Dense Enough To Be Brilliant: Patents, Urbanization, and Transportation in Nineteenth Century America*

Market Access

Elisabeth Ruth Perlman, perlmane@bu.edu

UNFINISHED AND INCOMPLETE

Click Here For the Current Version.

December 31, 2014

Abstract

This paper explores the geographical distribution of patenting in the nineteenth century United States, as it evolves in response to improvements in access to transportation. I revisit the Sokoloff (1988) hypothesis that increasing market access, caused by the spread of transportation infrastructure, led to an acceleration of innovation. I find that twenty years after the arrival of the railroad in a county, the number of patents per capita has doubled. Using cardinal detection lines from the most important ports in 1826 as an instrumental variable suggests that 30-70% of the increase in patenting between 1850 and 1860 was caused by the spread of the railroad in this period, and 15-30% of the increase between 1850 and 1870. These results are driven by the area of a county that is close enough make a round trip to transportation with in a day, and not by area further away. A 1% increase in the area of the county that is within 1.5 miles of some form of transport corresponds to a to a 1.5% increase in patenting. These results are robust to controls for urbanization. Much of the effect comes from patenting in counties that had not previously patented, suggesting that new access to existing markets spurs development and leads to integration into broader markets for innovation.

*Acknowledgments: I am extremely grateful to Robert Margo, Tom Nicholas, Jeremy Atack, Carola Frydman, Daniele Paserman, Claudia Olivetti, Michael Meurer, Suresh Naidu, Thomas M. Darby, Peter Sims, Matthew C. Berntsen, Henry Gross, Ross Corliss, and the volunteers at the Directory of American Tool and Machinery Patents and to seminar participants at the BU Development Reading Group, the BU Labor Reading Group, the Essex FRESH Meeting, the Harvard Economic History Tea, the All-UC Group in Economic History, the LSE Economic History Thesis Workshop, the EBHS Conference, the WEAI Conference the NBER Summer Institute, the SSHA Annual Meetings, and the NBER Productivity Lunch for their helpful advice, data, and comments. Finally, to Chloe Hite for excellent research assistance. All errors are my own.

1 Introduction

The movement of people, goods, and information among places increased dramatically over the nineteenth century, particularly within the United States, as the pace of technological change accelerated. This paper investigates the hypothesis that the "transportation revolution" facilitated innovative activity, as posited by Sokoloff (1988), who suggested that the increase in innovation in the United States in the early nineteenth century was driven by increasing market access. It uses data from U.S. patents to understand the geographic distribution of innovations. Patents are by no means a perfect measure of innovative activity, (Trajtenberg, 1990; Moser, 2004), but their paper trail is extensive, including descriptions of each invention. These records allow for systematic use as evidence for historical innovation.

Before the twentieth century, transport costs were of the utmost importance. Before railroads, waterways were by far the most efficient way to transport goods. Moving goods over land without mechanical power, even on the best highways, was costly.[1] In 1800, a message traveling between Boston and New York either moved by sailing ship, or over land by unpaved (and often dubiously maintained) roads. By the 1860s, the telegraph had become the fastest way of sending a message, and physical packages moved overland by train or over water via steam ship. These changes made transportation faster, cheaper and safer, effectively reducing the distance between locations. This reduction in distance was even more dramatic for places on the periphery of the transportation and communication network. By 1900, railroad tracks densely intersected much of rural America, promoting economic growth by linking far-flung factor and product markets, encouraging the exploitation of regional comparative advantage (Fogel, 1964; Atack et al., 2011; Donaldson and Hornbeck, 2013).

[1] It has become a commonplace to note that it cost about the same amount to ship goods between London and Boston as to travel 30 miles over land in the U.S., or about the distance from downtown Boston to Concord, MA.

