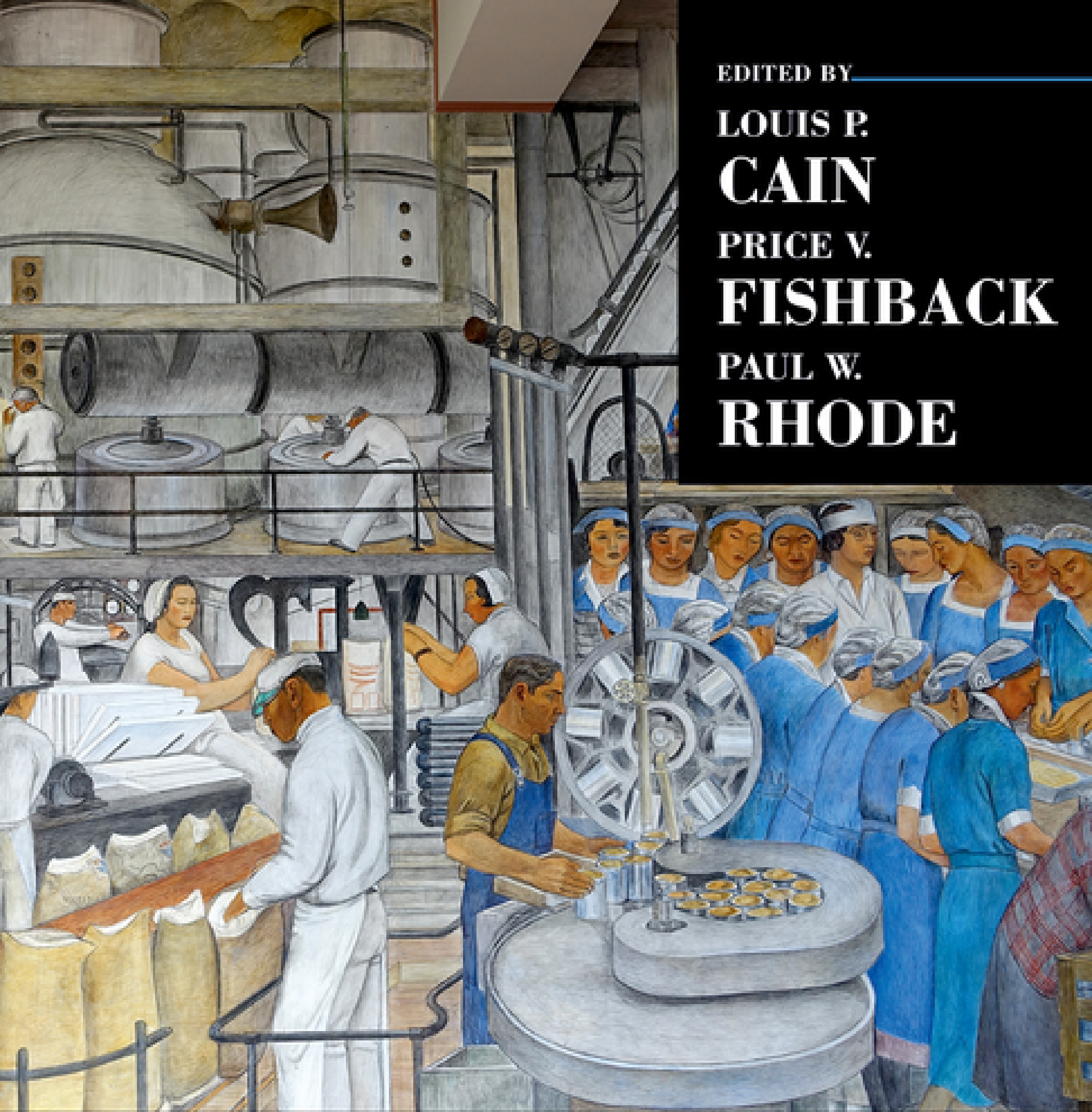

EDITED BY
LOUIS P. CAIN
PRICE V. FISHBACK
PAUL W. RHODE

The Oxford Handbook *of* AMERICAN ECONOMIC HISTORY

VOLUME 2

THE OXFORD HANDBOOK OF

AMERICAN ECONOMIC HISTORY

VOLUME 2

THE OXFORD HANDBOOK OF AMERICAN ECONOMIC HISTORY

VOLUME 2

Edited by
LOUIS P. CAIN
PRICE V. FISHBACK
and
PAUL W. RHODE

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide. Oxford is a registered trade mark of Oxford University Press in the UK and certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America.

Library of Congress Cataloging-in-Publication Data
Names: Cain, Louis P., editor. | Fishback, Price Van Meter, editor. | Rhode, Paul Webb, editor.
Title: The Oxford handbook of American economic history /
edited by Louis P. Cain, Price V. Fishback, and Paul W. Rhode.
Description: Oxford ; New York : Oxford University Press, [2018] |
Series: Oxford handbooks | Includes bibliographical references and index.
Identifiers: LCCN 2017052624 (print) |
LCCN 2017053491 (ebook) | ISBN 9780199983148 (online component) | ISBN 9780199947980 (updf) |
ISBN 9780199947973 (2-volume set : alk. paper) | ISBN 9780190882617 (volume 1 : alk. paper) |
ISBN 9780190882624 (volume 2 : alk. paper)
Subjects: LCSH: United States—Economic conditions.
Classification: LCC HC103 (ebook) | LCC HC103 .O98 2018 (print) | DDC 330.973—dc23
LC record available at https://lccn.loc.gov/2017052624

1 3 5 7 9 8 6 4 2

Printed by Sheridan Books, Inc., United States of America

Contents

VOLUME 2

PART IV. TECHNOLOGY AND URBANIZATION

PART V. GOVERNMENT AND ECONOMIC POLICY

PART IV

TECHNOLOGY AND URBANIZATION

CHAPTER 19

INNOVATION AND PATENTS IN AMERICAN ECONOMIC HISTORY

PETRA MOSER

Intellectual property rights are at the center of heated policy debates that have been triggered by litigation in "patent wars," which threaten to paralyze critical parts of the world economy. The outcome of these debates are crucial to the welfare of future generations, as they set the parameters within which firms create, adopt, and share innovations that range from new medicines to mobile phones, tablet computers, and web content. Among the critical question are the following: (1) How strong should patent rights be to provide the optimal incentives for the creation and adoption of new technologies? (2) Are patents equally effective and necessary across industries? For example, should improvements in software, business methods, or living organisms be patentable? (3) Do patents encourage or discourage the diffusion and adoption of new technologies?

Much of the existing empirical evidence on the effects of patent laws comes from analyses of historical US and European patents. The first article of the US Constitution instructs Congress to "promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors to the exclusive Right to their respective Writings and Discoveries." The history of patent laws, however, begins in 1474, when the Venetian Republic first offered exclusive rights to inventors and entrepreneurs who had invented or imported new technologies to Venice. The Venetian system was originally adopted to attract skilled artisans by offering temporary property rights for their ideas, rather than to encourage domestic innovation. Other European countries began to copy the Venetian system not only to promote economic development, but also to reward political and financial support (David 1994, 134).

The US system modified existing patent systems laws by creating the world's first modern patent law, with the explicit goal of promoting the creation and commercial adoption of new technologies. A key difference between the US and earlier systems was its examination system, introduced in 1823, by which technically trained bureaucrats established the patentability of each new invention. The US system also offered patents at substantially lower fees, roughly 5 percent of fees in Britain, which, at least in theory, would allow for broad access to the protection offered by intellectual property rights. These features were widely copied by other countries, and have been credited with encouraging the invention and the commercialization of technologies in the United States (e.g., Khan and Sokoloff 1998, 2001).

Patents as a Decisive Factor in the Industrial Revolution

In 1623, Britain's Statute of Monopolies replaced a system of royal monopolies—which King James I and monarchs before him had exploited to extract rents for their citizens—with a system of legal property rights in ideas. North and Thomas (1973) credit this change with helping to encourage Britain's industrial revolution, arguing that properly executed, legal property rights in ideas create incentives to invent, to promote innovation, and to encourage economic growth.

Similarly, the rapid industrial development of the United States in the nineteenth century has been interpreted as providing additional evidence for the importance of patent laws (e.g., Khan 2005; Khan and Sokoloff 1993; Lamoreaux and Sokoloff 1999). This provision established the foundation for the world's first modern patent system, which Khan and Sokoloff (1998, 2001) argue was instrumental in encouraging technological progress and economic growth in the United States. Compared with the British system, the US patent system was substantially cheaper and easier to navigate.

In 1851, European visitors to the Crystal Palace Exhibition in London first noticed the rapid technological advances that US inventors had made and began to consider the link between the characteristics of the US patent system and the country's surprisingly high levels of innovation. Until the 1851 fair, the United States had been considered a backwater, with low levels of technologies. In fact, the US Commission to the Crystal Palace had initially struggled to fill the exhibition space at the world's fair (Moser 2005). Soon after the fair opened, however, visitors were amazed by US innovations in agriculture, such as McCormick's grain reaper, which promised enormous productivity gains in agricultural production. Samuel Colt's revolver demonstrated advances that US inventors had made in developing a specific American system of manufacturing, in which machines and final products, such as the revolver, were made with such uniformity and precision that they could be treated as interchangeable parts (Hounshell 1988).

Soon afterward, Britain's Parliament established a commission to investigate the US system of manufacturing (Rosenberg 1969, 2). In his 1855 report on the "New York Industrial Exhibition," Commissioner Joseph Withworth emphasized the importance of the US patent system, with low fees and easy accessibility to the public (Rosenberg 1969, 377).

Alternative Determinants of the Rate and Direction of Technical Change

Alternative accounts have emphasized the importance of advances in science, differences in factor prices, and cultural differences in determining the rate and direction of technical change. In the United States, for example, high costs of labor relative to the abundance of natural resources are widely credited with encouraging mechanization and the development of a specifically American system of manufacturing (Habakkuk 1962; Hounshell 1985; Rosenberg 1963, 1969, 1972; Rothbarth 1946). Accounts of innovation in Britain and the United States also indicate that innovation occurred independently of patents as a result of knowledge sharing (Allen 1983; Nuvolari 2004; Thomson 2009). For example, technical performance in the Cleveland (UK) iron-smelting district improved rapidly between 1850 and 1875, as iron producers shared technical details about the construction and performances of blast furnaces (Allen 1983). Mokyr (2009) explains how knowledge sharing and a shift toward science-based experimentation during the Enlightenment helped set the stage for Europe's industrial revolution.

Most empirical analyses use data on *patented* inventions as a measure for innovation and technical change because patent data are the most readily available source of systematic data on innovation. Following the pioneering analyses of hand-collected data in Schmookler (1962, 1966), and Sokoloff (1988), the availability of electronic data sets, such as the National Bureau of Economic Research's (NBER's) Patents and Citations data set for 1976–2002 (Jaffe, Hall, and Trajtenberg 2001) and optical character (OCR) files of US patents since 1920, has further contributed to the popularity of patent data.[1]

By comparison, innovations that occur outside of the patent system are more difficult to measure, especially across industries and across countries. Studies of individual industries, such as Allen (1983), Nuvolari (2004), and Lampe and Moser (2010), have addressed the data issue by collecting measures for improvements in the performance of specific technologies, such as the height of blast furnaces, the temperature of iron blasts, or the speed of sewing. There are, however, few alternatives to patents as a measure for variation in technical change across (and within) industries, across countries, across regions, and over time.

Patent data, however, are unsuitable for measuring variation in innovation in response to patent laws, because they cannot measure technical change in countries without patent laws, or in industries in which inventors rely primarily on alternative mechanisms to protect their intellectual property. For example, respondents from 634 American research and development labs in 1983 (Levin et al. 1987), and from 1,478 American firms in 1994 (Cohen, Nelson, and Walsh 2000) report that, in most industries, secrecy is more effective than patents as a mechanism to protect intellectual property. If, however, the use of patents varies systematically across industries or over time, raw patent counts, without controls for technology-specific differences, are a poor measure of innovation.

Exhibition Data as an Alternative Measure of Innovation

Alternative data to measure innovations with and *without* patents are available from a series of prominent international technology fairs, which for the first time in history systematically collected and displayed technologies from across the world, and remained the leading showcase for new technologies up to World War I. In addition to the Crystal Palace Exhibition in London in 1851, these fairs include the US Centennial Exhibition in Philadelphia in 1876, the World's Columbian Exhibition in Chicago in 1893, and the Panama-Pacific International Exposition in San Francisco in 1915 (Moser 2003, 2005, 2011, 2012). In 1851, the Crystal Palace, a 1,848-foot-long greenhouse of cast iron and glass, was the largest enclosed space on earth; it housed 17,062 exhibitors from forty countries; at a time when London had less than 2 million inhabitants, more than 6 million tickets were sold (Moser 2005). Its US counterpart, the American Centennial Exhibition in Philadelphia in 1876, attracted more than 10 million visitors (Kroker 1975, 146); these visitors could have walked more than twenty miles to see all 30,864 exhibits from thirty-five countries. In 1893, the World's Columbian Exposition covered 717 acres of land and water in Jackson Park by Lake Michigan in Chicago; it attracted 27.5 million visitors. In 1915, the organizers drained part of the San Francisco Bay and converted it to a fairground, which was to welcome 30,000 exhibitors from thirty-two countries and 19 million visitors.

Detailed information on exhibits is available from contemporary reports and from the catalogues that helped visitors navigate the exhibition grounds. For example, Britain's *Official Catalogue of the Great Exhibition of the Work of Industry of All Nations* (1851) became a "book of reference to the philosopher, merchant, and manufacturer" (Auerbach 1999, 94) to an exhibition that was a "veritable acting industrial encyclopaedia" (Tallis 1852, 234; *The Times*, December 5, 1850). Catalogue entries include the exhibitor's name, location, and a description of the exhibit:

> 406, Fourdrinier, E. N., 38 Barclay St., Sunderland, Patent safety apparatus, for preventing loss of life and destruction of property when a rope or chain breaks in shafts of mines and collieries.

In addition to the prize-winning reaper and Colt's revolver, examples of exhibits include power-loom lathes, and the first sewing machines.

A key benefit of data on the world's fairs of 1851, 1875, 1893, and 1915 is that exhibits were subject to a systematic and careful process of selection, while selection processes at smaller regional fairs and even the later world's fairs became substantially more haphazard. In 1851, governments set up a comprehensive national organization of commissioners who systematically visited cities and rural areas to set up local selection committees and to advertise the fairs to ensure that exhibits were representative of a country's innovations (Auerbach 1999, 55, 87; Bericht 1853, 40, 64).[2] In Britain, sixty-five local commissions formed more than 300 subcommissions (Auerbach 1999, 32). Local committees selected exhibits in the first round; their submissions competed at the national level. National committees selected innovations to represent their country in a hard-fought "contest of industries," to which Britain had "challenged the world" (Kretschmer 1999, 46–48; Auerbach 1999, 68–69, 78, 189). Less than 30 percent of applicants were admitted to the fair (Bericht 1853, 50, 117). After the fair had closed, universities, scientists, and museums competed to purchase exhibits (Auerbach 1999, 105, 120).

