security was threatened from the north, were totally unreceptive to the idea, as was America's good friend, Saudi Arabia.

Any security system in the Gulf region, whether linked to outside powers or limited to local states, must overcome or neutralise the sources of local conflict if it is to be effective. There is probably no way of preventing all conflict among Gulf states, in view of the many historical territorial disputes (*inter alia* Iran-Iraq, Iraq-Kuwait, Iran-Saudi Arabia, Abu Dhabi (UAE)-Saudi Arabia, Oman-Saudi Arabia) which can flare up even though some have been nominally settled by negotiation, to say nothing of unrest and rebellion within states (for example the possible resumption of guerrilla war in Oman's Dhofar province or the outbreak of revolt by ethnic minorities such as the Kurds in Iran or Iraq) which can bring intervention from outside.

The interlocking of regional conflicts with the superpower competition maximises the dangers for all concerned: it encourages local states to look outside the region for support against their neighbours; it encourages outside powers to play on local disputes in order to gain advantage over rival powers; it tends to blow up local conflicts into larger ones by the operation of prior commitments and the fear of losing friends and allies. All those dangers would be reduced if the states of the region could move purposefully and in concert to assert their non-alignment in relation to the superpowers and to organise their own system of security; and if the superpowers, by express agreement or in practice or both, would respect those decisions and would not seek bases or special security ties within the region. Local conflicts would not disappear, but they might have a greater chance of being settled or contained by collective procedures or by the operation of a regional balance of power, with a good chance that outside powers would see their own interests better served by remaining aloof than by intervention.[10]

These possibilities have an air of unreality in the present day and age, for both local states and outside powers have other conceptions of what is required for their interests, their destiny or their security. For those in the Gulf or the West who think of a revived CENTO or for those in Moscow or among the radical groups of the region who see emerging a new socialist order under Soviet leadership, the idea of a non-aligned regional system is not compatible with their own. For the local states the issues are complex and not always presented in clearest form. None wishes to cut outside ties when a reliable regional system does not yet exist. Actually, both these trends, that of interlocking and that of regional non-alignment, are present in the

thinking of political leaders and others in the leading Gulf countries, and the two will be contending, probably without sharp and definitive resolution, for some time into the future.

VI

We have dwelt long, perhaps over-long, on the security problems because they are complex, they involve the ultimate questions of peace and war, and because decisions of all parties, local states as well as outside powers, are based so heavily on them. Nevertheless, the real problems of the years ahead, having in mind the global setting of the affairs of the Gulf, are likely to be economic and political. The oil revolution of the early 1970s had important effects other than on military strategy, military ambitions, the balance of power and the security ties of the Gulf states with outside powers.

The main result of the rise in oil prices was to give almost all the Gulf states, large and small, more economic power and more influence on affairs beyond their own region. At home they had large sums available for development. The mere presence of so much money and so much new economic activity tended to take the edge off the appeal of liberation movements and other revolutionary causes. Of course, social change wrought by economic development was bound to generate pressure for change in political institutions as well, and in Iran the combination of conspicuous waste, economic stress, cultural shock and the rigidity and repressions of the political system produced a popular religious-political revolution that overthrew the Shah and brought new elements into contention for power. Although Iran's economy was severely shaken and oil production was at half the pre-revolution level, that production remained crucial to Iran itself, to fellow members of OPEC, and to the oil-importing countries of the world.

The big questions for the future revolved round the positions the Gulf states would find themselves in the world economy, and separately with the developed countries of the West (including Japan), with the Communist countries, and with the Third World. These questions have to do mostly with oil and gas: how much and how rapidly the Gulf states will produce it; to whom they will sell it; what they will do with the money it brings; how long the oil reserves will last; and what situation they will face as the oil age declines. These are indeed big questions, the answers to which do not depend on the Gulf states alone. They depend partly on the main consuming countries, which will be dependent on the oil of the Gulf for the next decade and probably longer. They may depend also on the Soviet Union and Eastern Europe as their demand for oil approaches the

point where it will exceed their own supply.

The first thing to note is that oil power is not absolute. OPEC, and especially those of its members in the Gulf region with their very large reserves, have the economic power virtually to dictate the world price of oil, but they cannot do so in total disregard of the damage that may be done to the economies of those states that depend on imported oil, for they would thereby damage their own interests in development and investment and might provoke violent reaction. The surplus oil money certain Gulf states are generating cannot be used indiscriminately; it cannot be wasted (as much of it is now) without heavy cost to the interests of both consumers and producers. By the logic of their complementary economies the Gulf oil producers and the oil-consuming industrial states have come to see that their common interest is a trade-off between oil and development. Co-operation does not always proceed easily and evenly, however, for the terms of the trade-off are by no means agreed, either as to supply or as to price.

Both the Gulf states and the West must look ahead to situations of strain. Their nature and timing are not wholly predictable because the oil supply itself, from the Gulf and from other sources, is not predictable. Yet we can be fairly sure that there will come a time in the 1980s, before the West has developed adequate alternative sources of energy, when its demand for oil will go beyond what OPEC is willing or able to supply, when the strong interest of the consumers will be to get more and more oil from the Gulf and that of the Gulf states will be to raise the price, limit production and keep oil in the ground as an asset for the future. It is at such a point that the hypothetical scenes of dispute and conflict talked about in 1973 and 1974 may become reality.

This is an aspect of the global setting of the future that both consumers and producers would do well to foresee and to guard against. Unless agreements have been reached in advance, within each group as well as between groups, the strain could easily unleash situations of cut-throat rivalry and conflict from which all would suffer. It is not useful to speculate on just what those possibilities would be, but there can be little doubt that Saudi Arabia would be at the centre of them. As the country with the largest oil resources, it would have the most influential position in determining OPEC's decisions on supply and on price or, if OPEC did not hold together, in determining whether the West's economy would face readjustment or disaster. Because of its key position, however, Saudi Arabia might be exposed to pressures, both from within the region and from outside, that its political structure could not withstand.

The present American relationship with Saudi Arabia is a very comprehensive one that seems to satisfy the interests of both parties.

Although uncertainty and mistrust exist on both sides, the 'special relationship' gives both some sense of assurance about the future. It could be a false assurance on the side of the Saudis if they are engulfed in political turmoil at home or in the region, or on the American side if the Saudis find they have a stronger interest in limiting than in expanding oil production. Potential trouble lies also in the thirst of other consuming states for the oil of Saudi Arabia, which some see as having been for too long the near-exclusive preserve of the United States. It will be important for the United States to be in harmony with the other consumers and for Saudi Arabia to be in harmony with the other producers *before* the strains of a supply-demand crisis appear, and that is by no means assured.

The Soviet Union by that time may be much more concerned with the oil of the Gulf than it is today. Now that the Gulf states have virtually completed their take-over of the properties of the Western oil companies and demonstrated their power over the West, the Kremlin can no longer make much of the old theme of unstinting Soviet support for the struggle against the oil monopolies. The Soviet Union is left out of the rapid expansion of development and of foreign trade taking place in the region since it cannot be a major market for Gulf oil and cannot provide the goods those countries want. It may become more involved if its general economic decisions bring it into greater participation in the world economy, but it does not have the hard currency or the advanced equipment to make itself a sought-after partner. The established Soviet pattern, that of bilateral barter deals, is not responsive to the desires of the Gulf states, not even Iraq. But the Soviet Union has a special energy problem of its own that foreshadows new perspectives.

The general phenomenon of shrinking oil supply in relation to demand may affect the Soviet bloc as well as the West. Some Western estimates foresee an inadequacy of Soviet oil to meet domestic demand in the USSR and Eastern Europe as early as 1985.[11] Other estimates put the date later, and the Soviet government itself makes no such prognostication at all. We can be sure that the Soviets will do what they can to avoid a situation of dependence on imported oil. Their entire approach to security and economic policy is to retain control of vital economic decisions. For so critical a resource as oil they do not want to be in the position in which the United States now finds itself. Accordingly, they can be counted on to make great efforts to develop the oil and gas of Siberia, as they are now doing, even at great cost and at the expense of other parts of their overall economic plan. While doing so they may wish to increase imports of oil and gas from Iran, Iraq and certain other Arab states along existing patterns, in order to supply certain areas of their own country and to keep up exports to Eastern and Western Europe, but

probably not beyond modest levels.

Questions will arise both for the Gulf states and for the West if the Soviet Union cannot rapidly develop its own energy resources and has to become a net importer of oil. If the Gulf states begin to export more to the Soviet Union, they will have less to export to the West. If they refuse to do so — which is quite likely since the Soviets will be short of hard currency and will want to make their own terms — will they expose themselves to Soviet action more drastic than economic bargaining? The West and Japan will face the question the Soviets have already put to them during past years: whether to help with capital and technology to develop Siberian energy resources, getting a share of the new production to ease their own energy shortage. It would increase total world oil production, presumably to the benefit of all. But this course would raise some questions. If capital is to be poured into development of energy, why not in the Western world rather than in the Soviet Union? Should the West help that country to become stronger and give hostages to it in the process?

If the Soviets should turn their attention to the Gulf as the answer to their energy problem, not to bargain but to use military force, they would have no guarantee of long-term success. They would probably have to cope with sabotage of equipment, non-cooperation, inability to transport the oil and counteraction from the West. Enough risks and problems, perhaps, to rule out military action, but it is useful to take account of 'worst-case' contingencies, for preventive diplomacy as well as for military planning. The oil states of the Gulf have gained tremendously in prestige and in power in the past few years, and they have largely succeeded in keeping their relations with the outside world on the economic plane, where they are strong, rather than on the plane of military power. But in a world of rival superpowers and blocs hungry for oil to keep their economies from running down, the life of nations sitting on the world's largest reserves would not be tranquil. Some time before the mid-1980s the Gulf states and the Western consumers and the Soviets will have to talk about this problem.

In confronting the military giants and their conflicts of interest and ideology that spill over on to other continents, the Gulf oil countries, though now richer than the rich, still have much in common with the poor developing countries of Asia, Africa and Latin America. The relationship is a complex one, not explicable by economic interest alone or any single factor. Historically, the Gulf countries are a part of the Third World. They have had the experience of weak peoples dealing with the strong. They have gone through the anti-colonial revolution and have a feeling of solidarity with others who have done the same. Their assertion of oil power, in turn, was acclaimed all over the Third World as a blow for the

common cause, much as Japan's victory over Russia early in this century was widely hailed by coloured and colonial peoples as a great blow struck at the hitherto dominant European powers.

How far have the oil states seen their new role as one of leadership of the developing against the developed countries? Algeria, in its public stance at least, displays no doubt that it is their appointed role to bring about a fundamental change in world economic relations, a 'new international economic order'. Algeria has less oil and less economic power than the rich oil states of the Gulf and has a relish for high-flown rhetoric. Wealthier states like Saudi Arabia do not need a new economic order. They are doing quite well under the present one. Yet the strength of their conviction is not to be taken lightly. They are not cynical but serious in lecturing the industrial countries on their obligation to create a new international economic order based on a more equitable distribution of wealth.

There are two ways in which the oil states can help the Third World to improve its status. The first is by direct use of their own expanded resources to save poor countries from the effects of higher oil prices or to finance their development programme. The second is by using newly gained bargaining power to force the developed countries to make the basic concessions on raw materials, prices, access to markets, development funds and all the other matters that Third World spokesmen had been pushing for at the UN General Assembly, at UNCTAD and elsewhere for a good many years.

The aid which Gulf oil states have given to other countries has been pictured by them as generous and described by critics as inadequate, unfairly distributed, or unduly political. But it has been substantial.[12] Although they have not been willing to sell oil at concessionary prices, they have been willing to help with grants and loans, have created investment agencies, and have contributed to the special facility established by the International Monetary Fund for the purpose of cushioning the effects of the higher price of oil. Except for those monies provided to the IMF, however, the extension of aid has been very much coloured by the national purposes of the donor states and their political relations with the donees. Grants and loans from Saudi Arabia have gone largely to Arab and Islamic countries. The principal lending agencies of Kuwait and Abu Dhabi have limited their loans to the Arab world. Iran under the Shah has paid particular attention to the Indian Ocean region, where he hoped to play a leading political and military role. The Arab states have negotiated aid arrangements with the states of black Africa, but without coming close to satisfying their wants and needs.

Without denigrating the generosity of the oil states of the Gulf or playing down the political importance to them of maintaining goodwill among the developing countries not blessed with oil, it is not to

be expected that with their own financial resources they will or can have a major impact on rates of growth or standards of living in the Third World as a whole. Their own needs take priority. The urge for rapid development, concern for investment for their own long-term future and concentration on their own region are so compelling that they alone will not have the resources to pull the rest of the Third World along with them.

It is natural for them to put the monkey on the West's back and join the call for a new international economic order. At the long and inconclusive negotiations at the Conference on International Economic Cooperation in Paris in 1975–7 they made common cause with the other Third World countries, refusing to make direct arrangements with the developed countries on oil matters in the absence of progress on the long set of disputed issues connected with the new economic order. But the deadlock merely illustrated the fact that the only hope for really effective aid to the Third World as a whole over the next two decades lies in a co-operative effort of OPEC and the developed countries of the West to keep a functioning world economy going in which more and more Third World countries can gradually come to play an active part on their own.

The shape of the world economic order, in fact, is being determined not by any negotiations in Paris or at the United Nations, not by the pressure of OPEC on behalf of less fortunate friends in the Third World, but by the actions of the developed countries among themselves and by the growing economic interchange between them and the oil states of the Gulf.

In the global setting of today, both in security matters and in economic prospects, the Gulf region finds itself more and more at the centre of world politics. American officials have ranked it with Europe and the Far East as an area of vital interest, a judgement confirmed by President Carter in his message to Congress on 23 January 1980. The leaders of the Soviet Union have consistently described the region as of primary security importance because of its location on the southern frontiers of the USSR, and the move into Afghanistan was dramatic proof of that conviction. This clash of global interests and ambitions takes place in an area of endemic instability and fragile political institutions. When there has been apparent calm, it is deceptive, masking problems of internal stress and regional conflict that leave no room for complacency or short-sighted leadership.

The revolution in Iran has created shock waves. Despite the increasing importance of the Gulf nations as independent actors, their future is not entirely in their hands. No region of the world lives in isolation. The superpowers and all the countries in need of Middle East oil have had to learn the lessons of interdependence. The Gulf nations, having gained new power and prestige because of the fact of

interdependence, also remain subject to forces generated by it, forces that may be beyond their own will and control.

Notes

1. *United States Foreign Aid Programs in Europe*, Hearings before a Subcommittee of the Senate Committee on Foreign Relations, 82nd Congress, 1st session, 7–23 July 1951 (GPO, Washington, DC, 1951), p. 277.

2. Interview with Secretary of State Kissinger, *Business Week*, 13 January 1975.

3. Oles Smolansky, 'The Soviet Union and the Middle East' in William E. Griffith (ed.), *Critical Choices for Americans: the Soviet Empire, Expansion and Detente* (D.C. Heath, Lexington, Mass., 1976), pp. 269–71.

4. See Geoffrey Kemp, 'Scarcity and Strategy', *Foreign Affairs* (January 1978), pp. 396–414, for contingencies of this nature.

5. Address of Secretary of Defense Harold Brown to the Council on Foreign Relations, 6 March 1980.

6. There is a large body of writing, both Soviet and Western, on this subject. My own summary conclusions on it are contained in two articles: 'Soviet Policy in the Middle East' in Roger E. Kanet and Donna Bahry (eds.), *Soviet Economic and Political Relations with the Developing World* (Praeger, New York, 1975), pp. 85–99; and 'The Soviet Union in the Middle East', *The Middle East Journal* (Winter 1978), pp. 1–12. For a recent detailed and balanced estimate, see Shahram Chubin, *Soviet Policy towards Iran and the Gulf* (Adelphi Paper No. 157, IISS, London, 1980).

7. The two sides held three rounds of talks in 1977. Both said publicly that an agreed limitation on their military presence and activities in the Indian Ocean would be in their own and in the world's interest. However, the United States was unwilling to cut naval forces along the lines the Soviet Union proposed, nor was it willing to ban nuclear-armed ships. The talks lagged and then were discontinued altogether following the Soviet military involvement in Ethiopia in 1978. Then came Afghanistan.

8. See John C. Campbell, 'The Soviet Union and the Middle East, Part I', *The Russian Review* (April 1970), pp. 152–3.

9. It should be noted that the non-alignment stance of the PDRY was patently fictitious. At the Havana conference and elsewhere it was one of those following Cuba's lead in trying to turn the non-aligned movement in pro-Soviet directions.

10. See R. K. Ramazani, 'Security in the Persian Gulf', *Foreign Affairs* (Spring 1979), pp. 821–35; *idem*, *The Persian Gulf and the Strait of Hormuz* (Sijthoff and Nordhoff, Alphen aan den Rijn, 1979).

11. For example, Central Intelligence Agency, *Prospects for Soviet Oil Production* (Washington, DC, April 1977).

12. UNCTAD, *Handbook of International Trade and Development Statistics*, II.D.2 (1979), p. 441; International Bank for Reconstruction and Development, *World Development Report, 1979* (Washington, DC, 1979), p. 157. The latest complete figures, for 1977, indicate aid of more than $5 billion from Middle East oil states in that year.

2 THE STRATEGIC BALANCE AND THE CONTROL OF THE PERSIAN GULF

Geoffrey Kemp

Introduction

For at least the next decade the control of the Persian Gulf will be a key factor in calculations about the relative global balance of power between the Western industrial countries and the Soviet Union. As a result of the revolution in Iran, the Western countries have lost great political, economic and military leverage over the local countries of the Gulf and the adjacent regions. If the Soviet Union, by direct or indirect means, were able to outmanoeuvre the West and gain strategic control of Iran, the impact upon the security of the rest of the Gulf and ultimately the Western alliance system would be extremely grave, in many ways as serious as overt Soviet hostilities in Europe and the Far East.

In this essay it will be argued (a) that the major threat to vital Western interests in the Gulf is the potential for Soviet control; (b) that, while this may be a temporary phenomenon owing to the unique condition of the energy market over the next ten years, the West's vulnerability and lack of consensus on energy policy coincides with a window-of-time during which Soviet military power projection capabilities are growing and its overall military power peaking; (c) that only a coherent and well orchestrated policy by the United States can deter the Soviet Union from adventurism in the region; and (d) that without a coherent Western strategy the temptation to the Soviet Union to interfere in Gulf affairs will grow.

Despite the fact that now it is a truism to say that the Gulf is vital to the West, what does this statement mean? Does it mean that the West will continue to need Persian Gulf oil into the indefinite future and without that oil its industrial furnaces will die down and the Western economic system, as we know it, will collapse? Clearly, the answer is no. Alternative energy supplies, including new oil sources, can be developed and, if the political price is paid, conservation methods can be enacted and economic growth can be slowed down, thereby reducing demand. While a slow-down in growth would have serious political and economic repercussions for Western countries, it should not spell the end of Western civilisation. Gulf oil is important

because it is permitted to be important. Unfortunately, the political costs of reducing this dependency have, until recently, been too great for Western, especially American, leaders to pay.

The denial of Persian Gulf oil by local states to the West through revolution, embargo or production cut-backs could cripple the Western economies if *all* the oil flow was reduced for several months. However, to imagine a situation when *all* Persian Gulf countries were prepared to deny oil to the West assumes either that there has been a dramatic change in their policies or that, with the exception of Saudi Arabia, the Gulf countries are prepared to forgo very important revenues for political purposes. Based on their behaviour since 1973, this seems unlikely unless the cut-backs are due to internal problems such as those experienced in Iran during 1979. *In extremis*, the West has the capacity to hurt the oil-producing countries *provided* its leaders are prepared to take risks and ultimately be prepared to stand up to challenges from the Soviet Union. Freezing Persian Gulf oil assets in Western capitals and denying arms and technology could have a serious impact upon development and military programmes in the Gulf. A total embargo might happen in the event of another Arab-Israeli war but that might depend on the role taken by the United States during that war. If the United States did not resupply Israel, the excuses for implementing an embargo would be slender.

In military terms, Persian Gulf oil is not vital to the war-fighting capability of NATO. NATO's requirements for oil in event of a 60-day conflict are small in comparison to the total oil demand for the peacetime domestic economies. By rationing programmes and a reordering of priorities, there is no reason why the NATO countries could not 'oil' a 60-day war. If, however, a general war with the Soviet Union were to be protracted, similar to the Second World War, then the denial of Persian Gulf oil would have a serious impact upon industrial wartime production. However, if one looks at the German and Japanese examples of the Second World War, remarkable things can be done with substitutes under wartime conditions but at an enormous price.

With these caveats in mind, it should be stressed that the strategic importance of Persian Gulf oil stems from its relationship (a) to the peacetime economies of the Western world, especially Europe and Japan, and (b) the possibility that under certain circumstances the oil could fall under Soviet political and military control. If one removes the Soviet Union from the power equation in the Persian Gulf, then the West's security dilemma is more manageable.

The Soviet Threat

There are at least three hypothetical circumstances under which the Soviet Union might threaten Persian Gulf oil: first, in a situation of general war with the West; second, as an outgrowth of conflict in the Gulf if the Soviet Union intervened to protect, help, or hinder one or more local countries; third, in a contingency where neither general war nor local war were under way but where the Soviet Union felt threatened elsewhere and was prepared to disrupt the economic power of the Western world, to redress what it felt were unacceptable threats to its security, for example a major crisis between the Soviet Union and China or a crisis in Eastern Europe.

General War Scenarios

In event of a major non-nuclear war breaking out with the Western allies, there might be advantages to the Soviet Union in threatening Western oil supplies from the Gulf. This could be done in two basic ways: first, by *attacking* the oil supplies themselves either at the source or in transit with nuclear or conventional weapons or by using military power to gain control of all, or part, of the Gulf in order to *threaten* oil supplies. From the Soviet point of view, there are costs and benefits of these options. Attacking oil supplies would be a much easier *military* task than invasion of the Gulf. Yet the risks of escalation would be great. Direct interdiction of the Persian Gulf oil facilities, in particular the Saudi and Iranian terminals at Ras Tanura and Khark Island respectively, would be seen as a deliberate act of war as serious, perhaps, as an attack on Berlin. If nuclear weapons were used, it would be a clear sign that the Soviet Union was prepared to fight a major war with all the weapons in its inventory. This would undoubtedly spur reprisals, if not an all-out conflagration with the West. If conventional weapons were used, the Soviet Union might be able to destroy the oil or put it out of commission for many months but it would also stand a risk of precipitating World War III or major non-nuclear reprisals against Soviet ports and ships on the high seas or even its own highly vulnerable domestic energy system.

On the other hand, a carefully orchestrated non-nuclear invasion could *avoid* direct attacks on oil supplies. However, if such an invasion were taking place in the context of a general war, it can be argued that the West would assume that sooner or later the Soviet Union would either attack or capture the oilfields and, therefore, the risks of escalation, although in the initial stages less severe, would eventually have to be faced. It is possible that a general war with the Soviet Union might be non-nuclear and result in an impasse in NATO and the Far East. In these circumstances, the Soviet Union

might have great incentives to capture Persian Gulf real estate in order to have a bargaining chip during eventual negotiations. The assumption here would be that neither the Soviet Union nor the West would want to escalate to nuclear war and would, therefore, be prepared to bargain with whatever territory they held at the 'end' of the conventional general war.

Nevertheless, all general war scenarios involving the Persian Gulf, although interesting and important, would probably take second priority to other tasks that both the Western and Soviet forces would be given. For instance, based on current Soviet doctrine, it is unlikely that they would devote a large portion of their submarine and naval aviation to the interdiction of oil supplies out of the Gulf because they would need these forces for more important tasks in the Atlantic and Pacific. Another major constraint on Soviet interdiction of the sea lines in the Gulf and Indian Ocean relates to constraints on their infrastructure. Until such time as they have adequate basing facilities, their capacity to sustain a naval campaign against the SLOCs would be very limited and they would be vulnerable to Western anti-submarine warfare. If, however, they had a major base facility in the southern seas, for instance in South Africa, this would change the balance of maritime power in the southern seas, at which point oil SLOC interdiction would be a more serious problem.

Similarly, the requirements for an invasion force of the Gulf could be high and very competitive in a general war *provided* the West and the regional countries were prepared to resist and not give the Soviet army a free ride to the Gulf from the Caucasus. The terrain in northern Iran and eastern Turkey is very difficult and a Western defence of these northern and eastern borders is, in theory, possible. However, it would require the co-operation of Iran and the neutralisation of Iraq and the participation of Turkey. Furthermore, until such time as the United States increases the size and lift capabilities of its rapid deployment forces, even in the best circumstances the Western powers might arrive with too little, too late.

Local War Scenarios

The second and more likely case in which Soviet military power might be used in the Gulf relates to local war scenarios. In this case it is assumed that the Soviet Union would have an interest in exploiting opportunities if local conflicts presented it with the chance to win an important victory over the West. A review of the past twenty years in the Gulf region reveals that although there have been a great many *coups d'état* and political assassinations resulting in changes in leadership, and numerous border disputes, some of which remain unresolved, there have been few incidences of inter-state war. This,

in part, was due to the fact that until 1971 the UK retained a military presence in the region and, therefore, acted as policeman. It also should be noted that following the British withdrawal the local countries have made serious efforts to resolve outstanding disputes which in the past might have led to conflict: for example in 1975, Iraq and Iran resolved the Shatt-al-'Arab and Kurdish conflicts; Saudi Arabia resolved most of its territorial claims with neighbours; Iran has given up its historic claim to Bahrain; and, although there are still several offshore disputes, especially between Iraq and Kuwait, the track record would indicate that the probability of local countries going to war for territory is not very likely *unless* accompanied by ideological conflict.

There are two basic types of ideological conflict in the Gulf: those that stem from ethnic and religious sources, as have been amply demonstrated in Iran, and those that have their origins in radical-conservative conflicts. Examples of the latter type of conflict have taken place in Oman, where the radical regime in South Yemen has aided radical guerrillas in the Dhofar province of Oman with Soviet arms, and in Afghanistan, where conservative guerrillas have been fighting the Marxist regime.

Although the fear of left-wing radical regimes is a major source of concern throughout the Gulf today, radical threats come from both extreme left *and* extreme right. The revolution in Iran in 1978 and 1979 represented a right-wing religious backlash against the Shah's modernisation programmes. The extremists not only believed that the Shah was 'illegitimate', but that his regime was changing the nature of Iranian society by liberating women, downgrading the power of the religious leaders and introducing decadent Western customs. Iran's present problems highlight the dilemma that all conservative regimes face in adapting to modernisation. The Iranian tragedy is further compounded by a long history of terrorism on the left, ethnic separation and the close proximity of the Soviet Union. Perhaps because the revolution was built upon a coalition between the right and left, internal violence continues to pose a more serious danger to the regime than a *coup d'état* or assassination.

The Soviet Union's capacity for interfering and manipulating the internal Iranian situation is an ever-present factor. Were local ethnic conflict to get out of control in northern cities such as Tabriz or Mahabad, there is the possibility that the Soviet Union could use this as an excuse for intervention. If the Soviets occupied the northern provinces, it would embarrass, perhaps fatally, the Khomeini regime and would not automatically lead to universal condemnation in the West in view of the excesses of the revolutionaries and the ambiguity of the situation.

The question of regime stability in Saudi Arabia poses different

problems. There have not been upheavals similar to those experienced in Iran over the years culminating in the 1979 revolution, yet, because the Saudi regime is traditional, it faces potential socio-economic pressures from within the country. If there is to be internal change in Saudi Arabia, it will probably come from within the army, or some faction of the royal family aligned with elements of the armed forces. It is not impossible to conjure up scenarios under which this might occur. For instance, a crisis involving Palestinian terrorism and threats to oil supplies could polarise the elite. More radical junior officers might rebel, especially if the regime were inclined to be ruthless against terrorism. Because of the nature of Saudi society, a change of government could come about very quickly and in those circumstances intervention by external powers, including the Soviet Union, is a possibility, especially if 'invitations' were extended by a radical group that claimed power or which has temporarily ousted the royal family.

As for the other Persian Gulf states, all have strengths and weaknesses. Traditional leaders in Oman, Kuwait and the United Arab Emirates all face potential domestic opposition. This is particularly the case in some of the richer countries where there are very large foreign minorities who are not granted any of the rights normally afforded to people resident in a foreign country. The less rich Gulf countries, including the Yemens, face severe economic problems and, although Iraq has a lot of oil, it needs credits and technology from the West necessary for development. It seems unlikely that the Iraqi regime will work for closer ties with the Soviet Union at this stage, although that possibility should not be ignored.

This overview of the sources of instability in the Gulf region is not meant to be predictive. Although it is not possible to isolate the precise point at which domestic conflict may erupt, domestic upheaval in at least one of these countries will probably occur in the near future. In other words, although the probability of any *particular* violent event occurring is low, the probability of at least one event occurring is very high.

Soviet Intervention and the Global Balance of Power

The third possibility for Soviet intervention would be in circumstances when the Soviet Union was in conflict elsewhere, possibly with China, and was prepared to use its strategic proximity to the Persian Gulf to drive a political wedge between the Western allies, including their own relationships and their relationships with China. For instance, the close co-operation between Japan and China is an anathema to the Soviet Union given its hostilities with China and its own needs for the advanced technology and capital that Japan can provide. In event of a Chinese (or East European) crisis, the Soviet

Union would face great risks in terms of its own future stability as a Communist state unless it could isolate and divide its adversaries with a swift *coup de main*. Threatening direct military intervention against those 'supporting' the enemy would be one obvious route to take but this would have dangers of escalation. Occupying the Gulf or, more specifically, Iran in such circumstances would be a risky operation too, but if it were achieved with maximum surprise and speed the fighting would probably be over before the West could respond. In fact, given the current (November 1979) status of the Iranian armed forces and the US arms embargo against the Khomeini regime imposed during the hostages crisis of that month, it is reasonable to argue that the Soviet army could occupy northern Iran in a matter of days rather than weeks. In these circumstances, the mere presence of the Soviet Union on the Persian Gulf would probably be sufficient to signal to Europe and Japan that support for China was a losing game. The Soviet Union would not need to destroy or control oil supplies; it would be the *implicit* threat of future Soviet control of oil that would have the most powerful impact upon the political leaderships of Europe and Japan.