There is a clear intuition that increased market access leads to patentable innovation, and this has been explored both in theoretical and empirical papers. The expected return to a line of research is directly related to the size of the market in which the new product (or product related to the new process) might be sold. A number of studies (Pavcnik, 2002; Amiti and Konings, 2004; Van Biesebroeck, 2005; Becker and Egger, 2013; Deloecker, 2007; Fernandes, 2007; Foster et al., 2008; Topalova and Khandelwal, 2011), using modern plant- or firm-level data, have noted that exporting firms become more productive after trade liberalization, suggesting that greater market access encourages innovation. Extensions of the Melitz (2003) model made by Bustos (2011) and Lileeva and Trefler (2010) give a clear motivation as to why firms facing a larger market might want to invest in technology that potentially improves their marginal cost of production—larger markets allow for more units to be sold, thus providing for larger returns when firms reduce their marginal costs of production.

In a much older literature, economic historians focused on the role the expansion of demand played in Britain during the Industrial Revolution, while other scholars have emphasized the importance of the stock of knowledge in determining the direction of technological change.[2] Thus, there are many reasons to expect that increased connectivity would lead to an increase in both the quantity and dispersion of patenting. Transportation brings together ideas and people, which in turn creates the potential for greater division of labor. It also provides access to formal credit institutions, and decreases the price of transporting both inputs and outputs. All of these factors might affect the patenting rate.

Sokoloff (1988) hypothesizes that increasing market access caused the increase in innovation seen in the early nineteenth century. Noting the change in which areas of New York and Pennsylvania patented between 1790-1846, Sokoloff (1988) argues that in the eastern United States counties that had previously had poor transportation links along the newly-built

[2] These two the views by no means capture the full extent of this debate.

canals, particularly the Erie, saw a sharp increase in patenting activity. He finds that there is a positive relationship between canal (but not ocean) access in several time periods (in cross-section), thus concluding that "the record suggests that low-cost transportation to market was a necessary rather than sufficient condition for high levels of inventive activity" (Sokoloff, 1988, pg. 837).[3]

Sokoloff's study was limited by both data and computing power restrictions. He could only observe patenting in detail for a narrow band of time in the early nineteenth century, and only for a sample of patents. His results are suggestive of the relationship between market access and innovation for America in this period. This paper, by using the universe of patents registered in the United States from 1790 to 1900, can expand the temporal and geographical scope of the analysis. The period sees a much greater increase in transportation infrastructure, and under more varied conditions, providing a superior test of the Sokoloff hypothesis.

This paper is part of a larger project researching the relationship between innovation, as represented by patents, and transportation in nineteenth century America. This paper fo-

[3]Sokoloff (1988) notes many reason a transportation link might increase patenting. Sokoloff (1988) mentions that "expansion of markets, changes in cultural attitudes, learning-by-doing, the propensity to invest, information flows, the organizations of production, output mix, the degree of specialization, and the amount of resources available for inventive activity" all might be reasons for the increases observed. However it is clear that Sokoloff's main hypothesis falls in line with the intuition gained from the Bustos (2011) and Lileeva and Trefler (2010) extensions to the Melitz (2003) model, which suggest that, upon gaining access to larger markets, it is worthwhile for the firms with the lowest marginal costs of production to invest in productivity upgrades. The increase in units sold allows them to recoup a larger investment. Or, as Sokoloff (1988) puts it: the effect is likely caused by "individuals and firms choosing to commit additional resources to a search for useful discoveries in response to the increase in the expected return to such investment that stems from their integration into a larger pool of both potential customers and competing suppliers." However, Sokoloff (1988) also gives space to alternative reasons for patting to increase, noting that "gaining low-cost access to a large market could alter behavior through changes in the prices of goods or in the returns to activities, and thus nurture cultural attitudes more favorable to invention, enhance learning-by-doing, improve the flow of information to potential inventors, raise the amount of resources available for allocation to invention, increase the propensity to invest in general, or foster such changes in methods or in the extent of factor specialization as to facilitate the discovery of possible refinements in technique or other inventions.... as well as the improvement in the stock of knowledge... would help account for the patterns in the data."

cuses on the question of how increased market access affects innovation; the other papers will examine the effects of population density, urbanization, and information flows.

