Measuring Inequality

A Methodological Handbook

Philip B. Coulter

MEASURING INEQUALITY

MEASURING INEQUALITY

A METHODOLOGICAL HANDBOOK

Philip B. Coulter

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

To Mignon

First published 1989 by Westview Press

Published 2018 by Routledge
52 Vanderbilt Avenue, New York, NY 10017
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

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

Library of Congress Cataloging-in-Publication Data
Coulter, Philip B.
Measuring inequality : a methodological handbook / by Philip B. Coulter.
p. cm.
Bibliography: p.
Includes index.
ISBN 0-8133-7726-9
1. Equality. 2. Proportional representation. 3. Income distribution. 4. Social classes. I. Title.
JC575.C68 1989
323.42—dc19 89-5442
CIP

ISBN 13: 978-0-367-01301-1 (hbk)

CONTENTS

TABLES AND FIGURES

TABLES

FIGURES

PREFACE AND ACKNOWLEDGMENTS

The impetus to write this book grew out of curiosity and frustration. For a research project in which I was involved, I wanted to select an appropriate index to measure inequality, so I searched for a book that comprehensively reviewed the available indexes, identified their operational similarities and differences, and clarified their theoretical underpinnings. Discovering that no such book existed, I became increasingly frustrated and curious.

It became evident that I would have to undertake my own systematic review of the literature, presumably in my own discipline, in order to identify the alternative measures and choose an appropriate one on the basis of proper theoretical and methodological criteria. This effort led to additional frustrating discoveries. First, I encountered a bewildering abundance of inequality indexes—well over fifty distinguishable measures. Second, my review of the methodological literature on inequality measurement took me through the issues of literally scores of professional journals in five academic disciplines—economics, geography, political science, sociology, and statistics. Third, although I found some cross-disciplinary referencing of inequality measures, by and large each discipline's inequality measurement remained insulated from that of other disciplines.

Fourth, with few exceptions theory and technique remained separate. Methodologists who devised the measures seldom grounded them thoroughly in theory. Similarly, persons who wrote about inequality theory usually did not actually invent measures nor discuss their theoretical properties. And no comprehensive theoretical discussion could be found in the literature; inequality theory lay in bits and pieces in numerous journal articles. Fifth, the many researchers who used inequality measures to analyze actual data usually appeared to be unaware of either the theoretical or technical properties of their chosen index. Most chose an index on the basis of habit and convenience and made no effort to justify their choice.

What began as a decision to spend an afternoon reading up on a few things in one of the books on the subject turned out to be an extraordinary professional adventure over a period of several years. When it became patently clear that the multidisciplinary field of inequality measurement stood in a confusing state of disorganization, I decided to do something about it. This decision led me to create a new measure, the index of inequity, and to write this book. *Measuring Inequality: A Methodological Handbook* is the result of my attempt to organize

the field, across several social science disciplines, in terms of its mathematical foundations, theoretical properties, and technical characteristics.

Because the field contains an astonishing array of unique indexes, establishing taxonomic order with reference to these three sets of fundamentals seemed to be the highest priority. The book's intended purpose, however, is very practical. Now, presumably, researchers who want to conceptualize inequality more skillfully and choose, use, and interpret an appropriate inequality index on the basis of theoretical and technical criteria can do so more easily.

I am grateful for the encouragement, advice, and help of numerous friends and colleagues, who read all or parts of the manuscript. In particular, I profited enormously from the help offered by Rick Hofferbert of the State University of New York at Binghamton; Dorothy James of American University; Majeed Alsakafi, Charles Leathers, and Martha Griffith of the University of Alabama; Chong-Hwa Leeper; and an anonymous reviewer for Westview Press. Lennie Schmandt converted the manuscript from ordinary typing to camera-ready with precision and dispatch. In addition, I want to thank the University of Alabama for providing me with a one-year administrative leave, which gave me the time to read, think, and write almost without interruption.