Security Options for the West

Given the above points, what are the realistic security options for the West in planning for its security in the Persian Gulf region? Because of the importance of the Gulf to the entire Western industrial system, it is difficult to see how the region can continue to be treated as a discrete area decoupled from the military issues of Southern Europe, the Mediterranean and Indian Ocean. The stability of the Gulf is linked to the future stability of Europe and Japan, since threats to one area pose threats to the others. Certainly a major change in the balance of power in Europe would have profound effects upon the Persian Gulf countries themselves; in particular, Saudi Arabia and other 'conservative' regimes. These countries, therefore, have a long-term interest in maintaining a strong Western commitment to a favourable balance of global power. Equally, the Europeans have a major vested interest in the Gulf. These mutually reinforcing interdependencies suggest that a new concept of Western security should be formulated. This should go so far as to extend the mandate of an alliance such as NATO to include the Gulf countries, since this would be politically inappropriate in peacetime despite the fact that from a logical strategic point of view it has a lot of merit.

The relationship between Japan and the Gulf is somewhat more tenuous. A collapse of the Japanese economy would probably not have the same political impact in the Gulf as a collapse of Western

Europe but it would, nevertheless, represent a blow to economic development and the potential loss of a major oil market. In contrast, Japan's dependency on the Gulf is great and, therefore, any change in Gulf policy with respect to oil could have far-reaching domestic repercussions in Japan. Yet Japan has neither the capacity nor the will to consider extending its security perimeter beyond the South China Seas, let alone into the Indian Ocean. However, it does have the capacity and, perhaps, the will to expand its local military capabilities and thereby 'release' elements of the US Eighth Fleet for duties in the Indian Ocean area.

In reconsidering the global map of Western interests to take greater account of the Gulf, the first conclusion to be faced is that a new area of responsibility has been added to Western military commitments at the very time when the overall presence and projection capability of the United States and other allies have been weakened. The immediate impact of this is increased pressure upon the already thinly spread American forces. For the past few years the European commitment has received the greatest attention at the expense, some feel, of the Gulf and Far East. If the new priorities, such as the Gulf, cannot be assumed at the expense of other commitments, then one of two things will have to happen. Either the Western powers will have to increase the overall size of their forces or local powers will have to plan a more active role in joint Western defence efforts. What this might also mean is a new distribution of responsibilities throughout the Western alliance that could include trade-offs, such as a substitution of *some* American presence in the Pacific for a greater Japanese Pacific contribution and, in return, a greater American presence, at least for part of the year, in the Indian Ocean-Persian Gulf region.

Beyond these options, there have been suggestions that NATO should redraw its boundaries to include the Gulf and South Atlantic and Indian Ocean. There are great political difficulties in these suggestions in any peacetime context. Even with NATO's current boundaries, different priorities pertain: the Norwegians have a different view of the importance of the Yugoslavian contingency than the Italians; similarly, it is difficult for Turkey to get too exercised about the problems of Spitzbergen in view of its own immediate territorial proximity to the Soviet Union. Nevertheless, there is, within the NATO hierarchy, a general agreement that certain common goals and force levels must be pursued irrespective of local antagonisms that remain unresolved, such as the Cyprus conflict between Greece and Turkey. In wartime, NATO would obviously redraw its map and extend operations into the Middle East and Indian Ocean if necessary. In peacetime the alliance, as such, can do little other than talk about the need for a more global perspective. However, individual allies, such as France, the United Kingdom and

Turkey, are capable of providing certain military assistance in peacetime to *deter* Soviet adventurism in the Gulf.

Furthermore, it is not out of the question that some local Gulf countries, which have otherwise conflicting views on certain issues such as oil pricing and some territorial claims, might be persuaded to support Western initiatives designed, in part, to help their security needs. The most radical suggestion would be for powerful pro-Western countries, such as Turkey, Israel, Egypt and Saudi Arabia, to formulate some concept of defence co-operation. This is unlikely to happen in the short run given the emotions and history of the region and Saudi antagonism on the issue of the Egyptian-Israeli peace treaty but, if the Soviet threat becomes more serious, then it may be Saudi Arabia rather than Israel that will have to decide whether it is prepared to give up ideological claims for security. The real question is whether the local Gulf countries themselves are prepared to consider more formal military relations with the West. In the case of Iran, this seems unlikely for the foreseeable future. The involvement of Saudi Arabia in any such system is more promising, though proposals for an American military presence in the region, including, for instance, an air base at Dhahran, have made little progress. The Saudi regime wants an American commitment but not one that is overtly visible. In peacetime, this means a US maritime presence in the Indian Ocean is more acceptable and more likely than a forward base involving the deployment of American forces in the Gulf itself.

In the absence of any formal treaty arrangements, the alternatives would be to continue current policy, which is for the West to develop strong bilateral relations with pro-Western Gulf states. However, as has been seen in the case of major military programmes in Iran and Saudi Arabia, this can run into serious problems in the United States. For instance, in 1978, the Carter Administration attempted to justify the sale of F-15 Eagles to Saudi Arabia on a broad geopolitical threat assessment to the regime over the next ten years. The opposition came primarily from Israel's supporters, who argued that, irrespective of threats posed by hypothetical Russian aeroplanes in South Yemen, the most likely use for the F-15s would be in a renewed Arab-Israeli conflict. There can be no doubt that the Israeli argument has to be taken seriously, for the record suggests that, in event of war, weapons will be used irrespective of restraints imposed upon them before the war began. Thus there is likely to be continued controversy over the Saudi Arabia build-up quite apart from questions as to whether the Saudis have the capacity to absorb and operate effectively all this modern equipment.

The absorptive capacity argument is, of course, most relevant in the Iranian case; namely, that although Iran under the Shah made a

good case for building up a modern military force, it did not have the infrastructure or technical manpower to operate a modern military force effectively in all dimensions. This seems to have had the overall effect of weakening rather than strengthening its capacity to fight, at least in the short run. It is, after all, the capacity to fight that counts and some have argued it would have been better for Iran to set its sights more lowly for force levels in the first instance before embarking on projects which were unlikely to come into fruition for decades.

What this suggests is that in military terms the contribution of local powers — with the exception of Israel and elements of the Turkish and Egyptian forces — to regional defence is likely to remain small although the participation of countries such as Saudi Arabia is vital in view of their geography.

In conclusion, if Soviet adventurism is to be deterred in the Middle East-Gulf, it will be necessary for the Western powers, especially the United States, to demonstrate the capability to project military forces into the region that *are capable of fighting* Soviet conventional forces and those of its surrogates. This is a difficult task, especially in view of the fall of Iran and the geographical advantages open to the Soviet Union. However, it is not impossible in view of the constraints that the Soviet Union would face in projecting power beyond Iran and the inherent regional advantages the West may have if it carefully orchestrates a strategic policy in the region. What this means in practical terms is that the proposed US 'rapid deployment force' will, indeed, be expensive if it is to be effective. Merely taking forces earmarked for other contingencies will not solve anything. What is therefore required is *additional funding* for this force which, incidentally, is something the United States needs irrespective of its particular interests in the Gulf. This, in turn, suggests that further increases in the European and Japanese budgets will be necessary to preserve some semblance of political equilibrium within the alliance. Unless these steps are made, the Middle East-Gulf will remain a military vacuum and, as such, a highly unstable, dangerous region.

Table 2.1: Gross National Product and Military Expenditure (millions of US dollars)

	Iran		Iraq		Saudi Arabia		Kuwait		Oman		Qatar		Bahrain		UAE	
	GNP	MILEX	GNP	MILEX	GNP	MILEX	GNP	MILEX	GNP	MILEX	GNP	MILEX	GNP	MILEX	GNP	MILEX
1972	16,304	1,189	4,145	473	5,447	941	3,716	106		44		27		4	1,070	18
1973	26,502	2,097	5,235	467	8,130	1,478	5,926	116		77.5		24		5	1,430	26
1974	46,574	3,225	10,591	803	23,197	1,808	10,808	162		169		41		6	172	56
1975	52,835	8,800	13,196	1,191	37,096	6,771	10,934	230	1,595	359						
1976	66,703	9,500	15,400	1,417	44,319	9,938	12,227	320	1,875	768						
1977	75,732	7,894	16,300	1,660	55,400	7,539	12,000	322	2,500	687	1,000		1,700		7,700	100
1978	—	9,942	15,500	1,695	64,200	13,170	—	336		768		61				661

Source: *The Military Balance* (IISS, London, 1979); *OPEC Yearbook 1978.*

Table 2.2: Military Establishments of the Gulf States

	Iran	Saudi Arabia	Iraq	Kuwait	Oman	Qatar	Bahrain	UAE
Total Armed Forces								
Regular	413 (1978)	44.5	222.0	11.1	19.2	4.7	2.3	25.15
Reservists (000)	300	—	250.0	—	—	—	—	—
Army								
Personnel (000)	285	35	190.0	9	16.2	4	2.3	23.5
Tanks	1,735 MBT	350 MBT	1,800 MBT	280 MBT	—	12 MBT	—	30 Lt.
Air Force								
Personnel (000)	100	8	28	1.9	2.1	0.3	—	0.75
Combat aircraft	447	178	339	50	35	4	—	52
Navy								
Personnel (000)	30	1.5	4	0.2	0.9	0.4	(0.2)	0.9
Major craft	28	4 (+ 120 small craft)	48	28	13	6	9	9

Source: *The Military Balance.*

Table 2.3: Revenue and Expenditure of the Gulf States (millions of local currency)

		1972	1973	1974	1975	1976	1977	1978
Iran	Revenue	383.3	448.6	1,427.3	1,628.8	1,895.5	2,097.0	1,612.0
(000 mil.)	Expenditure	380.3	504.4	1,287.3	1,614.2	1,933.3	2,511.0	2,068.8
Iraq	Revenue	975.5	1,654.9	2,983.5	3,202.6	5,045.0	5,998.4	—
	Expenditure	1,117.9	1,399.9	2,604.9	3,640.6	5,045.0	6,339.1	—
Saudi Arabia	Revenue	13,200	22,810	98,247	110,935	110,935	146,493	—
	Expenditure	13,200	22,810	45,473	110,935	110,935	111,400	—
Kuwait	Revenue	597.7	675.3	2,271.4	3,960.4	3,033.1	2,272.7	—
	Expenditure	396.7	536.7	1,085.2	1,032.6	1,375.3	1,988.0	—
Oman	Revenue	53.0	65.0	303.2	459.4	505.4	613.3	509.0
	Expenditure	69.4	91.7	329.3	495.5	580.7	534.3	560.0
Qatar	Revenue	1,230.4	1,719.7	5,496.6	7,134.7	8,927.2	8,154.5	—
	Expenditure	958.6	1,514.8	1,931.3	5,302.3	5,808.9	7,318.6	—
Bahrain	Revenue	—	—	123,746	137,450	201,862	263,897	280,000
	Expenditure	—	—	77,856	121,768	203,191	256,200	280,000
UAE	Revenue	2,180.8	4,001.8	4,211.6	13,083.7	18,401.0	—	—
(Abu Dhabi only)	Expenditure	1,735.6	3,346.5	3,391.1	12,877.1	18,605.0	—	—

Source: OPEC Yearbook 1979.

Table 2.4: Balance of Trade (millions of current US dollars)

	1972	1973	1974	1975	1976	1977
Qatar						
Exports: total	397	618	2,016	1,808	2,210	2,048
of which oil	382	601	1,979	1,757	2,138	2,031
Oil exports as % of total exports	96.2	97.3	98.2	97.2	96.7	99.1
Imports: total	141	195	271	413	775	—
Trade balance	256	423	1,745	1,395	1,435	—
United Arab Emirates						
Exports: total	1,082	1,801	6,392	6,879	8,543	9,489
of which oil	1,035	1,740	6,306	6,727	8,238	9,135
Oil exports as % of total exports	95.7	96.6	98.7	97.8	96.4	96.3
Imports: total	482	821	1,705	2,669	3,327	—
Trade balance	600	980	4,687	4,210	5,216	—
Iraq						
Exports: total	1,108	1,944	6,600	8,297	9,272	9,664
of which oil	1,022	1,836	6,505	8,177	9,114	9,507
Oil exports as % of total exports	92.2	94.5	98.6	98.6	98.3	98.4
Imports: total	705	894	2,371	4,215	3,470	3,541
Trade balance	403	1,050	4,229	4,082	4,902	6,123
Kuwait						
Exports: total	2,984	3,829	10,964	9,183	9,830	9,832
of which oil	2,833	3,593	10,564	8,593	9,093	8,880
Oil exports as % of total exports	94.9	93.9	96.4	93.6	92.5	90.3
Imports: total	797	1,053	1,552	2,390	3,324	4,460
Trade balance	2,187	2,776	9,112	6,793	6,506	5,372

Table 2.4: continued

Iran						
Exports: total	4,088	6,249	21,516	20,226	23,435	23,974
of which oil	3,637	5,614	20,906	19,634	22,917	23,260
Oil exports as % of total exports	89	89.8	97.2	97.1	97.8	97.0
Imports: total	2,570	3,737	6,616	11,696	12,567	18,394
Trade balance	1,518	2,513	14,900	8,530	10,868	5,580
Saudi Arabia						
Exports: total	4,517	7,630	30,992	27,746	36,125	40,929
of which oil	4,502	7,600	30,955	27,698	36,076	40,881
Oil exports as % of total exports	15	30	37	48	49	48
Imports: total	1,136	1,975	2,859	4,213	8,694	10,770
Trade balance	3,381	5,655	28,133	23,533	27,431	30,159

Source: OPEC Yearbook 1979.

Table 2.5: Area, Population and Population Density of the Gulf States

	Area (Square miles)	Population	Population Density (inhabitants/sq. miles)
Iran	628,000	35,210,000	56
Saudi Arabia	927,000	7,870,000	8.5
Iraq	172,000	12,330,000	71
Kuwait	5,000	4,200,000	216
Oman	82,000	840,000	10
Qatar	4,000	200,000	500
Bahrain	230	350,000	1,521
UAE	32,000	710,000	22

Table 2.6: Volume of Crude Oil Exported 1972–8 (1975 = 100 base year)

	1972	1973	1974	1975	1976	1977	1978
Iran	96.5	112.9	114.9	100.0	111.6	104.4	96.5
Saudi Arabia	82.6	106.3	120.1	100.0	122.0	130.4	116.7
Iraq	69.0	95.0	93.0	100.0	111.0	102.0	116.0
Kuwait	164.0	148.0	123.3	100.0	100.4	90.4	98.5
Oman	82.7	85.7	84.8	100.0	107.7	97.8	92.7
Qatar	112.4	132.7	119.0	100.0	115.9	100.1	112.6
Bahrain (refined)	118.6	124.8	126.3	100.0	108.8	126.5	119.1
UAE	71.0	89.1	97.7	100.0	114.0	116.3	107.3

Source: International Monetary Fund.

3 THE SOVIET UNION AND THE PERSIAN GULF

Shahram Chubin

The persistence of Western debate about Soviet intentions in the Third World has been in some respects overtaken by the general consensus now prevalent that the USSR has, at a minimum and for the first time, achieved military parity with the West. The implications of this too are the subject of debate; some see military power today as too specific to translate into more general influence and see parity and a growing economic interdependence between East and West as a stimulus for conservatism on the part of the USSR. In this view the Soviet Union is becoming a *status quo* power with a stake in the international economic order. An acknowledgement of the USSR's role as a superpower, and a recognition of economics as the primary determinant of its foreign policy[1] would, in this view, contribute to a 'relaxation of tensions'. This school of thought tends to emphasise that Soviet policy is risk-averse, that it is constrained both by the limitations of the USSR in regard to forms of power (other than military) and to the complexities of regional politics and the dynamic nationalism of Third World states inherently resistant to the appeals of Communism and the lures of Moscow. It sees a world of complex interdependence, a *détente* coexisting with competition while the dynamics of regional and national politics stimulate uncertainty and instability in most of the Third World. Recognising the limits of the effectiveness of military power as a solution to these trends this school looks to a Western policy that promotes respect for certain inalienable principles, and to a search for areas of compatibility with local states. It seeks security multilaterally, with proven allies in Europe (and Asia) and by the encouragement of regional co-operation, formal and informal, elsewhere.

A second school of thought exists in the West that differs both in perspective and in emphasis from the preceding one. It cannot be said to be in direct or complete disagreement with much of the thinking of the first, but its focus is sharper, and its concerns more defined. It is more Hobbesian in its view of the international system, less impressed by the limits of military power than by its continued utility (particularly in affecting third parties' judgements) and less disposed to look at the constraints on Soviet power and the alleged paucity of the gains of Soviet foreign policy, than to note the gradual

transformation of the strategic environment in a direction adversely affecting Western interests. It is more concerned by the degree of Soviet effort in expanding its influence than by its failure to achieve irreversible breakthroughs. Unremitting Soviet pressure, persistent Soviet opportunism, incrementalism, probing and the setting of precedents have, or could have, in this view, a cumulative effect in transforming regional environments in directions favourable to Soviet influence. Just as the emphasis on 'breakthroughs' or 'setbacks' obscures the long view of history, the focus on regions as discrete entities unrelated or dimly related to the global strategic balance ignores interconnections between regional politics and Western interests. This school of thought, seeing no stable balance of power, is more concerned with its delicacy and with its fluctuations; nuclear deterrence is seen as neither automatic nor comprehensive, other balances are viewed as imperative, especially regional balances, which might include the need for an extra-regional balancer.

These two approaches[2] (admittedly greatly simplified) contend for dominance in US foreign policy and even within the Carter Administration itself. They account for the fluctuations of this administration; for its genuine optimism regarding the US world role in the long run, and the sudden fluctuations and emphasis on military power (which is rediscovered as it were) during crises. This ambivalence is a sign of the contradictions of the times; competition coexists with limited *détente*; public opinion wants limited foreign entanglements and maximum influence; domestic, economic and social issues compete for scarce resources with international demands; leadership is required by domestic needs, yet a troubled and turbulent international order also requires time, effort and patience, with little promise of tangible reward. A complex international system has replaced the more predictable loose bipolar one; hierarchy has vanished, and the capacity of superpowers to manage or 'order' international affairs has correspondingly declined as the agenda of the Cold War has been replaced by a myriad of local issues. Small wonder then that the two approaches outlined should coexist within one administration. Indeed, there are few analysts who do not simultaneously hold elements of both in their thinking. Nevertheless the two approaches and emphases lead to quite different interpretations of Soviet foreign policy and to advocacy of quite different remedies in dealing with it, and the problems of a particular region.

The Persian Gulf and the International System

Although lying outside the two principal blocs and hence constitu-

ting a part of the Third World, the Persian Gulf is unusual in several important respects. For the USSR it is an area — through Iran — directly contiguous to it which is assuming greater importance daily both for Soviet (and Eastern bloc) energy needs in the 1980s, and for its growing value to NATO and Japan, which might lend itself to political exploitation. Before assessing the record of Soviet policy in the region it is worth tracing the international context both in terms of the region's importance and of Soviet military power. The vulnerability of the West to interruption of its oil supplies needs little emphasis here.[3] The October 1973 embargo and the revolution in Iran in 1978 demonstrated this in their different ways; cut-offs in oil supplies from even one major producer for several months can damage the health of Western economies, so narrow is the margin of supply and demand. Political pressures within OPEC to reduce production for conservation (or not increase it in the case of Saudi Arabia) have recently increased. So too have the incentives for the manipulation of the 'oil weapon' for political purposes. Quite apart from the purposeful use of oil (prices and production) for political ends, there exist a wide range of factors which threaten the continued supply of oil; prolonged domestic instability; inadequate technical expertise and inter-state war. None of these need be the result of Soviet policy, though presumably some are susceptible to its exploitation.

Whether or not Western dependence on a vulnerable region is exploited by Moscow, the region itself is volatile and subject to sudden and unpredictable change. As a result of mainly domestic transformations, Iran and Pakistan have left the Western camp, while Turkey, dissatisfied with the fruits of its alliance, is only loosely hanging on. The demise of CENTO — formally dissolved in September 1979 — and the threat of disintegration in Iran reflect a transformation of the political environment weakening Western influence. While not attributable to either Soviet power or policy, the effect has been in practice to increase Soviet opportunities for influence. The fact that neither the course nor the direction of change is predictable (or irreversible), and that all Western losses are not clear Soviet gains, does not detract from this. It only underlines the volatility of this region, the evanescence of influence, and the need to see events in both their regional and global contexts.

The growth of Soviet military power, the attainment of strategic parity and a global reach interact with both the fact of Western vulnerability and the instability of politics in various regions. In the Persian Gulf Soviet physical proximity underscores the potential political utility of military power. It need not be a specific tool to be effective. But its value as a specific instrument is undeniable, for example to maintain a friendly regime in power in Afghanistan. And

though it need not be immediately transferable into sustained influence across the board in a political relationship, it can as a *new* factor affect the perceptions of local actors. It may be argued that for many states what is worrisome is not the growth of Soviet military power but the decline in Western power and will, and especially the loss of a measure of predictability about US responses. This may indeed be the case but the growth of Soviet military power is, I believe, an independent and concrete source of concern for many littoral states. Soviet naval power and mobility — an ability to shrink the 'effective distance' from the southern USSR to the Gulf — endow the USSR with a new military capability. This makes the USSR a concrete reality for the historically insulated Gulf states, and a new factor which has accordingly to be taken into account. It also endows the USSR with the capacity to use shifts in the regional balance to manipulate these perceptions, and to capitalise on regional discontents by offering its services as ally, protector or arms supplier.

In combination with growing military power the Soviet presence is felt in the Gulf primarily by its establishment of footholds on its peripheries. Soviet friendship treaties have been signed with Afghanistan (December 1978) and the PDRY (October 1979). The Soviet presence in Ethiopia, and in the Red Sea opposite Aden, is also large.

The relentless growth of Soviet military power and its visibility in terms of political commitment to the PDRY and Afghanistan, especially since 1978, may be seen as the acquisition of strategic footholds to exert psychological pressure on governments within the Persian Gulf. They also serve as a reminder to the West of its vulnerability in crises or in wartime. But Soviet interest in South-West Asia, and especially Afghanistan, should also be seen in terms of Sino-Soviet competition. In the event of a war with China the USSR would be hard pressed to supply its navy in the Far East. To escape the geographical constraints which bolster up and contain its navy, and to speed up supplies, there is a certain logic in the acquisition of a strategic infrastructure in this region.[4] Soviet geopolitical investments in Aden and in a port in Baluchistan (Gwadar) on the Indian Ocean, would make sense for this purpose: similarly the north-south Afghan railroad. (Iran with a long Gulf coastline and a long border with the USSR would also be important in such a scenario.)

Whether for 'defensive' reasons *vis-á-vis* China or for leverage *vis-á-vis* the West, Soviet commitments on the periphery of the Gulf have undeniably expanded. It is not yet clear whether military parity has stimulated a propensity for risk-taking. It may be that Western 'restraint' has indeed assured that there would be no risks attendant on this probing. However, in the past year public and repeated

emphasis in Washington on the need for interventionary forces, for the assurance of allies, etc., etc., may have ended this policy of indiscriminate abstention. If so, the momentum of Western setbacks may be somewhat arrested. The Soviet Union's willingness to assist in covert and indirect ways — for example through Cubans[5] and East Germans — which has facilitated 'plausible denials' in the past — may be influenced by more concerted Western response.

It seems likely, however, that Soviet military power has had and will continue to have a distinct effect on Soviet policy in the Persian Gulf, first by expanding the definition of her legitimate interests, and second by attempts to convert military power into 'commensurate political influence'.[6] Soviet policy, or rather style, may continue to be blunt and opportunistic; for example, it has not resisted the temptation to exploit setbacks for the West, even if they simultaneously endanger Soviet interests. The expansive definition of 'security' and an over-compensation (or over-insurance) militarily, combined with a prickly *amour propre* contribute to her problems with neighbours whose security and sovereignity are thereby threatened. The result may well be a diplomacy that succeeds only in alienating states who have no illusions that 'equality' in their geopolitical context can mean anything other than Soviet preponderance.

The Persian Gulf: the Local Context

Whereas in almost any other region of the Third World local politics, factionalism and domestic disputes are isolated from world politics, even the smallest tremors in the Gulf potentially impinge on great power interests and raise major issues. For example, the victory of a rejectionist faction in the Saudi leadership, or a shift in Saudi oil policy as a result, say, of a security threat in the PDRY, could have devastating consequences for the West. NATO's cohesion could be undermined by cut-throat competition among oil-consuming states in a tight oil market. Yet the instruments of influence available to outside powers are limited, primarily due to the nature of the threats to Western interests. Many of the 'threats' are in fact inchoate: uncertainty about the nature and direction of social change in societies undergoing rapid and disruptive economic, cultural and social transformation. The case of Iran has shown that rapid economic growth brings with it social and cultural dislocation and that indiscriminate modernisation not only generates demands for political participation but also risks a backlash — a reaction against obtrusive foreign influences and impersonal forms of social order. For all the Gulf states the priorities are local, the remedies

structural. Ayatollah Khomeini, for all his appeal in Iran, is a transient figure representing the mystique of charismatic leadership of a bygone era. For most of the Gulf states, be they Iran and Iraq with their fractured societies and minorities, or Kuwait and Bahrain with immigrant populations and diverse religious sects, the pressing concerns are those of national integration and societal cohesion, of meeting the growing demands of ever more restive and expectant peoples. The increased load on governmental structures is matched by a singular asset — cash — but whether it will be a sufficient tool by itself remains doubtful.

The ebb and flow of politics is thus from indigenous sources which do not lend themselves to clear and easy manipulation by outside powers. Indeed it may well be that the greater the involvement of these powers in the development process, the greater the likelihood of xenophobia engendered by them. The stresses of adapting cultures and institutions to meet new demands in novel ways will vary, but the outcome is likely to be both a synthesis (rather than a sterile imitation of any particular model), and a lengthy process characterised by fits and false starts.

In the Arab world, to be sure, there persist conflicts that are exploitable, yet here as in Iran the dominant, transcendent, issue is modernisation.[7] The core of the Arab world has shifted in recent years with this shift in political salience, making the Arabian Peninsula its centre. Iran's revolution underscored this reality of the dominance of local concerns, yet it also unleashed uncontainable social forces throughout the entire region. Quite apart from any policy changes afterwards, the revolution in Iran was itself profoundly destabilising. It provided both a model and an incentive for dissidents throughout the Gulf; for those unhappy with their lot, their political freedom, their religious freedom, the 'pollution' of traditional culture, or whatever. Its reverberations persist particularly in neighbouring states possessing Shi'ite communities (Iraq, Bahrain, Kuwait and Yemen). The revolution also altered the political balance of power in the Middle East, undermining moderate pro-Western governments, giving them the choice between political isolation and a tactful adjustment of policy. Iran's upheaval also demonstrated the intractable problem of containing or 'managing' change for powers outside the region. Quite apart from the palpable defeat for Western (particularly United States) policy, and the inevitable costs for their interests, the revolution provided no simple lessons for handling similar events in other countries. While over-identification with the Shah may have been costly after the revolution, it had provided tangible benefits for several decades. The hasty withdrawal of US support, combined with public equivocation and attempts to accommodate to Iran's

new leaders, who despise the West, only succeeded in undermining and alienating Washington's other allies in the region. The upshot was a renewed focus on relations with Saudi Arabia. But is she any more stable and, if not, are there any alternatives to the US cultivation of ties with the House of Saud? It appears doubtful that the West in the next decade can avoid heavy dependence on Saudi oil. So policy choices still need to be made, and Saudi security concerns must be met if there is to be a solid basis for co-operation. The Western response to Iran did much to undermine this vital, though precarious, relationship.

The Soviet response to Iran's revolution is instructive as an example of her more general approach to change in the region. First, there is little evidence of any direct Soviet role in the unrest, although indirectly she assisted the opposition forces through the PLO and the Iranian Communist (Tudeh) Party. Iranian leftist forces, the fedayeen trained by the PLO with Soviet arms, played an important military role in the last stages of the revolution. Second, although adversely affected by the replacement of the Shah by Khomeini (in terms of both economic relations[8] and political predictability), the West was hit harder. Soviet discomfort was thus lessened by the setbacks for the West which Moscow insisted on exploiting and emphasising. The same Soviet ambivalence which arose from the tension between welcoming Western setbacks and accepting similar, if not equivalent, losses characterised Soviet policy in general. Soviet support for the revolution was not publicly articulated until after the Shah's departure in mid-January 1979. On the other hand, Prime Minister Brezhnev's warning to the US against military intervention (of 18 November 1978) in Iran was subsequently used by Moscow as an illustration of the USSR's 'protection of the Iranian revolution'. Once Khomeini was installed, Moscow welcomed him as Iran's leader. Despite the repression of Iranian leftists and Tehran's problems with Soviet friends in the region (Afghanistan and Iraq), despite Iranian allegations of Moscow's suppression of its own Muslims, and despite accusations of Soviet support for Iranian Kurds in their quest for autonomy *vis-a-vis* the central government in Iran, the USSR forbore from any official criticism of Khomeini. Indeed the Tudeh Party in Iran, following Soviet instructions supported Khomeini in *all* his domestic policies, including his demand for an Islamic Republic. Even Tehran's seizure of American diplomats in November 1979, a precedent that could create problems for the USSR, elicited an equivocal Soviet response.

The same ambivalence persisted, and doubts about Khomeini's actions and the future of Iran were more than counterbalanced by the potential benefits to be gained from exploiting anti-US sentiments, particularly if these could be fanned throughout the region.

The Iranian revolution, an upheaval with profound geopolitical repercussions, and a débâcle for Western interests of major proportions, was essentially the result of domestic trends and policies. It set into stark relief the interrelationship between domestic and international politics by underscoring both through its impact the consequences for other states and the limitations on their influence to affect its outcome. The relationship between the strategic realm and the forces at work in regional politics, though not omni-present, or easy to quantify, is none the less present. Outside great powers, though unable to create trends in particular regions, can modify them; support for moderate states and opposition to revisionist states can affect both the risk-calculus of friend and foe. Soviet persistence, patience and incrementalism can give the impression of an implacable adversary inexorably expanding its influence. If the response to this by the countervailing power is one of timidity and sudden fluctuations, it will diminish the regional states' incentives for identification with it. The United States' performance throughout Iran's crisis and its aftermath succeeded only in narrowing to virtually nil the narrow margin for influence that existed.[9]

Apart from the strategic environment which affects the costs of risk-taking, the regional political context itself is a major conditioner of outside powers' influence. In the Persian Gulf states' foreign policy as often as not is a product of elite orientations. It is seldom buttressed by a deep historical consensus. It is therefore volatile and subject to rapid fluctuations dependent on the perceptions of these elites as to the best course to assure their security. The impression of a militarily strong USSR, self-confident and willing to probe in new regions, if not offset by an equally strong West, is likely to lead to a recalibration of policy to take this into account. One repercussion of Iran's revolution was to accentuate doubts in the Persian Gulf about US willingness to stand by her allies. Saudi Arabia in particular responded by lowering her pro-Western profile and hinting at possible relations with the USSR.