By combining the new and broader data set with insights from gravity models of trade, this paper seeks to establish a richer empirical basis for considering market access as a driver of innovation. This paper uses data from Atack (2013) on the spread of transportation networks to construct two measures of transpiration access. The first is a measure of the density of transportation access in a location, the percent of a county that is within a distance (e.g. 5 miles) of transportation. It relies on the fact that non-mechanized, overland transportation is orders of magnitude more costly than other modes, so the last few miles might easily to make up the bulk of an item's transport cost. The second is a cost-weighted measure of the market access; I take inspiration from Donaldson and Hornbeck (2013) to construct this single measure for every county in each census year. These two measures are is linked with two databases of geo-located patents, one from Tom Nicholas and the other my own, in order to analyze the relationship between improvements in market access due to transportation improvements and patenting activity in each county.

The density measure of transportation retains a positive relationship in cross-section, both for the time period examined by Sokoloff (1988) and for later ones, when the railroad had displaced water transport as the dominant mode of transport. However, the market access measure only shows a positive relationship in the later period. Both Sokoloff (1988) and the results in this paper suggest that access to the ocean is not related to patenting. This suggests that dynamics are at work beyond the intuition that contact with larger markets increases innovation. In first differences market access and patenting show a U-shaped relationship—in graph with change in market access on the horizontal axis there is an area of counties that exhibit a large positive change in patenting on the left side of a graph, a dip in the middle of the graph, and then an area of counties that exhibit a large positive change in patenting on the right side. The most developed counties are on the left side of the U, these are eastern counties that are losing relative market access. In a regression, controlling for population reduces the coefficient on this side of the U to zero. On the right side of the U are western counties that less developed and gaining relative market access. Here, controlling for population leaves a positive coefficient on the market access variable. Once a county is relatively populous, losing relative market access doesn't have a detrimental effect on patenting growth. This is consistent with transportation bringing locations into the larger economy.

The fixed effects estimates using both density measure and market access are positive. These estimates suggest that, for a county that is patenting, a 1% increase in the area of the county that is within 1.5 miles of some form of transport corresponds to a to a 1.5% increase in patenting. The estimated elasticity using the percent of the county within 1.5 and 5 miles of transportation are of similar magnitude. However, using the percentage of a county within 15 miles of transportation leads to much smaller estimated elasticities. This suggests that being close enough to transportation to make a round trip with in a day is important to the observed increase in patenting. The positive effect persists, even when controlling for the percent of population that lives in an urban area (2,500 people), or a metropolitan one (25,000 people). Results using cardinal detection lines from the most important ports in 1826 as an instrumental variable for the presence of the railroad suggest notable downward bias in these estimates. These estimates suggest that 30-70% of the increase in patenting between 1850 and 1860 was caused by the spread of the railroad in this period, and 15-30% of the increase between 1850 and 1870. The effect is much smaller if counties that do not patent are omitted from the sample.

A fixed effects estimate of the number of years until the arrival of a canal or railroad suggests that, after arrival, there is a notable but slow increase in patents per capita. Each year after the arrival shows an increase over the last, up to 20 years for the canal, and 40 for the railroad. This amounts to doubling the number of patents per capita after 20 years. The arrival of transportation causes gradual patenting growth, over and above that caused by population growth.