A concern with exhibition data is that transportation costs may have influenced what was exhibited at the fair and that exhibits from host countries may be overrepresented. For example, three-fourths of the thirteen acres in Machinery Hall was taken up with American machines in 1876, including the colossal Corliss steam engine, which provided power to the fair (Andrews 1894, 73). Historical accounts, however, indicate that foreign nationals presented many large and heavy innovations at the fairs in spite of transportation costs. For example, at the 1876 fair in Philadelphia, "Great Britain led the foreign nations in the size and number of articles (in machinery) exhibited. Canada, France, Russia, Sweden, Brazil, and other countries sent ingenious or powerful machines" (Andrews 1894, 73). The influence of transportation costs may also have been mitigated because exhibitors only paid for transport to local collection points (Bericht 1853), and heavy or fragile inventions could have been presented as models. In 1851, for example, of 194 British exhibits in the "Civil Engineering, Architecture, and Building Contrivance," 88 exhibits were represented by models, including a model of the 770-meter-long chain suspension bridge that was constructed in Kiev across the Dnepr between 1848 and 1853 (Rolt 1970, 157). Similarly, T. Powell of Monmouthshire, Britain, exhibited a "[m]odel for apparatus used for shipment of coals from boats or waggons at Cardiff dock" at the Crystal Palace, and A. Watney of Llanelly, Wales, exhibited "[m]odels of anthracite blast furnaces."[3]

Another potential source of bias stems from restrictions on exhibition space, which were set by the host country. Anecdotal evidence, however, indicates that such restrictions may not have been binding. For example, when the US commission to the Crystal Palace thought that US exhibitors would be short on exhibition space, it asked the British commission for more room and was granted its request (Haltern 1971, 150). Floor plans for the Centennial exhibition show that eleven countries, including Germany, Britain, Spain, and Sweden, built additional exhibition space on the Centennial grounds (Visitor's Guide 1875, 18).[4]

British Innovations

Consistent with accounts that emphasize the importance of alternative mechanisms, exhibition data indicate that the majority of nineteenth-century innovations occurred *outside* of the patent system. In 1851, 11 percent of British exhibits were patented. This low share of patented inventions may be considered a challenge to accounts, such as North and Thomas (1973) and Lamoreaux and Sokoloff (1998), which emphasize the importance of intellectual property rights as a driver of innovation and growth. Instead, it lends support to alternative accounts, which emphasize the importance of a culture of entrepreneurship (Landes 2003), experimentation (Mokyr 2009), the free exchange of knowledge (Allen 1983; Nuvolari 2004), and science (Arora and Rosenberg 1998; Mowery and Rosenberg 1998).

Low patenting rates are particularly remarkable considering that exhibition data cannot capture all innovations that are protected by alternative mechanisms. For example, exhibition data may underestimate innovations that were protected by secrecy because inventors were reluctant to display such innovations at a fair. To address this problem, the UK government created a registration system to offer "Protection from Piracy to Persons exhibiting new Inventions" (Protection of Inventions Act, April 11, 1851). Only 600 exhibitors, less than 3.3 percent of Crystal Palace exhibitors, however, took advantage of this system. Instead, many exhibitors chose to advertise their finished products, rather than the machine that had produced the product, to protect their intellectual property in the machine. For example, Drewsen & Sons of Silkeborg, Jutland, exhibited "[s]pecimens of paper, glazed by a machine constructed by the exhibitor," instead of their paper-making machine (Official Catalogue 1851, 210).[5]

Were Patent Laws a Necessary Condition for High Levels of Innovation?

In addition to analyses of the share of patented inventions, analyses of variation in the number of exhibited inventions, the distribution, and the quality of innovations can help to shed light on the effectiveness of patent laws as a mechanism to encourage innovation. Nineteenth-century patent laws offer a particularly useful setting for such analyses, because no other time period offers an equal amount of variation in national system. For example, data from instructions manuals on foreign patent laws for US and British inventors (Coryton 1855; Godson 1840; Kingsley and Pirsson 1848) indicate that the length of national patent laws varied from zero years in Denmark and Switzerland to fourteen years in Britain and the United States, and fifteen years in Bavaria, Belgium, and the Netherlands.

Another benefit of this setting is that nineteenth-century patent laws had been adopted according to the idiosyncratic allegiances of national rulers (Penrose 1951, 13), and generally had not been altered in response to lobbying from individual industries or countries. Data for the Crystal Palace Exhibition cover 12,694 total exhibits and 5,833 from fifteen foreign countries; three of these countries offered no patents. For the US Centennial Exhibition in Philadelphia in 1876, the data cover 19,073 exhibits, including 14,608 from fourteen foreign countries; two of these countries offered no patent protection, and three offered only short-lived patents. Even though exhibition data cannot capture all innovations without patents, they capture a large number of innovations that previous analyses had to omit because they have relied on patent data. Most important, exhibition data make it possible to compare inventive activity in countries with and *without* patent laws. Even within countries that have patent laws, exhibition data make visible a new set of innovations, which inventors choose to protect through alternative mechanisms, such as secrecy.

Comparisons of exhibits and prizes across countries with and without patent laws in 1851 and 1876 suggest that the existence of a national patent system may *not* be a necessary condition to encourage innovation (Moser 2003; Moser and Zimring 2012). In 1851, Switzerland contributed 110 exhibits per capita (measured in millions of people), compared with a mean of 55 and a median of 36 across all countries. Switzerland also won a disproportionate number of prizes for being especially innovative (47 compared with a median of 11 and a mean of 20 prizes per one million people).

Small Differences in Patenting despite Large Differences in Patent Laws

As an alternative test for the importance of patent laws, exhibition data also make it possible to examine variation in the inventors' use of patents across different types of patent systems. Most prominently, Khan and Sokoloff (1998, 2001) have credited the superior design of the US patent system with encouraging technical progress and economic growth through the "democratization" of invention in the United States.

The British system reflected its origin in royal privilege and effectively limited access to a select class of patentees. In contrast, the United States established the world's first modern patent institution—one that was consciously designed to stimulate participation in invention across a wide spectrum of the population (Khan and Sokoloff 1998, 295).

Consistent with this argument, comparisons of the process of patent application reveal striking differences in costs. Patent fees were roughly $30 for domestic inventors in the United States in 1851, compared with 100–120£ in Britain, "or approximately 4 times per capita income in 1860," and 350£ for a patent that covered Scotland and Ireland (Khan 2005, 31).[6]

British inventors faced a drawn-out and expensive process, with exorbitant legal fees and bribes (Dutton 1984; MacLeod 1988, 76). In "A Poor Man's Tale of the Patent

System," Charles Dickens characterizes the trials of the fictional artisan and inventor Old John:

> Look at the Home Secretary, the Attorney-General, the Patent Office, the Engrossing Clerk, the Lord Chancellor, the Privy Seal, the Clerk of Patents, the Lord Chancellor's Purse-bearer, the Clerk of the Hanaper, the Deputy Clerk of the Haaper, the Deputy Sealer, and the Deputy Chaff-was. No man in England could get a patent for an Indian-rubber band or an iron hopp, without feeing all of them. Some of them, over and over again. (Dickens 1925, 64–70)

By comparison, US inventors could submit their patent applications by mail. Mokyr (2002, 295), however, cautions:

> Whether this difference demonstrates that a well-functioning system of intellectual property rights is truly essential to the growth of useful knowledge remains an open question. For one thing, the American system was far more user-friendly than the British patent system prior to its reform in 1852. Yet despite the obvious superiority of the U.S. system and the consequent higher propensity of Americans to patent, there can be little doubt that the period between 1791 and 1850 coincides roughly with the apex of British superiority in invention. The period of growing American technological leadership, after 1900, witnessed a stagnation and then a decline in the American per capita patenting rate.

Exhibition data on rural-urban variation in the use of patents support the hypothesis that the US system increased participation in patenting by rural inventors. For British exhibits, patenting rates were substantially higher in London, compared with rural areas, but there is no evidence of significant differences in patenting rates between urban and rural patenting rates in the Unites States (Moser 2012, 13–14, Table 4). Results from the United States, however, are based on a small sample of 574 exhibits, even though British data include more than 6,000 exhibits. Results, however, are suggestive and highlight the need for additional analysis to examine whether and how institutional differences in patent laws may have influenced participation.

Importantly, however, US patenting rates were only slightly higher overall compared with patenting rates for British exhibits. Only 15.3 percent of all US exhibits at the Crystal Palace Exhibition were patented, compared with 11.1 percent of exhibits from Britain (Moser 2012, 54). These findings challenge the view that the design of the US patent system played an important role in encouraging participation in the US system—and innovation more broadly—and call for additional analyses.

Did High Litigation Risks Discourage Patenting in the United States?

A potential explanation for the surprisingly low rates of patenting in the United States is that the costs of enforcing may underestimate the total costs of patenting in the United

States. Specifically, the costs of *enforcing* patents may have been substantial, if the risks of litigation were high. Historical accounts indicate that US courts have been more likely to uphold the patent rights of early generations of inventors, while British courts tended to be more anti-patent (Dutton 1984; Khan 2005). Proponents of stronger patent systems, such as Khan (2005), have argued that these differential attitudes favored invention in the United States.

Historical accounts of US innovation reaching back to Eli Whitney's saw gin, however, illustrate the important role of courts and litigation in determining the boundaries of US patents. The US Patent Act of 1793 specified that later generations of inventors could only patent the improvement that they had added to an existing invention, but not "make, use, or vend" the whole machine. The law also specified that the original, first-generation inventor could not use or sell the improvement and that "simply changing the form or proportion of any machine . . . shall not be deemed a discovery." It was, however, left to the courts to determine where the original inventors' improvements ended and the contributions of later generations began. In 1793, the same year as the passage of the Act, Whitney had patented an improvement of the saw gin, which made it possible to remove seeds from commonly used short-staple Upland varieties of cotton (Lakwete 2003, 67–68; Olmstead and Rhode 2008, 1130). Whitney's 1793 patent, however, remained subject to litigation until December 1807, one month after it had expired.

More generally, providing stronger intellectual property rights to early generations of "first inventors" may reduce incentives to patent and invest in research and development (R&D) by reducing the range of topics that are available for improvements and by increasing litigation risks. Returns to future generations of inventors may have been particularly low when the US Patent and Trademark Office (USPTO) issued broad patents that were subsequently enforced by US courts. In 1846, for example, the USPTO issued patent 4,750 to Elias Howe for an Improvement in Sewing Machines. Howe's patent was broad enough to cover most commercially viable sewing machines at the time. Like a twenty-first-century "patent troll," Howe used his patent to threaten litigation, instead of commercializing his invention. In 1852, a District Court upheld Howe's patent, and he began to collect license fees of $25 per machine, roughly one-fifth the average price of a sewing machine (Lampe and Moser 2010). Other firms sued based on their own patents, and production came to a near halt in the 1851–1856 "sewing machine wars" (Bissell 1999, 84). By 1867, Howe had received $2 million in license fees (Parton 1867), roughly $27.8 million in 2011 dollars.

In an ideal setting, a patent system that protects ownership in ideas independently from production may allow creative individuals—without the means to produce their inventions—to specialize in the activity of invention. It is, however, also possible that the necessity to defend patents in court precluded this career path for a large portion of the population. Moreover, the granting of broad and strong rights may have laid the seeds for high litigation risks and socially wasteful expenses for strategic patenting, a key problem of twenty-first-century patent law (e.g., Bessen and Meurer 2008). Additional analyses of the US legal system, and of decisions to enforce—or abandon—broad patents, are needed to further our understanding of the historical effectiveness of the US system in encouraging invention.

The Use of Patents across Industries

Exhibition data also reveal a large amount of variation in the use of patents across industries. In 1851, fewer than 5 percent of Britain's chemical exhibits, 10 percent of scientific instruments, and 8 percent of exhibits in food processing were patented, compared with 20 percent of manufacturing machinery (Moser 2012, 46). The data also show that—despite the most substantial differences in national patent systems—US inventors appear to have relied on patents, and to have avoided them in the same industries as British inventors (Moser 2012, 58). Similar to the results on substantial similarities in overall patenting rates, these findings suggest that policies, which reduce the costs of patent applications, failed to substantially alter the patterns of participation in the US patent system.

Instead, historical accounts suggest that variation in the effectiveness of secrecy, as an alternative to patents, was instrumental in determining variation in patent use. With crude tools of chemical analysis, secrecy was an effective mechanism to protect mid-nineteenth-century improvements in chemicals, because competitors could not yet reverse-engineer improvements. Valuable dyes such as indigo and madder red proved impervious to industrial espionage until the late nineteenth century (Haber 1958, 83). Secrecy was also effective in protecting improvements in scientific instruments, including the rectangular prisms of Swiss glassmaker T. Daguet of Soleure and the optical instruments of Danish makers (*Bericht I*, 1852, 813–819, 930–941).

Scientific breakthroughs, such as August Kekulé's model of the benzene ring in 1865 and Dmitrii Mendeleev's publication of the periodic table in 1869, transformed chemical analysis in the second half of the nineteenth century, and increased the risks of secrecy (Haber 1958, 81). Exhibition data indicate that these advances may have increased inventors' dependency on patents. At the world's fairs in 1851 and 1876, 0 and 5 percent of US chemical innovations were patented. By 1893 and 1915, 19 and 20 percent of US chemical innovations were patented. By comparison, advances in analytic methods had no effects on innovations in manufacturing machinery, which had always been easy to copy. Exhibition data show that inventors' use of patents stayed roughly constant between 44 and 49 percent for manufacturing machinery (Moser 2012, 62–67). These comparisons suggest that scientific breakthroughs, which make it easier to analyze and reverse-engineer new technologies, may cause inventors to rely more heavily on patents.

Should Living Organisms Be Patentable?

Another set of policy issues revolves around the question of whether modifications of living organisms should be patentable. The first US law to create intellectual property rights in living organisms was passed with the Plant Patent Act of 1930, which created intellectual property rights for plants that propagate asexually, through roots rather than seeds. In the aftermath of World War I, Congress had passed this law as a means to assure independent, domestically produced food supplies for the United States. In deliberations leading up to the Act, breeders had argued that they required intellectual property rights to recover the large costs of developing improved plants because competitors could easily propagate improved plants after a breeder had invested many years of research. In the absence of property rights, Stark Brothers Nursery, a prominent US breeder, had built a cage with a burglar alarm to surround its improved Golden Delicious apple tree to prevent competitors from stealing cuttings. Such measures, however, would not be effective to protect production on a large scale, in addition to being socially wasteful. By providing *legal* protection through plant patents, Congress hoped to encourage domestic innovation and the development of a domestic US plant-breeding industry.