The Soviet Approach to the Region

Three factors stand out in Soviet relations with the Gulf states in the 1980s. First, the proximity of the region to the USSR diminished by technology, accessible to Soviet power by land and sea, and increasingly seen as falling within a zone of 'legitimate' Soviet interest. (Because of her contiguity with both the USSR and the Persian Gulf and Indian Ocean Iran must assume top priority in Soviet planning.) Second, Western dependency on Persian Gulf oil will increase not diminish in the 1980s. As a result the region assumes importance as a

point for exercising political (and psychological) pressure against NATO and Japan and perhaps for stimulating inter-alliance feuds. Third, though not as vulnerable as the West, the Eastern bloc as a whole will become increasingly dependent on the region for oil in the next decade.[10] This could result in either commercial competition between the two blocs for oil, or political-military pressure to obtain oil on concessionary terms. In combination the three factors suggest a heightened Soviet interest in a region in which Western interests are vulnerable, and Western influence declining.

The record of Soviet policy, in this as in other regions, is consistent. As a superpower it has sought political parity, a position as co-guarantor with a right to an equal say in the region's affairs. In practice Soviet geographical propinquity would render any notional parity theoretical, giving it a political preponderance analogous to that of the US in the Caribbean. On one level Soviet policies in the region are unimpressive. Persistent interest has not been rewarded by irresistible advances. Growing military power has not seen equal expansion of political influence; Soviet diplomatic assets and liabilities appear equally balanced. An inventory of the former would include proximity and persistence in a turbulent region; a body of doctrine and organisational support; and a revisionist political orientation at least partially in tune with the political opponents of existing regimes in the region. The handicaps, though, are equally impressive; proximity and military power stimulate caution as much as affection among allies; Communism, whatever its theoretical merits, is atheistic and anti-nationalist, unappealing attributes for both Islamic and secular nationalists; and Soviet technical assistance is hardly a currency for influence in a rich region able to buy the best.

On another level Soviet policies appear more threatening, namely the sheer persistence of effort, the hedging of bets, the dominance of the pragmatic over the doctrinaire, and the absence of doubts and guilt feeling which constrain Western policies. For the USSR history is not periodised into four-year cycles, policies do not ebb and flow according to presidential campaigns. While ideologically optimistic, the Soviet leadership has been selective and realistic in its policies. As the Iranian Revolution again demonstrated, the chief criterion for support for a regime is whether it is anti-Western (even if it is also anti-Soviet). This illustrates the Soviet goal of transforming the strategic environment in a direction hostile to the West. Since the USSR's influence in the region has in the past been limited, any reduction of Western influence can benefit it, even if it does not bring with it a direct or conspicuous increase for the USSR immediately. The side-effects of local crises can be exploited to undermine Western alliances and credibility, thereby increasing the Soviet role, at least potentially. As noted, incentives for seeking equidistance

between the superpowers increase as the protecting powers' credibility is undermined.

The record of Soviet policy has been uneven. With only two states well disposed towards the Eastern bloc (Iraq and the PDRY), the formal position of the USSR in the Gulf region is unimpressive. Formal diplomatic relations exist with only those states — Iran and Kuwait. Even as footholds in the region, Iraq and the PDRY have been of limited value. The former, politically isolated from Gulf politics until recently, has been, in any case, jealous of its independence. The latter, small and poor, has been on the periphery of Gulf politics politically as well as geographically. Nevertheless progress has been made with both states: in the signing of treaties of friendship; in the encouragement of anti-Western sentiments; and in the utilisation of the anti-Israeli line for a 'strategic alliance'. However, Soviet relations have been impaired by a refusal to take sides with Iraq against Iran (or Syria) or to provide the degree of political support and arms *vis-à-vis* Israel that the Ba'thist leadership in Baghdad has wanted. Cross-cutting regional rivalries, in short, have complicated Soviet choices in the region. The result in Iraq has been a setback for Soviet influence after a high tide in the mid-1970s. In the PDRY exists the only Arab state tightly organised on Marxist lines genuinely bent on regional expansion. This relationship is more promising. The client is poor, weak, dependent and ideologically militant. By providing arms and a sanctuary to the seccessionists in the Dhofar province of Oman it was able to transform a local separatist movement into an ideologically tinged 'national liberation' war. The Soviet Union was willing to assist in this, but not to expand the investment to confront all the Gulf states, including Iran or the US. Since 1978 PDRY pressure on the Yemen Arab Republic and periodic border skirmishes have succeeded in eliciting a strong US commitment of arms to the Sana regime, financed by Saudi Arabia.[11] The PDRY remains a potent instrument for pressure on Saudi Arabia in two ways — first, as a reminder of her continuing military impotence, second, as a threat to her politically. For a united Yemen — whether achieved peacefully or by force — constitutes a major threat to Saudi dominance on the Arabian Peninsula. Thus whether under the leadership of either Sana or Aden, it would be resisted by Saudi Arabia.

Hitherto the Soviet Union has given priority to relations with governments in the region, often at the expense of local Communist parties (for example Iran and Iraq). This policy may well be reversed in the future as relations with governments reach a stalemate and as social strains within their societies give the Communist parties a new lease of life. In the Arabian Peninsula there are no significant Communist parties, although they may become attractive as vehicles for

secular opposition to regimes. In the PDRY (as in Afghanistan and Ethiopia) the ideological-organisational element is strong. In the PDRY (as in these other two states) weakness and poverty will combine to make a reduction in the Soviet connection much more difficult.

In sensitive parts of the world, the size or military might of a state is not an accurate gauge of its capacity for disruption. A small injection of force can upset a regional 'balance' and make a difference between success and failure. Externally assisted *coups*, the provision of sanctuary, political assassination and the like make it difficult to discriminate between indigenous and external causes. Surrogate forces and advisers further blur the attribution of responsibility for specific actions which may be plausibly denied.[12] This is not to argue for exclusive attention to delicate regional military balances. Arms and security guarantees can at best provide time, not political solutions. How the time is used depends largely on the regional states, for by timely reforms and attention to demands for political participation many of the roots of instability can be severed before they grow into major security threats which can be harnessed by other states.[13]

Soviet optimism about the prospects for the expansion of her influence derives in large measures from the nature of the political-social forces unleashed by rapid modernisation, and the attendant possibilities for its exploitation. A catalogue of these forces includes:

(1) weak or non-institutionalised means for succession;
(2) increasingly politically mobilised populations with few outlets for self-expression;
(3) heavy and growing demands on weak administrative structures;
(4) societies fractured in differing degrees along religious, ethnic, linguistic lines vulnerable to sectarian feuds, disintegrative tendencies and secessionist movements;
(5) the vulnerability of most states to transnational appeals, be they Ba'athist, Arab nationalist, Pan-Shi'a, or Palestinian nationalist.

In addition to these there persists the Arab-Israeli issue which threatens to polarise the Arab states or to encourage their militance (to assure their security, as in Kuwait). The spill-over effects or chain reaction of instability in this region, demonstrated most recently in the events in Iran and the Arabian Peninsula in 1978–9 illustrate the vulnerability of the region.

How far these indigenous weaknesses are conducive to the extension of Soviet influence is as yet unclear. So far the USSR has

managed to convert its attraction as a potential restrainer of Iraq into an arms sales arrangement with Kuwait.[14] That this will translate into influence is not yet proven, although it does provide a precedent. Iran's revolution followed by its move to non-alignment and the fluid political conditions existing in that country have simultaneously undermined stability in the region, constituted a defeat for the West, and opened up the prospect for greater Soviet influence there. A spin-off has been Saudi Arabia's encouragement of informal ties with the USSR, the commencement of Aeroflot visits, and an incipient reluctance to be identified only with the West. The 'revival of Islam', though potentially damaging to the USSR with her forty million Muslims and her avowed atheism, has so far been more damaging to the West. Whether anti-Western feeling can be harnessed to Soviet purposes seems unlikely. Nevertheless to the extent it weakens Western influence and dissolves Western alliances it does benefit the USSR.

The outlook for the USSR in the region is promising. The end of Western primacy in the region closes an era begun by Britain nearly two centuries ago. The expansion of the Soviet diplomatic presence in the Gulf appears likely within the next few years. This will both reflect and accentuate pressures on the littoral states to show greater sensitivity to Soviet interests. Together with a naval presence in the Indian Ocean and quick-reaction interventionary forces positioned in the southern USSR, the Soviet Union will be a major new factor that must consistently be taken into account and even appeased by these states. A willingness to support local allies against their domestic enemies (as in Afghanistan) may indeed make Moscow an attractive partner for regimes worried about internal security. The desire to neutralise such possible interventions in future domestic crises may well constitute a powerful incentive to accommodate Soviet interests by states that are currently lukewarm, if not opposed to it.

Although the prognosis for the region is one of domestic instabilities and turbulence spilling over in some cases into inter-state conflicts, this does not necessarily augur well for external powers. While a breeding-ground for enmities and for outside-power exploitation, these conflicts are unlikely to be clear-cut cases providing them with easy choices. Social and political chaos will not be conducive to systematic exploitation. The balance sheet of Soviet assets and liabilities is as equivocal as the record of Soviet success to date is unimpressive. Nor is the regional environment especially suited to the expansion of Soviet influence. The growth of this influence in time will largely depend on Soviet opportunism, forcefulness and propensity for risk-taking. These may or may not increase as a result of a change in Soviet leadership, a larger Soviet bloc stake in the

region's oil, or a new reading (or misreading) of Western powers' responses.

Nowhere in the world is *détente* likely to be tested as keenly as in the Persian Gulf in the 1980s. Whether tacit agreement on a code of conduct of mutual restraint (of reciprocal access) emerges, or whether *détente* is limited only to Europe, the West will not be able to escape a dilemma. In an era of growing political and economic demands for priorities *within* the industrialised world there remains a very strong requirement for an active foreign and security policy in the Persian Gulf region. And while many local forces are converging to undermine stability in that area, the military dimension of security policies remains a powerful instrument to assure interests. Soviet policy will be powerfully affected by Western policies; timidity, guilt feelings, scapegoating and wishful thinking about the utility of military power will continue only to narrow its utility — asymmetrically at that. Security in the Persian Gulf and access to oil will surely not be achieved only by military means, but equally surely they cannot be buttressed without them. Above all, Soviet policies which encourage if they do not instigate disruptive forces can be deterred and a more cautious and moderate policy instilled only through an adequate policy that includes a military counterpoise, combined with a reputation for its judicious use.

Notes

1. Marshall Shulman, 'Toward a Western Philosophy of Co-existence', *Foreign Affairs*, vol. LII, no. 1 (October 1973).

2. Admittedly highly stylised here, few individuals would subscribe to either fully, and most would agree with elements in each; nevertheless this summary is intended to highlight the differing emphases of two approaches.

3. The US depends on Saudi Arabia for 19% of her imports, roughly 10% of total consumption, on Iran for 4% of total consumption. This dependence on Gulf oil is much higher for continental Europe and Japan, ranging from 50% to 70%. For the Western alliance as a whole dependence on the Gulf region will continue to grow throughout the 1980s at least. For a detailed discussion of Soviet policy towards the Persian Gulf, see the author's *Soviet Policy towards Iran and the Gulf*, Adelphi Papers, no. 157 (IISS, London, 1980).

4. See Michael McGwire, 'Naval Power and Soviet Global Strategy', *International Security*, vol. 3, no. 4 (Spring 1979), pp. 134–89.

5. There are 500 Cubans in South Yemen, according to official US sources, and 150 in Iraq (*IHT*, 22 October 1979).

6. See Helmut Sonnenfeldt and William Hyland, *Soviet Perspectives in Security*, Adelphi Papers, no. 150 (IISS, London, 1979), p. 21.

7. This includes its pace, scope and direction, especially the reconciliation of modernisation, with what Patrick Seale calls the 'moral basis of society', a consensus in the roles of church and state.

8. Especially in the cut-back of gas sold on concessionary terms to the USSR and

the cancellation of a larger project involving a three-way swap of gas in the 1980s, 1GAT-II.

9. See my 'The Strategic Repercussions of Iran's Upheaval', *Survival* (May/June 1979); apparent US acquiescence in Soviet definitions of the 'permissible' in relation to assisting allies starkly contrasts with the failure of US protests to make any impression on Soviet intentions.

10. The USSR holds 11% of the world's oil reserves. Space precludes a detailed discussion of this here but the assumptions are that Soviet oil production will decline, that the USSR will then face a choice between cutting exports (which provided 50% of her foreign exchange in 1977 of $6.7 billion) accepting a cut in economic growth, or limiting subsidised oil exports to East Europe. The last two have serious domestic political and alliance-weakening implications. Even if all three measures are followed, however, Soviet interest in foreign oil will increase; the natural area of interest will be the Persian Gulf.

11. But this was followed by an equal $300 million commitment of arms by the USSR to Sana in late 1979. Soviet arms are likely to be more effective because they have been the principal supplier and therefore can count on some knowledge among the Yemenis.

12. For example, is foreknowledge or acquiescence in a *coup* part of abetting it?

13. I have Oman in mind here.

14. The arms sold in 1976 were SAM-7 missiles. Soviet advisers did not come with them.

4 THE UNITED STATES-IRANIAN RELATIONSHIP 1948–1978: A STUDY IN REVERSE INFLUENCE

C.D. Carr

The conventional view of US-Iranian relations during the reign of the Shah is that of Iran as the dutiful pro-Western ally of the United States under the aegis of a ruler who owed his position to direct US intervention in 1953 and who remained a US 'puppet' throughout his tenure of office. To support this view, factors such as the acceptance of US intelligence gathering facilities on Iranian soil and the unwavering support of the Shah for US actions in South-East Asia are proferred as evidence. Indeed, the major reason given for the Shah's fall from power is that he was closely identified with the policies and politic-economic philosophy of the United States, with the inference that the United States exercised control over the Shah's actions and dictated policy from Washington. However, available data from primary sources now indicate that far from taking a lead from the United States and being controlled by the US executive, the Shah himself influenced and often controlled the US policy-making process in pursuit of the goal that dominated his philosophy throughout his reign; the pursuit of the total physical security of the state through the expansion and equipment of the Iranian armed forces to the highest qualitative level.

Variously described as his 'personal quirk', obsession or pre-occupation, the Shah's interest in the development of the Iranian armed forces and especially in the acquisition of advanced weaponry caused to create an environment of coercive diplomacy between himself and successive US administrations that occasionally lapsed into evident crises, with US-Soviet competition being forced into a confrontation by the deliberate actions of the Shah. The trauma induced by such crises, coupled with a carefully constructed campaign of diplomatic attrition on US-Iranian security matters emanating from the Shah, so effectively narrowed US options in regard to Iran that by 1972, when President Nixon allowed the Shah unprecedented *carte blanche* on arms procurement from the United States, this was simply the recognition of the *status quo*. The Iranian tail had been wagging the US dog for nearly two decades before that date and the control of US-Iranian relations, especially on security matters, had been held in Tehran and not in Washington.

The Iranian Control System

Richard Cottam, in his analysis of competitive interference, identifies two basic types of leverage available to states indulging in coercive diplomacy: active and manipulatable and passive and less manipulatable.[1] The latter is aimed at creating and maintaining a bilateral relationship whose environment is generally conducive to a beneficial, to the applier of the leverage, alliance. The former lever is used specifically to achieve a fixed and obvious goal through the exertion of threats and pressures against the antagonist. The Shah used both forms of leverage throughout his rule and, despite the manifest crudity of many of his actions, succeeded in reducing the constraints imposed upon him by the US foreign policy-making process.

At an early stage the Shah was forced to realise that his own perceptions as to the strategic importance of Iran to the United States and to the West and the means of safeguarding this 'shield of the Middle East' were at odds with US policy.[2] Even in 1947, barely a year after the Soviet 'puppet' regime had been driven out of Azerbaijan, the US Department of State had considered that, for military assistance purposes, Iran could not be considered amongst those states that were vital to US security or 'under direct and immediate danger'.[3] Under the original legislation on the US military assistance programme, the Mutual Defense Assistance Programme of 1949, Iran was lumped together with the Philippines and Korea under Title 3 status, to share a total of $27 million, whilst Greece and Turkey, with Title 2 status, were to receive more than $211 million. This disparity especially in regard to US military assistance to Turkey, was to prove the major bone of contention for the next decade and was to be the basis of the crisis in US-Iranian relations that emerged in January 1959.

The realisation that even a thousand-mile border with the Soviet Union and a history of Soviet interest and interference in Persia/Iran was not enough to persuade the United States to treat Iran as a special case caused the Shah to construct the passive leverage system that he hoped would persuade the United States to underwrite his security programme. Primarily designed to inhibit pragmatic, nuts-and-bolts issues, such as the ability of the Iranian armed forces to absorb large amounts of sophisticated weaponry or for the economy to support excessive defence expenditure, the system concentrated upon a campaign of continuous advertisement of the vulnerability of Iran and of the danger that such vulnerability posed to US security.

The Shah needed to achieve certain goals in relation to the US foreign policy-making body: he had to establish the importance of

Iran to the United States; the importance of himself to Iran; and the importance of Iranian security, in terms of numbers of armed forces personnel and the quality of their equipment, to himself. To achieve this he took every opportunity to stress the instability of Iran in relation to external threat and to promote his own plan to reduce that threat. By personalising his demands, the Shah ensured that his requests for US assistance reached the highest levels of the US executive, often to the President himself. Thus, US-Iranian relations are often a direct insight into the mind and functioning of the President of the day and as such provide valuable knowledge of the presidential management of foreign affairs.

Two letters from the Shah to Presidents Eisenhower and Kennedy respectively illustrate the Shah's technique of applied leverage. Both were ostensibly to congratulate the President on an election success (Eisenhower's re-election in 1956 and Kennedy's election in 1960), but each used the pretext to press the Iranian case for a special relationship with the United States, especially on military assistance matters.[4] The Eisenhower letter begins with hyperbolic flattery, as does the letter to President Kennedy (to the latter the Shah writes: 'We all rejoice at the prospect of a young and vigorous personality taking in his hands the reins of government at a crucial juncture in the affairs of a troubled world.') The letter to Eisenhower continues with the Shah particularising current regional and global problems (the Hungarian uprising and the Suez crisis) to the situation in regard to Iran, stressing that 'regional armed conflicts with conventional weapons are not to be ruled out as a thing of the past'. Taking the argument further, the Shah postulates that 'aggression can be forestalled, if countries occupying key positions are well prepared and their military as well as financial and economic needs supplied'. He asserted that Iran was such a key state and that 'the present military weakness of Iran would. . .constitute a danger not only to herself, but also to the Middle East region and as a corollary to the whole free world'. Re-emphasising Iran's 'unique position', the Shah requests the personal intervention of the president on his requests for military assistance and signs the letter: 'Believe me, my dear Mr. President'.

The letter to President Kennedy is almost identical in tone and content. The Shah stresses Iran's role as oil supplier to the West and as the 'key to Asia. . .and the key to Africa in the near future' (remarks tailored to fit events in Laos and the Congo). The Shah then proclaims the physical insecurity of Iran 'in these troubled times' and notes that Iran is in need of assistance that 'only America can furnish'. The letter finishes with a reminder to the president that the Iranian budget had to be prepared by March 1961 and therefore could the President inform the Shah of how much Iran could expect in the way of military and economic assistance for the coming year?

The letters are simply illustrations of a process that was continually present within US-Iranian relations, particularly during the Eisenhower, Kennedy and Johnson administrations. The effect was to create an acquiescent response from the US executive through attrition; it was easier to conform to the Shah's perceptions as to the security of Iran than to attempt to stem the torrent of requests through argument or counter-proposals. However, occasionally the Shah's demands for assurances and *matériel* from the United States could not be met, and it was on these occasions that the Shah utilised the active leverage device.

Prior to 1953 the United States had not committed itself wholeheartedly to the Shah as the most propitious leader for Iran. Their initial support for him was based upon the 'least worst' thesis; in October 1949, US ambassador John Wiley wrote to President Truman: 'We must perforce hope for much from the Shah since we may expect so very little from others in [Iran].'[5] A year later, Secretary of State Dean Acheson ventured: 'was this not a case of corrupt and incompetent government [in Iran] which no matter how much equipment or money [the US] put in it would be doomed to fail?'[6] Clearly the United States was less than committed to the Shah at the time and therefore he was not in a position to apply leverage. But after the United States assisted him in regaining control of Iran in 1953 he recognised an important corollary of their action that was to form the essential basis of his leverage system.

After the return of the Shah in 1953 his opponents, both within and outside Iran, proclaimed him to be a puppet of the United States and that he would be forever in their debt. But the Shah recognised that the United States, through their support for him personally, had signalled that they regarded both the stability of Iran and his continued presence as ruler important enough to undertake politically risky active intervention on his behalf. Through their actions the United States policy-makers had 'crossed the Rubicon' in their relations with Iran and had handed the initiative to the Shah. As a by-product of this commitment the Shah also learned the important lesson of never allowing any likely candidate to emerge that the United States might possibly be inclined to support in preference to himself. As a result, when the Shah finally fell from power there did not exist any group or individual with the experience, training and unblemished record capable of representing the middle ground in Iranian politics.[7]

In the years following the counter-revolution in Iran the United States found that it had hitched its wagon to an uncertain and volatile star. The Shah's 'appetite for soldiers and military hardware was unrealistically unlimited' and the pressure on the US exchequer for assistance on a global scale precluded accession to all of the

Shah's demands. Even after the 1958 *coup* in Iraq had persuaded the United States to promise more support for the Shah, this did not satisfy the Iranian leader and as a result he made use of the active lever for the first time.

The active lever was comprised of a 'controlled crisis', with the Shah choosing the issue, the time scale and the options. The controlled crisis itself required a number of factors to be present: the Soviet Union had to be visibly prepared to play a catspaw role; the issue had to be clearly defined and simple to implement; the US target had to be isolated (in the 1959 case it was President Eisenhower and Secretary of State Dulles) and capable of unilateral policy-making if pressed to do so; the Shah had to be convincing in his threats to turn to the Soviet Union if the United States did not acquiesce to Iranian demands.

Threatening to turn to the Soviet Union for aid if such *matériel* is not forthcoming from the United States was (and is) a fairly standard tool in developing state diplomacy. This threat, however, tends to lose much of its potency when uttered by a conservative autocrat whose political philosophy is the very antithesis of Soviet socialism. The Shah recognised this impediment and overcame it in two ways. First, he never threatened to ally himself distinctly with the Soviet Union but rather threatened to 'go neutral'. In the bloc-building atmosphere of the 1950s and 1960s to lose an ally to neutrality was tantamount to a 'win' by the competing ideology and this the Shah realised. He therefore sought to make it quite clear that he was capable of going 'neutral'; the Department of State opined in 1961 'that the Shah is capable of making such a switch [to neutrality] — one of our Ambassador's main tasks has been to dissuade him, to soothe him, and to reassure him of US support'.[8] The other device that he used to ensure that the United States was convinced of his earnestness when negotiating with the Soviet Union was through the projection of himself as a 'non-rational actor' on matters concerned with Iranian security. With the US executive convinced that the Iranian armed forces and military security were an 'obsession' resulting in an 'insatiable appetite', the Shah destroyed any rationalist approach to a controlled crisis by robbing the US policy planners of any attempt to construct lines of logic that might defeat his objective.[9]

The controlled crisis also relied upon the imposition of a time-urgency factor to further confuse and constrict the US policy-makers. In respect to President Eisenhower, who was a man who relied upon a broad spectrum of canvassed opinion before making decisions, this was particularly important, and the Shah ensured that the crisis of January 1959 had to be resolved extremely quickly (within twenty-four hours), with the resolution being handled by the

President, Secretary of State John Foster Dulles (who was ill) and the US ambassador in Iran. As a result, the President took the line of least resistance and acquiesced to the principal demands of the Iranian leader.

The controlled crisis device is an element of coercive diplomacy. The United States termed it 'blackmail' and the Soviet Union, recognising that they had been used for political purposes, accused the Shah of being a 'liar and a tool of the US'.[10] The Shah considered it to be a legitimate process, capitalising on the 'power of the weak' to overcome the interference of a major power in the maintenance of security of a sovereign state. Whether viewed from Washington, Moscow or Tehran, however, the willingness to use such a device illustrated that Iran was far from being the pliant, unquestioning ally of the United States that formed the official public image of Iran. Indeed, so successful was the Shah in applying leverage against the Eisenhower Administration that he became a victim of this success, for when the Kennedy Administration required a state upon which to practice its own more tightly controlled foreign assistance philosophy and programme, Iran became the natural choice.

Like many propositions of dubious validity, the Shah's assertion that Iran deserved special treatment from the United States was plausible if taken at surface value. Iran did have a long border with the Soviet Union and that country had traditionally coveted Iran as an outlet to the Indian Ocean. The Iranian armed forces were not of the size nor well enough equipped to overcome a Soviet attack. And Iranian oil was becoming increasingly important to Western economies, including the United States. But whilst these arguments obviously carried some weight with the US foreign policy-makers it was the scale of the Shah's demands and the manner in which he conducted his bilateral relationship with the United States that placed him in dispute with successive administrations. In 1949, he requested US support for a 300,000-man army, with at least 150 tanks and three wings of jet aircraft.[11] In 1958, he requested that the United States furnish equipment for five extra divisions, even though the twelve divisions already present in the Iranian Army averaged less than 40 per cent in strength.[12] He also requested F-100 aircraft, even though the Imperial Iranian Air Force had trouble maintaining the F-84s in its inventory, and further demanded NIKE and HONEST JOHN missiles.[13] This illustrates the 'most of the best' ethos that was to pervade the Shah's weapons procurement programme from the start and which was to culminate in six years of acquisition, from 1973 to 1978, when the Shah ordered more than 10 billion dollars' worth of arms from the United States alone.

Apart from being the largest purchaser of US arms under the Foreign Military Sales (FMS) programme, Iran achieved a number

of other 'firsts' in relation to US foreign assistance. Iran received more, in dollar terms, under the Military Assistance Programme (MAP) than any other country not in formal alliance with the United States. Under the transitional FMS credit scheme, Iran received more than twice as much as any other state, including more than $200 million under the US Export-Import bank's 'Country X' scheme. In qualitative terms, Iran was the first state to receive a number of sophisticated weapon systems; the F-5 aircraft, the F-4D and the F-14 being just three.[14]

It was to attempt to constrain this invasion of Iranian funds from the civil to the military sector that the United States, up until the Nixon presidency, attempted to influence the Shah's policies. Only the Kennedy Administration succeeded in overcoming the restrictions imposed by the Iranian control system and the premature termination of that presidency allows there to be some doubt as to the long-term success of the administration's countermeasures.

In order fully to assess the extent to which US-Iranian policy was controlled by the Iranian leader it is necessary to review certain events during the period 1953–78 and to establish certain facts. By following the train of circumstances it is possible to understand why the United States allowed itself to fall victim to the 'king's new clothes' delusion in regard to Iranian stability and why it was so unprepared for the revolution of 1979.

The Eisenhower Administration

The period of the Mossadegh government had two important effects upon US attitudes towards Iran. Apart from forcing the United States to recognise the Shah as the only Iranian capable of maintaining the pro-Western orientation of the country, the confrontation between Mossadegh and the British government over the nationalisation of British oil interests in Iran saw the beginning of the eclipse of British power in the Gulf region which was finally to come to pass with the decision, in 1968, to withdraw the British military presence from the area. The United States, in supporting the Iranian stance on the issue, assumed the mantle of guarantor for the security of the Gulf and thus Iran passed from the UK sphere of influence to that of the United States.

When the United States ensured the Shah's return to power in August 1953 it did so with the knowledge that Mossadegh had been a popular leader backed by the 'vast majority' of the Iranian people.[15] The Shah, therefore, would need tangible support from the United States, in terms of both economic and military assistance, if he was to build a power base within the bureaucracy and the armed forces.

As a result, US military assistance to Iran increased nearly fivefold from 1952 to 1953, with a similar rise in non-military aid. Much of the aid was in the form of grants under the various provisions of the Mutual Defense Assistance Programme (MDAP) and the Mutual Security Act of 1954. But one category of funds, the defence support grant, existed in an appropriations limbo that allowed it to be used with some flexibility by the US executive.

The defence support grant allowed the budget of the recipient government to be balanced by directly injecting US funds into the military appropriations sector of the budget. In the two years 1956–7, Iran received more than $60 million in defence support funds from the United States and this enabled the Shah to pursue his security goals whilst ensuring that the civil sector received enough funds to continue for a further year. Even with this support, however, the Iranian government announced a budget deficit of $80 million in June 1957 and the United States was made aware that Iran expected the US to 'pick up the tab'.[16]

This seeming ingratitude by the Iranians stemmed, in part, from what the Shah and many others of his countrymen felt was an implicit promise by President Eisenhower fully to support the Iranian armed forces development programme if Iran joined the Baghdad Pact.

During the Eisenhower years the philosophy of collective regional security agreements as a bulwark against Communism was zealously promoted by both the President and his Secretary of State, John Foster Dulles. For the Northern Tier states, of which Iran was one, the United States promoted the Baghdad Pact, a formal military alliance with the United Kingdom as a full member but with the United States only maintaining an associate membership. Iran joined the Pact in 1955 and thenceforth claimed that the United States had lured Iran into abandoning her traditional neutrality with promises of increased military assistance that was not forthcoming. As was reported by the US ambassador in 1959: 'The Baghdad Pact has meant nothing to the people or government of Iran other than the strong hope of massive aid and/or territorial guarantees from the US in return for Iranian adherence to the Pact.'[17]

On a number of occasions the US government had attempted to convince the Shah that it was economic madness to attempt to build up the Iranian forces to a level whereby they might hope to hold off the Soviet Union.[18] Iran could not hope to conduct a *tous azimuts* defence policy and therefore the best guarantee of Iranian sovereignty lay in 'the deterrent strength of the United States'.[19] This was somewhat undermined by the unwillingness of the United States to commit itself to full membership of the Pact but, in compensation, President Eisenhower proposed to equip the armed forces of Iran to

a level that he considered to be advanced, in terms of quality of equipment if not in numbers of personnel.[20]

President Eisenhower's generosity was overtaken by events when, in July 1958, the government of Iraq was overthrown by a group of radical army officers. In an attempt to counter unease amongst the other Pact members, Dulles and Eisenhower proposed an 'executive agreement' between each of the Pact members (apart from Iraq) and the United States, on a bilateral basis, which guaranteed the security of each state in the event of external threat.[21] For the Shah this was not good enough. He desired a firmer commitment by the United States, which he wished to be partially in the form of an immediate increase in US budgetary support and partially in the form of assistance for the formation of an additional five divisions for the Iranian Army. The United States felt itself unable to conform to this demand and the situation developed rapidly into a confrontation.