The Bustos (2011) and Lileeva and Trefler (2010) extensions to the Melitz (2003) model both suggest that, when facing a newly-opened market, only firms above a productive threshold will invest in innovation (with many parameters, these are the same firms that will invest in exporting). The establishment of a sufficiently productive firm could therefore be a prerequisite for increases in patenting. In the early part of the nineteenth century, most mechanized manufacturing relied on water power, which limited establishment location. Firms using water power could only locate where the geography is favorable for waterwheels: areas with a river of the correct grade. This meant firms had limited ability to cluster. Steam power changed this dynamic, making it beneficial to locate near places to which coal could be transported inexpensively. This permitted more diversity of locations, which in turn allowed for greater clustering (Gutberlet, 2014). The 1840s and 1850s saw an explosion of steam power use in manufacturing (Fenichel, 1966), use of steam power is correlated with large establishments and higher (labor) productivity (Atack et al., 2008). Choosing a production method that uses a steam engine as a source of power would be an investment in improving productivity , in the sense of Bustos (2011) or Lileeva and Trefler (2010). The railroad also increased establishment sizes in U.S. manufacturing (Atack et al., 2011), likely because of access to larger markets. These larger establishments, producing for larger markets, correlate with higher productivity and a greater division of labor, both of which may be important for increases in innovation.

2 Intuition from Trade Theory

2.1 The Gravity Model and Market Access

The empirical predictions made by the gravity model of trade are some of the most robust in the literature on the economics of international trade. This model uses a few simple variables to predict bilateral trade flows, while remaining agnostic on the reason why this relationship matches the data. The basic gravity relationship describes bilateral trade flows as between two places i and j:

$$trade_{ij} = \frac{y_i y_j}{y_w} (\frac{\tau_{ij}}{P_i P_j})^{1-\sigma} \tag{1}$$

where y_i , y_j and y_w are the incomes of i, j, and the world (total market), τ_{ij} is a bilateral resistance term, P_i and P_j are location specific resistance terms. Note that changing τ_{ij} has the same effect on trade as changing P_i or P_j, so lowering transportation cost is qualitatively similar to opening a county to trade. This is often simplified by noting that y_w is a constant for all pairs of counties and thus not needed in the estimation and taking $\frac{\tau_{ij}}{P_i P_j}$ to be the physical distance between the locations, because of the robustness of the empirical evidence.

The resistance term between countries incorporates anything that makes trade less likely including language barriers and cultural differences. Feyrer (2009) estimates, by using closing of the Suez Canal as a natural experiment that provides distance variation between counties, that about half of the resistance term between countries in the 1970s was due to transportation costs that varied by distance. In the nineteenth century U.S. where there were no formal trade barriers, except during the Civil War in the early 1860s, and few language barriers one would expect transportation cost to be a larger fraction of this term.

The size of a market i is the sum of all goods sold in it, the market in i captured by a trading partner j is the sum of its imports to i. Thus one can think of the market available to firms in location j as:

$$MA_j \approx \sum_i trade_{ij}. \tag{2}$$

If one labels the goods sold by j in j as $trade_{jj}$ this is a full description. Thus, market access refers to the areas with which a given region can effectively trade as adjusted for transportation costs.

For the purpose of the gravity model, however, it is irrelevant whether reductions in tau come from reduced distances or from decreased trade costs, as they have the same effect on total trade. Thus, the steep reduction in transport costs over the 19th century should have had an effect on trade analogous to a similar reduction in trade barriers.

2.2 Melitz Model and Investment in Process Improvements

Recent trade literature has examined the question of inventive activity, linking it to market openness. One of the facts motivating the Melitz (2003) model was that more productive firms export, while less productive firms only serve the domestic market. This is modeled with a section mechanism. A number of studies (Pavcnik, 2002; Amiti and Konings, 2004; Van Biesebroeck, 2005; Becker and Egger, 2013; Deloecker, 2007; Fernandes, 2007; Foster et al., 2008; Topalova and Khandelwal, 2011), using modern plant- or firm-level data, note that exporting firms often do become more productive after trade liberalization, suggesting that the increased market access motivates firms to invest in innovations that lower their marginal cost of production.