Data on plant patents, however, show that nearly half of all US plant patents between 1930 and 1970 were for improvements in roses, suggesting that the 1930 Act had failed to secure domestic US food supplies (Moser and Rhode 2012, 418–420). Anecdotal evidence indicates that the creation of plant patents supported the creation of a research-based domestic rose-breeding industry. Until World War II, US nurseries had depended on imported nursery stock from Europe. In the 1940s, however, roughly a decade after the Plant Patent Act, commercial nurseries, which account for the majority of plant patents, began to build mass hybridization programs for roses. Similar to pharmaceutical R&D, it takes up to twelve years to develop a new rose, and fewer than one in 1,000 seedlings typically lead to a commercially successful new rose (Moser and Rhode 2012). Once a breeder has developed a new rose, however, competitors can easily free-ride on his investments by propagating cuttings from an improved plant.

Historical evidence on registrations of newly created roses between 1916 and 1970, however, indicate that intellectual property played at best a limited role in encouraging the development of a research-based US rose-breeding industry. Less than 20 percent of new rose varieties that were registered after 1930 were protected by a plant patent (Moser and Rhode 2012, 429–434). Registration data also show that US breeders created *fewer* new roses after 1931. Moreover, data on the lineage of roses indicates that most roses that are commercially successful today are descendants of roses that public- rather than private-sector plant scientists developed *before* the creation of plant patents. Instead, historical records suggest that the US rose industry received a boost when World War II cut off rose supplies from European competitors and US breeders began to produce their own nursery stock based on roses that they had licensed from European breeders.

Patent Laws and the Direction of Technical Change

If the usefulness of patents varies across industries, as the example of chemicals, machinery, and plant breeding suggests, and inventors are heavily dependent on patent protection in some industries (such as machinery), but are able to protect innovations even without patents in other industries, radical changes in patent laws may influence the *direction*, rather than the level, of technical change (Moser 2005). These patterns are borne out in exhibition data: at the Crystal Palace, one-fourth of exhibits from countries without patent laws were scientific instruments, compared with one-seventh of exhibits from other countries (Moser 2005). Countries without patent laws also had larger shares of innovations in textiles, especially dyes, and in food processing.

The history of innovation in margarine illustrates the effectiveness of secrecy in food processing. Believing himself protected by a patent in his home country, the French chemist Mège Mouriès disclosed the process of producing margarine from suet to the Dutchmen Jurgens and van den Bergh. Jurgens and van den Bergh began to manufacture margarine in the Netherlands in 1871—two years after the country had abolished patent laws. When the partners split after a dispute, van den Bergh pursued independent research to improve taste and make margarine commercially viable. Van den Bergh protected the new formula through secrecy, and Jurgens was unable to match the superior taste of van den Bergh's margarine until 1905 (Schiff 1971).

More generally, the share of Dutch innovations in food processing increased after 1869, when a political victory of the free trade movement led to the abolition of patent laws. In 1851, 11 percent of exhibits from the Netherlands were related to food processing. In 1876, 37 percent of Dutch exhibits, including a disproportionate amount of award winners, originated from this industry (Moser 2005). Many other innovations in the field, including milk chocolate, baby foods, and ready-made soups, were made in Switzerland and the Netherlands when neither country offered patents (Schiff 1971, 52–58).

In addition to encouraging the creation of new technologies, patent laws are also intended to encourage the geographic diffusion of innovative activity, which is essential as a means to encourage growth. In an analysis of US patent counts, Lamoreaux and Sokoloff (1999) link the increase in US patenting in the late nineteenth century with the emergence of professional patent agents, whose role was to facilitate the trade in patented ideas. Early patent laws, such as the Venetian system, emphasized disclosure and teaching a new set of firms about the "mysteries" of more advanced technologies (David 1994). Effective patent laws are, however, often regarded as a mechanism to prevent, rather than encourage, the diffusion of patented ideas (e.g., Abramovitz 1989, 39–40). For example, Dutton (1984) credits the *in*effectiveness of the British patent system protection with facilitating the diffusion of ideas, and Allen (1983) and Nuvolari (2004) document the benefits of a free exchange of idea without patents.

Anecdotal evidence for individual inventors, such as Mège Mouriès, suggests that inventors may be more willing to disclose technical information if they believe themselves to be protected by a patent. In another example from early nineteenth-century England, the iron founder Robert Ransome began to advertise his plowshares to all ironmongers in Norwich and fifty outlets in East Anglia after he received a patent in 1803 (MacLeod 1988, 100). By contrast, inventors have fiercely guarded knowledge from spreading to people outside their social network in the absence of intellectual property.

For example, silk weavers in seventeenth-century Bologna hanged Ugolino Menzani for sharing the knowledge of a new silk-twisting machine with Venetian weavers (Belfanti 2004, 581), and mechanics in the nineteenth-century Pennsylvania cotton industry relied on family relations to exchange technical knowledge (Wallace 1986, 211–246). Even in the case of the collective invention for the (Cornish) steam engine (Nuvolari 2004) and the iron founders in Britain's Cleveland district (Allen 1983), the benefits of knowledge exchanges are extremely localized.

The shift toward patenting in the nineteenth-century chemical industry, as a result of scientific breakthroughs, creates an opportunity to explore whether patents may in fact help to encourage the diffusion of innovations, as is their intended purpose. A geographic analysis of exhibition data confirms that the shift toward patenting in chemicals was followed by a significant weakening in the geographic localization of inventive activity in chemicals. Data from the US decennial census records between 1840 and 1920 indicate that this shift toward decentralization cannot be explained by changes in the localization of production. While chemical innovation, measured through exhibits, became less localized over time, the localization of chemical production remained relatively stable after 1876. This suggests that patents may in fact help to encourage the geographic diffusion of innovation, even if they may fail to increase overall rates of innovation.

Conclusions

Crippling litigation and patent wars, which stifle innovation, have created an urgent demand for systematic research to guide reforms of the current system. Research topics that address key policy questions cover a broad range of questions, including the following: How strong should patent rights be for early generations of inventors, and how should patents best define free "space" to invent for later generations? Should patent rights be extended to biological innovations, such as living organisms? When does it make sense to violate existing intellectual property rights and allow a new set of firms to enter? What types of property rights, patents, copyright, or contractual types of intellectual property are most effective, and in which industries? And can policies that *weaken* intellectual property rights be effective in encouraging innovation?

Economic history offers an unparalleled wealth of data sets and drastic changes in protection, which allow systematic analyses of the effects of patent policy. For example, the US response to German aggression in World War I and II subjected all enemy-owned property—including patents and production facilities—to confiscation, and all foreign-owned patents became subject to compulsory licensing to US firms. Empirical analyses of these confiscations make it possible to examine the effects of violations of intellectual property on innovation in the country that benefits from foreign-owned property and in the country whose intellectual property is violated. On the opposite extreme, a weakening of antitrust enforcement under the New Deal in the 1930s extended to patents, and allowed competing firms to cooperate—and strengthen the monopoly position created by patents—by sharing patents in so-called pools. These pools have become a popular instrument to mitigate litigation risks today, but empirical evidence from modern pools is limited, and empirical analyses of historical pools can help to inform antitrust policies toward them.[7]

Historical episodes also make it possible to examine the effects of complementary policies—other than patent laws—to encourage innovation. For example, the racial policies of Germany's Nazi Party created an outflow of scientists and high-skilled workers that was largely independent of their realized or anticipated contributions to innovation in sending and receiving locations. This movement of high-skilled workers makes it possible to examine the effects of receiving high-skilled immigrants on US innovation, and to investigate the potential effects of a brain drain of high-skilled professionals, such as medical doctors, on innovation and welfare in the sending country.[8] Empirical analyses of historical policy changes can also improve our understanding of alternative intellectual property regimes, such as copyright laws. These laws, which set the parameters for innovation in a broad range of sectors ranging from literature and music to web design, are intended to encourage creativity, but there is little empirical evidence to date that they fulfill this role. In fact, excessively strong copyright terms may discourage creativity, similarly to excessively strong patents, and hinder the diffusion of new ideas. Research in economic history can shed light on these issues and is critically needed.

Notes

1. See Moser and Voena (2012) for an application and detailed description of the Google/USPTO data on US patents since 1920, and Lampe and Moser (2016) for historical data on patent citations since 1920.
2. In addition, national newspapers, including the UK *Times, Economist*, and *Sun*, and the US journal *Scientific American*, advertised the fairs in regular features (e.g., *The Times*, November 14, 20, December 6, 9, and 30, 1850).
3. To check for bias due to transportation costs, Moser (2012) compares the locations of British exhibits at the Crystal Palace with the locations of all 835 British exhibits at the Centennial Exhibition in Philadelphia in 1876. If transportation costs bias the data toward London, London's share of innovations should be lower at the US fair, when differences in transportation costs within Britain were negligible relative to total transportation costs. London's share of exhibits, however, was almost identical in 1851 and 1876, with 40.7 and 39.1 percent, respectively.
4. The mean area per exhibitor was approximately equal at both fairs, with 0.00118 acres (4.7753 square meters) in 1851 and 0.00125 acres (or 5.0586 square meters) in 1876.
5. Data on prize-winning British exhibits help to shed light on the interaction between the quality of inventions and inventors' decision to use patents. Existing theoretical models indicate that firms may decide to keep important innovations secret because patents require disclosure, which is risky if patents are ineffective at blocking competitors from using a patented invention (Anton and Yao 2004; Horstmann, MacDonald, and Slivinski

1985). Exhibition data, however, indicate that high-quality innovations are slightly more likely to be patented: in 1851, 15 percent of British exhibits that won prizes for exceptional usefulness and quality were patented, compared with 11 percent of average-quality exhibits.

6. Lerner (2000) calculates 1851 patents fees in year 2000 US dollars as $618 for the United States, compared to $37,000 for Britain.
7. See Moser and Voena (2011) for more detail on compulsory licensing under the US Trading with the Enemy Act, and Lampe and Moser (2010, 2012) for detail on patent pools.
8. See Moser, Voena, and Waldinger (2013) for an analysis of the effects of German-Jewish immigrant chemists on US innovation.

Bibliography

Abramovitz, Moses. (1989). *Thinking about Growth: And Other Essays on Economic Growth and Welfare*. Cambridge, MA: Cambridge University Press.

Allen, Robert C. (1983). "Collective Invention." *Journal of Economic Behavior and Organization* 4(1): 1–24.

Andrews, Benjamin E. (1894). *History of the United States*, Vol. 4. New York: C. Scribner's Sons.

Anton, James J., and Dennis A. Yao. (2004). "Little Patents and Big Secrets: Managing Intellectual Property." *RAND Journal of Economics* 35(1): 1–22.

Arora, A., and N. Rosenberg. (1998). "Chemicals: A US Success Story." In *Chemicals and Long-Term Economic Growth*, edited by A. Arora, R. Landau, and N. Rosenberg, 71–102. New York: John Wiley.

Auerbach, Jeffrey. (1999). *The Great Exhibition of 1851: A Nation on Display*. New Haven, CT: Yale University Press.

Belfanti, Carlo. (2004). "Guilds, Patents, and the Circulation of Technical Knowledge: Northern Italy during the Early Modern Age." *Technology and Culture* 45(3): 569–589.

Berichterstattungs-Kommission der Deutschen Zollvereins-Regierungen. (1853). *Amtlicher Bericht über die Industrie-Austellung aller Völker zu London im Jahre 1851*, Vols. I–III. Berlin: Verlag der Deckerschen Geheimen Ober-Hofbuchdruckerei.

Bessen, James, and Michael J. Meurer. (2008). *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk*. Princeton, NJ: Princeton University Press.

Bissell, Don. (1999). *The First Conglomerate: 145 Years of the Singer Sewing Machine Company*. Brunswick, ME: Audenreed Press.

Clark, Gregory. (2006). *A Farewell to Alms: A Brief Economic History of the World*. Princeton, NJ: Princeton University Press.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. (2000). "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)." NBER Working Paper No. 7552.

Coryton, John. (1855). *A Treatise on the Law of Letters Patent; For the Sole Use of Inventions in the United Kingdom of Great Britain and Ireland; Including the Practice Associated with the Grant; To Which Is Added a Summary of the Patent Laws in Force in the Principal Foreign States*. Philadelphia: T. & J. W. Johnson.

David, Paul A. (1994). "The Evolution of Intellectual Property Institutions." In *Economics in a Changing World,* Vol. 1: *System Transformation: Eastern and Western Assessments*, edited by A. Aganbegyan, O. Bogomolov, and M. Kaser, 19–62. London: Macmillan.

Department of Justice and Federal Trade Commission. (1995). *Antitrust Guidelines for the Licensing of Intellectual Property*. Available at https://www.justice.gov/atr/speech/1995-antitrust-guidelines-licensing-intellectual-property.

Department of Justice and Federal Trade Commission. (2007). *Antitrust Enforcement and Intellectual Property Rights: Promoting Innovation and Competition*. Available at https://www.ftc.gov/sites/default/files/documents/reports/antitrust-enforcement-and-intellectual-property-rights-promoting-innovation-and-competition-report.s.department-justice-and-federal-trade-commission/p040101promotinginnovationandcompetitionrpt0704.pdf.

Dickens, Charles. (1925). "A Poor Man's Tale of a Patent." *Houshold Worlds* 2(30): 73–75.

Dutton, Harold I. (1984). *The Patent System and Inventive Activity during the Industrial Revolution, 1750–1852*. Manchester, UK: Manchester University Press.

Godson, Richard. (1840). *A Practical Treatise on the Law of Patents for Inventions and of Copyright: Illustrated by Notes of the Principal Cases, With an Abstract of the Laws in Force in Foreign Countries*. 2nd ed. London: Saunders and Benning.

Habakkuk, H. J. (1962). *American and British Technology in the 19th Century: The Search for Labour-Saving Inventions.* Cambridge, MA: Cambridge University Press.

Haber, Ludwig F. (1958). *The Chemical Industry during the 19th Century: A Study of the Economic Aspect of Applied Chemistry in Europe and North America*. Oxford: Oxford University Press.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper No. 8498.

Haltern, Utz. (1971). *Die Londoner Weltaustellung von 1851: Ein Beitrag zur Geschichte der burgerlich-industriellen Gesellschaft im 19. Jahrhundert*. Munster: Aschendorff.