In the second week of January 1959 the Shah contacted the Soviet Union and indicated that he was willing to enter into negotiations with the Russians over a non-aggression treaty. The Soviet ambassador and his immediate superiors in the Soviet Foreign Ministry were aware of the Shah's motives and were extremely wary in their response. It was Khrushchev who personally decided to send a negotiating team to Iran and who decided to 'give the Shah almost anything he might want'.[22] The Shah began negotiations by stating that Iran would remain within the Baghdad Pact and although the Soviet team were later to agree to this point, their deliberations delayed matters long enough for the United States to initiate counter-moves.

The US ambassador, reporting directly to Secretary Dulles and the President by airgram, interpreted the Shah's move as 'blackmail' for more US security support and proposed a number of options:

(1) the substitution of a territorial guarantee or treaty for the proposed bilateral executive agreement;
(2) provide sufficient budgetary aid to satisfy the Shah, to be 'topped off with a showy offer of rockets, destroyers or other hardware';
(3) hope that the Shah will take fright at the last moment or that the Soviet negotiators would 'overplay their hand'.[23]

In the event, a compromise was reached through which President Eisenhower, in a personal letter to the Shah, pledged to reduce the Iranian budget through an increase in the US defence support fund, to increase the level of military assistance to Iran and covered all with a warning as to the dangers of negotiating with the Soviets.[24] The Shah received the letter and shortly thereafter became diplomatically indisposed and the Soviet negotiating team dispersed with some rancour. Upset at having been misused in so visible a manner, the

Russians began an anti-Shah propaganda campaign that was to last for more than three years and which, when the Shah found it necessary to use the Soviet Union for leverage purposes in 1966, made a sufficient impression on the Iranian leader as to encourage him to compensate them with a small order for Soviet military equipment.

Despite the warning from the US ambassador that 'repeated experiences with such appeasement show that its adoption and execution would foreordain another, and probably more serious, crisis within one year at the most', the Shah did succeed in his immediate aim of levering more budgetary support from the United States. His knowledge of the US policy-making process and of fiscal arrangements within the executive had made him aware that if the President was sufficiently pressed he could, through judicious manipulation of certain funds, find additional monies. With the controlled crisis providing the incentive, President Eisenhower diverted $13 million from the Presidential Contingency Fund and thus met half the cost of the Iranian budget deficit.[25]

Whether the Shah received promises of extra hardware is unclear. The lead-time on such items as NIKE missiles would have precluded their arriving in Iran before the end of Eisenhower's term of office and the assumption of President Kennedy to the position of chief executive ushered in a completely different attitude towards Iran. The armed forces were, however, expanded during the period 1959/60 and US military assistance to Iran also peaked during these years, to an average of $90 million per year. The Shah had manifestly succeeded in applying both passive and active leverage during the Eisenhower years and this served further to fuel both his grandiose military plans and his own belief in his ability to manage the US policy-making process. Unfortunately for the Shah, President Kennedy and his advisers also believed in their own capacity to manage and to dictate and the Shah and Iran were chosen as a primary target for a re-emphasis on the aims of US foreign assistance policy.

The Kennedy Administration

In March 1961, President Kennedy announced his foreign assistance policy in the form of a special message to Congress.[26] After announcing changes in the administration of the US foreign aid programme, the President outlined a philosophy that centred upon long-range planning for the economic development of the recipient states and upon the implicit right of the United States fully to involve itself in all matters pertaining to utilisation of American aid. For the recipients, and especially for the Shah, this augured the beginning of the most active period of US interventionism, a time in which 'you

[the US] tried to impose your type of regime on other people'.[27] For the supplier, the United States, the Kennedy initiative, as laid down in the Foreign Assistance Act of 1961, sanctioned the right of the donor to participate in 'social engineering' on the grounds that grant aid represented US public money and should be administered and accounted for as such.

President Kennedy personally ordered a review of the US aid programme to Iran in early 1961 and a presidential task force was forced to formulate a long-term programme for Iran.[28] After reporting upon the immediate internal political problems within Iran during the first months of 1961, President Kennedy charged the task force with investigating the possibility of the US providing further support for the Iranian armed forces for counterinsurgency purposes (this being at the onset of the 'counterinsurgency era') and also to report on the size and composition of MAP aid to Iran.[29] To achieve the latter the task force canvassed the opinions of those agencies most connected with military aid to Iran: the Joint Chiefs of Staff, the Department of State, Bureau of the Budget, Department of Defense and the Central Intelligence Agency. The replies of each of these agencies provide an interesting mosaic of bureaucratic response to contemporary attitudes on US-Iranian relations.

The State Department report was an amalgam of 'facts to be faced', arguments based upon the advantages of the *status quo*, and insubstantial proposals for development.[30] After admitting that the Iranian elections of early 1961 were 'largely rigged', the position paper painted a bleak picture of mob-ruled, xenophobic conditions if the United States was to withdraw its support from the Shah. Adopting a pragmatic stance, the paper then concluded that the United States had little or no control over the actions of the Shah:

> It is often suggested that the United States, using its aid programs as leverage, could issue orders to the Shah which would, by their implementation, result in political tranquility. It has been suggested that some of these reform measures would include the ending of corruption, the establishment of genuine democratic institutions, the downgrading of the military. . . These suggestions presuppose that the Shah is a creature of the United States. . .,a common misperception in Iran. Any United States ultimatums or even heavy-handed hints would be regarded by the Shah as an intolerable interference in his affairs and would probably result in corresponding moves on his part toward the USSR and neutralism.[31]

On this basis, the Department of State recommended to 'continue its present policy of reassurance to the Shah. . .along with delicate

inferences by our Ambassador to the effect that the Shah should devote his attention to his internal political problems rather than to foreign and military affairs'.[32]

The Joint Chiefs of Staff, somewhat predictably, supported an increase in US assistance to the Shah, couched in military terms that included full US membership in the Central Treaty Organization (CENTO), the replacement for the Baghdad Pact, and the stationing of 'atomic weapons' in Iran.[33] Also predictably, the Bureau of the Budget report accused the US bureaucracy of 'procrastination. . .irresolution. . .and the apparent inability of the country mission to project an impression of competence'.[34]

The Bureau of the Budget report was important, since it called into question the impartiality of the US mission in Iran, and claims that it had lost its objectivity resulted in the necessity of the President having to send his own representative, Chester Bowles, to Iran in early 1962. The BoB contended that:

> the preoccupation of the Shah with military forces appears to have influenced US policy posture to such a degree that support of the government has been equated with efforts to satisfy the demands of the Shah for military assistance activity.[35]

The CIA report hedged its bets, identifying the principal US problem as being 'how to give the Shah sufficient support to preserve his present pro-Western policy without encouraging excessive demands for aid'.[36] It also warned that if the Shah was 'convinced that the US was withdrawing or significantly reducing its support for him, the chances of his working out an accomodation with the USSR would be much greater'.[37] Thus the consensus opinion within the bureaucracy was to continue to support the Shah at least at the level that had been achieved during the Eisenhower Administration.

In June 1961, President Kennedy met Premier Khrushchev in Vienna and the subject of Iran was raised. As a result of the Soviet leader's interest in Iran and because of the doom-laden nature of Khrushchev's comments, Kennedy returned to the United States determined to investigate further the stability of the Shah and his country. The result was a determination to redirect Iranian efforts away from military matters and towards the socio-economic development of the state.

The bureaucratic input into the task force analysis of 1961 had demonstrated that the President required a more detached and objective judgement on the situation in Iran if he was to formulate a feasible long-term programme that he could present to the Shah. He therefore sent Chester Bowles to Iran in February 1962 to assess for himself the nature and extent of the social, economic, political and

military problems and to report back directly to the President. Bowles met the Shah and other Iranian leaders and, despite fierce opposition from the US ambassador, recommended that the military element of the US aid programme should be downgraded and supported the proposal made by the task force entirely to eliminate the budgetary support programme.[38]

By March 1962 President Kennedy and his staff had prepared a proposed approach to the Shah on the question of reducing his armed forces to 150,000 men (down from 170,000 and 90,000 less than the projected level that had resulted from the January 1959 crisis).[39] Stressing the interrelationship between military, economic and political factors, the proposal took the form of classic 'carrot and stick' approach to the problem. In return for accepting the cut in forces and the termination of defence support funds, the Shah was offered a military assistance package that stressed the qualitative aspect of procurement that the Shah held so dear. At a cost of some $330 million, the MAP proposal included two squadrons of advanced aircraft, a squadron of medium transport aircraft, the complete replacement of all the soft-skinned vehicles in the Iranian inventory and the completion of a military airfield.[40] But the Shah was unimpressed and for more than six months 'difficult and delicate negotiations which involved the President, the Secretary of State, the Secretary of Defense and the Joint Chiefs of Staff' were conducted between Washington and Tehran.[41] Finally, in September 1962, the Shah accepted the inevitable and signed the agreement.

The question must be asked as to why the Shah did not invoke the controlled crisis leverage system in order to modify the US programme in his favour. It is possible that he did in fact endeavour to do so but was foiled by the resolution of President Kennedy and his advisers. Evidence, albeit much of it circumstantial, indicates that on hearing of the US proposals the Shah began to inject uncertainty into the US-Iranian relationship through such devices as publicly implying that he would abdicate if the United States would not make concessions and warning the United States that 'if America laid down the foreign assistance burden, he was sure that the communists were willing to assume it'.[42] But these statements were simply precursors to his central coercive argument.

Since the events of January 1959 the Soviet Union had been conducting a vigorous propaganda campaign against the Shah and this had disturbed him greatly.[43] Any gesture that he might make to appease the Soviet government would have been welcome and therefore the Shah entered into negotiations on a matter which, in the climate of 1962, was as potentially injurious to US foreign policy objectives as the non-aggression treaty had been in 1959. Aware that Iran was considered by the Soviet Union to be 'analogous to United

States relations with Cuba', the Shah reacted favourably to a Soviet proposal to reach an understanding which would ban all 'foreign rocket bases' from Iranian soil.[44]

The Shah ensured that the United States was made cognisant of his negotiations during his visit to Washington in April 1962 and that he was continuing his dialogue with the Soviets after his return.[45] He was forced to realise the intransigence of the Kennedy Administration during the visit of Vice-President Lyndon Johnson to Tehran in August 1962, and with the termination of supporting assistance he appeared to capitulate. But on 16 September 1962, as foreign missile bases were beginning to take on a deep significance, the Soviet Union announced that Iran had pledged not to have foreign missiles within its borders and that it would 'never become the instrument of aggression against the territory of the Soviet Union'.[46]

Whether the Shah's actions were the result of pique or whether he found it necessary to underline his independence at this particular juncture is unclear. Relations with the Soviet Union did improve after this agreement, so it did have the desired effect upon Iranian-Soviet relations. But it failed to impress the Kennedy Administration, which pressed ahead with the oversight of the Iranian reduction of forces and reorganisation of budgetary priorities.

The premature ending of the Kennedy Administration leaves a number of questions unanswered. Would the Iranians have accepted the run-down in their forces or was the Shah preparing to exert more pressure in order to lever more assistance from the US? Would the resolve of the United States have continued at a high level to resist the hostility of its own bureaucracy to a plan formulated and largely implemented by the President's personal staff?

Whatever may have been the case, the Kennedy Administration was able to prove that with sufficient political will and with a viable, long-term proposal it was possible to impose restraint upon the Shah without Iran 'going neutral'. Since it was the only period in more than thirty years that the United States ever attempted to modify the Shah's ambitions and priorities, this ensures that the Kennedy proposal is somewhat of a milestone in US-Iranian relations. Whether or not one agrees with the contemporary assertion that the United States had not only a right but an obligation to intervene in the internal affairs of sovereign states in order to safeguard its own security, the important fact is that whilst all US administrations from Truman to Johnson believed in this right, only the Kennedy Administration attempted to apply it to Iran.

The Johnson Administration

If the Kennedy Administration represented the period of greatest

presidential involvement in US-Iranian relations, then the Johnson Administration represents the period of least presidential interest. Important decisions were made in respect to the development of US-Iranian relations during this period, but it was the traditionally administrative sections of the US bureaucracy that took charge of this development, rather than the upper echelons of the executive.

President Johnson remained preoccupied with affairs in South-East Asia and with the formulation of his own domestic social programmes throughout his time in office. Only when events reached crisis proportions, as they did in respect to US-Iranian relations during 1966, did the President personally take an interest. For the most part, policy remained in the hands of 'middle-level officials', with the all-important task of co-ordinating US arms sales being orchestrated from the small office for International Logistics Negotiations (ILN) within the Department of Defense.[47]

One of the major decisions in which the President did take a personal interest was the conclusion that, with oil revenues increasing, Iran could afford to purchase the defence *matériel* that it required. Against apparent Iranian opposition, the United States and Iran signed a Memorandum of Understanding on 4 July 1964 which committed Iran to purchasing military equipment for cash, through credits provided by the United States.[48] As a result of this agreement, Iran was to receive $504 million in credits to purchase US military services and equipment during the years 1965–9, of which more than $200 million came from the Export-Import Bank (ExImbank) through the 'Country X' loan scheme.[49] This was a device whereby the Department of Defense could utilise ExImbank funds without that agency knowing what was to be purchased or even which country would be using the credit. The availability of these funds and of other sources of cash and credit enabled the ILN to control an arms purchase from 'the cradle to the grave', without the necessity of having to involve superiors in the department or in other bureaux. In this way, during the formative years of US arms sales policy the control of both formulation and implementation rested with a small group of bureaucrats within the State and Defense Departments, whose chief motivation was to increase the dollar value of US arms sales for parochial reasons.[50]

Even with the Iranian parliament prepared to sanction the Shah's regular requests for funds to support military credits, the cost of the most modern weaponry mitigated against the purchase of such systems as the most advanced combat aircraft, of which, in the mid-1960s, the F-4 represented the apotheosis. Nevertheless, even with procurement cost of more than $2 million (1965) per unit, the Shah determined to obtain at least two squadrons of the aircraft and in early 1966 succeeded in having a team from the Joint Chiefs' staff

go to Iran to investigate the legitimacy of the Shah's claim on the aircraft.[51]

The Joint Chiefs' mission returned to Washington in March 1966 and reported that there was military justification, in the light of the threat posed by the Iraqis to the Iranian oil installations, for the acquisition of two squadrons of F-4D aircraft. But not everybody within the US executive was able to concur with this decision (Secretary of Defense Robert McNamara and the Agency for International Development both objected to the drain that the purchase, which would total $200 million, would place upon the Iranian economy).[52] The Shah was pressed to acquire more F-5 aircraft (a design specifically related to the needs and purchasing power of developing states) which had been in the Iranian inventory, under MAP, since January 1965. The Shah remained adamant on the F-4s, however, and the impasse began to develop into another controlled crisis confrontation.

Since the *rapprochement* with the Soviet Union that had begun with the missile bases agreement of April 1962, Iran had made demonstrable moves towards a closer relation with the Soviet Union and its Eastern European allies. Trade agreements had been signed with Poland, Czechoslovakia, Hungary and Romania. Most significantly, the Shah had signed an agreement for Soviet assistance in the construction of a steel mill near Isfahan, with payment to be made in natural gas, which up until that time had been flared off by the Iranians.[53] In essence, the Iranians were receiving the plant as a gift from the Soviet Union.

With this new background of cordiality between Iran and the USSR, the credibility of an Iranian threat to turn to the Soviets for advanced weaponry, which began to emerge in early 1966, was enhanced. After the standard preliminary warnings (in a speech to the Iranian parliament, the Shah stated: 'We take orders from no-one. . . We shall now see what our oil revenues will be and what will be the terms of sale of. . .military commodities. We shall then act according to Iran's best interests'[54]), the Iranians let it be known that they were considering a purchase of Soviet missiles and aircraft if credits for the F-4s were not forthcoming from the United States.

Aware that the decision on the financing of the F-4s was to take place in Washington in mid-July, the Shah injected his own time-urgency factor through the device of a publicly stated, if somewhat veiled, threat. In a newspaper interview on 14 July 1966, the Shah stated: 'We don't have the money to pay cash [for the F4s] and we will know in a week to ten days whether the United States understands our problem.'[55] With the United States aware of the distinct possibility of Iran purchasing advanced weaponry from the Soviet Union, with the political and strategic implications that such a move

would have upon the regional balance, there began a week of intensive consultation. The President, the National Security Council and the US ambassador were all involved in the discussions which resulted in acquiescence to the Shah's demands.[56] In September 1966, Iran contracted for a squadron of F-4s, with training and support, for $160 million, with options (subsequently taken up) on a second squadron for $40 million.

Apart from being another instance in which the controlled crisis was used with evident success, the sale of the F-4s was to have wider implications for US foreign policy. The Shah, fully cognisant of the extreme reaction within the Soviet Union engendered by his use of the USSR as a bargaining tool in the January 1959 negotiations, sought to militate against such a reaction in 1966 by offering the Russians a consolation prize. In February 1967, the Iranians announced the low-technology defence *matériel*, to the value of $110 million, was to be purchased on credit from the Soviet Union.[57] By this move, the Shah not only appeased the Soviet Union but also signalled to the United States that his threat of July 1966 was not completely hollow.

As an unfortunate by-product of the announcement of the agreement with the Soviet Union, the press and public within the United States were made aware of certain inconsistencies in policy-making and management surrounding arms transfers. Congress, which had been kept in a deliberately unenlightened state on arms sales matters, became angered that it had not been informed of the sale of the F-4 to Iran until a press announcement in December 1966 and instituted hearings into arms transfer policy, which resulted in the ending of the Country X loan scheme, the curbing of the power of the ILN and, through the Foreign Military Sales Act of 1968, legislation aimed at bringing arms sales within the realm of accountability to Congress.[58] But all this was to have little effect upon Iranian arms procurement ambitions, for by 1969 the Shah was beginning to pay for most of his arms from the United States in cash and the constraints imposed upon him by credit purchase were no longer present.

Although stymied in the early stages of the Johnson Administration by the hostility that was made evident by the termination of grant aid in 1964, the Shah was able to establish his autonomy very successfully during this period. He emerged out of the interventionist shadow of the Kennedy era and, by the time of the election of Nixon in 1968, had begun to symbolise a permanency and power that was not only extremely rare amongst developing states but was becoming increasingly rare amongst all of America's allies.

The Nixon Administration

Prior to 1969 the Shah of Iran had been forced to contend with a US foreign policy philosophy which was basically at odds with his own ambitions for the development of Iranian security programmes. The US belief in the necessity of interventionist policy, coupled with what it considered to be the right and obligation to play an active role in the formulation of Iranian foreign and domestic policy given to it through the extensive civil and military assistance programmes, imposed severe constraints upon the Shah's plans for a militarily and politically strong and autonomous Iran. As outlined in the Shah's 'independent national policy', the Iranian leader, dubious as to the continued willingness and ability of the United States to guarantee the security of Iran (and of himself as the ruler of Iran), decided upon a course of action that, whilst still exploiting the benefits to be gained from Western technology, aspired towards a military and economic strength that precluded external interference, especially from the United States.

Iran was greatly assisted in its desire for recognition as a regional power through the announcement by the British, in 1968, that their military presence in the Gulf was to be terminated. The Shah immediately promoted Iran as the new 'gendarme' to fill the vacuum and to stop the Soviet Union, together with the Iraqis, from turning the Gulf into a 'Russian lake'. It was to this problem to which President Nixon addressed himself in the early months of 1969 and which, in May 1972, was to lead to the end of all attempts by the United States to influence the plans and ambitions of the Shah for a 'Great Iranian Civilization', founded upon a physically secure state underwritten by large amounts of sophisticated military hardware.

In July 1969, Nixon's security adviser, Henry Kissinger, had ordered a review of the Gulf situation by the staff of the US National Security Council (coded NSSM-66).[59] In the previous month President Nixon had outlined his own security policy in the Guam Doctrine, which stressed heavily the requirement for 'self-help' amongst the allies of the United States. As a result, the NSC study reached the conclusion that Iran, together in partnership with Saudi Arabia (fulfilling a minor military role), should be fully supported in its desire to fill the vacuum left by the British withdrawal.[60] With Iran and Saudi Arabia acting as the 'twin pillars' of Western interests in the area, the threat of Soviet interventionism could be minimised without the need for direct US involvement. But if Iran was to act as a proxy of the US, it required that it be given the proper tools to do the job.

The United States, in concert with the British, pledged to underwrite the Iranian arms procurement programme in an effort to build

Iran into the undisputed military power in the Gulf and this gave the Shah the freedom of movement and action that he had pursued, through the use of the control system, since 1948. With credits supplied by the US government and with increased oil revenue resulting from successful negotiations with the international oil companies, the Shah began to order military equipment of a qualitative and quantitative nature that hitherto had been unimaginable. Between 1969 and 1971, orders for US equipment included 30 C-130 transport aircraft, more than 200 helicopters (US-designed but produced under licence in Italy) and hundreds of US armoured vehicles. But, despite the support given to him by President Nixon and Henry Kissinger, the Shah still felt constrained by the process of control through which all arms requests were forced to pass through within the US bureaucracy. These controls, especially in regard to the state-of-the-art aircraft that the Shah desired to procure for his air force, appeared to be thwarting him in his grand design for Iranian security, and he appealed to the President to be allowed to submit his own list of requirements without the necessity of each request having to pass through the standard review process.

In May 1972, President Nixon and Dr Kissinger visited Tehran. During that visit the Shah was told that henceforth all Iranian arms requirements would be allowed to be formulated by the Iranians themselves and that the United States would act simply as an implementational agency.[61] This decision was transmitted as a memorandum to the US bureaucracy in July 1972, instructing them to allow the Shah to have virtually any conventional weapons that he desired and for which he could pay.

Not every agency within the United States had supported the President's decision. Prior to the May 1972 trip, the Department of Defense had 'specifically advised the president in writing' not to commit the United States to a *carte blanche* policy on arms sales to Iran.[62] The President and Dr Kissinger, however, chose to disregard this advice and indeed evidence suggests that prior to this 'unprecedented step for a non-industrial country' there was no 'major interagency review' of US arms sales policy to Iran before the decision was made.[63]

Both the President and his security adviser may have been relying upon the financial constraints imposed on the Shah through the high cost of sophisticated weaponry to moderate the Iranian demand for arms. However, even by November 1972, the Shah was accelerating his arms procurement programme to previously unimagined limits (he ordered 400 helicopters from the United States in that month alone, bringing to 700 the number of such aircraft on order by that date).[64] The quantum jump in oil revenues brought about by the October 1973 crisis did not itself engender the Shah's $12,000 million

'shopping spree' for arms. The process had begun decades before, had been accelerated by the decision to sell the F-4D to Iran in 1966, and sanctioned by the presidential decision of May 1972. The revenue from the October 1973 crisis simply allowed the Shah to place orders from his list of optimal requirements.

Although by the onset of the Nixon Administration the Iranian control system had been made redundant, the Shah continued to make the US policy-makers aware of its latent threat through periodic references to the option of turning to the Soviet Union for military and non-military trade.[65] Apart from keeping the US 'honest', these utterances were aimed at overcoming any questioning of a relationship that the Shah now considered to be one of a dialogue between equals. Challenge to his right to indulge in what an increasing number of observers considered to be a frivolous waste of Iranian resources on unnecessary weaponry was answered with paternalistic hyperbole and allusion as to the power of the oil weapon. In this manner, as in the past through allusions to the East-West balance of power, the Shah was able to inhibit debate and co-opt disbelievers.

To argue whether the Iranian adoption of an aggressive stance on oil price rises was the result of the Shah's predilection for advanced weaponry or whether the increase in revenue stimulated that desire is a 'chicken and egg' debate of dubious value. But nevertheless, the Shah, the supposed staunch ally of the United States and the West, was a prime mover in the OPEC campaign for higher oil prices that brought about considerable hardship to that group of nations to which the Shah had always proclaimed a special bond, the developing states, and which sent the developed states into economic recession. This factor tended to be ignored within the United States when Iranian orders for US goods and services (especially the $15 billion bilateral trade agreement signed in March 1975) appeared to be 'sopping up' the petrodollars that had been accumulated as a result of the 1973 oil price rise. Even the *New York Times*, often a forceful critic of the Shah and his policies, proclaimed, in 1975, that Iranian orders for arms and other US products was 'a stepladder out of the recessionary cellar', forgetting that it was Iran that had helped place the United States in that cellar.[66] But, as with many other aspects of US-Iranian relations, the argument that the United States was actively benefiting from Iranian oil wealth was an illusion. In the period 1972–8, US imports from Iran, almost wholly of oil, rose from a current dollar value of just over $2 billion to nearly $29 billion. In the same period, US exports to Iran, also at the $2 billion level in 1972, rose to just under $18 billion, leaving a deficit of $11 billion in the favour of Iran.[67] Even the great capacity of US industry could not absorb the damage done by the increase in oil prices.

The Nixon and Ford administrations' unquestioning support for the Shah did not constitute aberrant behaviour in regard to US-Iranian relations but was merely the logical extension of the way that those relations had been developing since 1948. Worn down by decades of conflict with the Shah's regime, the US bureaucracy had been forced to rationalise its acquiescence to Iranian *diktat* in the same inflated terms that the Shah himself had used to justify increased aid to Iran in the early stages of the relationship. What had been unacceptable then had been made acceptable by the attrition of decades of applied leverage, assisted by what has been termed the inertia quality of the 'curator mentality' within the US Department of State and by the presidential seal of approval that was the Nixon Doctrine. What had been the 'Iranian problem' during the Truman, Eisenhower and Kennedy administrations had, by 1969, been transformed into an 'alliance and partnership' which precluded even constructive criticism.

Although the question of US support for the Shah's national security programme was not the only issue between the United States and Iran in the post-war period, it so dominated the relationship as to force all other issues into insignificance. In its simplicity, the clash of perceptions and philosophy over the Shah's concept of what constituted a reasonable outlay on defence should have resulted in a reasonable compromise being found at some juncture during the thirty years of the Shah's rule. The imposition of the will of either side, attained by the United States during the Kennedy Administration and by the Shah after May 1972, was not a matter of compromise but of victory and defeat, and the medium-term effects of both of these confrontations was to inhibit the free exchange of ideas and criticisms that a genuine alliance and friendship should foster.

After the Fall: Where did the US Go Wrong? A Post-mortem

In January 1979, when it became obvious that the Shah's regime was close to the finish, the United States began to question why it had been so unprepared for the Iranian Revolution. The US intelligence community was blamed for not warning the President and the executive of the impending chaos and investigations were instituted.[68] The reality, of course, lay far deeper and was far more fundamental than a failure of intelligence. The years of appeasement that the US ambassador had complained of in 1959 had robbed the United States of the ability to deal honestly with Iran and its problems. More important than its inability to conduct an honest dialogue with the Shah, however, was the inability to conduct an honest appraisal of US-Iranian relations within the policy-making and implementa-

tional process of the United States. Years of what was seen to be necessary and expedient acquiescence to Iranian demands became an exercise in justification and self-propagandisation that was to result in the facade of US-Iranian friendship and symbiosis becoming the accepted reality.

With pragmatic economic and political arguments providing support for an unquestioning US attitude towards the Shah, the 'king's new clothes' ethos that dominated US-Iranian relations in the last decade almost inevitably contributed to the build-up of pressure within Iran. With the United States championing human rights as a cornerstone of US foreign policy but ignoring the history of abuse of such rights within Iran for pragmatic reasons, the result was to ensure a general and violent dislike of the United States as a hypocritical nation without which the Shah would have been unable to remain in power. This build-up of pressure, which because of its seeming dormancy but cataclysmic results might be termed the 'Krakatoa effect', had vented on a number of occasions (in 1953, 1961 and 1963), but had been capped through aggressive action by the Shah and by the United States. But to suppress such a force is not necessarily to defeat it, but may simply ensure that when it does finally 'blow', the result is an explosion of undirected, destructive energy.

In justifying his support for the Shah's demands on defence matters during the debate in 1962, US ambassador Holmes stated:

> Because of the Shah. . .we have [in Iran] a government aiming. . .at the betterment of the people without having to resort to demagoguery, hysterical xenophobia or foreign adventures of the sort of thing that plagues us in many other underdeveloped countries. It is the Shah's authority which has saved Iran from being so plagued. . .I agree that the Shah is obsessed with the question of military security but believe that by alleviating these worries and increasing his confidence. . .we will more readily obtain full cooperation on other matters of great importance to us.[69]

A year earlier, before President Kennedy was to meet with Khrushchev in Vienna, an anonymous State Department official was charged with preparing, on a 'devil's advocacy' basis, the likely views of the Soviet Union on the Shah and his regime. Two such likely views were:

(1) if the Iranian government does not heed the voice of the people and alter its policies, the people will take matters in their own hands and make such changes themselves;

(2) irrespective of the changes in foreign policy by the present government, it is inevitable that the people of Iran will replace it with a government truly responsive to their desperate needs and demands, because the Shah cannot carry out the reforms demanded by the people without destroying himself — something he will never do.[70]

From both of the above can be divined, in the light of hindsight, a large measure of irony. The ambassador's plea for appeasement takes on the attributes of a self-fulfilling prophecy, whilst the anonymous State Department official, if he had been reporting, in early 1978, in his own guise, rather than that of a Soviet counterpart, might have been credited with oracular powers. But both, and most particularly the ambassador's communication, touched upon the essential error of US policy towards Iran: the belief that *ad hoc* responses could be substituted forever in place of a thoughtful, long-term plan for the mutually beneficial development of US-Iranian security interests. In the face of a single-minded monarch with fixed ambitions, the looseness and inconsistencies that form many of the safeguards of a democratic state became, for the Shah, exploitable foibles.

Perhaps if anything is to be learned from the US-Iranian experience it is to be found in the comment made by Lincoln Bloomfield, in testimony before a subcommittee of the US Senate's Foreign Relations Committee during hearings into US arms sales policy in 1976. Speaking specifically in regard to US arms sales policy towards Iran, he stated: 'But we now also know that few. . .countries remain either United States or Communist allies. Like all countries, they have no permanent friends or permanent enemies, only permanent interests.'[71] The question to be posed on the matter of US-Iranian relations during the tenure of the Shah is this: was thirty years of appeasement that achieved, in security terms, relative stability during that period, worth the 'inevitable' result? It is a question that has wider implications than those simply relating to the US-Iranian experience.