Bustos (2011) and Lileeva and Trefler (2010) both consider models that examine the decision of a firm in an open economy to invest in process innovation. Both of these studies then pair their Melitz (2003) style model with firm- or plant-level data investigating if when exposed to a wider market, firms invest in process innovation. The Sokoloff (1988) hypothesis is neutral on the sort of innovation that was induced by increased market access, as the U.S. has always allowed process patents, and among product patents there is no differentiation between final consumption goods and intermediate product, or other machines used in production.[4] Dhingra (2013) examines the choices firms make with regard to the introduction of a new brand (a new product). In this model, trade liberalization decreases the number of brands a firm offers; the model does not consider the question of new product entry via new firm entry.

As in the Melitz (2003) model firms enter a market (in this case, into existence) after paying a fix cost f. Firms' productivity is then drawn from a known distribution $G(\varphi)$, where φ indexes firms by productivity, from least to most productive. Consider only two states of the world: a closed economy or an open economy. After the entry decision, firms in closed economies can upgrade their production abilities, and thereby lower their marginal cost, by process innovation, by paying a fixed cost f_i. In Bustos (2011) the effect of the investment is certain, while in Lileeva and Trefler (2010) a firm draws a productivity increase from another known distribution (choice is made to match the investment heterogeneity seen in the data); in both firms are restricted to one product. Firms in an open economy can also choose to export if they pay a fixed cost f_x. Firms face an iceberg, per-unit cost of exporting τ. The market structure in the home and foreign market is monopolistically competitive, with consumers holding CES preferences, leading to constant markups. All markets clear (full employment), and there is free entry and exit. This leads to a heterogeneous productivity effect when a closed economy is opened to trade: some firms start exporting without investing, while others start exporting and invest. In autarky, firms above a productivity cutoff will invest in process upgrading. However, after the area is opened to trade, the investment

[4] In the nineteenth century patents were issued for both product and process innovations. For instance, of 13 patents issued to Boston, MA residents between January 1, 1850 and April 30, 1850 five were "product" patents—three of which were chimney-caps—and the rest were "process" patents. Thus, while process patents were by far the more important, product patents were also important.

cutoff decreases. Thus, if applying for a patent is an important part of process innovation, this model provides motivation for looking for increased patenting activity with the introduction of improved transportation infrastructure (note that Melitz (2003) shows that in his model reducing transport costs is very similar to lowering other costs of trade).

3 Background

3.1 The Patent System

The present system of state-created and -enforced monopolies in intellectual property, patents, developed from of an older tradition of state monopoly grants (Bracha, 2004). In Britain and its colonies, by the late eighteenth century, the process for requesting grants promoting the development of new industries and technologies in a location had become somewhat routinized. Still, the 1790 patent act, which outlined uniform standards for what was patentable and a uniform process for obtaining a patent at a low fee (about $5), was the first of its kind (Khan, 2005). The U.S. system became a model for other countries as they introduced or reformed their patent systems during the nineteenth century.

It is striking that, as a new country on the periphery, the 1790 patent act specified that a patentee must be "the first and true inventor" anywhere in the world. Most nineteenth century patent systems allowed grants to go to those who were the first to introduce the technology into the country, as colony and then state patents had done (Hrdy, 2013). The U.S. system fluctuated in its enforcement of this mandate, most notably dropping any attempt to examine patents for novelty in 1793 (but increasing the fee to $30) and then reinstating examinations in 1836 (see appendix A.1 for more details on changes in patent law). In 1887 the U.S. joined the International Convention for the Protection of Industrial Property that had been formed in Paris in 1883.

The creation of a monopoly over the patented invention is generally justified in two ways: by the increased incentive to engage in innovation that this monopoly creates and by the value of the information that inventors are forced to disclose as part of the application process. This public disclosure sets patenting apart from other methods that inventors might use to capture gains from innovation, most notably trade secrets (Moser, 2004). The effectiveness of the disclosure requirement varied, though inventors were required to describe their invention so that "person having ordinary skill in the art" would understand. Until 1871, anyone looking to copy the information contained in the patent would have needed to travel to Washington, DC or pay a substantial reproduction fee in order to read a patent specification. The patent office didn't publish summaries of issued patent until 1872. Therefore, investors interested in newly issued patents relied on the summaries provided by private periodicals; the Journal of the Franklin Institute published its first issue in 1826, and Scientific American Magazine in 1845. Despite the nationalistic name of the latter, both devoted substantial space to new inventions of British origin. The creation of a new, patentable innovation requires learning where the technological frontier is: what problems are interesting, what the existing best solutions are, and what lines of research have been or are being explored. Thus, the role of institutions that disseminate this knowledge is potentially large.