I also want to thank the Institute for Social Science Research at the University of Alabama for its support of this project and my colleagues at ISSR for their friendship and professionalism. The generous contributions of all these individuals and institutions made the book possible and better. Finally, however, I accept full responsibility for the book's flaws.

Philip B. Coulter

1

DISTRIBUTION IN SOCIAL SCIENCE

Distribution is one of the most important concepts in social science. It has been a fundamental concern in several academic disciplines for well over a century, and some contemporary social scientists argue that their disciplines actually originated in early philosophers' efforts to explain the origins of inequality in the distribution of valued resources. Naturally, social statisticians have devoted considerable effort to measuring this important concept.

The purpose of this book is to encourage the proper study of inequality. It explains the multidisciplinary inequality theory and careful conceptualization necessary to inequality measurement. It discusses specific techniques to operationalize inequality and stresses both their underlying mathematical logic and their statistical procedures. Most critically, it emphasizes the importance of carefully linking theory and measurement, conceptual definition and operational definition. What are the alternatives in conceptualizing inequality? How should we conceptualize inequality, given a specific research problem? Once we have appropriately conceptualized inequality, exactly how should we measure it? This book provides a virtual encyclopedia of the logic, procedures, performance characteristics, interpretation, and applications of inequality measures. It seeks to integrate the principles of inequality theory and statistical measurement techniques from several disciplines and, thereby, to facilitate more and better analysis of inequality.

This introductory chapter focuses on the importance, definition, and use of measures of inequality. The next section argues that distribution has been a major concern of several social science disciplines practically from their origins as disciplines one or more centuries ago. Then, the discussion turns to some basic definitions and standard analytical tools appropriate for measuring distribution in social science. Chapter one concludes with an overview of the remaining chapters and alerts the reader to the logic underlying their organization.

THE IMPORTANCE OF DISTRIBUTION

Distribution has been a perennial concern of political scientists, sociologists, and economists virtually since their origins. All three academic disciplines include reference to distribution in their definitions.

Probably the first scholar to examine inequality empirically and critically, Aristotle, in about 350 B.C., proclaimed that "in all states there may be distinguished three parts, or classes, of the citizen body—the very rich; the very poor; and the middle class which forms the mean" (Barker, 1958:180). Justice, he argued, "is found in the distribution of honors, of material goods, or of anything else that can be divided among those who have a share in the political system" (Ostwalt, 1962:116). Whether contemporary political scientists define political science as the study of "changes in the shape and composition of value distributions in a society" (Lasswell, 1935), "the authoritative allocation of values" (Easton, 1981:129), or the similarities and differences in the uneven distribution of political resources and political influence among members of a political system (Dahl, 1984:50-51; see also Brams, 1968), distribution and inequality are central concepts.

Among sociologists, Dahrendorf argued that "the first questions asked by sociology" were "why is there inequality among men? Where do its causes lie? Can it be reduced, or even abolished altogether?" (1968:152). He further claimed that the entire history of sociological thought could be written in terms of attempts to answer these questions. Since Marx (1963 and 1976) and Weber (Gerth and Mills, 1946), social stratification and class conflict based on economic class, social status, and political power have preoccupied contemporary sociologists. In fact, Blau argued that sociological theory "centers attention on the *distribution* of people among different positions either because they are members of different groups or because they differ in hierarchical status" (1977:1).

Classical economist Ricardo wrote that income distribution is the "principal problem" in economics and that economics should be devoted to "enquiry into the laws which determine the division of the produce of industry among the classes who concur in its formation" (Sraffa, I, 1926). Contemporary economists agree. For example, Samuelson, argued that the problems of distribution are "fundamental and common to all economies, but different economic systems try to solve them differently" (1970:16). Scott and Nigro defined economics as "the study of how people behave to allocate and distribute scarce resources among competing ends over time" (1982:18), clearly emphasizing the importance of inequality.