The Iranian Revolution has served to break the continuity of US-Iranian relations not simply because the Shah has departed, but because the revolution is a rejection of the value system that the United States represents and which the Shah used as a yardstick of Iranian development. Whatever the disputation that had emerged between the United States and the Shah over the priorities for that development, the Shah's admiration for Western technology and the symbols of the advanced industrial society remained an unwavering constant throughout his rule. When conflict arose between Iran and the United States it was on matters of time-frame and applicability

of such symbols rather than the advisability of applying essentially alien values upon a traditional culture. Under the Shah, progress was equated with numbers of miles of multi-lane highway under construction, visible public works on a grandiose scale and, most importantly, with armed forces developed to the limits of defence technology. It is these representative icons of Western interference that have borne the brunt of the revolutionary fervour and such is the force behind the rejection of these symbols that any Western attempt to explain or justify their presence in Iran will, for the foreseeable future, meet with suspicion, hostility and repudiation.

During the Shah's rule the United States craved for the public identification of Iran with US ideological and foreign policy programmes and goals. So successfully had the facade of amity and symbiosis been maintained that with the departure of the Shah the wrath that the revolutionaries vented upon the deposed leader inevitably overflowed to include the United States. All revolutions need readily identifiable enemies as well as manifest heroes and the twin 'evils' of the Shah and the United States provided that convenience within the Iranian Revolution. With no colonial heritage or experience of an independence struggle, it is perhaps understandable that the target of the Iranian revolutionaries should be what many in the developing world consider to be the arch-promoter of neo-colonialism and the purveyor of illusions of prosperity, the United States. But the fervour will diminish in time, the United States will be forced into reaching an accommodation with Iran and there will be obvious pressures to return to nostalgic policies.

Keeping Iran from 'going neutral' or 'going Communist' will inevitably dominate future US policy in Iran as it has done in the past. How the United States might actively or passively ensure that such an eventuality does not take place will require the United States not only to forget the slights, real and imagined, that have been directed towards it during the revolution but also to rid itself of all vestiges of the old, paternalistic order within the US policy-making and implementational process. Advice from elements within the US government departments that dwell upon the 'success' of the *coup* of 1953 and of the 'stability' imposed by the Shah upon Iran can only be of disservice in the formulation of any renewed US-Iranian relationship. Honest counsel from flexible minds uncluttered by embitterment or by past loyalties will be at a premium when the situation stabilises to the degree whereby normal diplomacy prevails. After thirty years of self-indoctrination such minds will be difficult to find, but must be a prerequisite to a mature and balanced attitude towards post-revolutionary Iran.

If Iran is as important to the United States as the actions of successive US administrations have indicated, then a fundamental adjust-

ment in the way that the United States should approach and negotiate with any future Iranian government should not be an unacceptable price to pay. It is an adjustment that the great colonial powers had to make on the dissolution of their empires and if it is made correctly it does not necessarily mean a diminution in the global status of the United States. Indeed, by adopting a less obviously cynical approach to relations between the developed and the developing states the United States could steal a march upon its rival, the Soviet Union. Above all, the United States must demonstrate that involvement does not have to be synonymous with intervention. The temptation to manipulate and dictate must be avoided.

The sour after-taste of the Iranian Revolution will linger for some time in both Iran itself and in the United States, and it will feed off a distillate composed of the thirty-year rule of the Shah. It is, therefore, extremely important that in an environment that encourages instant mythology and diatribal hyperbole that the factual evidence should be approached and utilised with as much candour as possible. The history of post-war US-Iranian relations has a validity for post-revolutionary relations if it is used to highlight pitfalls, rather than to score political points. It may not be able to tell the United States how to conduct future diplomacy with future Iranian governments, but it might be able to save the United States from making some mistakes for a second time.

Notes

1. Richard Cottam, *Competitive Interference and Twentieth Century Diplomacy* (University of Pittsburgh Press, Pittsburgh, 1967), pp. 82–3.

2. The Shah to President Eisenhower, letter dated 20 November 1956, Dwight D. Eisenhower Library (hereafter DDE Lib.).

3. Memorandum of Conversation on Iranian Purchase of Military Supplies, from Edwin Wright, Assistant to the Director of the Office of Near East and African Affairs, dated 14 April 1947, in *Foreign Relations of the United States 1947*, Vol. V (GPO, Washington, DC, 1971), p. 905.

4. Shah to DDE, letter 20 November 1956. Also letter from the Shah to President Kennedy, dated 26 January 1961, John F. Kennedy Library (hereafter JFK Lib.).

5. Letter: US Ambassador in Iran to President Truman, dated 25 October 1949, President's Secretary's File, Truman Library (hereafter Truman Lib.).

6. Summary of daily meeting with the Secretary of State (Dean Acheson), dated 1 March 1950, in *Foreign Relations of the United States, 1950*, Vol. V (GPO, Washington, DC, 1978), pp. 482–3.

7. 'We should of course be on the alert for the rise of competent and creative alternate leadership [to the Shah], in or out of the military, which might allow a reconstruction of alternatives. . .[though] this latter is very difficult in Iran. . .since all such contacts run the risk of alienating the Shah', from 'The Current Internal Political Situation in Iran', Department of State Position Paper, dated 11 February 1961, JFK Lib.

8. Ibid.

9. 'I agree with Ambassador Bowles that the Shah is obsessed with military security,' US Ambassador in Iran in letter to the Secretary of State, dated 19 February 1962, JFK Lib.

10. Telegram, US Ambassador in Iran to the Secretary of State, dated 30 January 1959, DDE Lib., and Telegram, US Ambassador to Secretary of State, dated 16 February 1959, DDE Lib.

11. *Foreign Relations of the United States 1949*, Vol. VI (GPO, Washington, DC, 1978), pp. 471 and 488.

12. Memorandum from the Joint Chiefs of Staff, dated 3 February 1959, DDE Lib.

13. Ibid.

14. 'Foreign Military Sales and Military Assistance Facts, 1975 and 1978', (Defense Security Assistance Agency, US Department of Defense, Washington, DC).

15. Briefing Memorandum for the President re meeting with Premier Mosadeq [*sic*], dated 22 October 1951, signed by Dean Acheson, Truman Lib.

16. *New York Times* (hereafter *NYT*), 8 June 1957.

17. Telegram, Ambassador to President, 30 January 1959.

18. 'If we attempted to build up our military establishments to a level comparable to that of the Soviet Union, we would wreck our economies and leave ourselves so weakened that we would collapse without even being subjected to military attack': Secretary of State Dean Acheson, in Memorandum of Conversation, dated 18 November 1949, Truman Lib.

19. Background Information on the Visit of the Shah of Iran, dated 28 June 1958, John Foster Dulles to the President, DDE Lib.

20. 'The President said he felt strongly that we should assign first priority to increased military and economic aid for Turkey and Iran. They should have all the assistance they can absorb. The Secretary said we had trouble with Defense about going further than 10 divisions for Iran and the President replied that we should get those 10 decisions fixed up fine [*sic*]'; Telephone Log, dated 16 July 1958, DDE Lib.

21. President Eisenhower was worried that the Baghdad Pact states might misinterpret the term 'special agreement' proposed by Dulles at a meeting of the Pact in London in July 1958 as being in the form of a treaty, which would have had to be ratified by the Congress. Dulles informed him that he told all of the Baghdad Pact states that the United States would *not* sign treaties with them, only make the less binding 'executive agreement'; Telephone Calls, the President to J.F. Dulles, dated 28 July 1958, DDE Lib.

22. Telegram, US Ambassador in Iran to Secretary of State, 16 February 1959, DDE Lib.

23. Telegram, Ambassador to Secretary of State, 30 January 1959.

24. Ibid., also Telegram, J.F. Dulles to US Ambassador in Iran, 30 January 1959, DDE Lib.

25. Summary of Obligations and Expenditures, Assistance Program for Iran, International Cooperation Administration, undated report (1962), JFK Lib.

26. Special Message to the Congress on Foreign Aid, 22 March 1961, *Public Papers of the Presidents*, John F. Kennedy, 1961 volume (GPO, Washington, DC), pp. 205–13.

27. *US News and World Report*, 27 January 1969.

28. Arthur M. Schlesinger, *A Thousand Days* (André Deutsch, London, 1965), p. 380.

29. Memorandum: Robert Komer to McGeorge Bundy, dated 22 May 1961, JFK Lib.

30. Position Paper, 'The Current Internal Situation in Iran', dated 11 February 1961, JFK Lib.

31. Ibid., p. 7.

32. Ibid., p. 9.

33. Ibid., p. 10.
34. 'Some Notes on the Situation in Iran', Bureau of the Budget, dated 20 March 1961, JFK Lib.
35. Ibid.
36. National Intelligence Estimate NIE 34-61, Central Intelligence Agency, 1961, JFK Lib.
37. Ibid.
38. 'US Ambassador Holmes proceeded to take a series of positions that were directly opposed to my own. I was trying very hard to push the Shah to move ahead with a land reform program and other long-overdue reforms. But [the Ambassador] wanted to bring all these planes in all the time to please the Shah. So we were on opposite sides'; Chester Bowles, Oral History Transcript, JFK Lib., p. 92.
39. Memorandum: National Security Council, Subject: NSC Action 2447, Proposed Approach to the Shah of Iran, dated 8 March 1962, JFK Lib.
40. Schedule of Alternative Categories for the Modernisation of Iranian Armed Forces, Possibilities for FY 1962-67, Department of Defense, JFK Lib.
41. US Military Assistance to Iran — Position Paper, Department of Defense, dated 22 April 1963, JFK Lib.
42. *NYT*, 13 and 14 April 1962.
43. CIA, NIE 34-61.
44. The President's Meeting with Khrushchev: Department of State Position Paper, dated 25 May 1961, JFK Lib.
45. *NYT*, 16 September 1962.
46. Ibid.
47. 'There is evidence to suggest that the type and destination of American military equipment is determined by middle-level officials in the Department of Defense'; statement by Senator Symington, *NYT*, 24 July 1967.
48. In a statement in 1967, Lieutenant-General Khatami, Commander of the Imperial Iranian Air Force, stated, when asked about the US policy of changing from grant aid to arms sales to Iran, 'I think it is a mistake. I am really sorry to see it. Once you make a mistake like that you pay and pay a lot more to make up for it,' *Aviation Week and Space Technology*, 7 August 1967.
49. *Foreign Military Sales Facts 1975*, p. 19.
50. In a letter to the National Security Council, Jeffrey Kitchen, a State Department official responsible for the integration of US arms sales with foreign policy objectives, complained strongly about the heavy-handedness of the ILN: 'the roughness of ILN pressure tactics is leaving a legacy damaging to the long-term local contacts [of the US embassies in countries to which the ILN had sent representatives] and that it is denigrating the role of the local US representation because the real decisions are being made by the ILN/Treasury team from Washington,' Memorandum from Jeffrey Kitchen to Alexis Johnson, National Security File, dated 24 March 1964, Lyndon Johnson Library.
51. Henry Kuss (Deputy Assistant Secretary of Defense for International Logistics Negotiations), in a statement before the Subcommittee of Near East and South Asian Affairs of the Senate Foreign Relations Committee, Hearings on Arms Sales to the Near East and South Asia, 90th Congress, 1st Session, 1967 (GPO, Washington, DC), p. 12.
52. US Armament and Disarmament Problems: Hearings before the Subcommittee on Disarmament, Senate Foreign Relations Committee, 90th Congress, 1st Session, 1967 (GPO, Washington, DC), p. 76.
53. *Keyhan International*, 6 June 1966.
54. *Keyhan International*, 2 March 1966.
55. *NYT*, 14 July 1966.
56. 'In 1966. . .I negotiated on behalf of my government with Iran to determine whether it could possibly afford to borrow 50 million dollars from the United

States. . . It took a week to negotiate that and we finally concluded; yes, Iran could barely handle a 50 million dollar loan'; Testimony of Leonard Alne, before the Special Subcommittee on Investigations, of the House International Relations Committee: *Hearings on The Persian Gulf, 1975: The Continuing Debate on Arms Sales*, 94th Congress, 1st Session, July 1975 (GPO, Washington, DC, 1976), p. 176.

57. In mid-1966, Senator Fulbright, Chairman of the Senate Foreign Relations Committee, heard that the Shah was negotiating for F-4 aircraft and sent a member of his staff to Iran, under the cover of an Interparliamentary Union group, to assess the need for such an aircraft for Iran. The staff member reported that he could see no such need and the resulting staff study, *Arms Sales and Foreign Policy*, served to start the controversy that ended with intensive investigation of US arms sales policy and the entering of the Foreign Military Sales Act on to the statute books — interview with Dr William Bader, Washington, 3 November 1978.

58. Ibid.

59. Tad Szulc, *The Illusion of Peace* (Viking Press, New York, 1978), p. 167.

60. Ibid.

61. *US Military Sales to Iran*, A Staff Report to the Subcommittee on Foreign Assistance of the Committee on Foreign Relations, US Senate, July 1976 (GPO, Washington, DC), p. VIII.

62. Multinational Corporations and United States Foreign Policy: Hearings before the Subcommittee on Multinational Corporations, Senate Foreign Relations Committee, Part 17, 94th Congress, 2nd Session (GPO, Washington, DC, 1977), p. 176.

63. *US Military Sales to Iran*, p. 41.

64. *Aviation Week and Space Technology*, 13 November 1972 and *International Herald Tribune*, 14 May 1973.

65. See the remarks of the Shah in an interview in *Business Week*, 24 September 1977: 'if the US were to cut off arms supplies. . .so hurt for hurt, if we have no other alternative, we could buy arms from the Soviet Union.'

66. C.L. Sulzberger, 'The Biggest Deal of All', *NYT*, 15 March 1975.

67. *International Monetary Fund Directory of Trade Yearbook* (IMF, 1979), p. 22.

68. Hearings on the Performance of US Intelligence in Iran, House of Representatives Subcommittee on Intelligence (GPO, Washington, DC, January 1979).

69. Telegram: US Ambassador in Iran to Secretary of State, 19 February 1962, JFK Lib.

70. President's Meeting with Khrushchev, Position Paper.

71. Lincoln Bloomfield, Testimony before the Subcommittee on Foreign Assistance, Senate Foreign Relations Committee, 94th Congress, 2nd Session, *Hearings on US Arms Sales Policy*, September 1976 (GPO, Washington, DC, 1977), p. 98.

5 PERSIAN GULF NUCLEARISATION: PROSPECTS AND IMPLICATIONS

Lewis A. Dunn

The overthrow of the Shah of Iran has not fundamentally changed the prospects for the spread of nuclear weapons to the Persian Gulf. By the mid-1980s, many countries of that region, including the Iran of Ayatollah Khomeini or his successors, will have available the technology for acquiring at least a rudimentary nuclear weapon capability. More important, the change of regime has removed only one possible political impetus to widespread proliferation in the Gulf: an Iranian nuclear weapon programme driven by the Shah's pursuit of regional hegemony and global prestige. Two other triggers — an expanding nuclear arms race between India and Pakistan and Iraqi acquisition of nuclear weapons in an attempt to upset the Middle East *status quo* — remain. The not improbable result well could be the dangerous entry of nuclear weapons into some or even many Gulf countries' arsenals by the late 1980s.

Eroding Technical Barriers

During the next decade, the technical barriers to nuclear proliferation in the Persian Gulf will gradually erode. A growing number of the countries there will acquire the rudiments of a civilian nuclear infrastructure — and with that knowledge begin to ascend the nuclear ladder towards the technical capability to acquire nuclear weapons. In addition, the probability of expanding nuclear grey marketeering will further erode those technical barriers and allow countries to compensate for the gaps of indigenous technologies.

Up the Nuclear Ladder

Of the Persian Gulf countries, Iran until the overthrow of the Shah clearly had the most ambitious civilian nuclear energy programme. Building on the modest nuclear infrastructure of the early 1970s — which consisted solely of the operation of a small research reactor; the availability of only a limited pool of trained Iranian nuclear technicians, engineers and physicists; and a semi-moribund National Iranian Atomic Energy Commission established in 1958[1] — Iranian plans had called for the rapid and ambitious

development of an indigenous civilian nuclear industry. To achieve that goal, the former regime began construction of two West German reactors,[2] initiated negotiations with the French firm Framatome to purchase additional nuclear power plants,[3] began sending students and others abroad for training,[4] undertook to build with assistance from the French Commissariat l'Energie Atomique a nuclear research centre at Isfahan,[5] and purchased a 10% share in the French-sponsored Eurodif uranium enrichment plant.[6]

These nuclear energy development plans have been drastically curtailed under the Khomeini regime. With the exception of possibly finishing the two West German facilities at some future date, there are no plans for building additional nuclear power plants.[7] None the less, the still existent even if weakened domestic nuclear research establishment, students still training abroad, and those near-completed plants provide a basis for later resumption of Iran's ascent of the nuclear ladder.

By contrast, Iraqi nuclear energy activities have accelerated. Organised in 1957, Iraq's Atomic Energy Commission since 1968 has operated a small 2 MWe Soviet-supplied research reactor. Successful negotiations have been completed for French assistance in the creation of a new Iraqi research centre outside Baghdad, with its centerpiece to be a French 70 MWe research reactor.[8] It appeared initially that that reactor would be fuelled with a material directly usable for bombs; but in fact such a reactor destined for shipment to Iraq was sabotaged mysteriously in April 1979.[9] France has agreed to replace the reactor but apparently hopes to use the available time to redesign the core to permit use of other than high enriched uranium.[10] Finally, Iraq is negotiating to purchase a 600 MWe nuclear power plant from France.[11]

Elsewhere in the Gulf, the Kuwaiti government also has indicated that it has plans for developing civilian nuclear technology. British, French and German companies have been requested to bid on helping Kuwait set up a nuclear research and training centre including a 40–50 MWe research reactor and desalination facility. In addition, statements by Kuwaiti officials indicate a longer-term interest in acquiring several dual-purpose power-generating/water-desalinating plants in the 1980s and 1990s.[12]

Similarly, Saudi Arabia has decided to set up an atomic research centre and has evinced an interest in using nuclear reactors for desalination.[13] Nor should Saudi Arabian financial support for Pakistan's aborted purchase of a French reprocessing plant be overlooked in considering its interest in nuclear technology. And though its future civilian nuclear programme goals remain vague, there are reports that Saudi Arabia may send some of its future engineers and physicists to Pakistan for training.[14]

Thus, the dissemination of civilian nuclear technology in the Persian Gulf, though still limited, will continue in the next decade. Even in Iran's case, the one example of diminished activities, some students trained abroad will return home and some limited ongoing nuclear research programme will remain. And as that process of mastering nuclear know-how and technology continues, these countries will move closer to the capability of producing rudimentary nuclear weapons, whether by misusing civilian nuclear research or power facilities or by seeking to build small semi-indigenous facilities dedicated to production of nuclear explosive material. Their gradual ascent of the nuclear ladder — as well as that of key countries in neighbouring regions — is depicted by Table 5.1.

Grey Market Assistance

The emergence of more extensive nuclear grey marketeering in the early 1980s would accelerate the erosion of technical barriers by compensating for the gaps in indigenous nuclear capabilities. Although reasons of space preclude a detailed discussion of this threat,[15] several aspects of this phenomenon should be noted.

First, following Pakistan's recent example, discussed more fully below, a country might circumvent the controls on the major suppliers' nuclear exports by purchasing 'grey area' components or materials — items that have legitimate civilian uses as well as being necessary for acquiring a nuclear weapon capability. Reliance on purchasing agents and various similar covers in order to hide the ultimate destination of the purchases also could be a part — as they were in Pakistan's efforts to acquire centrifuge technology — of these grey area transactions.[16]

In addition to such purchases of grey area components and materials not fully covered by existing nuclear exports controls, a Gulf country could use its financial resources to hire nuclear mercenaries. Especially if the domestic market for nuclear power in advanced countries continues to shrink, there may be a growing number of out-of-work nuclear engineers, technicians, scientists and administrators whose backgrounds would run the gamut from reactor design and operation to plutonium reprocessing and handling. They could provide aspirant Persian Gulf or Middle East proliferators with useful technical know-how and practical experience as well as needed organisational skills for setting up and running a nuclear weapon programme. Further, others among them might well have had experience with nuclear weapon design itself. Depending on the state of the indigenous technology and the particular incentives of the recipient, the assistance of nuclear mercenaries might speed up completion of a semi-indigenous nuclear weapon programme not dependent on misuse of the fuel cycle,

Table 5.1: Growing Technical Capabilities

		[as of] 1980	1985	1990
6.	Capability to build enrichment facility	Israel	Israel	Israel
5.	Capability to build small plutonium production reactor (2–5 bombs/year)	Egypt (?); Iran (?); Iraq (?); Israel	Egypt; Iran; Iraq; Libya (?); Saudi Arabia (?); Syria (?)	Egypt; Iran; Iraq; Libya; Saudi Arabia; Syria
4.	Effective access to significant quantities of divertible fissile material from civilian nuclear fuel cycle		Iraq	Egypt (?); Iran; Kuwait; Syria
3.	Operation of nuclear power plant(s) and start of nascent nuclear infrastructure		Iran (?): Saudi Arabia	Egypt (?); Iran; Kuwait; Syria
2.	Nuclear power plants under construction or planned	Egypt; Iran, Iraq; Kuwait; Saudi Arabia Libya; Syria	Egypt; Kuwait; Libya; Syria	Libya
1.	Assimilating basic theoretical knowledge			

Sources: See 'World List of Nuclear Power Plants', *Nuclear News* (August 1977); *Nucleonics Week*; John R. Lamarsh, 'Dedicated Facilities for the Production of Nuclear Weapons in Small and/or Developing Nations' and 'Level II Dedicated Facilities' in *Nuclear Proliferation and Safeguards*, Appendix Volume II, Part Two, Appendix VI-A and Appendix VI-B (US Congress, Office of Technology Assessment, Washington, DC, 1977); Albert Wohlstetter *et al.*, *Moving Towards Life in a Nuclear Armed Crowd*? (Pan Heuristics, 22 April 1976); and Ted Greenwood, George W. Rathjens and Jack Ruina, *Nuclear Power and Weapons Proliferation*, Adelphi Paper no. 130 (International Institute for Strategic Studies, London, 1977).

permit a more ambitious indigenous programme, or facilitate illegal diversion from those civilian nuclear fuel cycle facilities. Nor is talk of nuclear mercenaries simply speculation: Pakistan's efforts to build a centrifuge enrichment plant were greatly aided by the return of a Pakistani national who had had access to sensitive centrifuge design information at the URENCO plant in the Netherlands, and according to a recent report several Dutch scientists pessimistic about the prospects for nuclear energy in the Netherlands are considering leaving for jobs in other countries.[17]

Grey market nuclear collaboration with another country is a further means by which a Persian Gulf country could acquire a nuclear weapon capability. Whether involving the barter of oil for technical assistance or a joint venture in which each partner supplied a component or material that the other lacked, such collaboration could readily compensate for technical weaknesses. From this perspective, Saudi Arabian ties with South Korea and Taiwan could take on added significance if political pressures on the Saudis to move towards a nuclear weapon option intensified in the 1980s.[18]

Thus, in a variety of ways grey market transactions could supplement the capabilities of prospective proliferators in the Persian Gulf, further reinforcing the conclusion that the technical barriers to proliferation in that region are likely to erode in the next years.[19] But that erosion is a necessary but not sufficient reason for envisaging the spread of nuclear weapons to the Gulf. More important are the political dynamics that could drive some of those countries to 'go nuclear'.

The Political Dynamics of Persian Gulf Proliferation

Before the fall of the Shah, most speculation about proliferation in the Persian Gulf centred on Iran's incentives and disincentives for 'going nuclear'. An Iranian nuclear weapon programme motivated by the Shah's idiosyncratic pursuit of regional hegemony, global status and Iranian grandeur was regarded as the most likely trigger to proliferation in the Gulf. The Shah's overthrow — and the turn of Iran in the direction of the more circumscribed goals of its traditional foreign and security policy consensus[20] — has removed, therefore, that impetus to the spread of nuclear weapons in this region. None the less, forces already gaining momentum in both South Asia and the Middle East could set in motion expanding proliferation chains eventually spilling over into the Persian Gulf.

An India, Pakistan, Iran, . . . Proliferation Chain: Late 1980s Persian Gulf Proliferation

Building on nearly two decades of atomic energy-related research

and development, India detonated a nuclear explosive device in May 1974. The longer-run precipitates of that device may include successively a matching Pakistani test, a fully fledged Indian nuclear weapon programme, nuclear arms racing in South Asia, heightened pressures for a matching Iranian decision to acquire nuclear weapons, and proliferation elsewhere in the Gulf.

In the aftermath of India's 1974 test — claimed to be that of a 'peaceful nuclear explosive' — there was considerable pressure from Indian scientists to test additional nuclear devices. Partly responding to hostile foreign reaction, the then Prime Minister Indira Gandhi resisted those urgings. Then, in 1977, Mrs Gandhi's government was ousted in parliamentary elections and Morarji Desai became Prime Minister. Desai reiterated his earlier opposition to an Indian nuclear weapon programme and made that opposition government policy.[21] However, Desai's forced resignation as Prime Minister in July 1979 removed that constraint and once again restored a measure of fluidity to Indian nuclear weapon policy. Of greater importance, there are several pressures that could result in a fully fledged Indian nuclear weapon programme by the mid-1980s.

Evidence is increasingly available that soon after that Indian test the Pakistani government of the then Prime Minister, Zulfikar Ali Bhutto, initiated a two-pronged attempt to acquire a nuclear explosive capability. The most publicly visible part of that attempt was Pakistan's effort to purchase a spent-fuel reprocessing plant from France, a deal which ultimately fell through. That plant would have permitted Pakistan to separate plutonium from spent fuel taken from its one operating nuclear power reactor; and, as Bhutto admitted in 1978 during his trial by the military government that replaced him, Pakistan intended to use illegally that separated plutonium for a nuclear weapon test.[22] However, even while it was attempting to complete the purchase of that reprocessing plant, Pakistan was also pursuing another route to nuclear explosive material. It became clear in 1978 that the Pakistani government had been engaged in building a centrifuge enrichment facility to supply nuclear weapon-grade uranium. Much of that activity, as noted earlier, relied heavily on nuclear grey market transactions to avoid nuclear exports controls. As a result, though many uncertainties remain, by 1980–1 Pakistan may have sufficient weapon-grade uranium for a nuclear test.[23]

As these Pakistani efforts became increasingly more visible in the late 1970s, pressures on the Indian government to reactivate its by then shelved nuclear research intensified. At first Desai and then his interim successors took a moderate line of response, stating only that India was closely watching developments in Pakistan but as yet had found no reason to change its current nuclear weapon policy.[24] Thus

if Pakistan's nuclear activities stop short of actually testing a nuclear explosive device, it may yet be possible to head off an Indian response and reactivation of its nuclear programme. However, should Pakistan go ahead and test, the chances of prevailing on the Indian government not to move towards a fully fledged nuclear weapon programme would be minimal. Indian security and status as well as domestic political considerations would supply the motivations for answering what could be seen in Delhi as Pakistan's challenge, if not affront. That is, in the early 1980s a spiral may emerge in which Pakistan's responses to the initial Indian nuclear test will trigger new Indian actions and further Pakistani responses.

Even setting aside Indian concern about Pakistan nuclear activities, there are other factors that could contribute to an Indian decision to proceed with a fully fledged nuclear weapon programme in the mid-1980s. One such factor — India's continuing quest for international recognition and the concomitant desire of her leaders to assert Indian political independence — partly motivated the decision to test in 1974 and could re-emerge as an incentive for resumed activities. That deeply ingrained demand to be more than a 'pawn on the global chessboard',[25] to use Mrs Gandhi's phrase, is likely to continue influencing Indian policy. Moreover, those who contend that 'going nuclear' constitutes a necessary symbolic assertion of both India's unwillingness to be 'pushed around' and her claim to due respect cite the American recognition of the People's Republic of China as an apt example of nuclear weapons' political utility. Further, just as the 1971 decision to manufacture a nuclear device stemmed partly from adverse reaction to the United States' decision to sail the aircraft carrier *Enterprise* into the Bay of Bengal in that year during the Indo-Pakistani war, a future decision to resume nuclear weapon activities might be encouraged by a bitter clash with the United States in the early 1980s over India's unwillingness to accept full-scope nuclear safeguards.[26]

India's relations with the People's Republic of China in the mid-1980s also could increase the status and security-related incentives for nuclear weapon acquisition. Some within India had hoped that the May 1974 test would produce Chinese acceptance of India's political equality. The exchange of ambassadors and resumption of diplomatic relations may have been a step in that direction. But to the extent that in the future that recognition of what is considered full and proper Indian status is perceived to be incomplete, incentives for nuclear weapons as a political symbol *vis-á-vis* China could intensify. Similarly, more security-related incentives entangled with Indo-Chinese relations also increase pressures for an Indian weapon programme. One objection of the Indian leadership to proposals to establish a South Asian nuclear-free zone is that the zone would

exclude China, the major nuclear threat to India. And even though until now they alone have proved an insufficient trigger, growing Chinese nuclear capabilities could reinforce the other Indian nuclear weapon incentives. In fact, increased Chinese conventional military capabilities, including the purchase of advanced combat aircraft from the United States, could also heighten Indian incentives in the 1980s to acquire nuclear weapons. In particular, erosion of the conventional balance with the People's Republic could partly weaken the Indian military's current scepticism about nuclear weapon acquisition and foster a 'more bang bang for the rupee' approach.[27]

In addition, scientific momentum and domestic politics also could enhance the build-up of support for an Indian nuclear weapon programme in the mid-1980s. The urgings of Dr Homi Bhabha, the first Chairman of the Indian Atomic Energy Commission, greatly contributed to the incremental Indian development of a nuclear weapon option; while during the period from the Chinese test of 1964 to the 1971 Indian decision to test, as well as since the 1974 test, key Indian scientists pushed for a nuclear weapon programme and had to be restrained by the political leadership. There is no reason to believe that pressure from these scientists will relent in the 1980s — at a time, moreover, when the technical options would be all the more available.[28] Alternatively, if Mrs Gandhi or her supporters emerge in strengthened positions from the current political turmoil, that, too, could affect the balance of forces on this issue. Thus, the fall of Prime Minister Desai removed the one key domestic impediment to a resumption of Indian nuclear weapon activities.