As was the general character of firms in the early nineteenth century, invention was primarily done by lone individuals early in the century. Later in the century, larger firms explicitly investing in research and development became common. It was not until 1910, outside of the period of this study, that in-house R&D became the dominant mode of financing innovation (Lamoreaux and Sokoloff, 2005).

3.2 Transportation Improvements

Before the twentieth century, transport costs were of utmost importance. Moving goods over land without mechanical power, even on the best highways, was tremendously costly.

In the eighteenth century U.S. coastal cities were in many ways more economic-entangled with Europe than their hinterlands, not only due to colonial links, but also the large, several orders of magnitude, cost differentials between sea and overland transport. The distribution of population in Eastern Canada reflects this importances of water transport; even in the placements of cities today, the importance of transport along the St. Lawrence to the Atlantic is clear.

This lack of access to the interior of the country drove investments in transpiration infrastructure. In the era before the steam engine, this meant the building of postal roads and efforts to increase ease of travel on rivers.[5] The early part of the nineteenth century saw a large investment in canals. The most notable of these was the Erie in upstate New York, which opened in 1825, but there was also significant construction in both Pennsylvania and Ohio (including the eponymous canal, which opened in 1840).

Railroad construction in the U.S. began in 1820 with the first constitution being short links to mines and cities connecting to nearby, (seasonally) water inaccessible hinterlands, but it was in the 1850s that the U.S. experienced its first great wave of rail expansion. Approximately 22,000 miles of track were laid that decade. By 1860, in addition to dense coverage in the Northeast where rail construction began, the railroad network had reached Illinois, Indiana, and Ohio, with significant penetration into Wisconsin and Iowa. The South saw less construction, but it too experienced substantial growth in rail access in the 1850s (Atack et al., 2010). By 1870 the transcontinental railroad had been completed, though the western market it served was mostly limited to the San Francisco Bay area. By 1890, the areas around Portland, Seattle, San Fransisco, and Los Angeles all were connected to the same national rail network that had covered the country east of the Dakotas.

[5]Concurrently there were large investments in turnpikes in Britain (Bogart, 2005).

4 Data

4.1 Patents

A patent is a legal document that grants a limited term monopoly. The U.S. Constitution gives Congress the power "[t]o promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive rights to their writings and discoveries." In 1790, as permitted by this clause, federal patents were introduced, requiring that patentable subject matter be novel, useful and that the patent disclose the patented invention.[6]

The requirement of novelty was backed by the declaration in the 1790 Patent Act that the Patentee was supposed to be "the first and true inventor" anywhere in the world. The enablement requirement specified that a patent application disclose a claimed invention in sufficient detail for the notional person skilled in the art to carry out that claimed invention and, vitally for the use of this data, that this description of the invention be made available to the public immediately upon issue. Thus patents themselves transmit information about new technological ideas. Thought the enforcement of these requirements has greatly fluctuated, today legislation requires many of the same things: novelty, non-obviousness, (nominally) utility, and enablement (see appendix A.1 for more details on the law changes in the 19th century). Thought patents are only issued to individuals and not to corporations, these intellectual properties can be, and are often, sold. If the rights a patent confers has been sold prior to its being granted it is "assigned at issue," this assignment is recorded on the patent specification.

[6] While the first patents may have been issued with an eye toward the legislative discretion that chartered the monopoly rights from the system evolved, patents were soon seen as themselves a "right" owed to any new invention, regardless of government judgment on criteria besides novelty and non-obviousness (Bracha, 2004).