Horstmann, Ignatius J., Glenn MacDonald, and Alan Slivinski. (1985). "Patents as Information Transfer Mechanisms: To Patent or (Maybe) Not to Patent." *Journal of Political Economy* 93(5): 837–858.

Hounshell, David. (1985). *From the American System to Mass Production, 1800–1932: The Development of Manufacturing Technology in the United States*. Baltimore, MD: Johns Hopkins University Press.

Hounshell, David A., and John K. Smith. (1988). *Science and Corporate Strategy: Du Pont R&D, 1902–1980.* Cambridge, MA: Cambridge University Press.

Khan, B. Zorina. (2005). *The Democratization of Invention: Patents and Copyrights in American Economic Development, 1790–1920*. Cambridge, MA: Cambridge University Press.

Khan, B. Zorina, and Kenneth L. Sokoloff. (1993). "Schemes of Practical Utility: Entrepreneurship and Innovation among 'Great Inventors' in the United States, 1790–1865." *Journal of Economic History* 53(2): 289–307.

Khan, B. Zorina, and Kenneth L. Sokoloff. (1998). "Patent Institutions, Industrial Organization and Early Technological Change: Britain and the United States, 1790–1850." In *Technological Revolutions in Europe*, edited by M. Bergand and K. Bruland, 292–313. Cheltenham, UK: Edward Elgar.

Khan, B. Zorina, and Kenneth L. Sokoloff. (2001). "History Lessons: The Early Development of Intellectual Property Institutions in the United States." *Journal of Economic Perspectives* 15(3): 233–246.

Kingsley, John L., and Joseph P. Pirsson. (1848). *Laws and Practice for All Nations and Governments Relating to Patents for Inventions; With Tables of Fees and Forms, Also an Editorial Introduction with Explanations of Practice and Proceedings Used in Procuring Patents throughout the World*. New York: Kingsley and Pirsson.

Kremer, Michael. (2002). "Pharmaceuticals and the Developing World." *Journal of Economic Perspectives* 16(4): 67–90.

Kretschmer, Winfried. (1999). *Geschichte der Weltausstellungen*. Frankfurt: Campus Verlag.

Kroker, Evelyn. (1975). *Die Weltausstellungen im 19. Jahrhundert: Industrieller Leistungsnachweis, Konkurrenzverhalten und Kommunikationsfunktion unter Berücksichtigung der Montanindustrie des Ruhrgebietes zwischen 1851 und 1880*. Göttingen: Vandhoeck Ruprecht.

Lakwete, Angela. (2003). *Inventing the Cotton Gin: Machine and Myth in Antebellum America*. Baltimore, MD: Johns Hopkins Univeristy Press.

Lamoreaux, Naomi R., and Kenneth L. Sokoloff. (1998). "Inventors, Firms, and the Market for Technology in the Late Nineteenth and Early Twentieth Centuries." NBER Historical Working Paper No. 98. Issued in April 1997.

Lamoreaux, Naomi R., and Kenneth L. Sokoloff. (1999). "Inventors, Firms, and the Market for Technology in the Late 19th and Early Twentieth Centuries." In *Learning by Doing in Markets, Firms and Countries*, edited by Naomi R. Lamoreaux, Daniel M. G. Raff, and Peter Temin, 19–60. Chicago: University of Chicago Press.

Lampe, Ryan, and Petra Moser. (2010). "Do Patent Pools Encourage Innovation? Evidence from the 19th-Century Sewing Machine Industry." *Journal of Economic History* 70(4): 898–920.

Lampe, Ryan, and Petra Moser. (2012). "Patent Pools and Innovation in Substitute Technologies—Evidence from the U.S. Sewing Machine Industry." *RAND Journal of Economics* 44(4): 757–778.

Lampe, Ryan, and Petra Moser. (2016). "Patent Pools, Competition and Innovation. Evidence from 20 Industries under the New Deal." *Journal of Law, Economics, & Organization* 32: 1–36.

Landes, David S. *The Unbound Prometheus: Technological Change and Development in Western Europe from 1750 to the Present*. 2nd ed. Cambridge, MA: Cambridge University Press, 2003.

Lerner, Josh. (2000). "150 Years of Patent Protection." National Bureau of Economics Research (Cambridge, MA) Working Paper No. 7478, January.

Levin, Richard C., Alvin K. Klevorick, Richard R. Nelson, and Sidney G. Winter. (1987). "Appropriating the Returns from Industrial Research and Development." *Brookings Papers on Economic Activity* 18(3): 783–832.

Li, Xing, Megan MacGarvie, and Petra Moser. "Dead Poet's Property—How Does Copyright Influence Price?" *RAND Journal of Economics*. Available at http://ssrn.com/abstract=2170447.

MacLeod, Christine. (1988). *Inventing the Industrial Revolution: The English Patent System, 1660–1800*. Cambridge, MA: Cambridge University Press.

Mokyr, Joel. (1990). *The Lever of Riches: Technological Creativity and Economic Progress.* New York: Oxford University Press.

Mokyr, Joel. (2004). *The Gifts of Athena: Historical Origins of the Knowledge Economy.* Princeton, NJ: Princeton University Press.

Mokyr, Joel. (2009). *The Enlightened Economy: An Economic History of Britain, 1700–1850.* New Haven, CT: Yale University Press.

Moser, Petra. (2003). "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World Fairs." NBER Working Paper No. 9909.

Moser, Petra. (2005). "How Do Patent Laws Influence Innovation? Evidence from 19th-Century World's Fairs." *American Economic Review* 95(4): 1215–1212.

Moser, Petra. (2011). "Do Patents Weaken the Localization of Innovations? Evidence from World's Fairs, 1851–1915." *Journal of Economic History* 71(2): 363–382.

Moser, Petra. (2012). "Innovation without Patents—Evidence from World's Fairs." *Journal of Law and Economics* 55(1): 43–74.

Moser, Petra, and Tom Nicholas. (2013). "Prizes, Publicity, and Patents—Non-Monetary Awards as a Mechanism to Encourage Innovation." *Journal of Industrial Economics* 61: 763–788.

Moser, Petra, and Paul W. Rhode. (2012). "Did Plant Patents Create the American Rose?" In *The Rate and Direction of Inventive Activity Revisited*, edited by Joshua Lerner and Scott Stern, 413–441. Chicago: University of Chicago Press.

Moser, Petra, and Alessandra Voena. (2012). "Compulsory Licensing: Evidence from the Trading-with-the-Enemy-Act." *American Economic Review* 102(1): 396–427.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. (2014). "German-Jewish Émigrés and U.S. Invention." *American Economic Review* 104(10): 3222–3255.

Moser, Petra, and Assaf Zimring. (2012). "Patent Laws, Market Size and Innovation." Available at http://ssrn.com/abstract=2010136.

Mowery, David C., and Nathan Rosenberg. (1998). *Paths of Innovation: Technological Change in 20th-Century America*. Cambridge, MA: Cambridge University Press.

North, Douglass C., and Robert P. Thomas. (1973). *The Rise of the Western World: A New Economic History*. Cambridge, MA: Cambridge University Press.

Nuvolari, Alessandro. (2004). "Collective Invention during the British Industrial Revolution: the Case of the Cornish Pumping Engine." *Cambridge Journal of Economics* 28: 347–363.

Official Catalogue. (1851). *Great Exhibition of the Works of Industry of All Nations*. London: Spicer Brothers.

Olmstead, Alan L., and Paul W. Rhode. (2008). "Biological Innovation and Productivity Growth in the Antebellum Cotton Economy." *Journal of Economic History* 68(4): 1123–1171.

Parton, James. (1867). "History of the Sewing Machine." *Atlantic Monthly*, May.

Penrose, Edith T. (1951). *The Economics of the International Patent System*. Baltimore, MD: John Hopkins University Press.

Rolt, L. T. C. (1970). *Victorian Engineering*. Harmondsworth, UK: Penguin Books.

Rosenberg, Nathan. (1963). "Technological Change in the Machine Tool Industry, 1840–1910." *Journal of Economic History* 23(4): 414–443.

Rosenberg, Nathan. (1969). *The American System of Manufactures*. Edinburgh: Edinburgh University Press.

Rosenberg, Nathan. (1972). *Technology and American Economic Growth*. New York: Harper.

Rothbarth, Edwin. (1946). "Causes of the Superior Efficiency of U.S.A. Industry Compared with British Industry." *Economic Journal* 56: 383–390.

Royal Commission. (1851). *Official Catalogue of the Great Exhibition of the Work of Industry*. Harmondsworth, UK: Penguin Books.

Schiff, Eric. (1971). *Industrialization without National Patents*. Princeton, NJ: Princeton University Press.

Schmookler, Jacob. (1962). "Economic Sources of Inventive Activity." *Journal of Economic History* 22(1): 1–20.

Schmookler, Jacob. (1966). *Invention and Economic Growth*. Cambridge, MA: Harvard University.

Sokoloff, Kenneth. (1988). "Inventive Activity in Early Industrial America: Evidence from Patent Records, 1790–1846." *Journal of Economic History* 48(4): 813–850.

Tallis, John. (1852). *Illustrated London: In Commemoration of the Great Exhibition of All Nations in 1851*. London: John Tallis.

Thomson, Ross. (2009). *Structures of Change in the Mechanical Age: Technological Innovation in the United States, 1790–1865*. Baltimore, MD: John Hopkins University Press.

Visitor's Guide to the Centennial Exhibition and Philadelphia. (1875). May 10 to November 10, 1876. Authorized by the Centennial Board of Finance, and approved by the Director-General. The Only Guide-Book Sold on The Exhibition Grounds. J. B. Lippincott.

Wallace, Anthony F. C. (1986). *Rockdale: The Growth of an American Village in the Early Industrial Revolution*. New York: Knopf.

CHAPTER 20

TRANSPORTATION IN AMERICAN ECONOMIC HISTORY

JEREMY ATACK

In 1816, the infamous "year without a summer," American farmers found a wide range of prices being quoted for their crops. In July the price of grain ranged from 75 cents to $2.10 a bushel. However, while the product being sold was the same, the location of the markets differed. Farmers selling in Cincinnati, the low-priced market, would have loved to sell in Philadelphia, the high-priced market, while consumers buying in Philadelphia would have loved to buy in Cincinnati (Cole 1938).[1] To do so, the grain would have to be moved from where it was produced to where it was to be consumed. In this case, 500-plus airline miles stood between the two markets but, of course, there were no airlines then. Instead, the journey involved a difficult, expensive, roundabout trip.[2] It also involved considerable risk to life or property.[3] When production and consumption are only rarely co-located, moving people and commodities represented an absolutely vital service in the economy.

So long as travel and transportation expenses kept markets separated, price differentials could persist. Indeed, on the eve of the Civil War, Philadelphia grain prices were still higher than those in Cincinnati, but the differentials had narrowed dramatically as these cities were linked by more and better transportation services (Berry 1943; Cole 1938; North 1961) (see fig. 20.1). Such differentials would narrow further after 1865 (Atack 2013; Harley 1980), and this pattern is observed between every city pair and every good as communities became integrated into local, regional, national, and (eventually) international markets (Jacks 2005, 2006; Jacks, O'Rourke, and Williamson 2011; Slaughter 1995) (see fig. 20.2). These price movements represented a profound change, albeit one that was reversible by war and other disruptions to trade (Hynes, Jacks, and O'Rourke 2012).

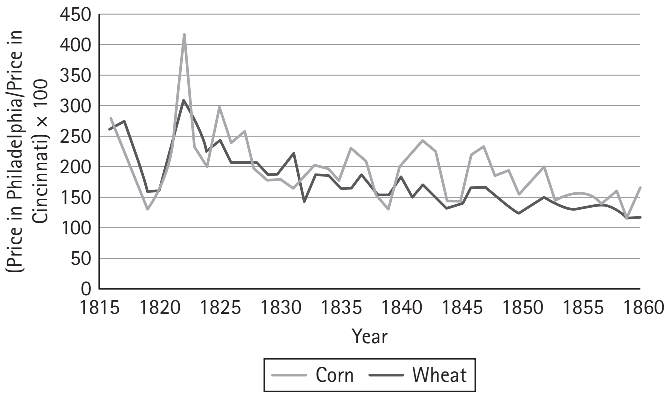

FIGURE 20.1 Relative price of corn and wheat in Philadelphia and Cincinnati, 1816–1860
Sources: Cole (1938, passim). See also http://www.vanderbilt.edu/econ/cipr/cole-historical-data.html. http://centerforinternationalprices.org/micro-price-data/cole-historical-data/.

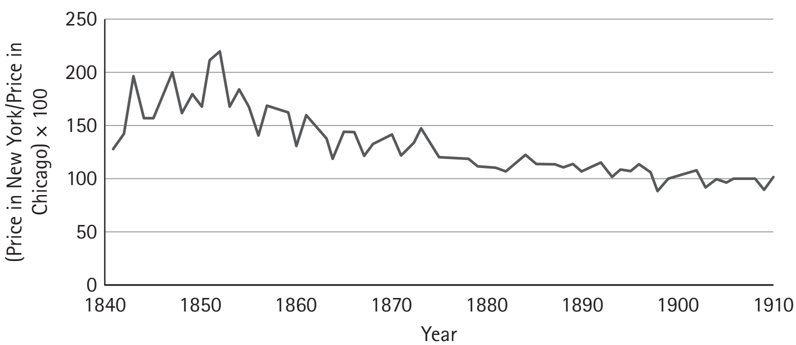

FIGURE 20.2. Relative price of wheat in New York and Chicago, 1841–1910
Sources: New York: Ronk (1936, Table 10, 46–47); Chicago: Boyle (1922, Appendix, 69–71).

A key issue to understand the development of transportation services in America, as well as the revolutionary impact of new modes of travel and transportation, is that as vital as transportation has been, it is undersupplied by the free market due to positive externalities in various guises that reduce the private returns below the social returns. Consequently, governments have often intervened in its supply (Callender 1902; Goodrich 1967; Haney 1968). Some of these positive externalities arise because the gains from reductions in shipping costs are divided between producers and consumers rather than being captured by the shipper, who cannot perfectly price discriminate; some arise from private property rights; yet others from network effects and other sources. Public sector efforts to remediate those externalities, such as land grants to railroads, however, created other problems that, in turn, demanded additional public policy solutions like the Hepburn Act to prevent railroads from gaining an unfair advantage over others in exploiting the mineral rights that they now owned as a result of those land grants.

Water-Borne Transportation

Shipment by water was—and remains—the cheapest mode of transportation. European explorers made extensive use of this continent's rivers and lakes. The early settlements were clustered along the Atlantic seaboard with easier access to the mother country and to other settlements. Indeed, the seas were such important avenues of commerce that specific legal rules—Admiralty Law—were adopted. These are assigned to the federal courts under the US Constitution.