Therefore, a range of status and influence-related, security-related and domestic institutional factors could readily lead India to resume its shelved nuclear weapon programme. What partial constraints that would exist, for example concern about outside sanctions, cost, domestic opposition and presently if temporarily limited technological capabilities, appear likely to be discounted or at most to influence only the programme's timing. Consequently, an Indian nuclear weapon programme in the mid-1980s is far from unlikely.

Turning to the next link in this South Asian proliferation chain, it is difficult to believe that India's emergence as a fully fledged nuclear weapon state in the mid-1980s would not place irresistible pressures on Pakistan to intensify its own nuclear weapon programme. Even a crude nuclear force could be estimated as essential for resisting Indian nuclear blackmail or unilateral Indian nuclear weapon use in a future conflict between the two countries. After Pakistan's experience in both the 1965 and 1971 Indo-Pakistani wars, reliance on an American or other countries' security guarantee — the main alternative to nuclear weapons — in all probability would be rejected. At the same time, these security-related pressures are likely to be rein-

forced by domestic elite and military pressures for an 'in-the-business bang' if that had yet to take place or, more likely, for the deployment of a crude nuclear force. Without a bomb, the leadership may fear a marked deterioration of military and bureaucratic morale. And, if civilian rule had yet to be restored in Pakistan, the ruling military faction might think that its failure to take those necessary steps to acquire Pakistani nuclear weapons would enhance the likelihood of its overthrow or at least greatly erode its internal position. Thus, notwithstanding the possible costs, failure to respond to India's bomb by seeking its own nuclear force would probably be estimated by Pakistan to entail too high domestic costs and external political risks.[29]

The emergence in the mid-1980s of nuclear arms racing between India and Pakistan, in turn, would generate both security and status-related pressures in Iran to resume earlier nuclear activities and to move to acquire a nuclear weapon capability. And, as is not improbable, had the Khomeini regime since been replaced, perhaps by a military-civilian government, these pressures would be all the greater.

Notwithstanding the current compatibility of many aspects of their security interests, the Iranian government of that day might conclude that once Pakistan had even a rudimentary nuclear weapon capability, Iran would have to match it. Simple balance of power calculations would be the most critical driving force. That is, prudence would be seen to suggest the rejection of a militarily weaker position in an anarchic international environment in which security and survival depended on self-help. Uncertainty about possible changes in the unstable Pakistani political situation would reinforce those dictates of prudence, for a future upheaval, it easily could be argued, might bring to power a government hostile to Iran. Though these motivations may appear insubstantial now, similar calculations have more than sufficed to support increased military spending and programmes in other situations.

Depending on the make-up of any future Iranian regime, status and prestige incentives for 'going nuclear' in the wake of Indian and Pakistani acquisition of nuclear weapons would also intensify. Though no longer committed to becoming the 'fifth great power', as was the Shah, a future Iranian government may still be reluctant to see India, let alone Pakistan, and not itself among the ranks of the nuclear powers. Such prestige considerations have already been a key factor in other countries' pursuit of nuclear weapons.

Though less acknowledged now than in the past, Pakistan's territorial integrity will continue to be important to Iran's security. Not only would Pakistan's dismemberment and the creation of an independent Pakistani Baluchistan increase unrest within Iranian Balu-

chistan, but the break-up of Pakistan would provide possible openings for greater Soviet involvement in the Gulf. In this context, India's emergence as a fully fledged nuclear weapon state could increase Iranian security incentives to follow suit. For a matching Iranian nuclear weapon capability could well be viewed in Tehran as a necessary precondition for exerting a restraining influence on a nuclearised India in its dealings with Pakistan. Similarly, such a matching capability could come to be regarded as necessary insurance and a bargaining buttress should events in Pakistan erupt into an Indian-Iranian confrontation.[30]

Taken together, if not sufficient separately, the preceding set of incentives could result in an Iranian decision to move closer towards a nuclear weapon capability. As a party to the Non Proliferation Treaty, Iran would probably withdraw from that treaty under Article X, citing a threat to its supreme national interests. Such a formal withdrawal, it would be hoped, would reduce the risk of a hostile external reaction. In addition, the regime might believe that Western dependence on Iranian oil exports could be used to bargain for a reduction of the scope of any sanctions. And paradoxically, the curtailment of the Iranian nuclear energy programme since the Shah as well as the diminished purchasing of advanced American conventional arms will have reduced Iranian vulnerability to these measures. In that sense, at least, and with a volatile and militant regime in Tehran, events since the fall of the Shah have adversely affected non-proliferation prospects by lessening potential Iranian disincentives to acquiring nuclear weapons.

Further, the possible fall-out of an Indian nuclear weapon programme in the mid-1980s would in all probability extend even beyond Iran. Specifically, Iran's decision to follow Pakistan and India into the nuclear age or to take steps in that direction could be expected to add to Iraqi incentives for 'going nuclear', as well as to reinforce greatly those of Saudi Arabia.

On the one hand, the long-standing rivalry between Iran and Iraq would provide a fertile climate for arguments that Iraq had to match Iran's nuclear weaponisation. These motives with their balance of power logic would be reinforced by domestic pressure, especially from military opinion within Iraq. Coming on top of those Israel-oriented incentives discussed below, the result is likely to be an Iraqi decision to acquire nuclear weapons.

On the other hand, given nascent Iranian and Iraqi nuclear weapon programmes, it is difficult to believe that Saudi Arabia would not seek to follow suit, although its technological weakness would probably force it either to purchase grey market nuclear mercenaries or assistance or to engage in grey market nuclear collaboration with another prospective proliferator to buttress its

ambitions. To recall, the latter collaboration could range from creation of a joint nuclear weapon programme with a politically compatible but more technically advanced prospective proliferator to a direct one-time purchase from a new proliferator of critical materials, know-how, expertise or other components. Particularly in light of Saudi financial support for Pakistan's purchase of a spent-fuel reprocessing plant, Pakistan could be one possible direct source of such supplies or even of a joint programme partner. Alternatively, Saudi Arabia could seek to purchase key nuclear weapon components and expertise from India whose leaders might be tempted to do so for financial reasons and pressured to do so by fear that Pakistan would complete the sale if they did not. Or, to reiterate, Saudi ties with South Korea and Taiwan could provide a framework for covert nuclear co-operation.

Thus, through a proliferation multiplier effect, the initial Indian decision to detonate a nuclear device could set in motion a chain reaction that might well lead during the 1980s to overt proliferation in the Persian Gulf. Of particular concern, moreover, is the fact that the initial elements of this proliferation chain are already unfolding. But even if events in South Asia take a different turn and the key linkages of this chain are severed, there is yet another route to Persian Gulf proliferation. This entails an Israeli-triggered proliferation chain encompassing Iraq, Iran and then others within the Persian Gulf. And here, too, the first links are already being joined.

The Middle East Path to Gulf Proliferation

Many observers have concluded that Israel already has covertly acquired one or more nuclear weapons, so-called 'bombs in the basement'. Even if that estimate is exaggerated, few doubt that since initiating its ascent of the nuclear ladder in the 1950s Israel has continually refined its nuclear weapon option to the point of now being within days of fully assembled nuclear weapons.[31] In the chain of reactions of several Arab states to Israel's status as a proto-nuclear weapon state are the seeds of more proliferation in the Middle East and beyond.

To begin, since the early 1970s Libya's Colonel Qadaffi has been attempting without success to acquire nuclear weapons. In 1973 he tried to purchase a nuclear weapon from the Chinese.[32] More recently he offered financial aid to Pakistan's nuclear programme in return for nuclear explosive material but apparently was turned down — if some Pakistani sources are to be believed.[33] These Libyan aspirations could be fulfilled in the 1980s, most likely either by ascent of the nuclear ladder or possibly by some type of nuclear grey marketeering.

There are reports that Iraq and Syria are also seeking nuclear

weapon capabilities.[34] Had that French-supplied research reactor not been sabotaged, its fuel (with reloads) would have provided sufficient weapon-grade uranium for several Iraqi bombs. Though under safeguards to deter misuse, Iraq well might have thought the cost of a hostile foreign reaction worth paying. Without that reactor fuel, Iraq will now have to seek to build up an indigenous programme for producing nuclear explosive material or grey market support to satisfy its nuclear weapon ambitions. And rumours of Iraqi interest in any plutonium separated by Brazil may be illustrative of how Iraq will proceed.[35] As for possible motivation, acquisition of a nuclear weapon would be a symbolic way to upset a Middle East *status quo* which already appears to have become increasingly hostile to Iraqi interests. Syria's recent efforts to acquire nuclear technology and reported assessment of the nuclear weapon option have been motivated partly out of concern with such recent Iraqi as well as Israeli activities.[36] But once again the apparent diplomatic utility of 'going nuclear' could also be playing a role as well.

Should any of these countries acquire nuclear weapons in the 1980s, others would probably intensify their efforts to follow suit. Particularly, either the prospect at that point of Iraqi acquisition of even a crude nuclear weapon capability or its actuality could readily trigger a matching Iranian effort. Past rivalries and disputes evaluated against the framework of an unstable Iraqi political system within which a sudden domestic change could bring back to power more radical anti-Iranian leaders would be sufficient incentive. Even for his successors the Shah's balance of power reasoning, that 'if ever a country [of this region] comes out and wants to acquire atomic weapons, Iran must also possess atomic weapons,'[37] may prove a surprisingly convincing guide to policy. Iraq's and Iran's emergence as nuclear weapon states would be likely then to trigger or greatly reinforce the incentives of others within the Persian Gulf.

Prospects for Stability in a Nuclearised Persian Gulf

Both greater regional instability and conflict are likely to follow the nuclearisation of the Persian Gulf. Mutual suspicion on the part of Iran, Iraq, Saudi Arabia and others within the Gulf about each other's nuclear weapon programme intentions would supply a new source of disputes. Moreover, to the extent that the key countries sought more than a simple, non-operational 'in the business bang', this effect would be exacerbated. In addition, the risk of local nuclear war could increase significantly. Much is likely to depend on the technical characteristics of these nuclear forces. At least initially,

some — for instance those of Iran, Iraq and Saudi Arabia — may justifiably be labelled 'serious but technically deficient' forces. In particular, these forces well might be characterised by a serious risk of unauthorised access to nuclear weapons by military factions or domestic dissidents, inadequate command and control systems and procedures, vulnerability to an opponent's first strike and so on. Given these deficiencies, these forces could have a high propensity to catalyse a nuclear war. Should a conflict or crisis arise, whether or not involving disputants solely within the Persian Gulf, the risk of inadvertent, miscalculated or unauthorised nuclear exchange due to these 'pre-emptive instabilities' may increase significantly.

Of equal, if not greater, importance, Persian Gulf nuclearisation may result in increased internal conflict within these new nuclear weapon states punctuated by incidences of the 'nuclear *coup d'état*'.[38] That is, with faulty protection against unauthorised access to the nuclear arsenal, nuclear weapons may be seized by an internal military faction or domestic group and used as bargaining levers during domestic upheavals or *coups*. Moreover, though there probably would be no intention to employ these weapons, such use cannot be precluded. Events might get out of hand or someone could act out of desperation. Further, in light of its past political instability, Iraq would appear especially vulnerable to such nuclear *coups* and upheavals. And now with the fall of the Shah, continuing domestic instability in Iran also appears a likely pattern; thus a struggle for succession to the Khomeini regime in the 1980s could pit against one another domestic forces which the Shah's strategy purposefully had kept divided and competitive. In that political environment control of the nascent nuclear arsenal could provide or be seen as likely to provide diffuse bargaining advantages.

Nor should the potential spill-over effects of such Persian Gulf nuclear *coups d'état* be missed. For example, assuming a domestic upheaval in a nuclear-armed Iraq which threatened to bring to power a far more radical leadership, Iran, Syria or even Israel might think it had no course but to intervene. Otherwise its failure either to destroy or remove the Iraqi nuclear arsenal or to support the more moderate leadership could result in nuclear weapons falling into 'the wrong hands'. Similar spill-over scenarios for other regional nuclear weapon states, as well as for those countries bordering the Persian Gulf in the Middle East and South Asia, could be delineated. Thus, the risk of nuclear *coups d'état* is of concern not only because it most typifies the possible increase of internal conflict within a nuclearised Persian Gulf but because the final result of such *coups* could well be regional intervention and conflict.

The change of regime in Iran thus appears likely to have had a

smaller impact on the prospects for proliferation in the Persian Gulf than might be thought at first glance. Rather, those prospects are most dependent on events emanating from outside the Gulf itself, events unconnected to the upheaval in Iran. To wit, if it proves impossible to head off a nuclear arms race in South Asia or to prevent Arab acquisition of nuclear weapons in response to Israel, a proliferation multiplier effect spilling over into this neighbouring region is not unlikely. The ultimate result could easily be greater internal conflict, local rivalry, and even the use of nuclear weapons.

Notes

1. Personal communication to author.
2. *Nuclear News* (Mid-February 1976), p. 55.
3. *Nucleonics Week*, 20 October 1977, pp. 13–14.
4. *New York Times*, 27 April 1975.
5. Personal communication to author.
6. *New York Times*, 6 January 1975; Paul Jabber, *A Nuclear Middle East: Infrastructure, Likely Military Postures and Prospects for Strategic Stability* (Center for Arms Control and International Security, University of California, Los Angeles, 1977), pp. 8–9.
7. *Financial Times* (London), 11 August 1979.
8. *Nuclear News*, October 1975; *Wall Street Journal*, 16 November 1976.
9. *New York Times*, 9 May 1979.
10. *Nuclear Engineering International* (July 1979).
11. Ibid.
12. Jabber, *Nuclear Middle East*, pp. 13–14.
13. *New York Times*, 9 April 1979.
14. *Nucleonics Week*, 16 June 1977, p. 2.
15. See Lewis A. Dunn, 'Nuclear Gray Marketeering', *International Security*, vol. 1, no. 3 (Winter 1977).
16. *Washington Post*, 2 May 1979.
17. *NRC Handelsblad* (New Rotterdam, Holland), 16 June 1979.
18. The ties with South Korea are typified by the presence of tens of thousands of South Korean construction workers in Saudi Arabia. And, apparently impressed by Taiwan's brand of anti-Communism, the Saudis have guaranteed low-priced oil shipments to that country since 1973. *New York Times*, 19 June 1978; *Washington Post*, 28 May 1979.
19. These countries already have access to some form of nuclear-capable delivery system, especially aircraft, so that as well is not a serious constraint on their ability to 'go nuclear'.
20. Shahram Chubin, 'Repercussions of the Crisis in Iran', *Survival* (May/June 1979), pp. 98–9.
21. See Paul F. Power, 'The Indo-American Nuclear Controversy', *Asian Survey*, vol. XIX, no. 6 (June 1979).
22. *Washington Post*, 8 December 1978.
23. *Washington Post*, 2 May 1979; *New York Times*, 12 August 1979.
24. *New York Times*, 20 April 1979 and 3 May 1979; *International Herald Tribune* (London), 21–22 April 1979.
25. Indira Gandhi, 'India and the World', *Foreign Affairs*, vol. 51, no. 1 (October 1972), pp. 65–78, p. 75.

26. For a well stated explanation which is equally an articulation of this quest for recognition, see Baldev Raj Nayar, 'Treat India Seriously', *Foreign Policy*, no. 18 (Spring 1975), pp. 133–54.

27. This was suggested to me by Onkar Marwah.

28. Ashok Kapur, 'India and the Atom', *Bulletin of the Atomic Scientists* (September 1974), p. 27; Onkar Marwah, 'India's Nuclear and Space Programs: Intent and Policy', *International Security*, vol. 2, no. 2 (Fall 1977), pp. 98–9.

29. Pakistani perceptions are reviewed in Anwar H. Syed, 'Pakistan's Security Program: a Bill of Constraints', *ORBIS*, vol. XIV, no. 4 (Winter 1973), pp. 952–74, esp. pp. 952–63. See also Alvin J. Cottrell, 'Pakistan: Internal Unrest and Military Weakness' in R.M. Burrell and Alvin J. Cottrell, *Iran, Afghanistan, Pakistan: Tensions and Dilemmas*, The Washington Papers, vol. 11, no. 20 (Sage Publications, Beverley Hills and London, 1974), pp. 53–68.

30. On Iran's interest in a Pakistani buffer state, see especially Shahram Chubin, 'Iran: Between the Arab West and the Asian East', *Survival* (July/August 1974), pp. 172–82, pp. 175–9; Cottrell, 'Pakistan', pp. 57–8; Sepehr Zabih, 'Iran Today', *Current History* (February 1974), pp. 66–9, p. 68.

31. See Robert E. Harkavy, *Spectre of a Middle East Holocaust: the Strategic and Diplomatic Implications of the Israeli Nuclear Weapons Program* Monograph Series in World Affairs, Volume Fourteen, Book Four (University of Denver, Graduate School of International Studies, Denver, 1977), pp. 3–19.

32. *New York Times*, 30 July 1977.

33. *New York Times*, 9 April 1979; *Washington Post*, 9 April 1979.

34. *New York Times*, 30 July 1977; *Financial Times* (London), 18 April 1978.

35. *The Middle East* (July 1979), p. 93.

36. *Nucleonics Week*, 23 February 1978, p. 2.

37. *Teheran Journal*, 16 September 1975.

38. See Lewis A. Dunn, 'The Nuclear Coup d'Etat', *The Journal of Strategic Studies*, vol. 1, no. 1 (May 1978).

6 THE IRANIAN MILITARY: POLITICAL SYMBOLISM VERSUS MILITARY USEFULNESS

Steven L. Canby

The fourfold increase in OPEC mandated oil prices in 1973 created the unusual situation whereby Iran could design and purchase a radically different military establishment than previously. Such a shift raises three questions: (1) the military's 'back-end' ability to absorb the new equipment and force increases; (2) the competitive 'sidewards' impact on the civil economy; and (3) 'front-end' military appropriateness.

Western countries have experienced and accommodated back-end and sidewards effects during periods of abrupt rearmament and mobilisation. However, few countries (including Meiji Japan) have ever attempted radical adoption of military structures that their own populace cannot produce, operate and maintain. The Shah attempted this. In addition his country was in the midst of a demanding social and industrial modernisation process. The back-end deficiencies in the Iranian military were well known.[1] The sidewards effect snowballed under the dual impact of over-ambitious economic planning and faltering oil revenues.[2] The third front-end question of military appropriateness has generally not been recognised, though it is a phenomenon that has historically debilitated many militaries, as for instance the French (and many others) in 1940. It is a question in the domain of general staffs and the interplay of tactical, operational and strategic factors.

There is a tendency throughout the world to equate sophistication in weapons with military usefulness and from there to military and political power. In point of fact, however, little relationship may exist between sophistication and usefulness and power may be more apparent than real. Except for the major powers who are fighting in a very special context (and it can be argued that even they sometimes unwittingly trap themselves into over-sophistication), it can be shown that weapons sophistication, while having a certain micro attractiveness, can in general be counter-productive.[3] Technological sophistication often substitutes for military thinking. Nowhere was this more apparent than in the Shah's armed forces. They were ill suited for all but one of their intended tasks: political symbolism.

Iranian Security: Problems and Solutions

Any Iranian government faces diverse military concerns such as military loyalty, internal security, suppression of separatism, border integrity, hegemony in the Persian Gulf, coastal protection, maritime presence, regional balance, air space protection and defence against Soviet intimidation and invasion. The Shah's solution was textbook American: armoured and mechanised divisions, an air force more sophisticated than any in Europe, and a pocket navy. This implies the Shah's military was not designed to cope with his specific concerns; and even more important, no attempt was made to address the special features of the Iranian context.

In particular, small powers such as Iran cannot fight the Soviets on their own terms. Instead, they must learn the peculiarities of the stronger opponent, identify weaknesses, and learn to take advantage of them. A 'relational' approach of this kind implies strategy and tactics which are shaped by the doctrine and character of the main enemy, the quality of his leadership and high command, the tactical skill of his field commanders and soldiers, his rigidities and flexibilities, his capacity to react under pressure, and so on. The aim is to apply specialised or localised strengths (for example terrain) against the identified weaknesses of the enemy in order to prevail by avoiding the direct clash of symmetrical capabilities. Iran violated these relational guidelines. The Shah could have learned much from the Israeli experience:

> Israel has realized that it cannot adopt foreign doctrines of the bigger powers which are incompatible with its material capabilities, political situation, and cultural milieu. It had to find its own solutions for its [own] problems. The reserve system, the weapons-acquisition and procurement processes, the logistical structure, and all other elements. . .had to be tailored to Israeli needs, and on occasion, had to be developed from scratch.[4]

The Former Strategy of Iran: Technological Superfluity

Although the goal of the Shah's strategy was to make Iran the leading military power of a region extending from the Levant to India and including the whole of Arabia, Iran's ground force deployments had an almost purely defensive and internal security character. In fact, three-fifths of its army was deployed near the Iraqi border, and half of the remainder was stationed in Tehran itself. Nearly all the armour was oriented towards the Iraqi border but the Shah's plans called for mechanising the remaining regular

infantry (three divisions and one independent brigade) including those forward deployed near the Soviet border (three brigades). The desolate eastern border facing Afghanistan was garrisoned by a single infantry brigade also to be mechanised. Only a single airborne and a Special Forces brigade were available for rapid reinforcement within Iran and also for use in the Gulf region. Except for the demonstration effect of his air and naval forces, the Shah actually would have had but a limited capability to intervene in areas not contiguous to his own borders.

The concentration of forces near the Iraqi frontier was consistent not only with the obvious need to face the Iraqi threat, but also with long-standing American notions on the containment of the Soviet threat. It was always believed that countries such as Iran could at best delay a Soviet invasion, until the arrival of large-scale US reinforcements. Accordingly, the decisive battle would then be fought on the central Iranian plateau, and Iranian forces deployed facing the Iraqi border would be readily available to fight there. In any case, it was believed that Iranian forces (and particularly armour) placed well forward on the Soviet frontier itself would only be lost early on, and could also be provocative in peacetime.

Strategically, this plan had a major flaw from the point of view of regional balance and hegemony. By implicitly leaving the entire northern rim open to Soviet occupation, the plan made it unlikely that Iran could really be an effective counterweight to Soviet-sponsored pressures like Iraqi incursions against Kuwait and the other Gulf states. A threatening Soviet posture on the common border would ensure the paralysis of the forces nominally deployed against Iraq and the two brigades of non-mechanised infantry which would have been the only Iranian forces capable of effectively delaying and defending in the mountains rimming northern Iran. Obviously some degree of Soviet coercion would be feasible with any pattern of deployment, but the dual-purpose deployment actually adopted by Iran provided the Russians with an almost mechanical tool for leverage. This of course, negated much of the Shah's purpose in building Iran into a major military power in its region.

The deployment plan was in any case based on two dubious assumptions: first, that Iranian armoured forces could actually wage war in the armour-mobile manner against Soviet forces, skilled as the latter are in true armoured manoeuvre, certainly as compared to Iranian forces. Second, the plan assumed timely and sufficient US reinforcements, in spite of the obvious difficulty of airlifting US forces (and particularly heavy units) to Iran.

But a defence of Iran cannot be based on armoured forces. Iranian armour cannot possibly match the general technological sophistication of Soviet forces,[5] nor their troop quality. Iranian armour can be

used to good effect in an attack *into* the Iraqi flatlands. In this mode, however, the Iranians do not require a homogeneous armour/mechanised force. As the Germans did in the Second World War and the Soviets do today with their Category III divisions, only the spearhead forces need be armoured and of high quality. The remainder can be a simpler infantry for consolidating the gains of the tank spearhead. But Iranian armour can *not* be used *defensively* against Iraqi mountain infantry invading into the Iranian mountains. A defence or delay in mountains cannot be mounted by armoured forces. It is too easily flanked and its line of retreat and sustenance cut by infantry, moving by foot or by helicopter.

An armoured defence in mountains is necessarily positional. Its positions may be strong, but the offensive actually has the advantage. To be sure, an attacker moving against a well entrenched defensive position is generally at a severe disadvantage; hence the well known 3:1 rule of thumb. But this is only true for *frontal* attacks against *alerted* defenders. The essence of tactics and of the operational art of warfare, however, is to subvert the defender's potential advantage, by two general means: stratagems and outflanking. The first depends upon deception and (then) surprise, the second upon opportune manoeuvre. The problem of a defender relying upon the firepower of fixed positions is typified by the dilemma he faces in countering outflanking moves. If he relies on well defended but isolated outposts, these will be vulnerable to turning movements across their line of retreat;[6] and the attacker need not attack them at all. To impose a disadvantageous frontal attack upon the enemy, the defender must therefore extend his lines; but in doing so, even a much more numerous defence force can then become over-extended. This then allows a competent attacker to develop stratagems to penetrate and separate the defender's extended line which can then be defeated in detail.

Because of the pervasive assumption that unassisted Iranian forces could not cope with the Soviet Union, the Shah placed very few of his troops near the Soviet border. Only small forces were readily available for forward defence and delay operations. It may be argued that the Shah could always have redeployed his forces from elsewhere for delay operations against the Soviet Army. But given the structure of forces, this was not in fact possible. Any attempt to use armour well forward would have led to the early loss of the expensive armoured/mechanised forces, and possibly also to the rapid disintegration of the entire army from induced panic. Armoured units cannot be used to defend or delay in mountains; they can only be deployed *behind them* to seal the exits. In mountain valleys their positions can be readily turned; and once cut off, armour cannot be extricated. Since armoured troops are psychol-

ogically unprepared for non-armoured combat, the result is inevitably defeat in detail and infectious panic.

For an armoured defence of the plateau to be workable, the Iranians would have had to delay significantly a Soviet advance across the northern mountains in order to gain time for the redeployment of their own armoured divisions arrayed on the Iraqi border and for US reinforcement to make up for a shortfall in forces. A delay in the mountains would have required light infantry. But subsequently the Iranians would have had to fight on the plateau and this would have required armour and not infantry, let alone of the light variety. The delay phase and defence of the central plateau would thus have required two very different force structures. Before the Shah's overthrow, Iran had four armoured divisions and was in the process of mechanising its four infantry divisions; this would have left only two non-armoured brigades. The problem thus arose that even if the Iranians could have fought in an armoured manner, they lacked the light forces necessary to gain the time to do so.

Assuming that the Iranians could have mounted the required delay operations by acquiring mountain infantry and by a correspondingly reduced emphasis on mechanisation, the outcome would then have depended on their ability to use tanks in combined-arms teams. It is questionable whether the Iranian Army had this capability; very few armies do. Many armies have tanks but few have the ability to co-ordinate tolerably well the branches of the combined-arms team. Even fewer have the mindset, tactics and communications needed to use armour in *fluid* manoeuvre, as opposed to semi-static setpiece movements. These capabilities would have been very difficult to acquire for the Shah's army. Aside from cultural factors, it was the Shah's practice to compartmentalise the branches of his armed forces, and to centralise control in his own hands. This, of course, reduced the chance of a military *coup d'état*, but it also militated against the joint training and co-ordination needed for true armoured warfare.

Accordingly, it is reasonable to assume that Iranian armoured forces would have been no match for Soviet armour once the latter reached the plateau. Iranian forces would probably have been deployed piecemeal, in separated blocking positions; consequently, they would have been quickly outflanked and defeated even by quantitatively inferior Soviet armoured forces. The situation could only have been stabilised by a rapid and large-scale US reinforcement. But until enough American forces arrived to form a significant share of the total allied force, any American units in the field would also have been jeopardised by the collapse of Iranian units on their flanks.

Finally, the conduct of a large-scale armoured defence on the

Iranian plateau would have been dependent upon Iraqi benevolence. Whether fluid or FEBA-like, a defence of the Iranian 'waist' would uncover the Iraqi border. This in turn would have created the possibility of Soviet passage through Iraq directly to the Gulf; the only rail LOC and the main road LOC from the outside to the plateau would thereby have been threatened.

In short, an armoured defence on the central plateau was not in fact militarily feasible. For the Iranians, American reinforcements were too uncertain, if only because of the NATO contingency. For the US, on the other hand, the option amounted to a foolish reinforcement race with the USSR. The time factor was not an independent variable; under the circumstances it was largely dependent on the quality of Soviet planning and the size of their forces committed. There was thus very little likelihood that the defence could have held long enough for significant US reinforcements to arrive. In the interim, US units in place would have been in great jeopardy. If the situation *did* stabilise because of US reinforcements, the United States would have been saddled with an unattractive protracted conflict in circumstances apt to favour Soviet rather than American persistence.

It can thus be concluded that for tactical, operational and strategic reasons, a defence based on expensive armoured forces is inappropriate for Iran. Armoured forces are suitable for attacking into the Iraqi flatlands, but not for defending against Iraq within Iran's own (mountainous) borders. But most important, a defence based on tank forces is not viable against the Soviets and leaves northern Iran hostage; the ensuing weakness vitiates the Iranian role in the regional power balance. Outside alliances do not materially offset this deficiency. For Iran's lesser military concerns, the armoured force structure is constricting. Armoured forces cannot be moved about readily and they lack the infantry content necessary for internal security, suppression of separatism, defence of the rough and thinly populated eastern border, and intervention among the Gulf states.

Similarly, it can be argued that Iran had no real need for a sophisticated navy and air force. Iran's vital importance as an oil exporter eliminates its need for an expensive 'blue-water' fleet, or an expensive air defence. And the shift from an armoured force to a defence based on agile mountain infantry eliminates much of the need for air defence of, and ground support for, ground forces. Thus strong political connections with the West simplify and even eliminate the most expensive and demanding (for Iran) military concerns. For the Western connection made it unnecessary for Iran to develop naval forces capable of coping with Soviet interference with Iranian seagoing commerce. Instead of trying to cope with this threat directly as it did, Iran actually needed only to control its own coastal waters and

to *circumvent* a Soviet naval blockade by transferring the burden of securing maritime access to the major Western powers — a task which the latter could not avoid, given their dependence on Persian Gulf oil. In this case, the linkage between Iran and the West would clearly work to the benefit of the Iranians.

The Western connection could also have been valuable in the other arena of high-cost technology — air defence, where the main purpose (other than the assertion of sovereignty over the air space) would have been to contain Soviet military intimidation by overflights and intrusions and to limit terror bombing. The Soviet Union could not exercise the upper range of intimidation (the threat of nuclear attack) for fear of the US reaction and the long-run effect upon world-wide nuclear proliferation. Nor could the Soviets have exercised the option of (non-nuclear) strategic bombing against Iranian cities without risking similar reactions. In any case, Soviet strategic bombing capability remains limited: Soviet fighter-bombers are still too restricted in payload and range to do more than mere terror bombing. A strategic bombing campaign would require Soviet use of vulnerable heavy bombers of 1950s vintage.