Distribution of valued resources is one of the most enduring practical problems of human societies and, not surprisingly, represents a strong confluence of theory from several social science disciplines. Leading thinkers in political science, sociology, and economics agree that distribution is a defining characteristic of polity, society, and economy. All polities, societies, and

economies are similar in that they are characterized by inequality, but their most important differences are the extent and types of their inequality. Measuring inequality well, therefore, is a compelling enterprise.

CONCEPTUAL SPECIFICATIONS

By this time, general definitions of distribution and inequality are probably fairly clear. Their meaning is implicit in the way social scientists incorporate the concepts in definitions of their own academic disciplines. It is useful, nonetheless, to establish specific conceptual definitions of distribution and inequality at this point, so that our subsequent discussion will refer to a common core of meaning. I offer the following generic definitions to illuminate the meaning of the terms distribution and inequality in social science:

> Distribution is defined as the division of units among components of a social system. Inequality is defined as variation in that division.

Each of the major terms in these definitions requires explanation in order to convey the precise meaning of inequality. Variation refers to the process of changing in form, condition, or substance from a former or usual state or from an assumed standard. It can exist with respect to differences among components at a given point in time or within one component across periods of time. Division refers to separation of a whole into parts to be shared and resembles dispersion in statistics.

The term units refers to whatever is to be divided into parts to be distributed among or possessed by components. A component possesses a share of the units, some frequency and proportion of the total units. Components can be individuals, groups, or places among which units are distributed. Or components can be categories or classes among which people are distributed. The term system indicates an arrangement of things so interrelated as to form an organic whole separate from things outside the arrangement. It provides the reason that the components logically go together and form an identifiable and coherent arrangement of people, places, or groups.

A brief review of some of the units, components, and systems studied by social scientists will illustrate the definitions. Researchers have investigated the division of units such as federal aid, financial transactions, income, industrial productive capacity, market share, land, native languages, parliamentary seats, political party preference, population, war-making capacity, public services, minority group members, religious preferences, social status, students, unemployment, and wealth. Possessing shares of any of these units is a characteristic of the components. Some of these units are scarce and valuable resources, such

as income, power, and status. Attitudes and language may be abundant and neutral.

Social scientists have studied the division of these units among numerous components. People studied as components include individuals and candidates for political office. Social groups that have been examined include business firms, communes, farms, households, income classes, occupational classifications, political parties, population groups, school classes, and sectors of the economy. Places that have been analyzed as components include legislative districts, nation states, urban neighborhoods, school districts, regions of a country, and American states. Some of the geographical places, such as communes and neighborhoods, have important social group characteristics and could be classified either way.

The primary, large-scale systems that give meaning and coherence to division of units among such components are, of course, society, polity, and economy. Social status (units) is divided among occupations (components) in a society (system). Votes (units) are divided among political parties (components) in a multiparty polity (system). Income (units) is divided among income classes (components) in an economy (system). All three can be combined. Power (units of social, economic, and political resources) is divided among nation states (societies/economies/polities) in the world (international system).

An important point emerges from this review of units and components. It is possible not only for an investigator to conceptualize the distribution of units among individuals or social groups, but also to conceptualize the distribution of individuals among components such as places or social groups. For example, dollars can be distributed among people, and individuals can be distributed among ethnic groups. The type of conceptualization that is appropriate depends on the problem at hand. Selecting an appropriate statistical technique to measure inequality, given the problem at hand, is critical.

Many studies of distribution have examined small subsystems of the national society, polity, or economy. Typical examples include division of (a) native language spoken among territorial units in a region, (b) police protection services among neighborhoods in a city, and (c) shares of the market among business firms competing in a particular type of business.

The type of variation determines the distributional property of the system. If all components possess the same number of units, the distributional property is equality. If all components do not have the same number of units, the system is characterized by some degree of inequality. But inequality is a generic concept with many variations and forms.

Investigators use the concept inequality for two purposes, description and explanation. Examples of description include describing inequality in one system, comparing inequality in two or more systems, and describing changes in inequality in one system over time. Examples of explanation include cross-sectional and time-serial analysis of inequality either as a cause or a consequence.