Though patents are imperfect indicators of technological improvement, they are the most accessible and detailed written records of innovation. As such, economists and economic historians have long studied them to probe the economics of technological development. While these studies have been wide ranging, they have been constrained by the availability of data. Until very recently, the difficulty of gathering a truly representative sample has necessarily focused research on narrow subsets, be they single industries or limited geographical areas.

This paper makes use of a geo-linked database of patents issued 1790-1900 and links together several data sources. In 1836 there was a fire in the patent office, which burned all the patents that had been issued to that date. In an attempt to recover from the damaged this caused the patent office put out a call for existent information on patents; in 1874 congress used the information the patent office had received to compile a list of patents issued 1779-1836 (a 1847 print volume presumably provided much of this information). This 1874 congressional list has been updated by volunteers, such as Jim Shaw and the maintainers of the Directory of American Tool and Machinery Patents, who have found patents that the 1874 congress was not aware of. I geo-located the patents by merging the town and county information with a database of historical town names from the AniMap 3.0.2 County Boundary Historical Atlas and the U.S. Board on Geographic Names (part of the Department of the Interior). This involved disambiguating many duplicate town names, and correcting a number of typographical errors in the patent database. This allowed each patent to be located with latitude and longitude coordinates. The rest of the patent data comes from Tom Nicholas' dataset of patents issued from 1836-1900, which has latitude and longitude coordinates of the listed places on these patents. These geo-located patents are then merged with the National Historical Geographic Information System (NHGIS) shape-files of U.S. county boundaries. This allows patent counts by county to be created.

This paper uses both contemporaneous county boundaries and a sample of consistent land area counties, harmonized to 1840 boundaries as suggested in Hornbeck (2010).[7] U.S. Census Data is from Haines (2010). Transportation data are from Atack (2013), these data are linked with shape-files of U.S. county boundaries to explore the spread of railroads and canals.

Figure 1 shows the total patents in each year as well as the US population from the census. Note that that number of patents per person does is not at all smooth over time, and the large increase in patenting activity that starts in the 1850s. Tables 1, 2, and 3 show the summary statistics for each year. In these table and in most of the analysis done in this paper, the number of patents refers to those issued in a two year period: the complete named year and the complete year before the named year. Years in which there is no data for a variable are years for which it is missing in my dataset.

Figures 2, 3, and 4 are maps that show the spread of patenting across the country. They present the number of patents issued per 1000 people in five year bins following the named year as well as the spread of the canal and railroad network by county for six benchmark years, 1790, 1810, 1830 (this picture is nearly equivalent to the later period presented in Sokoloff (1988)), 1850, 1870, and 1890. One can clearly see the increasing area that is involved in patenting, as well as the increasing amount of patenting per person.

Figure 5, give more sense of how the concentration of patenting has changed over time. The Herfindahl index falls substantially over the nineteenth century, reading a low in 1870, and being close to flat between 1880 and 1900.

[7]For more details on this computation see the boundary shifting files on my website.

4.2 Approximating Market Access

There are two ways I measure transportation access: the percent of the county that is within a given number of miles from transportation and approximations of equation 2. As noted by Fogel (1964) and contemporaries writing about the "transportation revolution" overland transportation was orders of magnitude more costly than other modes. Even if a place was relatively close to a railroad or waterway, the last few miles overland will likely make up the bulk of the transport cost. Thus knowing how much of the county is within easy range of transportation is a reasonable approximation of how much of the county has greater access to markets.