A key principle of Admiralty was "navigability"—a term that initially applied only to those waters subject to the ebb and flow of the tides (Story 1833). However, the Northwest Ordinance of 1787 extended the principle to navigable waters emptying into the Mississippi and the St. Lawrence, declaring them to be "common highways and forever free" (US Continental Congress 1787). Variations on this principle were subsequently enshrined in congressional acts that admitted new states to the Union. But what constituted navigability remained unspecified and thus contested because a "common highway" was a public good—subject to issues such as nonexcludability, congestion and wastage, and impairing private property rights.

When America was settled, ocean-going vessels were universally wind powered. Most were multipurpose vessels and, given the prevalence of war and piracy, were often armed for self-protection or to serve as military auxiliaries if needed. These considerations led to smaller cargo capacities and larger crews than strictly necessary to operate the vessel. Both factors raised costs. The Dutch gradually adopted less military-inspired designs that lowered costs by emphasizing speed, maneuverability, and carrying capacity. In response, other nations sought to exclude foreign ships to protect their inefficient domestic maritime industries. Nevertheless, as trade routes became more regular and familiar and as the needs for protection declined, productivity in ocean shipping gradually and averaged 0.45 percent per year (see North 1958, 1968; Shepherd and Walton 1972; Walton 1966). Although very modest, these productivity gains accumulated throughout the colonial and early national periods.

By the early nineteenth century, the pace of productivity growth in ocean shipping had accelerated to over 3 percent per year, driven by the switch from wind to steam power (especially the compound marine steam engine), and by the switch from paddlewheels to propellers (North 1958). But most important factor was the growth of trade itself, especially the reduction in back-haulage excess capacity.[4] As a result, North's American export freight rate index (fig. 20.3) had fallen to around 30 (1815 = 100) by 1830, was in the 20s by the 1850s, and had fallen below 10 on the eve of World War I.

In the early twentieth century, the driving forces for continued productivity growth were the adoption of the marine diesel engine, which decreased crew size while increasing cargo space and vessel range; increased vessel sizes, which realized scale economies; increased use of bulk carriers (specialized vessels designed for a single type of cargo like grain, ore, or oil, rather than a mix of general cargo shipped in bags, barrels,

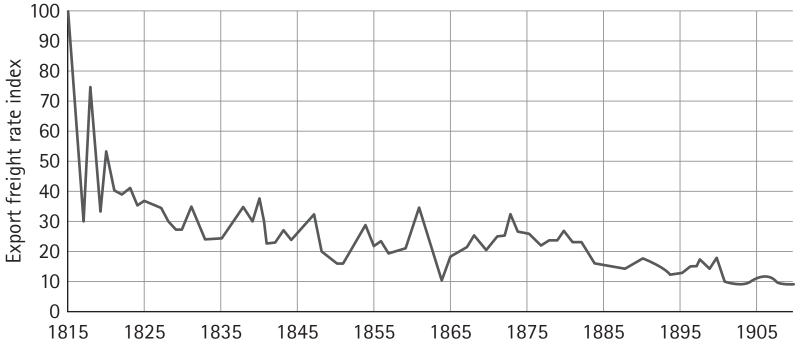

FIGURE 20.3. North's US export freight rate index (1815 = 100)
Source: North (1958, 554).

or other small unitized containers); and better port facilities. An even more important source of productivity growth was the adoption of standardized shipping containers beginning in 1955. These large reusable intermodal steel containers were shipped sealed from the source (thereby minimizing theft and damage), loaded and offloaded mechanically (simplifying handling), and could be tracked throughout their travels (Levinson 2006). By 1970, container ships delivered goods at half the cost of conventional vessels, primarily because of dramatically reduced cargo-handling costs (United Nations 1970). These costs have continued to decline as vessel sizes have grown and port facilities have developed to meet the specific needs of container ships (Hummels 2007).

The Great Lakes, as large inland seas, generally saw the same changes in shipping as ocean transportation (Williamson 1977). However, Great Lakes shipping faced an additional constraint on vessel size, particularly of bulk freighters, from the size of the lock chambers between lakes. For example, the Welland Canal around Niagara Falls has been enlarged many times but still remains smaller than the Soo Locks between lakes Superior and Huron. Moreover, early traffic on the Great Lakes was heavily influenced by the New York and Ohio canals that connected to it. For example, Cronon (1991) estimates that more than 90 percent of Chicago's grain trade went east by Lake in the 1850s, and the share remained around 50 percent in the 1870s. Reflecting these productivity gains in lake shipping, Chicago and New York wheat prices converged at a rate of about 1 percent per year, declining from a ratio of 2:1 in 1841 to rough parity by 1914 (Atack 2013; Boyle 1922).

Along rivers and streams, canoes were paddled and small boats were rowed or poled. Crudely built rafts could drift downstream. The desire for more reliable and less strenuous power sources led to experiments applying steam to river and coastal navigation, with the first successful demonstration of a steam-powered boat being made by John Fitch on the Delaware River before delegates to the Constitutional Convention in 1787. While his design was clumsy (the engine powered banks of oars via cranks), Fitch operated America's first steamboat service across the Delaware River between Philadelphia and Burlington, New Jersey.

Other steamboat designs would prove more successful. Robert Fulton's paddlewheel steamboat, the *Clermont*, operated on the Hudson River between New York City and Albany beginning in 1807, and New York State granted Fulton and his investors (including Robert Livingston) a monopoly on steam navigation in the state's waters. They subsequently sought similar privileges from other states, but only the Orleans Territory (Louisiana) complied. Indeed, New Jersey granted exclusive rights to steam navigation in its waters to a different group. Litigation between these rivals eventually forced the US Supreme Court in *Gibbons v. Ogden* to clarify that the Constitution's Commerce Clause included navigation (US Supreme Court 1824).

Early steamboats followed traditional sailing vessel designs and were built with massive timbers, broad-beamed and deep-hulled, and carrying their cargo below deck. Power was supplied by large, but fuel efficient, low-pressure steam engines of Boulton & Watt design. The first steamboat on the Ohio and Mississippi, the *New Orleans*, built by Fulton and his partners in Pittsburgh in 1811, followed these same principles.[5] However, these would prove suboptimal for the Western rivers, and a unique design better suited to shallower, swifter water and the primitive nature of the trade emerged. Like 30 percent of steamboats operating on the Western rivers, the *New Orleans* was soon sunk by one of the many hazards to navigation on the rivers, and the average boat survived less than five years (US Senate 1852). This rapid turnover meant that the capital stock was young and quickly incorporated successful new designs. Those designs used smaller timbers with cross-bracing for strength, lighter decking, and flimsy superstructure to help reduce vessel draft. Cargo was placed on the deck, as was a lighter, high-pressure steam engine, capable of sudden bursts of speed, and the boats became longer relative to their breadth, allowing use on constricted waterways (Hunter 1949). Moreover, growing familiarity with the rivers, the adoption of rights of way, the regularization of operation, and the removal of some hazards to navigation allowed for higher speeds and nighttime travel. Each enhanced productivity and reduced travel time and shipping costs (fig. 20.4) (Haites, Mak, and Walton 1975; Hunter 1949). While the *New Orleans* had

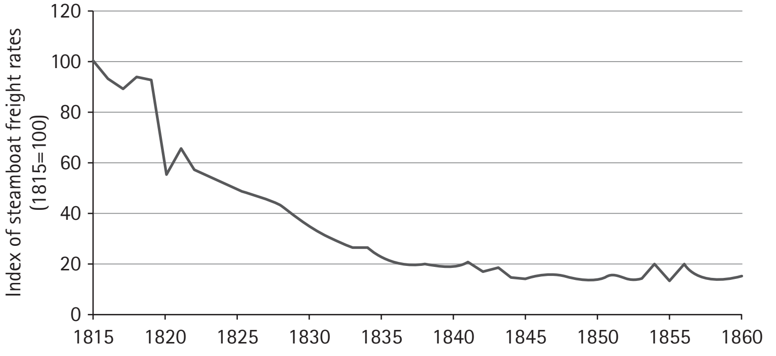

FIGURE 20.4. Western river steamboat shipping costs, 1815–1860
Source: Mak and Walton (1972, 639).

taken 82 days to reach New Orleans, Hunter (1949) reports that by the 1850s travel times had been cut to about 10 days, and the fastest boats made the trip in less than 4.5 days.

As more and better steamboats operated on the rivers, freight rates fell sharply in real terms, especially on upstream traffic, which was especially arduous and time-consuming. For example, Hunter (1949) reports that a trip by keelboat from New Orleans to Louisville around 1820, against the current, typically took three to four months, whereas a steamboat took less than two weeks at that time. Consequently, the steamboat quickly drove keelboats out of business. Flatboats and rafts were still able to compete on downstream routes on the lowest value-weight, least time-sensitive cargo. Indeed, steamboats reduced a major cost of flatboating—the upstream trip (Haites, Mak, and Walton 1975; Mak and Walton 1973).

During the Civil War, both sides blockaded the Western rivers, cutting off traffic. By the time the war ended, railroads extended throughout the Mississippi valley, diverting the higher valued, more time-sensitive cargo onto the more direct rail routes.[6] The rivers, however, remained important for grain and the like. Increasingly, postwar traffic was carried by multi-barge tows that moved far more cargo at much lower cost than a single steamboat. In the late 1940s, a single tug and seven barges carrying 4,000 tons of cargo represented an investment of about $100,000; a steamboat cost at least twice as much and carried only half the cargo (Hunter 1949). Today's larger tows on the lower Mississippi average thirty to forty barges with 40,000+ tons of cargo.

The first canals were built relatively late in the United States, beginning around 1800, because of low population density and the high cost of capital. None was successful, and they thus failed to spark imitation. Things changed with the Erie Canal. Although ridiculed when first proposed, it proved to be an immediate financial success.[7] Sections of the canal were opened as completed: the first between Rome and Utica in 1819, then Utica to Syracuse in 1820. Traffic on each opened section spurred construction of the next section, and almost $1 million in tolls was collected while the canal was still incomplete (Whitford 1906).

To help finance construction, New York had guaranteed the canal bonds, but never had to make good on the promise as the canal proved such a resounding financial success. It also reshaped trade flows (Condit 1980; Whitford 1906). For example, in the mid-1830s, over 80 percent of Midwestern freight went through New Orleans, but by 1860 Midwestern trade via the Erie to East Coast ports easily exceeded that through New Orleans (Haites, Mak, and Walton 1975). No other canal would prove as successful as the Erie; most were dismal failures, despite the best efforts of their promoters.[8] Nevertheless, tens of millions of dollars were plowed into 4,250 miles of canals from a mix of public and private sources (see Goodrich 1960).[9] Sunk costs were high relative to highways, averaging three to six times as much per mile as properly graded and constructed turnpikes. Canals also proved vulnerable to acts of nature, for example, when the rivers that maintained a canal's summit level flooded, washing them out.

Three canal building phases have been identified. Each was sparked by some encouraging promise or notable success, albeit at escalating costs; each was ended by a crisis or more sober reappraisal. The first wave was sparked by the success of the Erie, focused on the East Coast, and ended in 1834 with the disappointing results from Philadelphia's efforts to emulate the Erie's success with the Pennsylvania Mainline Canal to Pittsburgh. In that first wave, almost 2,200 miles were built. The second phase began with Ohio's ambitious plans to link the Ohio River and Lake Erie, was concentrated in the Midwest, and foundered with the overly ambitious Wabash and Erie Canal, which bankrupted the state of Indiana. The third cycle was less ambitious, encouraged by the opening of the Illinois and Michigan Canal, focused on shorter links and feeders to existing routes, and ended with the Panic of 1857 and what was seen as the inevitable success of railroads (Goodrich 1961; Segal 1961). In particular, canals competed directly with railroads both in terms of construction costs and topographical requirements, but railroads provided faster service year round, while canals provided cheaper service but were closed in winter.[10]

Overland Transportation

Roads were nonexistent in pre-settlement America since the indigenous peoples did not use the wheel as a mode of conveyance. Consequently, early settlers made use of narrow native and animal trails, even though these did not connect directly to points of importance to the European settlers. Moreover, forests covered most of the eastern half of the continent even after the Revolution, prompting one contemporary traveler to remark, "I scarcely passed three miles together, through a track of unwooded or cleared land. I always found the roads, or rather the paths, bordered and obscured by a copse or forest . . ." (Volney 1804). It took many years before these were replaced by wider roads that accommodated a wagon and team and connected centers of European population, commerce, and political power. Despite its name, even the "Great Wagon Road," extending southwestward from Philadelphia, through Lancaster, York, and the Shenandoah Valley to Augusta, Georgia, a distance of over 700 miles, was unsuitable for wagons during the colonial period except along its northern stretches (Rouse 1973).

Until properly paved, roads were often impassable in bad weather and were easily damaged by heavy traffic.[11] So long as road use was free and access unrestricted, no individual had an incentive to maintain them or bear the costs of constructing and maintaining roads elsewhere that he could not access. Consequently, early long-distance roads—especially those through sparsely settled land or fulfilling some strategic need—generally became the responsibility of the central government. King Charles II, for example, ordered the construction of a road—the King's Highway—from Boston to Charleston as a means of unifying his American colonies. Similarly, early in the French and Indian War, General Braddock had a road cut across the mountains from Fort Cumberland (Maryland) toward Fort Duquesne (what is now Pittsburgh). The eastern part of this would subsequently become a part of the National Road begun in 1811 (US Senate 1808).[12] Further west, the Natchez Trace was widened by the US military and proved a vital transportation route between Nashville, Tennessee, and Natchez, Mississippi, in the War of 1812. Similar concerns prompted the improvement of roads in the twentieth century, most clearly in construction of the National Defense (a.k.a. Interstate) Highway System.

Few public roads, other than the National Road, were properly paved before the automobile age, limiting the loads and trip distances of horse-drawn wagons.[13] Instead, it fell to the private sector to provide all-season roadways for which tolls were charged. Most turnpikes were built in the Northeast and Midwest under license from the state governments as joint stock corporations. Between 1792 and 1845, over 1,500 turnpike charters were issued, and they made up almost one-half of all incorporations in Pennsylvania before 1830 and more than a quarter in New York, New Jersey, and Connecticut. Most were locally financed, and their stockholder lists "reveal a web of neighbors, kin, and locally prominent figures voluntarily contributing to what they saw as an important community improvement" (Klein and Majewski 2008). Tolls were often regulated by the state and exempted local traffic and use for "public purposes"—including going to or from church—and were easily evaded. Revenues thus often failed to cover maintenance and did not generate an adequate return on capital.