The Western connection was thus most important and most useful to Iran in dealing with those threats that required a high-technology military response,[7] while a shift to a relational military structure further reduced the need for sophisticated weaponry.

Relational Defence: More Defence for Less Money

It is apparent that Iran's lesser military concerns can only be satisfied by an infantry-oriented force structure. The question to be established in this section is whether such forces could satisfy the more demanding military concerns, the most demanding being a Soviet invasion. This can be of two forms: a whirlwind *coup de main* and the more traditional in-strength invasion. If Iran could design its defence to resist both a surprise attack and the more deliberate invasion, northern Iran would no longer be hostage to the Soviets, Soviet intimidation would have little credibility, and Iran could play a more meaningful role in a regional balance. It also follows that if the Iranians could defend in the mountains against the Soviets, a similar capability would exist against lesser opponents, even if they were more appropriately structured for mountain warfare. Nor in the postulated design would Iran lose the capability to attack into the Iraqi flatlands with armour.

In designing a defence strategy against a conventional ground invasion by the Soviet Union, Iran cannot possibly expect to match such an enemy in terms of numbers, technology or troop quality.

Any Iranian solution presupposing a contest of material resources, firepower and Lanchesterian-type attrition models is doomed to failure. Soviet reinforcement capabilities are simply too overwhelming. To be sure, the ground transport net leading to the Iranian plateau from the USSR is constraining and is also potentially vulnerable; but other outside lines of sustenance on to the Iranian plateau are even more constrained. In a direct clash of armour and firepower, the Soviet Union is bound to prevail. Only a *relational* defence, designed very specifically to exploit both terrain advantages and the particular weaknesses of the Soviet force structure can offer hope for success. For example, Soviet armour-mechanised forces are plainly unsuited for combat in high-mountain terrain, while Iranian forces could be specifically so designed. To be sure, mountain terrain will not favour a *positional* defence against a more fluid offence; but mountain barriers do allow scope for a combined defence featuring light infantry blocking and counter-stroke tactics, complemented by positional elements, and by armour in the rear, to cover exit routes from the mountains. While the Shah's planned all-armoured force structure could not have implemented a relational defence (and was not meant to do so), the latter remains a feasible (and much cheaper) alternative and it calls for a force structure largely made up of mountain-oriented infantry.

In addition to providing for a defence against an invasion in the conventional manner, Iran, as other countries directly threatened by the Soviet Union, must also cope with *coup de main* operations. Among other things, this calls for some armoured cavalry for deployment in the deep interior, near the nerve-centres that would be the primary targets of Soviet air-landed operations. In what follows, it is argued that a relational defence strategy for Iran would have to be based on the barrier on the northern mountains; no effective defence can be achieved by armoured combat in the plateau, or by an enclave strategy focused on the Zagros mountains: nor is a territorial ('people's war') defence a realistic proposition given Iran's separatist tendencies.

A Neglected Aspect of the Threat: Coup de Main *Operations*

In conventional assessments, a Soviet *coup de main* against Iran is sometimes dismissed as a piece of worst-case analysis, not worthy of serious consideration. Yet from the Soviet perspective a rapid, low-casualty *fait accompli* is obviously desirable. Soviet planners do not believe in the measured and constrained allocation of forces in their offensives. Their model is likely to be that of the Manchurian operation. This becomes apparent when it is realised that strategic surprise is not in itself a critical requirement.[8] All that is necessary is that a sufficient uncertainty of intentions be maintained to inhibit the

victim from taking necessary precautions for fear of provocation. The critical factor is the ability to mount a high tempo *coup de main*. This calls for tactical surprise as to the *location* and *manner* of attack.[9]

A *coup de main* can take many forms. While the separate techniques are known, or discernible, their combination is not. The aim of the *coup de main* is to create an overwhelming sense of helplessness by combining confusion measures in the interior with the image of irresistible military forces soon to arrive. While no contingency plan for the defence can be foolproof, it must at a minimum contain the following elements:

(1) media arrangements to dispel disruptive rumours;
(2) a police/gendarme network to protect key installations and personalities, and to contain saboteurs and small air-delivered enemy raiding parties;
(3) an air-defence system capable of inhibiting movement by air at least *after* the initial surprise (the denial of the first air penetrations is not a realistic goal);
(4) ground forces capable of pocketing and then destroying large air-delivered groupings; and
(5) a capability to delay the *movement* of the main follow-up forces, invading across the border.

The characteristic instrument of *coup de main* operations is not force *per se*, but rather the *illusion* of force contrived by the attacker, which can lead to a paralysis in the defender's ability and will to resist. The classic German Blitzkrieg of 1939–41 was based on such effects, and even a full mobilisation is not a sufficient response to the phenomenon. But whereas in Europe a *coup de main* is easy to mount due to the small distances involved and the dense transport net, Iran as a relatively large country with few good roads does have some natural protection. The northern rim of rugged mountains can be a military asset for Iran, *vis-à-vis* this threat. Notably the geography could make airborne assaults into the deep interior of Iran a risky proposition *if* a credible delay capability were at hand against the land invasion threat. The Soviets are unlikely to accept such a risk to their elite airborne-assault forces (it must always be remembered that Iran is only of secondary strategic importance to the USSR).

In the logical sequence of a defence against a *coup de main*, Iran's first requirement is to deal with airborne assaults intended to create disarray in the interior. This task is made easier by a good air defence. But air defence is *not* a prerequisite. An over-emphasis on air defence — as in Iran's case — can be counter-productive and

weaken overall defensive capabilities.

The effective counter to airborne assault then is not a foolproof air defence system but rather quick-reaction *armoured cavalry* forces on the ground, and the effective use of the rural gendarmerie system. The task of Soviet airborne assault forces would not just be to seize and defend their air LOC for subsequent reinforcement: the air-head would only be a means with which to launch immediate raiding operations — operations that cannot be launched directly from Soviet territory due to the distances involved. In other words, the air-landed force would immediately send out small raiding groups to disrupt the Iranian reaction to the whole Soviet offensive. A few helicopters and some light-armour vehicles would be available to the air-landed force; but motor-cycle troops might also play a role. (Used as the lead, disruptive elements of the (1939–42) German *Panzer* divisions, motor-cycle troops are now being increased in the divisional reconnaissance battalions of the Soviet Army.)

The value of light elements are often discounted, since they do not rate highly by the twin attrition criteria of firepower and survivability. But light elements would act as 'dragoons' relying on prompt manoeuvre rather than firepower. Their purpose would be to move quickly to points of logistic and tactical importance, in order to seize them before the defender begins to react. This would create the impression of omnipresence, spread confusion and demoralisation among the population and in the military command system, and would force the defender to attempt to dislodge the attackers, who would then enjoy the advantages of the tactical defensive.

Against such raids and seizing parties emanating from an air-head, gendarmes and armoured cavalry units are the most effective antidotes. The gendarme system offers the possibility of intercepting the smaller Soviet 'desant' groupings *en route*, as well as of guarding key installations (and personalities) before their arrival. Armoured cavalry would provide an integrated combat team ready to deal quickly with larger 'desant' groupings. Cavalry units can be moved rapidly, they have both heavy elements for 'fixing' and light elements for 'working in' and — most important for a quick response — their battle team is already organically integrated. This avoids the time-consuming preparation that would be required to organise a co-ordinated response from distinct tank and infantry formations, oriented as these are for more formal methods of combat.

If the tentacles emanating from the air-head can be defeated, the air-head itself loses much of its purpose. In fact, because of the enemy's obvious need to achieve a ground-force link-up and do so quickly, the air-head then becomes an enemy liability. For reasons of prestige and morale, the attacker must strive to sustain it, but if the

ground link-up is prevented, and its tentacles destroyed, the air-head can no longer consolidate the initial gains — even if the necessary reinforcements can be delivered. The air-head thus becomes a trap and resource-sink for the attacker, which in due course can be annihilated.

Even though an air-head can be made useless by destroying its tentacles, it would still be necessary to contain it. In the case of Iran, Soviet air-heads would have to appear in the proximity of one of the crucial nerve-centres, and perhaps only Tehran itself is truly critical. An airfield is obviously a desirable target which would facilitate subsequent operations, but it is *not* critical. An airfield is only essential for the deployment of large forces and to sustain their prolonged combat and that of course would be the very negation of the purpose of the operation. Nevertheless it is obvious that one or more airfields could be one of the many initial objectives of the air-delivered element of a *coup de main*.

In combating the air-head force itself, the two important ingredients are rapid containment and air defence. At least for the initial sealing off and containment of the air-delivered force, armoured cavalry again offers the best prospects. It has two advantages over regular line formations. First, it is more suited to sealing off the approaches from the air-head because its combat teams are organic. Second, it is designed in part for the security role, which would be important since air-delivered forces would seek to 'flow' around containing blocking positions. Thereafter, and particularly if it is deemed desirable to destroy the airhead by close combat, regular line formations would be more suitable.

Air defence is an important element in containing and then eventually destroying an air-head. But neither fighters nor sophisticated SAM missiles are likely to be as important as automatic cannon in close proximity. These weapons could also serve as a major element in the containment and destruction of the air-delivered forces in ground combat. The aim of air defence against the air-head is obviously to prevent reinforcement and sustainment. In practice, this would be accomplished by making air landing and parachuting prohibitively expensive. This is a task for low-level air defences, still the recognised domain of the gun. Enemy aircraft flying above the gun envelope at 10,000 feet or so cannot reinforce the air-head and neither could they provide close support. Fighter-bombers flying in from their USSR bases would necessarily be payload limited; nor would 'smart' ordnance solve Soviet problems since seeker devices would not work well against the low-contrast containment forces (including optically directed AAA).

It can of course be argued that a good air defence alone could be sufficient to deter and/or prevent a *coup de main* through the

destruction of Soviet aircraft *en route* to the objective. This argument, however, overlooks the implications of surprise: (1) sophisticated air defences are not robust against one-time countermeasures; (2) air defence is a 'sieve' system, which takes time to tighten; and (3) the loss of surprise as far as the air-defence control centre is concerned does not automatically translate into a loss of surprise on the ground. The time lost in identifying the threat and disseminating warnings depends on command linkages between the ground and the air forces (in the Shah's time, *he* was the link), on the readiness of the ground forces concerned, and on the ability of the ground formations to form and deploy responsive battle groups. In the case of Iran, an air defence warning would not have translated into a prompt and effective warning on the ground, except during special periods of alert (which of course would themselves be cause for the postponement of a *coup de main*). Thus even if the air defence system worked well from the first, a *coup de main* could still succeed. On the other hand, it must be recognised that sophisticated air defences would set a higher 'entry' price for the Soviet Union, and they add important risk factors for them.

The worth of Iran's sophisticated air defences against the *coup de main* threat thus depends on two unknowns: the 'entry' price necessary to deter a Soviet attack, and the robustness (= degraded effectiveness) of the system itself, once the Soviets use all the countermeasures available to them. The combat effectiveness of highly centralised and electronically sophisticated air defence systems remains an open question. The US Air Force pins high hopes on such systems, but it is noteworthy that the RAF and Luftwaffe do not, at least in the cluttered conditions of European warfare.[10] An air conflict in Iran would of course be much less cluttered, but on the other hand much less scientific expertise has been devoted to the special problems of electronic-wave propagation in Iran. It is of course well known that the Soviets themselves are very active in electronic warfare while the Iranians are merely the recipients of electronic packages produced by others. Hence in a conflict against the Soviets, there is a very high possibility that Iran's air defence system would be temporarily neutralised.

The final aspect of a defence against a *coup de main* is the delay of the main invasion forces advancing to link up with the airborne 'desant' elements. Against this threat, Iran has a distinct geographical advantage as compared to most other neighbours of the Soviet Union. The Iranian plateau is protected by a high mountain barrier (on the Iranian side of the border) which cannot be outflanked, and must therefore be penetrated. While there is legitimate argument over the merits of a mountain-based defence in protracted war, it is nevertheless clear that such a system is unambiguously

advantageous as far as delaying actions are concerned. In mountains, defensive forces can significantly delay much larger and heavier forces through demolitions and direct combat.

Across the 1,000 miles of the Iranian-USSR border as the crow flies, there are only ten routes suitable for vehicular passage (including the Herat route from Afghanistan). The immediate Soviet aim in a *coup de main* offensive would be to reach and reinforce the air-heads before they are destroyed by the Iranians. Since they are not short of units, and since they would seek to maximise the appearance of strength, as well as to exploit allocation errors by the defenders (virtually inevitable in the confusion of a surprise attack), the Soviets would undoubtedly send forces into most if not all the approaches. Some of these forces may be small, but all will be of sufficient strength to push through (and secure) thinly guarded routes until exploitation reinforcements can be brought up.

The Soviet problem is to devise an approach that would allow them to get one or more land columns into position to relieve quickly the pressure on their air-heads in the Tehran region. The normal Soviet procedure for the advance would have two shortcomings: (1) their standard formation of march columns with advance guards and recce well forward is likely to act as a triggering mechanism for Iranian demolitions guards; and (2) the land routes from the USSR itself are too long, and worse, they converge into only three approaches while still at some distance from Tehran. The Soviet advance is therefore more likely to be opened by light precursor forces which would try to seize key points along several of the major routes. Lead columns aimed at Tehran might originate on the Caspian Sea, for example by amphibious landings at Chalus and Babol Sar. While most of the land approaches from the USSR converge most unfavourably from the Soviet point of view, a landing at Babol Sar would offer the prospect of divergent paths. This of course means that reinforcement echelons could be swiftly assigned to the routes that become open, thus multiplying the effectiveness of the invasion force as a whole. This consideration, plus the fact that the initial link-up column has as much symbolic as substantive military importance, enhances the attractiveness of the Chalus and Babol Sar approaches, even though this implies that the initial relief forces would be relatively small. As in any surprise operation, the safety of these columns would not derive from their own capabilities, but rather from the confusion generated by the offensive as a whole.

In this scheme, the light precursor elements would seize bridges, tunnels, passes and the like, in a variety of exercised 'special operations'. The key is obviously to secure critical locations before local gendarmes and army forces can react. In many situations, this would

merely mean overcoming unsupported guards; in some cases, it might simply mean posting Soviet guards before the gendarmes can move into their designated guard-posts.

While precursor forces can take various forms, the most obvious possibility is the use of helicopter-borne infantry. Helicopters are of course ideal for landing small 'desant' parties and flat-deck ships in the Caspian Sea could serve as platforms for helicopters. Helicopters based in the USSR itself would not have the range and carrying capacity to support the necessary surprise operations in the mountains bordering the Caspian Sea. Flat-decked ships (which could be provided without loss of technical surprise) would solve this problem and would also be of use to support main forces moving along the coastal routes before turning into the mountains. Helicopter-delivered motor-cycle troops[11] could also be landed with 'desant' teams to move quickly by road, thus adding another dimension to surprise.

The basic elements of the envisaged 'precursor' units are already present in the reconnaissance companies of the tank and rifle regiments, and in the reconnaissance battalions of the Soviet tank and motorised rifle divisions. The reconnaissance battalion TO&E includes a motor-cycle platoon and also a long-range recce company composed of small teams. The recce company of the regiments has a small motor-cycle scout section. In each unit, additional personnel are available to use motor cycles in scout-like operations. Thus in the standard formations engaged in an offensive against Iran, a considerable 'desant' force would be organic in any case and would be available for precursor duties even if all airborne and air-assault divisions and brigades would be assigned their primary role, and would thus be unavailable. Out of the 24 divisions in the three military districts adjacent to Iran, Soviet TO&Es would call for as many as 1,700 motor-cycle troops and 120 long-range reconnaissance teams. Even if divided among the ten routes into Iran, this force alone could be seriously disruptive. And, of course, there is no reason why this 'base' could not be enlarged by the temporary assignment of recce units from the remaining 137 Soviet tank/MR divisions.[12]

In addition to heli-borne elements, more old-fashioned means could also be used to seize and secure the entry points of the road net. A subversive network of Soviet sympathisers could be created to be activated when needed. Dissident ethnic groups could be exploited on a similar basis. Even commercial trucking (and tourism) from the USSR could be used for 'Trojan Horse' operations. Those means already provide ample opportunities for peacetime route reconnaissance. The techniques these groups would use would be quite similar to those of the 'desant' groups. The difference is mainly in the means

used to arrive on location; one uses visible transport; the other covert means.

There is a tendency to think that the solution to surprise lies in high technology, such as sophisticated early warning radars and remote surveillance techniques. It is obvious, however, that these would only be effective against high-signature modes of transport. Such technology is certainly not effective against approach and seizure techniques which rely on cover and deception. It would therefore be unwise to spend large sums on sophisticated technology. Moreover, there is also reason to doubt the effectiveness of high-technology systems in countering heli-borne penetrations. These systems can provide warning, but it is doubtful if such warning can be used in a meaningful military way, except to alert guards already in place. High-performance aircraft are not suitable for intercepting low-flying helicopters, and the early warning likely to be available (assuming Soviet deception to mask the operation until the launch of the aircraft) would not be sufficient to deploy ground air defences suitable for use against low-flying helicopters. One hardware solution of possible relevance in this case would be the anti-helicopter helicopter, such as the Soviet Union appears to have developed.

While a Hind-like anti-helicopter aircraft and early warning systems would be useful, the only robust solution to the problem remains a force of guards capable of rising to the demands of the situation. Certainly, only guards can cope with 'desant' operations based upon cover and deception. A system capable of coping with these threats is also capable of dealing with clandestine penetrations of the border areas. There would be some merit in dealing with the highly visible, high-technology end of the spectrum, but the real focus must be on organisational solutions, which are both cheap (as far as the hardware goes) and also militarily robust and versatile.

The solution to the precursor problem must be found in the *gendarmerie* system of internal security, which Iran has (or will at least have again, when order is restored). Iran's 74,000 gendarmes are now distributed in many platoon company and battalion size outposts. Their primary duty is the maintenance of law and order. But their dispersion also gives the central government (1) a useful means of keeping remote localities under observation, (2) a quick-reaction guard capability, and (3) a means of rapid alert and mobilisation. The gendarme system thus offers a ready-made framework for disrupting desant-type operations. Their numbers in the mountain belt would have to be augmented, and they would also require prompt reinforcement with army units. Assuming manpower of tolerable quality, given automatic weapons and suitable training, gendarmes could make the quick seizure of bridges, tunnels and passes rather difficult for the Soviets. Equally important, the

gendarmes have an inherent interception capability: to detect suspicious activity, to ambush the smaller 'desant' groupings in transit, and even to use REDEYE-type missiles as well as automatic weapons against over-flying helicopters.

Forward Defence and the Northern Mountain Barrier

It is widely believed that as between defence and offence, the defence is the stronger of the two. It would therefore follow that if the barrier strength of mountains is added, a mountain defence must become stronger still. This conclusion is often confirmed by eclectic examples drawn from history, such as the German success in containing the allies on the Italian Front in the Second World War. In point of fact, however, while the defence generally has had the advantage in *open* terrain, the reverse is true in a mountain defence. The Italian example is atypical for several reasons: the German command, and notably Field Marshal Kesselring, was exceptionally gifted; the Allied command, for its part, was cautious and schematic.[13] But most important — the Germans practised *fluid* tactics behind screening entrenchments rather than positional warfare, and most armies are quite incapable of fluid tactics in the compartmentalised terrain typical of mountain environments.

Clausewitz wrote that the defence is in general stronger than the offence, but not in the case of positional mountain warfare in which the attacker was strongly favoured. Clausewitz's line of argument is particularly important because it also contains a seminal outline of German Blitzkrieg theory.[14] Clausewitz recognised that in open terrain the defence has the emplacement advantage of static blocking positions and, more important, it has the advantage of the second move against an attacker's flanking attempts. In mountain warfare, the defender does have a greater static-position advantage *but the compartmentalisation of terrain forecloses the more important second advantage*. The attacker can therefore turn even strong positions, cut off their line of sustenance and retreat, and thus induce a general collapse unless the defence is distinctly more agile. In Clausewitz's thinking, the use of reserves and the relative tempo of manoeuvre are the keys to both the defence and offence. The firepower criterion is only germane to the defence of static positions, but in mountain warfare especially the attacker's objective is to outflank while the defender's proper reaction is to counter-flank. It is therefore wrong to size the forces required by means of firepower scores, or by the capacities that avenues of approach could support. It is not the *lead* echelons of the attacking force that the defender must worry about, but rather the follow-up echelons, ready to peel off in order to turn blocking positions.

It is thus to be concluded that the key to victory in mountain

warfare offensive or defensive is the same as in armoured manoeuvre: echeloned reserves, fluidity and an advantage in the relative tempo of action. It is clear that such a defence cannot possibly be based upon a combination of armoured formations and static entrenchments. Armoured formations can have no role in the mountain barrier itself. A light-armour cavalry screen and even some armour may be needed *forward* of the Northern Mountain barrier for peacetime presence functions. But these forces would almost certainly be unable to extricate themselves through the barrier and they would add little to the defence of the barrier itself. The more important role of armour would be to close the exits from the barrier to the central plateau, as well as in countering *coup de main* airborne landings.

On the other hand, a fluid and agile defence of the mountain barrier can be highly effective. This would still need some positional elements. Indeed fortifications are the indispensable anchor of the defence and they are critical to prevent surprise seizures of the road net. *What has to be avoided is the notion that positional defences in themselves can provide a meaningful defence.* Even major fortifications provide only a temporary check of the attacker's advance. The only *permanent* solution is a mobile reserve which can rapidly counter the attacker's flanking movements and attack his rear echelons. But this is by definition difficult to achieve in the compartmentalised terrain typical of mountains, though helicopter transport offers an apparent solution. On the other hand, if the defence is seduced into a cordon deployment by the apparent advantages of positional defences, it will lack the wherewithal to form sufficient reserves, and the troops will become not only too sedentary but also too dependent upon roads.

As in armoured warfare, positional defences should be viewed as no more than pivot points for the manoeuvre of fluid offensive forces. There are, however, four distinct differences.

First, the troops of the pivot positions must mentally condition themselves for eventual exfiltration, with the abandonment of whatever heavy weapons are used in the fortified position themselves. Only light weapons can be saved in each move; the attempt to bring back more will simply lead to the loss of the garrison and perhaps also to the spread of infectious panic. In today's world, the mortars and automatic weapons and light cannon which should be the mainstay of positional defences are relatively cheap. For the price of a single high-performance aircraft, one can buy hundreds of automatic weapons or mortars. Guns and such no longer have their former symbolic significance; they should now be viewed as expendable and their replacements can be pre-stored in successive fall-back positions. But garrisons cannot be sacrificed so easily, if

only because of the effect on the morale of the remaining troops.

Second, flanking counter-attacks do not obtain their effect from their size or firepower. Attacks may be co-ordinated with artillery, but platoon and company-size operations will be the norm. Their usefulness springs from the manner in which they attack: they must materialise out of the mountains to launch hit-and-run attacks. What protects them, and also magnifies their morale impact, is deception and surprise. Against enemy infantry attempting to out-flank pivot positions, their task would be to turn the enemy's own flank. This can serve to lengthen the life of the defensive position, but their main task is to attack the attacker's own deep flanks — in the area where the enemy's soft supporting forces are to be found. In mountain warfare, particularly against an opponent attempting to push armoured columns through, successful attacks against the supply units in the rear can quickly lead to the demise of the (heavy) main forces.

Third, if fluid forces are to obtain the requisite degree of deception and surprise in mountain warfare, they require logistical autonomy. Only this can allow them to 'spring' from the mountains, to acquire an aura of destructive evanescence. To the Western military mind, keyed to the heavy tonnage requirements of front-line troops, it seems that such autonomy cannot be achieved. It is, however, a virtue of genuine mountain infantry forces that they are inherently adaptable to logistical stringency. Theirs is not the world of mechanisation with its attendant demands on fuel, spare parts and maintenance. Stealth and the exploitation of terrain compart-mentalisation substitute for heavy firepower support. Thus their main logistical requirements are for the human body, for small arms and grenades, and for a limited amount of mortar support. These are all cheap, *long-life*, small-tonnage items. Most usefully, they could be cached throughout the mountains in peacetime. This would allow mountain troops to move about freely and it would reduce enemy efforts in coping with them to mere shadow-boxing.

Fourth, and finally, there is the wide opportunity for demolitions, a dimension to warfare that is not so important in the flat lands. The unusually large numbers of bridges and tunnels which can be demolished obviously enhances the defence and adds to its delay potential. Particularly important for the long-term defence is that demolitions could exhaust an enemy's engineering capacity, particu-larly if the obstacle gaps are well defended, and the enemy's engineers are made top-priority targets. Attacks against the enemy engineering troops are particularly effective given the unusually heavy demand of armoured forces for engineer support in mountain terrain, and the relative scarcity of engineers in the overall Soviet force-structure. For the Iranians this would have favourable

implications: either a Soviet underestimation of their engineering requirements, possibly leading to a Soviet inability to advance; or an increment to deterrence should the Soviets fully appreciate the heavy burden that would be placed on their scarce engineering units.

The complication of a Soviet advance is one goal of demolitions, but the latter could also contribute to the direct destruction of Soviet formations; in fact they could lead to demoralising Soviet defeats *if combined with fluid counter-attacks.* Troops advancing in mountain terrain, whether they are attacking positional defences or moving administratively in columns, are vulnerable to all the special stratagems of mountain warfare: induced rockfalls, contrived landslides and avalanches, and the like. These are not only deadly, but also terrifying to the survivors. Demolitions are also useful in congesting enemy units whose advance is temporarily blocked and in isolating them from assistance (by demolitions on their rear) while they are attacked suddenly and unexpectedly even by relatively small units. Non-combat supporting units are particularly vulnerable to such moves. The only protection against this tactic would be unpalatable for the Russians: the widespread dissipation, and stringing out of the relatively scarce Soviet infantry strength along the routes of march. The first weakens the Soviet thrust; the second presents an opportunity for the Iranians to focus their attacks on the Soviet infantry in the knowledge that once this component is destroyed, the remainder of the Soviet force becomes particularly vulnerable.

The *operational goals* of a mountain strategy are similar to those of a fluid armoured defence: to induce exhaustion and over-extension so as to set the stage for annihilation. The enemy should be allowed to exhaust himself by pursuing apparent success; he is to be enticed into attacking and flanking successive road blocks in depth, the final block being the defender's own armoured force covering the exits from the northern mountains. During this mountain phase, the defender's priority targets would be the Soviet engineers and infantry and *not* the armour. The main task of the position elements would be to prevent the breaching of obstacles, with a priority of effort against engineering equipment. It should also attempt to target as much of the enemy infantry as possible, ideally while still mounted or in the process of dismounting from its vehicles. But this obviously requires tactical surprise which is hardly possible if the attacker is properly preceded by light reconnaissance elements. The enemy response preferred by the defence would obviously be the frontal assault; it must be prepared for this eventuality, but much more likely the attacks will be flanking moves on the ground or by helicopter. Such attacks are vulnerable during the movement phase; it is the task of the fluid component of the defence to catch the enemy

in disadvantageous positions while his forces are still *en route.*

The exhaustion and over-extension of the enemy, desirable goals in their own right, also fit into the larger scheme of manoeuvre warfare by being the pre-conditions for decisive counter-strokes. Soviet tank and artillery — the prime targets in flatland warfare — should *not* in general be primary targets at this stage. These components are not well suited to mountain combat, and their destruction will come about anyway if the overall operation is successful. So long as these elements are intact, the enemy command will be impelled to press on. *Hence a narrow-minded focus upon tank-destruction would misdirect the efforts of the defending force.*

The critical decisions for the defence are to choose the timing and form of the counter-stroke. Should it be early and in the form of many repetitive attacks against the enemy column as it inches forward? Or should the defence content itself with pinpricking the columns by stand-off fire and demolitions until the invading formation is over-extended? From the purely military viewpoint, the latter course is more decisive. On the other hand, the political leadership may want to report early victories both for its own prestige and to bolster public morale.

For Iranians against Russians, it is important that successive blocking positions be able to hold for significant periods. It may therefore be necessary to initiate smaller counter-strokes from the beginning. The risk is that the defence will also be exhausting itself, by having to attack into (or in the proximity of) the armoured vanguard itself instead of softer targets. If the blocking positions are effective, the defence is much better served by biding its time. If the Iranian units prove of high mettle, the same overall strategy can be unfolded more rapidly and to more effect. In this case the positional element need mount little more than a delay, withdraws on to the flanks, and allows the mechanised column to move into a large-scale trap. In both cases the counter-strokes then aim at the isolation and destruction of the soft element in the enemy supply columns and of his logistical support in general. Thereafter, the columns could be progressively split up, and defeated in detail.

In counter-stroke operations, the over-extension of the enemy is an intermediate goal whose aim is to force the vanguard to out-run its logistical support (or suspend the advance) and also to stretch out the combat units beyond their range of mutual support. In mountain warfare, over-extension accomplishes three purposes. *First* (and most important), it allows the defender to cut off the hard mechanised elements by attacking soft spots in an extended chain of communications. If this line of supply is not quickly restored, the logistically demanding armoured elements may find themselves immobilised, or even forced to abandon their equipment. *Second,*

the enemy's *effective* frontage becomes 'U' shaped. His efforts are focused upon moving through the blocking positions but in doing so, his flanks become progressively weaker and more vulnerable to counter-attack. Time is then tactically on the side of the defenders.

The *third* reason why it is advantageous to aim for over-extension is because it allows the defender to launch remote and distracting attacks by fire prior to the counter-stroke. In particular, stand-off tactics can then be used to snipe at the attacker's infantry and engineers and generally wear down his strength. Demolitions can be repeatedly set, bringing down landslides and the like upon the attacker's columns. PGMs could also add a new dimension by picking off the enemy's vehicles at long range, in particular engineer and infantry-laden vehicles. One possibility would be to form amorphous observation and firing parties which would control and designate terminally guided rockets in an indirect fire mode. This system is based on the ability of the observation and firing parties to remain masked at some distance from the target area. If discovered, they would fade into the terrain, if possible ambushing the pursuing enemy infantry. A variant could consist of line-of-sight laser-designator parties a kilometre or so from the target with firing parties three to five kilometres from the target. Each laser spotter could target three or four vehicles per minute until the column's vehicles masked themselves with smoke or moved off the road to seek cover — if available. Low-powered radios and M-10 (mortar) plotting boards would be quite sufficient for command and control.[15] Another variant would consist of spotting parties with high-powered radios for communication to firing parties far away on the flanks, or well behind the line.[16]

It should by now be apparent that in mountain warfare the determining factor is the relative *fluidity* of the forces, i.e. *the very antithesis of the presumed advantages of positional defences as measured in relative firepower scores.* In flatland warfare, Clausewitz assumed equal fluidity; in the mountains he presumed an asymmetry in favour of the attacker. But assuming that the temptation of adopting a positional defence with the resulting operational passivity is avoided, the defender can then obtain two distinct advantages from his prior possession of the terrain.