Researchers sometimes seek to explain inequality as a dependent variable across several systems at a single point in time. Similarly, they often use inequality as an independent variable to explain some other phenomenon cross-sectionally. At other times, analysts explain change in a single system's distribution across time, in which case inequality is a dependent variable in a dynamic or time-serial analysis. One can also use change in a distribution across time as an explanatory variable.

Researchers have given many different names to specific types of equality or inequality, for example, advantage, apportionment, concentration, deprivation, deviation, dissimilarity, diversity, entropy, fluctuation, fragmentation, heterogeneity, (hyper)fractionalization, imbalance, inequity, instability, polyarchy, unanimity, uniformity, and variation. These are names of some of the distributional properties that scholars have investigated as well as names of the indexes of distribution they have devised to measure these properties.

The actual number of identifiable, measurable concepts of inequality is not as overwhelmingly large as the list implies. As subsequent chapters will demonstrate, however, social scientists use many different conceptualizations and operationalizations of inequality. We turn next to three graphic techniques that aid conceptualization of distribution.

USEFUL ANALYTICAL TOOLS

Three graphic analytical tools are essential to conceptualizing and measuring inequality: the frequency distribution, the cumulative frequency distribution, and the Lorenz curve. Each provides a different kind of visual analysis of distribution and serves as the basis of several inequality measures. Although I illustrate each through application to income distribution, remember that income is only one kind of unit that can be distributed among components. Social scientists have studied the distribution of many other important units, and we shall consider them throughout this book.

Frequency Distribution

As the curve or frequency polygon in figure 1-1 demonstrates, the frequency distribution plots rank-ordered income classes or groups on the horizontal axis and number of income receivers on the vertical axis. The curve begins very low in the extreme left of the graph and ascends rapidly to a peak. The peak is the income category with the largest number of income recipients and is called the mode or modal category. At the peak, the curve descends rapidly at first, but at one point (called the arithmetic mean) it's downward slope becomes less radical until it actually parallels the horizontal axis. The area around the mode is graphically well displayed; these are the middle-income recipients. The right-hand "tail" of the distribution, however, conceals much of what is happening to the

Figure 1-1
Frequency Distribution

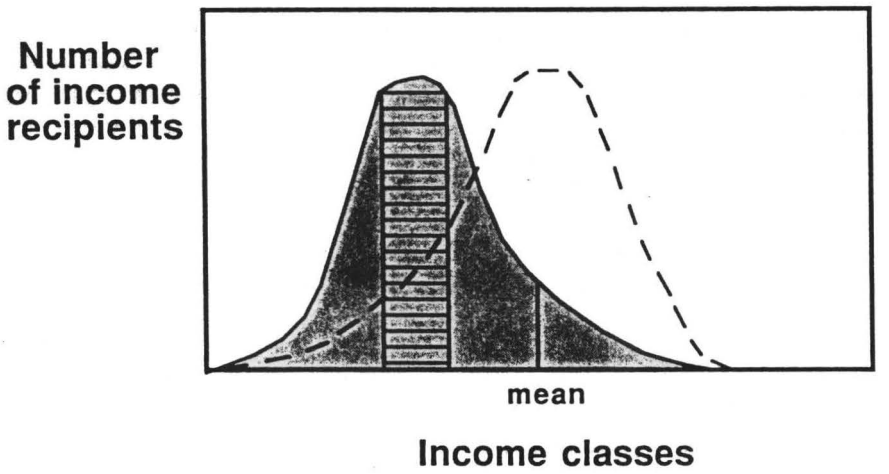
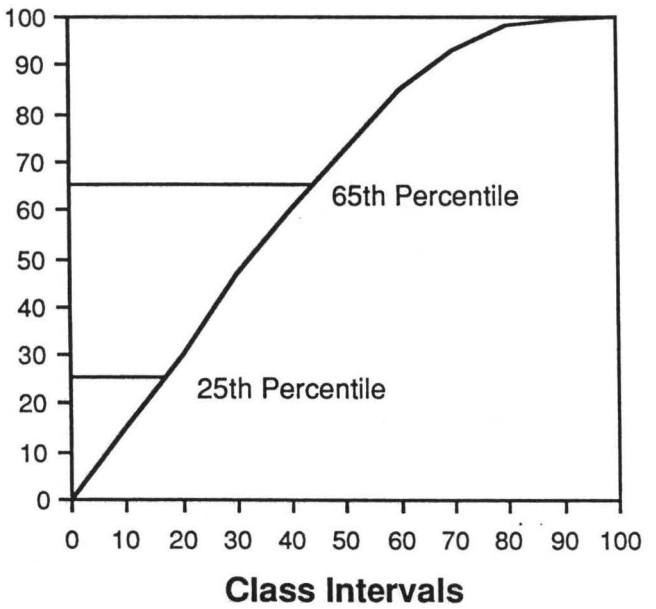