Turnpikes, like most canals, turned out to be mediocre investments in terms of their dividends, but because most investors were property-owning locals, they were able to capture a good fraction of the benefits that the improvement generated. Even after railroads and canals came along, turnpikes continued to supply transportation services, complementing these new modes by moving goods and people to them more efficiently (Callender 1902; Klein 1990; Klein and Majewski 1992). Indeed, the number of turnpike incorporations in Ohio did not peak until the early 1850s, well into the railroad era (Evans 1948). But, as canals and railroads siphoned off the more profitable traffic, many turnpikes fell into disrepair, were abandoned, or became a part of the state's highway system.[14]

The political economy of roads began to change when the bicycle craze swept through urban America in the late 1870s. The earliest bicycles, with their solid rubber tires, illustrated their nickname "boneshaker," especially on deeply rutted, often muddy, gravel roads. Pneumatic tires help cushion the ride, but smoother roads were the key. Urban cycling enthusiasts allied with the nation's farmers to lobby for better roads, beginning with the League of American Wheelmen in 1880. Their campaign met with some modest success, and the US Department of Agriculture in 1893 established an Office of Road Inquiry to disseminate information about the quantity and quality of roads. According to that Office, in 1904, there were about 2.15 million miles of public roads in the United States. Only 154,000 miles of these were surfaced; the rest were dirt and not suitable for year-round travel, especially by internal-combustion-engine-powered vehicles.

Improvements came only very slowly; one measure of how slowly is that the per ton-mile cost of shipping by road was essentially unchanged at 25 cents over the course of the nineteenth century (US Senate 1904). Even more dramatic is Duncan et al.'s (2003) description of the first transcontinental automobile trip when, in 1903, Dr. Horatio Jackson set out from San Francisco in his open two-cylinder Winton Tourer to win a $50 bet. He would arrived sixty-three days later in New York having been forced to take the same trails in the Western half of the country as pioneers before the Civil War. His trip, including gasoline, food, hotels, parts and repairs, cost $8,000—over $200,000 at today's prices.

Early automobiles were relatively expensive novelties until the development of cheaper automobiles like Ford's Model T (1908–1927) made them a mass-market product. As auto traffic grew, so too did complaints regarding the state of the nation's roads. Little, however, was done because most road funding was local and limited, but the needs were national and greater. Change required structural reform in highway funding and the redefining of responsibilities between local, state, and federal authorities. This would take time—and a national emergency.

In 1916, with World War I underway in Europe, Congress passed the Federal Road Aid Act, appropriating $75 million of matching funds spread over five years, for rural public post roads. A year later, every state had a highway agency to administer those funds and develop highways plans, but America's entry into World War I slowed progress. Indeed, after the war ended and sixteen years after the first transcontinental road trip, a military convey set out from Washington, D.C., for San Francisco to test whether the nation's roads could handle the transportation needs of a "modern" war, using the so-called Lincoln Highway, one of two early transcontinental routes suggested for motor vehicles.[15] Unfortunately, it was slow going, especially west of Iowa, arriving in San Francisco sixty-two days after leaving Washington. Over the entire route, the convoy averaged less than sixty miles in their ten-hour day.

Following this disappointing performance, Congress not only renewed but expanded its program of highway matching funds in 1921, and the Bureau of Public Roads asked the Army to provide a list of roads of "prime importance in the event of war." The resulting map, presented to Congress by General John Pershing, bears an uncanny resemblance to the present-day interstate highway map. Most of the roads on Pershing's map were eventually built, many as a part of the interstate system begun in 1956 during the Eisenhower administration. It is probably not coincidental that Eisenhower had participated in the 1919 transcontinental convoy.

Better roads became ever more urgent as automobile traffic grew. In 1920, 1.9 million automobiles rolled off the assembly lines, and there were a total of 8.1 million passenger cars registered in that year. By the end of the decade, car registrations had almost tripled (23.1 million), plus there were 3.5 million trucks. Distances traveled, however, grew more than fourfold, implying that people were traveling greater distances or more frequently. At the start of the decade, the nation had 369,000 miles of surfaced rural roads and streets. By its end, there were 662,000 paved miles, and by the outbreak of World War II, the surfaced rural roads and streets stretched almost 1.4 million miles (Carter et al. 2006).

While paved roads were suitable for year-round use by heavier traffic, their serpentine routes slowed travel. Intersections were frequent since, as public roads, the public had the right to connect as best suited their individual needs. Congestion in and around towns slowed long-distance traffic. The solution lay in the construction of an entirely new, limited-access network of roads that bypassed towns and other bottlenecks: the

Interstate Highway System. As of 2013, this system has 47,856 miles of roadway (out of about 4 million miles of public roads) and accounted for about a quarter of all vehicle-miles traveled. Data assembled by Klein and Majewski (2008) in table 20.1 suggest that the early nineteenth-century investment in private turnpikes was of a comparable magnitude in many states, although the lack of a coherent, coordinated network and coordination probably reduced the effectiveness of those turnpikes.

The mechanical advantages of running wheeled vehicles along a track were well known, having been used in mines and quarries. Instead, the railroad's key technological advance was the use of a steam locomotive. This was first successfully demonstrated in 1825 by the English engineer, George Stephenson, with the twenty-five-mile Stockton to Darlington railway, when a locomotive hauled a load of eighty tons and at one point reached a speed of twenty-four miles per hour—far beyond the capacity of any horse-drawn vehicle (Davies 1975, 2004). From there, the innovation quickly spread to America, where the leaders of Eastern port cities were desperately searching for a competitive answer to the Erie Canal and the rise of the port of New York. Philadelphia followed New York's lead and put its faith in the Mainline Canal. Baltimore's city

Table 20.1 Cumulative Turnpike Investment (1800–1830) as Percentage of 1830 GNP

State	Cumulative Turnpike Investment, 1800–1830 ($)	Cumulative Turnpike Investment as Percent of 1830 GDP	Cumulative Turnpike Investment per Capita, 1830 ($)
Maine	35,000	0.16	0.09
New Hampshire	575,100	2.11	2.14
Vermont	484,000	3.37	1.72
Massachusetts	4,200,000	7.41	6.88
Rhode Island	140,000	1.54	1.44
Connecticut	1,036,160	4.68	3.48
New Jersey	1,100,000	4.79	3.43
New York	9,000,000	7.06	4.69
Pennsylvania	6,400,000	6.67	4.75
Maryland	1,500,000	3.85	3.36
TOTAL	24,470,260	6.15	4.49
Interstate Highway System, 1956–1996	330 billion	4.15 (1996 GNP)	

Notes and sources: Table from Klein and Majewski (2008, Table 3), citing Durrenberger (1931, 61) for Pennsylvania turnpike investment; Taylor (1951, 210–211) for New England's turnpike investment; Fishlow (2000, 549) for data for New York, New Jersey, and Maryland. State GDP data from Bodenhorn (2000, 237). Figures for the cost of the Interstate Highway System can be found at http://www.publicpurpose.com/hwy-is$.htm.

leaders, however, recommended the construction of a railroad and, by February 1827, had secured a corporate charter from Maryland for the Baltimore and Ohio Railroad. Construction began in 1828, and the first passengers rode its rails in May 1830, albeit in horse-drawn carriages. The first steam-powered trip took place several months later (Dilts 1993). Boston's leaders arrived at the same solution as Baltimore, creating the Boston and Worcester Railroad with the goal of crossing the Berkshire Mountains and intercepting traffic on the Erie Canal at Albany (Guild 1847). Trade diversion lay at the heart of these early decisions to build railroads. Ultimately, however, trade creation probably played a much more important role in the success and economic impact of the railroad (Atack, Haines, and Margo 2010; Atack et al. 2010; Atack, Margo, and Perlman 2012; Atack, Jaremski, and Rousseau 2014; Atack, Jaremski, and Rousseau 2015; Atack and Margo 2011, 2012).

Railroad mileage expanded rapidly. By 1840, there were over 2,800 miles of railroad in operation, compared with just under 3,300 miles for canals, and while canals added only about 400 miles in the 1840s, rail mileage grew by over 5,000 miles (Chandler 1965). In the 1850s, more miles of canal were abandoned than were built, but railroad mileage tripled. Much of the new construction was in the Midwest, linking Midwestern cities and farms to the East (Fishlow 1965). The first railroad linking to the East Coast (the Michigan Southern) entered Chicago in 1852, and by 1858 the Pittsburgh, Fort Wayne, and Chicago Railroad was providing the first all-rail route to the Eastern seaboard from Chicago (Cronon 1991). Moreover, construction began in 1851 on the Illinois Central Railroad, which ran the length and breadth of the state and, when completed in 1856, was then the longest railroad in the world (Brownson 1915; Gates 1934; Stover 1975; Sutton 1951).

Railroad construction paused during the Civil War, but surged again in the postwar era. The first transcontinental railroad, the Union Pacific, opened in 1869. Others would follow, providing a series of trunk lines to which other branch railroads could connect.[16] Railroad construction itself reached its historic peak in 1872 when almost 7,500 miles of track were built—almost as much as had existed in the entire country in 1849. Thereafter, construction activity varied from year to year but stayed above 1,400 miles per year from 1866 through 1914, with several years when more than 5,000 miles were built (Carter et al. 2006; Chandler 1965).

Many economists (Fels 1959; Rostow 1948, 1963; Schumpeter 1939) have viewed railroad construction as the driving force for the aggregate economy. Each mile of single track, for example, represented at least 350 tons of iron or steel, and employed thousands, moving earth and ballast to grade the rights of way. Certainly, there is no denying that railroads were major investments and had a profound effect upon local and regional economies from their inception. Just a mile of track cost more than most fully equipped steamboats and more than ten times the average investment in a farm and twenty or more times the capital invested in the typical manufacturing firm at the time. However, as Fogel (1962, 1964, 1979) has argued, their historical importance to the aggregate economy is more nuanced.

By 1860, railroads had become the nation's first billion-dollar industry. Much of this capital was raised through publicly traded debt rather than equity. However, whereas roads and the early canals were typically financed locally, railroad financing was usually regional, national, and even global. Their securities played a major role in the broadening and deepening of financial markets. Indeed, the New York Stock Exchange was primarily a debt market and was dominated by railroad securities until the end of the nineteenth century, when waves of railroad bankruptcies began to tarnish the luster of those debt contracts as secure investments (Adams, Adams, and Walker 1871; Mixon 2008).

This financial leverage increased the returns to railroad equity holders in good years, but increased the risk of default in bad ones unless sufficient cash could be generated to meet short-term obligations, especially as new lines were extended into increasingly marginal areas. Thus, railroads and their securities played a major role in the financial crises of 1857, 1873, and in 1893 (when a quarter of the nation's railroads were in bankruptcy) (Fels 1959; Thorp, Thorpe, and Mitchell 1926). Moreover, the concentration of railroad equity holdings made railroads vulnerable to takeovers and consolidation, as in 1899–1900, when one-sixth of railroad mileage in the country was merged and the Pennsylvania Railroad eventually absorbed 600 other lines.

The cost structure of railroads with high ratios of fixed to variable costs, characteristics often associated with "natural monopolies," also meant that price wars during economic downturns could be savage. In the short run, a railroad could reduce its price so long as it covered marginal costs, but its long-run viability required that all costs (including opportunity cost) be covered. Collusion, bankruptcies, and takeovers were also common, and each prompted yet more state involvement. Railroads also needed the support of the state to cross public highways, exercise eminent domain to avoid being held hostage by property owners, and deal with liability issues. The states, for their part, feared hostage-taking by the railroads so that many of the charters contained time limits that granted a road limited exclusive rights to a particular route only if they completed the road within a set time.[17] Moreover, the whole process of granting charters and railroad construction was rife with corruption.[18]

Railroads differed from most other businesses in that their workforce and production were scattered across a very large geographic area and thus not easily subject to continuous supervision. Railroads therefore quickly evolved sophisticated reporting methods for monitoring performance, assigning responsibilities, and tracking expenses, including the multidivisional firm (e.g., passenger vs. freight, track maintenance and repairs vs. operations, etc.) and cost accounting (Chandler 1965; Louisville and Nashville Railroad and Fink 1874; Ramsdell, McCallum and the Erie Railroad Company 1856).

Steam locomotives had a voracious appetite for water and needed to stop frequently, perhaps every seven to ten miles, unless the train had a water tender (in which case its range might be extended to 100–200 miles). Refueling depots were also required along the track since engines consumed about one pound of fuel (initially wood, but increasingly coal after 1850) for every five to ten pounds of water. Ultimately, these mechanically complex, polluting behemoths were replaced by simpler, more efficient diesel engines that drove electric generators for electric motors that actually provided the tractive power. As these proved reliable and as their power increased, diesels entered mainline use, powering passenger trains in the 1930s (Stover 1970) before gaining acceptance as freight engines (Lytle 1968; Mansfield 1963; Marx 1976). While seemingly cumbersome and roundabout, this form of electric traction had the advantage of avoiding the cost of electrifying hundreds of thousands of miles of track—a major consideration.

Electric traction had already proved itself in city tramways and subways, as Most (2014) documents, and then spread to the suburbs with interurban electric railroads—indeed, it helped make those suburbs feasible. Interurban railways were quickly spread despite the very modest success of the pioneers (Bogart 1906; Hilton and Due 1960). By 1900, about 2,100 miles of interurban electric track had been built. On the eve of World War I, interurban trains ran on 15,500 miles of track concentrated in seven states, most of them in the Midwest (Hilton and Due 1960). Thereafter, however, the interurban electric declined in popularity as access to private automobiles increased.

The cost of shipping by rail has fallen dramatically and consistently throughout, due to the substantial productivity growth associated with the adoption of various inventions and improved procedures (Caves, Christensen, and Swanson 1981; Fishlow 1966). One of the first American railroad inventions were equalizing levers that spread the locomotive's weight equally across all driving wheels regardless of track contours. This maximized the pulling power of the locomotive while simultaneously reducing wear and fatigue on any given section of track. Similarly, American railroad engineers were pioneers in the application of movable trucks—the smaller leading wheels on locomotives—to help guide the train around curves. Innumerable improvements followed in boiler, flue, and fire grate design, automatic oilers, axle bearings, couplers (especially the adoption of standardized and automatic couplers in 1873), and brakes (especially the Westinghouse fail-safe air brakes in 1869). Track was also improved, first by the switch from strap iron to the "T" rail, then with the substitution of steel rails for iron after the Civil War. Each allowed for an increase in train weight and speed while reducing maintenance and lengthening the track-replacement cycle. Railroad ties have been improved, with treated wood replacing untreated wood and concrete eventually replacing both. Ballasting has improved. Signaling and the adoption of a universal standard time made for the more efficient use of single track. Stations have been unified ("Union stations"), as have freight yards. Moreover, whereas there were once many incompatible track gauges, these were eventually standardized in the 1880s (Puffert 1991, 2000).