First, he can cache supplies and thereby increase the logistical autonomy and fluidity of his mountain infantry. *Second*, by prior study of the terrain and skilful use of adjoining positional defences, the defender can shape in his favour the framework in which military operations will take place.

A small country arrayed against a large one can have a third advantage. It can design its military forces for the specific circumstances in which it would have to fight while large countries cannot.

Their forces and thinking are usually shaped to meet their own more dangerous adversaries. The Soviet Army has a force structure completely unsuited for combat in mountains. Soviet 'motorised rifle' divisions are not infantry at all, but in fact heavy armoured forces, a reality not much affected by the apparent training of parts of some of these units in mountain warfare. *It is unlikely that the Soviet Army would re-shape its standardised forces to match an Iranian shift to an agile mountain defence.* The Russians tend to be contemptuous of their smaller neighbours and expect them to oblige their might by a swift collapse. Further, Soviet planners tend to think in terms of handling smaller countries by *coup de main* operations. (The present Soviet force structure is very well suited for such operations). Additionally, even if Soviet planners were aware of a possible problem, it would be difficult for them to change their force structure. An Iranian change in force structure would be unlikely to generate a sufficient impulse for responsive Soviet change until a moment of crisis, when structural change would no longer be feasible.

Finally, structural change to accommodate local conditions is incompatible with the Soviet mobilisation system. The Soviet Army is oriented towards combat in open terrain against NATO and the Chinese. The cutting edge of their forces are the Category I divisions while their Category II and III divisions are follow-up and space-consolidation forces. Specialising a force for Iran would disrupt this scheme. Special Category I (and back-up Category II and III) divisions would have to be organised for Iranian (and similar) contingencies. This increases costs while reducing reserves suitable for NATO and Chinese contingencies. From the Soviet viewpoint, the latter two can threaten the security of the state while a certain degree of force inefficiency in a secondary theatre such as Iran does not.

It is for these reasons that there is the possibility that even a small power such as Iran could defend itself alone against the Soviet giant. But this is only possible if Iran can specifically organise its defence to take advantage of specific Soviet weaknesses, and of the contextual conditions. It should also be noted that a mountain-defence orientation is compatible with Iran's remaining security requirements. A force structure oriented for mountain warfare is cheaper, more easily redeployed and more effective in *defending* Iran's borders with Iraq and Afghanistan than the present flatland force structure. Iran would still need some armoured forces to implement this option and thus it would still retain a capability to launch punitive expeditions into the Iraqi flatlands.

There is finally the issue of internal security. A modern army with sophisticated weapons is much less suited for population control than a force consisting primarily of light infantry. In armoured

forces too much of the structure is support-oriented while the combat forces themselves are far too heavy in firepower and far too destructive for effective population control. (For the Shah, there was the special factor that the more sophisticated the force, the larger the proportion of potentially disaffected urban recruits within it, and the smaller the proportion of potentially more loyal rural recruits.) Most governments in politically unstable countries find it necessary to fractionalise their military forces and this makes it very difficult, as noted earlier, to exercise armoured forces in the proper combined-arms manner. However, a structure oriented towards mountain warfare is compatible with command compartmentalisation.

How Much for Tactical Air Power?

Iran will always require some air defence to affirm sovereignty, to protect its air space, and to inhibit intimidating terror- and counter-value bombing. Deterrence against countries like Iraq require Iran to have similar counter-value capabilities. But capabilities of this order do not require an air force of the size and sophistication of a RAF or Luftwaffe. A 'Belgian' or 'Dutch' air force would suffice for these purposes. If an even larger air force is nevertheless desired, consideration should be given to updated versions of such *strafing* aircraft as A-1s and Mustangs. These aircraft are cheap, easy to maintain, and could be particularly effective in Iranian conditions. They can assist the army in the lesser military concerns and in mountains where sophisticated high performance aircraft would not be effective. If dispersed on small strips on the deep flanks of likely Soviet approaches, they could also be effectively used against rearward supply columns and the precursor elements which precede and broaden a Soviet advance.

Above this minimum, the amount and type of air power is very much a function of the strategy and type of ground force. For example, an armoured defence on the central plateau allows full play for tactical air power. The armour must be protected and the enemy's own armour and resupply made vulnerable to air attack. A defence of the Zagros Mountains would reduce the need for friendly air defence while increasing the scope for offensive air against an invader stretched across poorly netted mountains and plateaus. On the other hand, defence of the Zagros also entails abandonment of more than half of the supporting air base structure. A territorial defence by definition requires no air support at all. But, as already suggested, these are not interesting strategies. What needs to be investigated is the requirement for tactical air power in support of

the postulated strategy of forward defence in the northern mountain barrier with agile mountain infantry.

The Role of Offensive Air in Mountain Warfare

It is widely believed that offensive air power is very effective in mountains. This thinking is derived from the belief that offensive air is a powerful instrument of interdiction and that its effect can be enhanced by the restrictive nature of mountainous terrain. But as is so often the case in conventional warfare, what is apparently true tactically is often not true operationally.

The problem with offensive air is that it does not relate well to the needs of a fluid mountain defence against a predominantly mechanised Soviet attacker. The object of the attacker is to push his columns through the barrier on to the plateau. The object of *delay* — in which offensive air could indeed be useful — is to cause the enemy to expend time and effort against relatively small forces by placing repeated blocks against the head of the columns. The tasks of air power are then to reinforce by fire the blocking positions, and more important, to provide its own delay by interdicting the road-bound columns. On the other hand, since even small forces can delay very large forces in mountains, it is also apparent there is a trade-off between prior preparation and offensive air. Prior preparation is obviously the more cost-effective; on the other hand, it is also to be recognised that unanticipated situations will arise even in the best case and that tactical air can be instrumental in coping with such situations until ground forces can be suitably deployed.

Delay in mountains by the nature of its dependence on static blocking positions falls into the category of a passive defence. This is a correct use for small forces, for only by relying on positional defences (and demolitions) can such a force impose a significant delay. The defender's main problem is the timing of his withdrawal to successive positions. Early withdrawal undercuts the purpose of the delay; late withdrawal leads to the possible loss of the forward forces to flanking movements by the enemy. But a defence in mountains should not be passive at all, and its purpose is no longer delay but rather the annihilation of the enemy's entire force. The focus is now no longer on the attacker's spearheading lead units but on the force as a *whole*, as explained above.

Offensive air can provide close air support and battlefield and supply interdiction. Close air support can be used to complement positional blocking defences; but it is not essential. The attacker is best stopped by an obstacle covered by fire. The tasks of firepower are to prevent the enemy from repairing or bypassing the breach and from flanking the defending force. Tactical air power is ill suited for these tasks. Its firepower lacks the continuity of a force on the

ground for covering an obstacle and for countering local flanking efforts. It can also be circumvented by night movement. Even with good visibility, spotting and targeting the enemy's flanking infantry requires a forward air controller, while weapons delivery and weapon effectiveness are degraded by the steep slopes and cover provided by the terrain.

In mountain warfare interdiction loses its normal content. In flatlands warfare, the purpose of supply interdiction is to reduce the *size* of the force that can be *sustained* in combat. Air power cannot be expected to completely stop the movement of supplies, but rather to impede and cut the supply tonnage to some fraction of full capacity. In practical terms, since even a thin transport grid can support army groups (for example the Eastern Front in the Second World War) and a modest road an army (for example Africa Corps and British Eighth Army), supply interdiction reduces supportable forces from 'armies' to 'corps'. Once the Soviet attacker has broken on to the plateau (or is attacking into the Zagros), this reduction is an important consideration, particularly given the 'reinforcement race' with US deployments. But in the mountains itself, road capacity is not a constraining factor. The force being fought in the latter case is only that at the very tip of the column — basically a battalion-sized task force. It is only this 'battalion' tip that will be demanding the full array of supporting services. In breaking through the mountains, the entire force stretched in 'administrative' column behind the tip battalion can be no more than 1–2 divisions. This means that the force itself will be operating well within the full tonnage capacity of the communication line. Thus supply interdiction is not likely to have a meaningful impact until *after* the break-out on to the central plateau when larger forces must be assembled for sustained combat.

Battlefield interdiction is also unlikely to be very useful in a mountain defence, given a fluid infantry defence against a predominantly mechanised Soviet attacker. Battlefield interdiction against an attacker has two functions: disruption and isolation while the attrition of enemy elements is a by-product. For a positional defender, the purpose of battlefield interdiction is to disrupt the attacker's planning and co-ordination and the flow of reinforcements necessary to exploit any success. The two subordinate aims are to reduce the effectiveness of enemy forces and to prevent the enemy from transforming tactical into operational success, thus in effect easing the front-line demands on friendly ground forces and providing time for positioning them advantageously. As in flatlands combat, tactical air power can be a useful complement for this. The difference is in scale: in flatlands, large forces are involved; there is thus not only a large pay-off in disrupting co-ordination and the flow of reinforcements, but considerable firepower must be

mounted and this can generally be done *only* by air power. In the mountains, however, the forces involved are smaller, co-ordination is accordingly simpler and presents less opportunity for disruption, while on the other hand even small groups of defenders can prevent the conversion of tactical into operational success.

For counter-attacks, the purpose of battlefield interdiction is again disruption and isolation, but with the emphasis upon isolation of the battlefield, physically and psychologically. The mountainous terrain reduces this task to the creation of blocks on the road so that column segments can be isolated from mutual support and attacked in detail. Tactical air power is suitable for this task; but it is done better by ground roadblocks and demolitions to provide continuity in the block and to create the impression for the offence that his forces are surrounded and attacked by overwhelming force. Tactical air cannot be used in direct support of the counter-attack itself, because of the difficulty of co-ordination and the proximity of the defender's attacking troops themselves to the effects of the air-delivered munitions.

In short, as long as there is an asymmetry between a fluid (infantry) defence and an attacker basically structured for armoured combat, there is little need for offensive air in support of the mountain infantry. The ground forces do not really need the supplemental firepower it provides. The problem of the ground forces is to cope with stratagems and the flanking movements — and for these tactical air offers little help. Similarly the disruption power of tactical air yields little that the ground forces cannot provide themselves.

Disruption itself is also less meaningful. It is to be recalled that its two benefits are to reduce the effectiveness of assaults and to impede the flow of exploiting reinforcements. In mountains, defending (as opposed to delaying) positional forces are not significantly helped by reducing the effectiveness of frontal assaults against them, while small units can be positioned in depth to block successful penetrations along the road axis.

It should be emphasised, however, that this conclusion is dependent upon an asymmetry, whereby the defence is designed for its special conditions while the attacker is not. If the attacker were to modify his forces, reducing their armour content to acquire mountain infantry, tactical air power (and helicopters) would then assume more importance due to the resulting compartmentalisation of the defending forces now compelled to distribute themselves across the mountain frontage. But it has also to be reiterated that the Soviets are unlikely to convert their forces in this manner. If they were to restructure, it would follow that while a better and more lengthy resistance could still be offered in the mountains, Iran would

be eventually overwhelmed by the weight of Soviet superiority. In that case, the best that Iran could expect would be a *conditional* defence, i.e. a defence based on *relative* performance in the hope that the offensive would become too costly and slow for the Soviet Union. A conditional defence implies a protracted capability, and this too rules out the possibility of meaningful role for Iranian tactical airpower. Against the Soviet Air Force, the IIAF would most likely be too fragile and would be lost early, thereby compromising a defence dependent upon tactical air power. American air power could of course fill the void, but it should be noted that active US participation would undercut a conditional defence since the Soviet Union would not be dissuaded by the embarrassment and cost of becoming bogged down in a secondary theatre against a putatively third-rate opponent.

Weapon Effectiveness in High Mountains

The discussion above was based on the general nature of mountain warfare and the role of tactical air power in the overall scheme of battle; in other words, air power was treated as more than just air-delivered firepower. In contradiction to the above is the view that air power is so destructive that its firepower alone justifies its existence. This raises the question of the destructiveness of air power in mountains. Targets can be classified on a spectrum of difficulty: the most difficult are small and dispersed infantry groups infiltrating in mountain terrain; the easiest targets are high-contrast positional defences. In between the two are bridges and vehicles on the roadnet. It is obvious that tactical air cannot cope with the first set of targets. It is equally obvious that tactical air can readily target high-contrast positional defences. The critical target category, however, is the roadnet. It is against this category of targets that tactical air is generally claimed to be highly effective.

There is some question, however, whether tactical air can in fact attack effectively vehicles or 'bridges' in highly mountainous terrain. In this type of terrain the roadnet and vehicles assume high contrast. But their location is generally masked by the terrain itself, thus seriously complicating target acquisition and weapons delivery. Where the road twists and turns, there is a problem of intervisibility in targeting vehicles. But the more general problem is that a road running inside the folds of the mountains can readily be masked by wafting smoke from generators or pots, without affecting visibility on the road itself. Even light masking can defeat target acquisition from high-flying aircraft which seek to deliver optically guided PGMs. (Similarly IR-guided PGMs can be defeated by decoy hot spots set along the road.).

Low-flying aircraft can, of course, fly right into the haze and

acquire their targets. But it is obvious that jet aircraft can only do this at considerable peril in reduced visibility. As it is, 'F-4s' are not manoeuvrable enough to fly in close mountains; and 'A-10s' lack the thrust to 'stand-on-tail'. In addition to flying into the mountainside, pilots must also be cautious of flying into a cross-fire from automatic guns and heat-seeking missiles shooting *down* from the ridges along the flight approaches to the more obvious targets.

Finally, there is the question of munitions for low-flying aircraft. As against tanks, only the A-10 has a suitable weapon with its 30mm GAU-8 cannon. Tanks and other armoured vehicles can be protected against cluster submunitions by 'chicken-wire' detonating rigs on top. Conventional 'iron bombs' have CEPs too large to be useful against armoured vehicles, but they can be useful for starting rockslides and the like.

In combination, these factors complicate the usefulness of offensive air. In particular, against heavy combat formations with considerable organic air defence, tactical air power may be effectively neutralised. Offensive air would be most effective against the attacker's softer logistic formations which generally lack escorting air defences. This, however, amounts to supply interdiction and implies that the Soviets would already have crossed the mountain barrier to reach the central plateau.

Air Defence in Mountain Warfare

While offensive air may not be of particular value for a defender, it is nevertheless true that air defence is of paramount importance. This follows from two differences between the attacker and the fluid defender: the defender's need to base his blocks on readily targetable positional defences and the attacker's need for helicopters to mount flanking movements and to cope with the defender's fluid counter-attacks (nourished by pre-positioned caches). This, however, is primarily a *ground* air-defence requirement. A fighter air defence provides a psychological uplift to ground troops and is sometimes instrumental in maintaining their resolve. Fighters, however, cannot cope with terrain-hugging helicopters and an Iranian fighter defence would in time be worn down by the superior weight of Soviet air power. It is therefore fortunate that the air defence requirement can be handled with light automatic cannon and with missiles of the REDEYE variety. For the positional defender the main task of air defence (as always) is to enforce stand-off distances in order to degrade further delivery accuracy (already degraded by terrain conditions). Large numbers of randomly dispersed small patrols with machine-guns and REDEYE missiles could make helicopter operations impractical given the degraded performance of the helicopter at high mountain altitudes and their inability to overfly the light

missile and light gun weapons envelope.

Unlike the positional elements, the defender's fluid infantry groupings are not vulnerable to tactical air power. The major threat to these groupings is from enemy infantry delivered by helicopter after their presence or intent has been discovered or disclosed.

There is, finally, a limited air defence requirement for the LOCs. This requirement is alleviated by the general difficulty of targeting where roads run within deep folds of the terrain, and also by the attacker's desire to capture the road more or less intact. In an agile mountain defence, dependence on the roadnet is minimised to begin with by the prior caching of supplies and by the prior readiness to withdraw cross-country, when the position becomes flanked and made militarily untenable (the main enemy tactic). In general, the defender will be more interested in allowing the LOC to deteriorate. His strategy calls for allowing the attacker to push into the depth of the mountains, absorbing his engineering capability and troop labour effort on a partially destroyed LOC, while withdrawing on to the flanks in preparation for later counter-attacks against the flanks and rear of an over-extended mechanised force.

In Conclusion: The Theory of the Relational Force Structure

While not widely recognised, the demand for large, complex and advanced high-technology weapons is a choice derived from that country's method of warfare. Even with identical estimates of the magnitude, composition and immedicacy of the threat, sharply different weapon requirements may arise, depending on the place of the chosen operational method in the attrition/manoeuvre spectrum.

In attrition methods of warfare, the enemy is treated as a mere array of targets; the goal is then progressively to reduce this enemy by administering upon him firepower of sufficient volume and accuracy. While tactics are undemanding and the results quite reliable — so long as the enemy chooses to operate in conveniently targetable mass formations — attrition methods generate a demand for combat aircraft with large payloads, battle tanks and much artillery. Above all, this method of warfare absolutely requires either superior technology or a net *matériel* advantage overall, hence the much alluded trade-off between force quantity and technological quality.

The attrition approach is familiar, for it is the style of warfare adopted by the Western Allies in the First and Second World Wars and taught and spread in their military asssitance and weapons-selling programmes. In fact, it may not always be realised that there is an alternative in the relational/manoeuvre approach. Associated

primarily with the German Army at various times and more recently with the Israeli armed forces as well as with irregular guerrilla forces, the starting-point of relational/manoeuvre methods is the assumption of material inferiority. Therefore the aim must be to exploit identified shortcomings in the enemy array and any locale advantages in order to win by manoeuvre; instead of applying strength to main strength as in attrition methods, localised and specialised strengths are to be applied to the identified weak points of an enemy array which may be greatly superior overall, in an inventory sense and even in troop quality. (Hence the paradox that the best regiments are British but the best armies are German.)

While attrition methods tend to be standardised and homogeneous, relational/manoeuvre must be place-specific, responsive to the cultural milieu, and highly attuned to terrain factors. As a result, a relational/manoeuvre response to any given threat generally results in a greatly reduced demand for large, complex, advanced high-technology weapons, as compared to an attrition response to the identical threat. For example, a relational/manoeuvre defence of Iran *vis-a-vis* the Soviet Union can be based on the use of agile mountain infantry to hold the northern barrier of mountains, instead of armour-mechanised forces held back to fight it out on the plateau. And given the first choice, thin air defences can suffice, since mountain infantry would not present many lucrative targets to enemy air attack; by contrast, once primary reliance is placed on armour-mechanised forces, a thick and sophisticated air defence is essential, since such forces present high-contrast targets of high individual value.

Needless to say, territorial defence against the Soviet Union is by no means the only perceived security need in Iran's defence planning. But it is the most demanding in terms of size, technology and cost, and it illustrates the potential of relational/manoeuvre methods in making possible a defence just as effective — and often much more resilient — even while generating a greatly reduced demand for expensive and provocative high-technology weapons.

Notes

1. See for instance *US Military Sales to Iran*, a Staff Report to the Subcommittee on Foreign Assistance of the Committee on Foreign Relations, US Senate, July 1976.

2. For a discussion, see Theodore Moran, 'Iranian Defense Expenditures and the Social Crisis', *International Security* (Winter 1978), pp. 178–92.

3. See for instance, Jack N. Merritt and Pierre Sprey, 'Negative Marginal Returns in Weapons Acquisition' in R. Head and E. Rokke, *American Defense Policy* (Johns Hopkins University Press, Baltimore and London, 1976); and Steven Canby, 'American Defense Policy: the Problem is Not More Money', *AEI Foreign Policy and Defense Review* (American Enterprise Institute for Public Policy Research,

Washington, DC, April 1979).

4. Michael Handel, *Israeli's Political-Military Doctrine*, Occasional Papers in International Affairs, Number 30 (Harvard University, July 1973), p. ii.

5. The Shah's forces were acquiring many highly visible items of sophisticated weaponry, but specific items do not equate to across-the-board sophistication in firepower, control, electronic warfare, etc. It is to be particularly noted that some of Iran's sophisticated weaponry would have been quite ineffective against the Soviets without the immediate reprogramming aid of American EW specialists.

6. The very poor showing of the Indians against the (outnumbered) Chinese in the Himalayan War can be explained by this phenomenon. The Indians nominally held the advantage of entrenched positions, but their overall defence was static while the Chinese operation was fluid.

7. In regard to other threats from the USSR, it was much less useful. For example, the best long-term insurance against the threat of direct Soviet support of insurgency would be a diplomatic balance and such economic linkages as the Soviet importation of Iranian gas and, later, oil.

8. Historically true strategic surprise is a rare phenomenon. The Soviets were well aware of German preparations in Spring 1941, as were the Japanese of the Russian in the summer of 1945. For an excellent discussion of the broader implications of the Manchurian model, see J. Despres *et al., Timely Lessons of History: the Manchurian Model for Soviet Strategy* (RAND, July 1976).

9. For instance, the Belgians knew in 1914, and in 1940, that the Germans had to capture the Liege fortresses quickly; yet in both cases the Germans successfully executed *coup de main* operations.

10. For an elaboration, see Steven L. Canby, 'Tactical Airpower in Armored Warfare'.

11. As much as one-third of the infantry element of German Panzer divisions in 1939–41 were motor-cycle troops. It is to be noted that these troops were mainly used to seek paths around points of resistance: they usually fought mounted but on occasion dismounted, as in the turning of the British positions on Mount Olympus in 1941.

12. *The Military Balance 1978–1979* (The International Institute for Strategic Studies, London, 1978).

13. For a polite but lucid critique of Allied performance, see Field Marshal Kesselring and General of Cavalry Westphal, *Questions Regarding the German Strategy During the Italian Campaign* (US Army Historical Series, MS. B-270, undated).

14. It is interesting to note that Austro-German mountain tactics in the First World War closely resembled the future 'Hutier' and subsequent German armour tactics. The parallel is most vivid in E. Rommel's *Infanterie Greift an*. As early as 1915 Lt. Rommel of the Wurttemberg Mountain Battalion was using tactics that would later be called 'Hutier' tactics. The similarity of mountain tactics with what later came to be known as Blitzkrieg tactics can also be been in Clausewitz's discussion of the defence in mountains; *On War*, Book Six, Chapters 15–17, pp. 417–33 in the Howard/Paret edition (Princeton University Press, Princeton, 1972).

15. Components for such systems are already available. Since US military services have no requirement for this approach, its potential has been ignored, and the US RDT&E community has been unwilling to invest in component integration. For example, the anti-tank rocket could be a standard 2.75-inch rocket with an anti-tank warhead and a terminal seeker. Rocket and warhead make up the system's bulk and weight. Large numbers of the rockets/warheads could be economically cached in the mountains; the high-cost seekers would be kept by the firing party.

16. With the use of the US Navy's SMARTROC, a 500 lb warhead could be fired from simple guide rails or commercial trucks at a distance of 26 kms for modest cost. Time lags would not allow this system to be fired at moving columns, but SMARTROCs could be fired against fixed targets to block roads at weak links (passes, culverts etc.). A second SMARTROC with an APAM warhead could then be fired at vehicles queuing up at the block.

7 ARMS TRANSFERS, INDIGENOUS DEFENCE PRODUCTION AND DEPENDENCY: THE CASE OF IRAN

Stephanie G. Neuman

Introduction

Manufacturing defence items is not a cost-effective enterprise for most countries. Few countries in the world, whether more or less developed, can produce military technology as economically as can the United States and a few European states. Why then are more and more governments clamouring for licensing rights to produce arms indigenously rather than buying arms directly from major suppliers?

This chapter analyses the relationship between the military industrial capabilities of Third World countries and their dependency on the states which supply them with arms. It focuses on the tension between the foreign policy goal of independence, which leads many governments to establish a domestic military industry, and the internal constraints which prevent them from achieving their goal. It discusses the kinds of trade-offs Third World defence planners must make between economic factors and perceived security needs.

Two main hypotheses serve as the organisational framework for this chapter.

(1) Arms transfers to less industrialised countries (LICs) initiate an evolutionary progression toward the indigenous production of military technology.
(2) Dependency for Third World states is no less a product of indigenous defence production than of arms transfers.

Although these hypotheses are formulated in general terms to apply to all Third World states, the Iranian Air Force (IIAF) in particular is used for the purpose of illustration. Iran was chosen for several reasons. First, during the 1960s and 1970s the Iranian military acquired a large number of arms which spanned the spectrum of sophistication, from simple to advanced. This is particularly true for the Iranian Air Force.

Second, Iran received arms transfers for a relatively long period of time, from a variety of suppliers, under a number of different financial agreements. Thus Iran, because of its relatively long and

diversified history of arms transfers, provides a unique research opportunity to study the relationship between economics and defence within the military sector over time. If there are common elements in the defence planning of Third World leaders, this chapter assumes they are most likely to appear in the policies of those countries which have had extensive procurement experience.

The Escalator of Development

Hypothesis 1: *arms transfers to less industrialised countries initiate an evolutionary progression towards the indigenous production of military technology.*

It is hypothesised here that arms transfers and the indigenous production[1] of military equipment are part of the same development continuum. As the first shipment of military equipment arrives, an 'escalator of development'[2] begins with the *support of* imported technology and ends with the *production of* components and/or total weapons systems. The stages of this developmental process, although varying widely in time intervals across countries, are so intimately connected to the arms transfer process that the dividing line between indigenous production and arms transfers is often indistinguishable.

Beginning with the delivery of the first weapon system, the sequence follows a predictable, although not necessarily orderly, pattern. It is characterised at each stage by increasing recipient participation:

(1) servicing and repair of imported weapons systems;
(2) overhaul of imported weapons systems;
(3) weapon system assembly under licence;
(4) fabrication of simple components under licence;
(5) co-production of weapon systems under licence;
(6) indigenous design and development.

The first two maintenance phases of the developmental sequence are generally the most difficult and time-consuming to achieve in any service. The US Air Force had identified three different levels of technical expertise necessary for the operation and maintenance of defence equipment.[3]

(1) *Base or Organisational Level Maintenance.* This, the most elementary kind of maintenance, is essentially a servicing operation (such as gassing, lubrication, etc.) which includes making minor repairs by removing and replacing spare parts. Work is performed

on, or in direct support of, an aircraft on the flight line or in the inspection dock.

(2) *Intermediate Level Maintenance.* At this level, work is performed on aircraft systems, subsystems and components which have been removed from the aircraft and brought to a shop (generally on the base). This is synonymous with 'off aircraft' maintenance and is often referred to as the 'shop' or 'field shop' part of 'base maintenance'. Mechanics at this level are trained to do minor internal repairs of components (for example hydraulic cylinders) which can be pulled apart if necessary, certain parts replaced, and then reassembled.

(3) *Depot Level Maintenance.* The most complicated repairs are done at this level, generally off base at a designated military or commercial facility where components are sent for complete teardown and reassembly. This kind of maintenance encompasses the major inspection and repair as necessary (IRAN) maintenance functions which include overhaul and modification of the airframe, the aircraft system, subsystem and components. It involves the repair of equipment that cannot be done at the base level and requires a higher technological skill than either of the other two forms of maintenance. As one industry employee commented: 'Once you learn how to tear down and reassemble an aircraft, the capability for domestic production of aerospace components is relatively easy to acquire.'

This ascending order of fundamental maintenance tasks suggests a close connection between support and production, and the kinds of basic skills necessary to master these capabilities. The limits to growth of the military sector in any particular country are clearly dictated by the number of complexity of systems to be supported or produced, and the technical skill level of the population. But the size of a country and its level of industrial development are equally important factors. In all military services, the decision to establish a depot-level facility is determined not only by the skilled manpower resources available, but by economies of scale. In many smaller countries, the force structure of the military cannot sustain a depot-level industry. In the air force, for example, there are often not enough planes to keep a depot busy, so it is more economical to have major repairs done elsewhere. For most developing countries, major repairs are sent to depot-level shops in either the US or Europe. Only Israel and Taiwan are known to have depots of their own.

To comprehend any country's technical ability and its potential for indigenous military production, one would need information about the level of competency in many skills throughout the military sector, and some concept of the existing industrial infrastructure.

The simple 6 × 6 matrix below includes only some of the data categories needed to estimate domestic military-industrial capabilities.[4] The 'escalator of development' can be thought of as operating across weapons systems, services, time and countries. Thus the aircraft industry in country X can be at one level of development while naval production is at another. The reverse kind of production capability may be characteristic of country Y. The discussion of Iran (below) will elaborate on the number of human and material resources necessary to acquire an indigenous production capability in one military service. The matrix included here is designed to help the reader graphically imagine the vast resources required for an 'across the board' capability in the other military services as well. Little wonder so few countries have been able to attain it.

The Iranian Air Force

The Iranian Air Force, between 1964 and 1978 was confronted with the task of modernising rapidly. Although the large number of arms transferred to Iran, their degree of sophistication and the short time span within which the deliveries were made is not typical for most less industrialised countries, the fact that so much happened so quickly affords us a condensed, or telescoped, glimpse of the stages of development within the military.

For all practical purposes, the modernisation of the Iranian Air Force began with the delivery of 106 F-5A/B supersonic jet fighters in 1964. Between 1968 and 1970, 39 more F-5s arrived — a total of 145 F-5s to be absorbed within seven years.[5] In addition to these transfers, 10 C-130 transports were delivered between 1968 and 1970[6] and 36 US F-4 fighter jets (the workhorse of the USAF in Vietnam and the state-of-the-art air system at that time).[7] By 1978, there were 64 C-130s in Iran's inventory, and 20 Boeing 707 and 747 to supplement the transport force. Iran had received over 200 F-4s;[8] 169 of the latest model F-5E/F; and 78 F-14A fighter jets,[9] the latter one of the most sophisticated aircraft in the US inventory, which the US Navy was still having problems absorbing. Thus, collapsed into a fourteen-year period, and not counting training planes, helicopters and other smaller air force programmes, the IIAF was tasked with the maintenance and operation of seven major weapons systems approximating 500 aircraft (see Figure 7.1).

To accommodate this equipment, the IIAF had grown to a strength of 100,000 men, with operational units which included 10 fighter-bomber squadrons, 10 ground attack squadrons, 1 reconnaissance squadron, 1 tanker squadron, 4 medium and 4 light transport squadrons and 5 surface-to-air squadrons.[10] By 1978, arms transfers from the US had made the IIAF inventory one of the most modern in the world.