Figure 1-2
Cumulative Frequency Graph

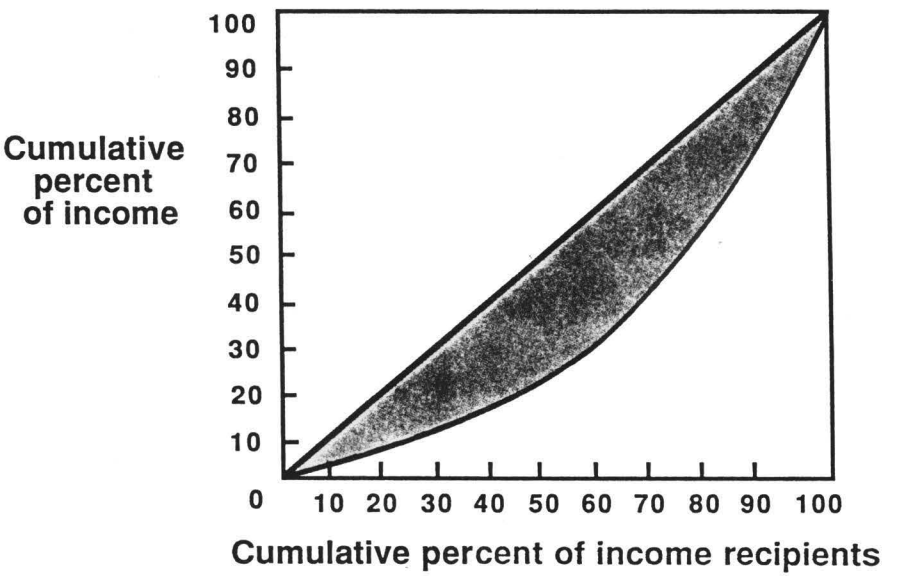


rich and super-rich. The horizontal axis would have to be about two hundred yards long to display the very rich accurately!

Cumulative Frequency Distribution

The cumulative frequency distribution is easily constructed from the frequency distribution. The horizontal axis remains the same. The vertical axis, however, expresses cumulative percentages, zero to one hundred. It gives the percentage of cases below (or above) each class boundary, as figure 1-2 indicates. It permits the frequency distribution to be partitioned at any point, according to the investigator's interests. For example, the lower twenty-five percent can be separated from the upper seventy-five percent of the income recipients. This partitioning represents the first quartile. Centile partitioning divides the total frequency into one hundred equal parts. Figure 1-2 includes designation of the first quartile and the sixty-fifth percentile.

Figure 1-3
Ordinary Lorenz Curve



The Lorenz Curve

The American statistician, Max Lorenz (1905), devised the most widely used diagrammatic display of distribution (Pen, 1971). Called the Lorenz curve, it appears in figure 1-3. First, the income-earning units, individuals or families (sometimes classes of individuals or families), must be arrayed in ascending order. The horizontal axis plots the number of income recipients, from lowest to highest income earners, in cumulative percentage terms. As we move along the horizontal axis to point twenty, we have the lowest twenty percent of income earners or income-earning families. At point fifty, we have fifty percent of all families (the fifty percent that earns the least). At one hundred, we have all of them. The vertical axis plots the cumulative percentage share of income going to each percentage share of income earning units (individuals or families). It also starts at zero and goes to one hundred percent.