Productivity gains in railroading have continued throughout the twentieth century (Caves, Christensen, and Swanson 1981) (see fig. 20.5). Not only did electric motors replace steam engines, but, beginning in the 1990s, AC-motors replaced DC-motors in diesel-electric engines, providing still-better performance. Moreover, as locomotives have grown more powerful, train size has increased, allowing railroads to realize additional scale economies. At the same time, crew requirements have been reduced with improved systems of braking and computerized systems managing, for example, multiple engines from a single location and controlling power delivery to limit wheel slip and optimize power delivery. Freight handling has been streamlined with the adoption of intermodal systems like the use of shipping containers and "piggyback" systems.

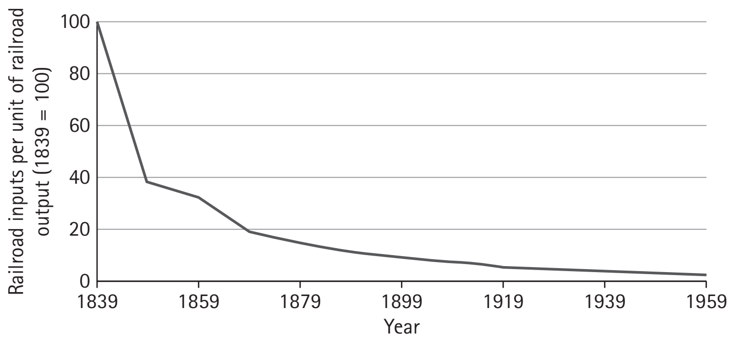

FIGURE 20.5. Railroad inputs per unit of output

Note and source: Calculated as the inverse of Fishlow's productivity index (1966, Table 10, 626) and rebased so that 1839 = 100 to facilitate direct comparison.

Coordination of operations—absolutely essential in a largely single-track rail system like that in the United States—however, has been more elusive. Mergers have helped—for example, BNSF (Burlington Northern and Santa Fe, where the Burlington Northern and the Santa Fe railroads were, themselves, the products of hundreds of railroad mergers over a 160 year period)—but many issues remain unresolved, such as whether freight or passenger traffic has priority. The government has also promoted cooperation and integration through regulation such as the Interstate Commerce Commission and legislation such as the Esch-Cummins Act of 1920. Sometimes the government has even intervened directly, as during World War I when the United States Railroad Administration (1917–1920) briefly took over the nation's railroads, and again in the 1970s when Congress created Amtrak and Conrail in the face of the imminent collapse of rail service in the eastern half of the country.

Air Transportation

A relatively long time elapsed between the first powered flight (1903) and the first commercial flight in January 1914, when the former mayor of St. Petersburg, Florida, paid $400 for a twenty-one-mile, twenty-three-minute trip across the bay to Tampa in a two-seat flying boat—the inaugural flight of the St. Petersburg-Tampa Air Line.[19] This was much less time than the alternatives: two hours by boat, four hours by train, and twenty hours by automobile. Nonetheless, early aircraft were slow, had limited range, were uncomfortable with open, unpressurized cockpits that exposed passengers to the buffeting wind and cold, flew low in the turbulence, and could not fly in bad weather or at night. The novelty of flight quickly wore off! Not surprisingly, commercial aviation was slow to take off, despite the flood of trained pilots produced by World War I and their barnstorming displays around the country. Beginning in 1918, the federal government decided to subsidize airmail delivery, and this played a central role in the emergence of commercial aviation. Early airmail contracts were awarded to the private sector. After several accidents, the contracts were reassigned to the US Army, but it proved unreliable, so by the mid-1920s the Postal Service had its own fleet of planes and routes.

In 1925, however, Congress passed the Contract Air Mail Act (the Kelly Act), directing a return to private contracts beginning with feeder routes and at rates that were attractive. A 1926 amendment to the Kelly Act set the subsidy at $3.00 per pound for the first 1,000 miles and 30 cents per pound for each hundred miles thereafter. Carriers thus preferred shorter routes under the 1,000 mile limit, and many of the early successful contractors would eventually evolve into major airlines.[20] The federal government further intruded into aviation with the Air Commerce Act of 1926, creating the Aeronautics Branch within the Department of Commerce, with regulatory authority over pilot certification, the airworthiness and safety of planes, investigating accidents, and charged with improving facilities and providing navigation aids—including lighted runways.[21]

The year 1925 also saw the debut of the Ford Motor Company's Trimotor airplane, which was well-suited to mixed use, carrying up to twelve passengers in its cabin plus cargo (like mail) and with a range of over 500 miles. Indeed, in 1927, Pan American Airlines was built around Ford Trimotors, carrying mail and passengers between Key West and Cuba. Soon thereafter, Pan American began to specialize in much longer distance routes to Central and South America using flying boats, expanding to Asia by 1934, eventually adding European routes in 1937 (Daley 1980). Most other early airlines, such as Eastern (1926), Delta (1928), United (1929), and American (1930), however, concentrated on domestic air travel, which began to gain traction in the 1930s with the appearance of ever-better passenger planes such as the Boeing 247—whose early production was committed to United, forcing United's competitors to seek an alternative. That (better) alternative was found in the Douglas DC-3, which proved to be a rugged, all-purpose workhorse that is still in use today (Heppenheimer 1995).

As air traffic grew, especially in the vicinity of major cities, rules regarding approaches, landings, takeoffs, and flight paths became increasingly important. Although such activity had been authorized in the Air Commerce Act of 1926, the Aeronautics Branch only stepped in reluctantly after private efforts to establish an air traffic control system had failed.

The most decisive event in the growth of commercial aviation in the United States was Franklin D. Roosevelt's sweeping reorganization and consolidation of civil air authority through the Civil Aeronautics Act of 1938 that created the Civil Aeronautics Board (CAB). This agency not only inherited the rule-making and safety-related regulatory powers from earlier legislation, as well as responsibility for airport development, they were also given broad economic authority over the operations of the airlines themselves. They used this power to set and regulate airfares and to parcel out routes between competing airlines. The CAB's authority, however, only extended to interstate traffic. This provided an opportunity for one local carrier—Southwest Airlines—to grow. By concentrating on local Texas routes with the highest potential traffic and by cutting costs (e.g., by using secondhand planes from just one manufacturer and in a single seat configuration), they could offer lower fares and began to make inroads into the business of other airlines serving the state. This created pricing tensions within what had been a stable cartel system (Farris and Pohlen 2006; Farris and Swartz 2006; Morrison 2001).

The Southwest Airlines model and Texas experience also made it clear that the CAB model of civil aviation worked less well for the general public than the airlines (Morrison 2001). The public's choice among carriers and flights had been limited by the CAB, and there was essentially no price competition between "competing" airlines. Thus, the system began to break apart, and Congress eventually passed the Airline Deregulation Act of 1978, phasing out the economic regulatory authority of the CAB.

In the years since, many airlines have disappeared through mergers and bankruptcies, airfares have fallen drastically (especially in real terms), and the commercial aviation market has become much more chaotic (Bailey, Graham, and Kaplan 1985; Kahn 1988a).

Economic Issues Affecting the Provisioning of Transportation

The positive externalities that transportation improvements generate partly explain the gap between private and social returns for various canal and railroad ventures that scholars like Fogel (1960), Fishlow (1965), Mercer (1969), and Ransom (1970) have found. In most cases, the private returns should have been more than adequate to justify the investment unless capital was severely rationed, but the relatively high social rates of return also have been used to rationalize government intervention to promote such improvements, for example, by Goodrich (1960, 1967). Contemporaries also appealed to these general welfare gains in making the case. For example, in his 1808 report to Congress on roads and canals, Treasury Secretary Albert Gallatin argued,

> Good roads and canals will shorten distances, facilitate commercial and personal intercourse, and unite, by a still more intimate community of interests, the most remote quarters of the United States. No other single operation, within the power of Government, can more effectually tend to strengthen and perpetuate that Union which secures external independence, domestic peace and internal liberty. (US Senate., 1808)

But Gallatin was concerned that the scale and scope of such ventures and the free ridership attending them required the "early and efficient aid of the *Federal* Government" (US Senate 1808, emphasis in the original). Indeed, providing transportation infrastructure like roads, harbors, and navigation aids funded out of general tax revenue has been one of government's functions since earliest times. Such investments improve the general welfare through trade, help maintain internal control, and promote security from external threats—each clearly a part of Gallatin's argument.

The nature of the free ridership, however, undermined Secretary Gallatin's case for federal involvement because the free riders were the fortunate property owners whose land lay close to the improvement but who neither used those services nor financed their construction (Edel and Sclar 1975; Gin and Sonstelie 1992; Spengler 1930).[22] The United States has a complex governance structure. While government had the capability to tax those gains, property taxes were the domain of state and local governments, not the federal government. Moreover, there were doubts about the constitutionality of federal government involvement in such activities, and President James Madison vetoed action on Gallatin's plan:

> I am not unaware of the great importance of roads and canals and the improved navigation of water courses, and that a power in the National Legislature to provide for them might be exercised with signal advantage to the general prosperity. But seeing that such a power is not expressly given by the Constitution, and believing that it cannot be deduced from any part of it without an inadmissible latitude of construction and a reliance on insufficient precedents; believing also that the permanent success of the Constitution depends on a definite partition of powers between the General and the State Governments, and that no adequate landmarks would be left by the constructive extension of the powers of Congress as proposed in the bill, I have no option but to withhold my signature from it. (Madison 1817)

Nevertheless, his successors, James Monroe and John Quincy Adams, not only continued work on the National Road, but also expanded the scope of federal assistance to transportation by approving the General Survey Act of 1824, which authorized surveys for roads and canals "of national importance, in a commercial or military point of view, or necessary for the transportation of public mail."[23] Their successor, Andrew Jackson, however, adopted a stance in line with Madison's. This pattern has continued to the present. Consequently, federal programs for internal improvements have progressed in fits and starts, with much of the progress being undone by traffic and nature during suspensions (Goodrich 1960, 1967; Haney 1968).

Taxation, and a general opposition to taxes, further complicates government investment in transportation infrastructure. Until the passage of the Sixteenth Amendment in 1913, the Constitution had mandated that direct federal taxation be on a per capita basis, but transportation infrastructure improves opportunities for trade and enhances property values, neither of which is on a per capita basis. Thus, it was difficult to ensure that the beneficiaries of improvements would cover their costs through taxes. Tolls met the requirement, but they were easily avoided or, if collected, stolen, and few improvements were interstate. Taxes on airline tickets also satisfy the original constitutional requirement regarding taxation, and state and federal highways have long been funded by excise taxes on fuel. States led the fuel tax push, beginning with Oregon in 1919, and by 1929 all states had imposed the tax. The federal government added its own excise tax beginning in 1932. No user fees, however, can be levied on navigable rivers because of the language in the Northwest Ordinance (US Continental Congress 1787).

Most early roads were local and were supposed to be maintained by local residents. Rather than impose unpopular general taxes to accomplish that goal, the obligation was met by compulsory labor services, a holdover of the medieval corvée. The resulting work was generally of poor quality. The most heavily traveled roads were often the least well maintained, and higher income residents could pay for a substitute to satisfy their obligation. Over time, however, with rising incomes, this option increased specialization and professionalization in road construction and maintenance, albeit at a glacial pace. While good transportation made moving goods easier, it also facilitated migration, and as population began to migrate westward, the tax basis of eastern communities eroded, making them even more reluctant to support transportation improvements (Barron 1984; Grossman 1989; Stiles 1996).

While tax dollars were scarce, the federal government was generous with grants of land—beginning with the creation of the state of Ohio, which in 1802 was promised 5 percent of the proceeds from public land sales to help build roads. Other road grants followed, and canal grants were even more generous, but both were dwarfed by the land grants to the railroads, particularly the transcontinentals. The railroad grants were made directly to the companies themselves, or were transferred to states for conveyance to the railroads, beginning with the Illinois Central in 1851. By 1871, the federal government had awarded 131 million acres of public land, while Texas had donated another 27 million acres. With improved transportation, the value of these lands appreciated and was captured by the railroads and is estimated at between $130 to $500 million (Ellis 1945, 1946; Fleisig 1975; Gates 1954; Henry 1945; Mercer 1969, 1972; Rae 1944; Sutton 1966; Traxler 1958; Wilson 1942).[24] Regardless of the gains realized by the land grant recipients, however, the federal government, as an even bigger landowner, gained so much more. These came with a cost: Railroads promoted settlement in areas poorly suited to agriculture, resulting in environmental degradation and social upheaval, and regulations like the Hepburn Act of 1906 were needed to prevent railroads from enjoying an unfair advantage in the market for minerals that they now owned.

Each mode of transportation, but especially railroads and waterborne traffic, experienced tremendous productivity gains since first innovated. Much of this productivity growth was due to technological change. Some were subtle, like changes in the ratios of hull dimensions; others more obvious, like steel rails or turbofan jets. There were also improvements in organization and coordination—including network effects—and the general growth of traffic that increased utilization of facilities. As a result, point-to-point shipping costs have fallen, often to a small fraction of their initial levels in real terms (see table 20.2).

Like rural Eastern communities, not everyone gains from new or cheaper shipping. In particular, the providers of older or unimproved transportation see demand for their services fall. With the coming of the railroad, stagecoach lines and the inns along their routes felt an immediate effect. Similarly, the owners of downtown businesses lost out when interstate highways bypassed the town. Less obviously, however, there are those producers who once were protected by distance and high transport costs, but who now lose out to lower cost producers located elsewhere. This was especially true of American

Table 20.2 Freight Rates 1815–Present (2010 Cents per Ton Mile)

Mode		1815	1860	1913	1950	2010
Road		439.0	542.0	567.5	58.9	16.5
River						0.7
	Upstream	90.0	27.1			1.8
	Downstream	20.0	27.1			1.8
Canal			27.1	18.2		
Ocean		12.2	9.5			0.4
Railroad			81.3	14.0	40.8	3.0
Air					199.3	101.0

Sources: Calculations based on quotes from various studies (Hunter 1949, Tables 20–21, 658–659; Taylor 1951, Appendix A.2, 442; Fogel 1964, see "rates, freight" in index; Fishlow 1965, Table 43e, 334) plus Department of Transportation and assorted unrecorded sources collected and adjusted to 2010 using MeasuringWorth.com CPI calculator.

farmers in the Northeast whose lands were of marginal fertility, many of whom abandoned their farms and the occupation, moving to cities or further West (Barron 1984; Field 1978;). Others would be forced to shift to more labor-intensive activities (i.e., increased labor effort) like dairying and market garden produce to stay in business (Boserup 1973; Thünen and Hall 1966).