The Lorenz curve gives the relation between the percentage of income recipients and the percentage of income they earn. The diagonal represents total equality, in other words, one percent of the families receives one percent of the income, five percent receives five percent, ten percent receives ten percent, etc. through all families and all income. The curve traced by a given empirical income distribution, however, will usually sag beneath the diagonal. For example, the bottom thirty percent of the families might earn only eight percent of the income. Similarly, the top ten percent may earn thirty percent of the income. The greater the sag in the curve below the diagonal, the greater the inequality. If all individuals or families had equal income, the curve would coincide with the diagonal. If one individual or family received all the income, the curve would coincide with the vertical and horizontal axes.

Because the Lorenz curve is a particularly useful analytical device, we will return to it often in subsequent discussions. It provides the theoretical basis for several important inequality indexes. It is also used to compare income distributions among two or more countries, in a single country before and after some tax reform, among occupational groups in a single country, etc. If the Lorenz curve of distribution A is everywhere above and no where below the Lorenz curve of distribution B (i. e., A is everywhere northwest of B), then distribution A is said to "Lorenz-dominate" distribution B. Distribution A is more equal or less unequal than distribution B. Two Lorenz curves that intersect, however, cannot be ranked with respect to their inequality. Comparing distributions with intersecting Lorenz curves requires use of other distributional characteristics. Nonetheless, the Lorenz curve conveys a graphically distinct, geometrically-derived impression of a distribution.

THE ORGANIZATION OF THIS BOOK

This chapter introduces inequality as a social science concept and illustrates three key graphical tools in distribution analysis. Chapter two details numerous conceptual and operational criteria that comprise inequality theory. These principles are also helpful in selecting an appropriate statistical index with which to measure distribution, given a substantive or conceptual definition of inequality. In addition, chapter two presents the statistical notation necessary to measuring inequality.

Inequality theory and the statistical notation discussed in chapter two are important and prominently reappear in all of the subsequent chapters. The principles of inequality theory are used to evaluate each of the inequality indexes presented, that is, to describe the way it performs under varying conditions and to explain why. In some cases, satisfying the criteria represents strength. In other instances, it simply represents an alternative. The main point of these criteria is to provide researchers with the tools necessary to conceptualize distribution properly and to make rational measurement choices, given the distributional problem to be analyzed.

The last inequality principle discussed in chapter two is the mathematical model from which a particular inequality index is derived. This criterion is so important that I use it as the organizing principle for the next four chapters that present intragroup inequality measures. Most measures of inequality are derived from one of four mathematical models: deviations, combinatorics, entropy, and the social welfare function. Chapter three presents measures of inequality based on a deviations model. Indexes based on combinatorics receive attention in chapter four. Chapter five examines measures utilizing entropy. Chapter six analyzes distribution indexes based on the social welfare function. Discussion of each measure is designed to highlight both its underlying mathematical logic and its step-by-step computational procedure. The mathematical model provides both the quantitative logic underlying a given equation and a means of interpreting its results.

Chapter seven extends the logic developed in the previous four chapters to examine measurement of intergroup inequality. Intragroup indexes measure inequality between persons or components in a single population. Intergroup indexes measure inequality between people in different groups, for example, occupational status differences between blacks and whites.

Chapter eight deals with inequity. In this case, equality is deemed to be inappropriate, and some other standard must be used. First, the investigator must decide how many units each component should possess—an unequal but "equitable" share of the units. Then, an index of inequity will measure the extent to which the actual distribution conforms to the chosen equity standard or equitable distribution. Inequity indexes are particularly useful in cases in which