On balance, however, the gains seem clear. Moreover, since few products in a complex economy such as the United States are consumed where they are produced, the impact of these changes on transport costs permeates the economy and contributes to economic growth. The question is, how much?

A simple way to measure how much is to calculate the increase in the consumer surplus and rents arising from the switch from high-cost to lower-cost transportation providers. The true measure is more complex, as transportation modes were never perfect substitutes for one another that differed solely in terms of the out-of-pocket shipping cost. Railroads, for example, were much speedier than other modes of transportation until aircraft came along. They were also less sensitive to the weather. Collectively, these benefits are referred to as "social savings," but there are also induced and general equilibrium effects to be considered—impacts even less perfectly measured.

Because of the importance of these issues and the hyperbole surrounding the coming of the railroad, economic historians have devoted considerable effort to measuring the social savings. According to calculations by Albert Fishlow (1965, 52) for railroads, the "direct benefits of $175 million . . . are just about 4 percent of 1859 gross national product. For an innovation still in its infancy, this clearly is an impressive contribution." Later, the social saving should have been even larger as their market penetration grew and because of their more rapid productivity growth: perhaps to as much as $1.8 billion or 15 percent of GNP in 1890, of which about $300 million is due to passenger savings. This latter figure, however, is larger than Robert Fogel's (1964, 223) estimate of social savings in 1890 that he put at "well below 5 percent of gross national product." Adding savings from passenger traffic raises Fogel's bottomline estimate to 5.9 percent of GNP (Boyd and Walton 1972) and recent work by Donaldson and Hornbeck (2016) makes a further, if modest, upward revision.

There are also estimates of the social savings from canals and steamboats (Fishlow 1965; Haites and Mak 1978; Ransom 1964, 1970; Segal 1961). These are substantially smaller than for railroads, perhaps 1.2 percent of GNP for steamboats and 2 percent for canals. The various social savings estimates are summarized in table 20.3.

Comparative statics analysis, however, can only go so far. What is needed is a dynamic general equilibrium model to capture those omitted interactions between markets. Jeffrey Williamson (1974, 1975), for example, has used a computable general equilibrium model to simulate the American economy from 1870 to 1892. Assuming that transport costs were unchanged over the period (when, in fact, railroad freight rate from Chicago to New York declined by about 60 percent), he concluded that real GNP in 1892 would have been 21 percent lower than it actually was. Since this figure measures the savings from reduced transport costs in general and includes handling charges and insurance

Table 20.3 Direct Social Savings Estimates for Different Modes of Transportation at Various Benchmark Dates, 1825–1890

Transportation Mode	Year	Source	Direct Saving	% GNP in Year
Canals	1845		$40m	2.2
Steamboats				
	1825		$1.83m	0.2
	1835		$11.88m	0.9
	1845		$22.07m	1.2
	1855		$41.65m	1.0
Railroads				
	1859		$175m	4.0
	1871	Williamson[a]	$690m	9.0
	1880	Williamson[a]	$1,674m	16.0
	1890	Fishlow[b]	$1,800m	11.8
		Fogel[c]	$560m	3.7
		Fogel+Boyd/Walton[d]	$904m	5.9
		Williamson[a]	$3,197m	21.0

Source: % GNP calculation made using the GDP series at MeasuringWorth.com by Johnston and Williamson (2016).
[a]Williamson (1974), [b]Fishlow (1965), [c]Fogel (1965), [d]the Fogel estimate plus a calculation for passengers by Boyd and Walton (1972). See text where source not otherwise given.

as well as marine freight charges on exports, it overstates the impact, but it is "close" to Fishlow's back-of-the-envelope calculation for roughly the same period. Other CGE estimates, however, are closer to Fogel's original estimates (Kahn 1988b).

Regardless of which figures on social savings are "best," the market has often been quite generous in its rewards to innovators, like investors in the Erie Canal, steamboats, many of the early railroads and even some of the airlines. These successes typically share one characteristic in common—they were market responses by prescient, risk-taking investors, whereas those investments that had to depend on some form of direct government involvement or incentive were generally less successful. However, even when these transportation projects came up short as purely private investments, the social returns generally justified their existence. Infrastructure, especially transportation infrastructure, has always been and remains a critical element in American economic growth and development.

The Future of Transportation

In the global economy of today, the tyranny of distance (and even the time element associated with it) has largely ceased to play a major role in economic activity. Markets are integrated as never before, and it matters less and less where production takes place relative to consumption, thanks to a succession of improving modes of transportation and an ongoing and sustained pattern of productivity improvements within each mode.

Modeling these changes and their impact has increasingly relied upon analysis using geographic information systems (GIS) to provide ever-better locational and location-specific data.[25] However, while the historical evolution of river and canal navigation and railroads have been mapped into GIS databases, the evolution of the nation's roads and air routes have not.[26] Roads, in particular, pose a challenge because of their varying characteristics (dirt versus gravel versus more durable surfaces like concrete and blacktop and the number of lanes) and their inaccurate mapping early on.[27]

With improvements in transportation has come an increased degree of interdependence that can lead to disruptions by shocks such as the failure of a key bridge, a business bankruptcy, or a war. The system, however, has proved sufficiently robust that such shocks have had only a temporary effect. One consequence of this success has been a growing complacency about the transportation infrastructure and the need to maintain and preserve it. This has been particularly true where that infrastructure is dependent upon the public sector and the public purse. For example, the Highway Trust Fund, the source of most federal highway funding since 1956 and funded by the fuel excise tax, no longer covers highway expenditures, yet the current federal tax of 18.3 cents per gallon is scheduled to be cut to 4.3 cents per gallon in September 2016, and there have been proposals to return highways and their maintenance to the states. There have also been proposals to increase the excise tax rate sharply and make it an ad valorem rather than a specific tax (Lowry 2015). In the current legislative stalemate, neither set of proposals has made much progress, even though the current status quo is untenable, and much of the infrastructure is approaching or past its life expectancy.

This reluctance to invest or reinvest in infrastructure is not confined to the public sector. For example, the nation's airlines are increasingly flying older planes, and the rail system has been slow to adopt improved signaling methods and positive train control or to replace uneven track.

Nevertheless, there remains the promise of significant future productivity gains in every mode of transportation, many of which will also contribute to the general welfare by reducing existing negative externalities. For example, autonomous driving vehicles on the nation's highways promise increased traffic density and higher average speeds but with fewer accidents (with implications for body shops, the health-care system, and insurers, as well as family welfare). Moreover, an increasing share of the vehicles on the highway are more fuel efficient thanks to hybrid or alternative fuel systems, lighter weight designs, turbochargers, and the like. This reduces pollution but exacerbates the shortfall in the Highway Trust Fund. Aircraft increasingly use composite materials to reduce weight and ever-bigger turbofan jets with higher bypass ratios to improve fuel efficiency and reduce noise, and modest increases in plane speed can be made before physical limitations like the sound barrier impose new constraints. Air traffic control is taking advantage of collision warning systems and other forms of automation to increase traffic density in the air by reducing plane separation. Similarly, the United States could vastly improve its passenger train system in higher density, shorter distance corridors on the East and West coasts to compete with air by separating freight from passenger traffic and adopting bullet train technologies like France, Japan, and China.

Notes

1. See also http://www.vanderbilt.edu/econ/cipr/cole-historical-data.html
2. There were several possible routes and modes of travel. For example, one could (1) go down the Ohio and Mississippi rivers to New Orleans and then by coastwise vessel to Philadelphia; (2) up the Ohio River to Pittsburgh, then overland to Philadelphia; or (3) overland to Lexington, Kentucky, where the Wilderness Road ended, then following the Great Wagon Road via Roanoke, Virginia, to Philadelphia. For an account of travel along some of these routes, see Michaux (1805) and Volney (1804).
3. For example, steamboating data presented to Congress (US Senate 1852, 115) praising the safety of such "modern" modes of travel indicates that one passenger was killed or seriously injured per thousand miles traveled.
4. That is to say, carrying close to a full cargo on all legs of a voyage (North 1958).
5. Leaving Pittsburgh on October 20, 1811, the *New Orleans* reached Louisville on October 27, where it was forced to remain until late November when a rise in the river allowed its passage over the Falls of the Ohio—a series of rapids and falls where the river drops 26 feet in 2.5 miles. By early December, it had reached the Mississippi where, cushioned by the river, it survived the massive New Madrid earthquake and reached New Orleans on January 10, 1812. The *New Orleans* subsequently operated between its namesake and Natchez.

6. For many years, however, few bridges crossed major rivers, especially their navigable, lower reaches. Rather, ferries were used to move traffic—road and railroad—across them. Where bridges were built, they were often subject to fierce litigation, allegedly for impeding navigation; for example, see US Supreme Court (1851, 1855); Monroe (1992). Indeed, railroads bridged the Mississippi River in 1856, years before the first rail bridge crossed the Ohio at Louisville in 1871 (US District Court for Northern Illinois 1857; US District Court for Southern Iowa 1859; US House 1858; US Supreme Court 1867).
7. For example, Thomas Jefferson purportedly described it as "little short of madness."
8. Pennsylvania coal canals like the Lehigh and the Schuylkill Navigation were exceptions.
9. Ohio canals received generous land grants, and the state initially tried to handicap railroads when they began to compete with canals; Indiana issued state debt and created banks to fund its canals. The Raritan Canal was wholly private.
10. Several factors handicapped canals and limited their capacity. Efforts to increase barge speed, for example, created wakes and washed out the canal banks, and the physical dimensions of lock chambers determined the maximum size of barges. Locks also delay canal transit, and the water that they use can create supply problems. This was particularly true at the highest point on most canals, where locks were required in both directions and streams and rivers had their least water. Locks on the Erie Canal were enlarged twice, first in 1834–1862 and then again when the canal was rebuilt beginning in 1903.
11. For example, Ohio (1804) stipulated that when first "built," tree stumps in the roadway could not exceed one foot in height and that "wet and miry places" had to be covered with timber and earth.
12. By 1818, the National Road had reached the Ohio River at Wheeling. Seven years later, Congress authorized its extension through Ohio and eventually to Vandalia, Illinois, in 1854, although the Ohio River was not bridged until 1849.
13. Draft animals could travel twenty to thirty miles a day. Pack animals could move about 250 pounds. Wheels enabled a single-horse to pull a 1,000-pound-cart; two horses 4,000 pounds and proportionately more for larger teams on hard level surface (Folwell 1916; Marcy 1859; US Army Ordnance Department 1862).
14. A similar fate befell planks roads. These enjoyed a brief period of popularity beginning in the 1840s that, by and large, lasted just as long as it took the first planks to rot. Over 1,300 companies built more than 10,000 miles of these roads, at least 3,500 miles of which were in New York. New construction waned after 1853, and the remnants were converted to dirt and gravel roads by the 1860s (Klein and Majewski 2008; Majewski, Baer, and Klein 1993).
15. The other route, the Dixie Highway, developed several years later, originally ran south from Chicago to Miami.
16. The Southern Pacific met up with the Texas and Pacific Railroad to the east of El Paso in 1881, linking to the Atchison, Topeka, and Santa Fe Railroad in New Mexico in 1882 and the Galveston, Harrisburg, and San Antonio Railroad in 1883.
17. Dilts (1993) has an extended discussion regarding the time limits imposed upon the B & O variously by the legislatures of Maryland, Pennsylvania, and Virginia.
18. With regard to political intrigue and corruption, see, for example, Adams, Adams, and Walker (1871); US House (1873).
19. $400 (equivalent in purchasing power to over $9,700 in 2015 according to MeasuringWorth.com) represented a significant premium for the right to be first; subsequent passengers paid $5 (equivalent to $121.25 in 2015) each way. However, despite the time savings and lower fare, the airline ceased operation just four months later.

20. TWA, which emerged from the merger of Western Air (the original expansion of "WA") and Transcontinental ("T") Air Transport, and ceased operations in 2001, was one of the early successful bidders.
21. These can be compared to the public provision of lighthouses along the coast and navigation buoys. See Coase (1974), but also Barnett and Block (2007).
22. Negative externalities are also possible—for example, noise, fumes and congestion. See, for example, Rosen (1995). The evidence, however, suggests that these are swamped by the positive externalities.
23. Other legislation, the Rivers and Harbors Act, authorized the first expenditures to improve navigation on the Ohio and Mississippi rivers.
24. Depending upon whether one writes down the value from the nominal $1.25/acre sale price of the land on the grounds that no one was (obviously, since the land remained in government hands) willing to pay that much for it, or computes the revenue that railroads realized from its sale. These figures do not include the vast mineral wealth in coal, oil, natural gas, etc., that railroads have realized from their land grants in the twentieth century.
25. See, for example, the numerous studies authored and coauthored by and with Atack (2010a, 2011, 2012, (with Margo) 2012, 2013, 2014, 2015) and Donaldson and Hornbeck (2016).
26. See Atack (2015a, 2015b, 2015c) for historical GIS SHP files for steamboat-navigated rivers, canals, and railroads. SHP files for today's transportation system may be found at United States Department of Transportation, Bureau of Transportation Statistics, National Transportation Atlas Database: http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/index.html.
27. There has been some effort to map early turnpikes by Klein and Majewski (2008) and plank roads by Majewski, Baer, and Klein (1993), but these are not in the form of GIS databases or SHP files.

Bibliography

Adams, C. F., H. Adams, and F. A. Walker. (1871). *Chapters of Erie, and Other Essays*. Boston: J. R. Osgood.

Atack, J. (2013). "On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited." *Journal of Economic History* 73(2): 313–338.

Atack, J. (2015a). "Historical Geographic Information Systems (GIS) Database of Nineteenth Century U.S. Canals." From author's personal WWW page at Vanderbilt University: https://my.vanderbilt.edu/jeremyatack/files/2015/09/19thC_Canals2.zip.

Atack, J. (2015b). "Historical Geographic Information Systems (GIS) Database of Steamboat-Navigated Rivers during the Nineteenth Century in the United States." From author's personal WWW page at Vanderbilt University: https://my.vanderbilt.edu/jeremyatack/files/2015/09/SteamboatNavigatedRivers2.zip.

Atack, J. (2015c). "Historical Geographic Information Systems (GIS) Database of U.S. Railroads." From author's personal WWW page at Vanderbilt University: https://my.vanderbilt.edu/jeremyatack/files/2015/12/RR1826-1911Modified122715.zip.

Atack, J., F. Bateman, M. R. Haines, and R. A. Margo. (2010). "Did Railroads Induce or Follow Economic Growth? Urbanization and Population Growth in the American Midwest, 1850–1860." *Social Science History* 34(2): 171–197.