Several approximations need to be made to tractably apply equation 2 to the data available for the nineteenth century U.S., where there are no good estimates of regions' incomes nor of the trade flows between counties. First, as implied above, $\frac{\tau_{ij}}{P_i P_j}$ can be estimated as the transportation cost between i and j. Second, since GDP is unavailable, population is used as a crude proxy for income, under the assumption that income distribution is roughly equivalent across the US. One can think of the market available to a firm as the number of people that it can reach with its product; people are out of reach if formal barriers prevent them from buying the product, if their income is too low, if cultural barriers are too high, or if transportation costs make the product unfeasibly expensive. A firm's reach declines with the size of a resistance term. The size of the market that a firm in any one location has access to is a function of its position and the resistance term. It can be estimated by a trade resistance weighted sum of the people in all locations (assuming income is homogenously distributed across regions).

Taken together, market access is modeled as a function of transport costs and the distance of a county from other populations. Thus market access for a county, MA_j, can be approximated as:

$$MA_j \approx \sum_i {}^{pop_i}/_{pop_{US}} \cdot c_{ij}^{-\theta} \tag{3}$$

where pop_i is the population of location i, pop_{US} is the total population of the U.S., c_{ij} is the resistance term between i and j, approximated as the transportation cost (cost times distance) between i and j, and $\theta = \sigma - 1$. An alternate way of deriving market access is described in Donaldson and Hornbeck (2013). They approximate it as:

$$MA_j^{DH} \approx \sum_i pop_i \cdot c_{ij}^{-\theta} \tag{4}$$

Both MA_j and MA_j^{DH} can rise either when c_{ij} falls, or when pop_i increases. However, MA_j rises only when the proportion of population that lives in i increases, rather than increasing for general economic growth; over the nineteenth century both the general cost of transportation and the population level will change substantially.

While the population that a county contains is an important component of market access, it independently affects the patenting rate.[8] When Donaldson and Hornbeck (2013), approximate market access they omit self-population out of endogeneity concerns; this paper does the same.

I compute an approximate market access based on equation 3. In order to simplify the calculation, for computation reasons, each counties approximate market access is only based on larger locations—as their effect will be so much larger than the effect of smaller locations.[9] Also, given the setting, it is being able to move new goods to the cities that seems to be

[8] Indeed, since I measure patenting as patents per capita, it mechanically effects this variable.

[9] In particular I compute the cost of moving from every county to: New York, NY; Suffolk, MA (Boston); Philadelphia, PA; Washington, DC; Albany, NY; DeKalb, GA (Atlanta); Baltimore, MD; Erie, NY (Buffalo); Mecklenburg, NC (Charlotte); Cook, IL (Chicago) ; Hamilton, OH (Cincinnati); Frankin, OH (Columbus); Wayne, MI, (Detroit); Marion, IN, (Indianapolis); Pulaski, AR (Little Rock); Jefferson, KY (Louisville); Davidson, TN (Nashville); Orleans, LA; Allegheny, PA (Pittsburgh); Henrico, VA (Richmond); Monroe, NY (Rochester); St Louis, MO. I tried to pick cities form ever region, and that rose to importance during the ninetieth century, as well as ones that were important at the begging of it.

the concerning problem. See figures 6, 7 and 8 that show approximate market access in 1810, 1840, and 1870, notice the high computed market access along the coast, the low computed market access in Appalachia, and the increasing importance of the railroad over time. As noted above, I use transportation from Atack (2013), while population data is from U.S. Census Data is from Haines (2010) with county boundaries harmonized to 1840 as in Hornbeck (2010).

Approximate transportation cost is computer by joining Atack (2013)'s transportation data with the 1840 U.S. county boundary shapefile from NHGIS; after which each county is denoted as having a railroad, river, canal, or port. If two adjacent counties have one of these features, they are assumed to be connected. Starting from the population center in questions I give approximate costs to every county adjacent to it using centroid to centroid distances (if less than 150 miles from each other) and rates taken from Donaldson and Hornbeck (2013) which in turn takes them from Fogel (1964).[10] If one of the adjacent counties has not been priced by this procedure yet, or if the newly calculated cost is lower than the previous, it is added to a queue of counties to be examined. In each examined county, the cost of transportation to its neighbor county is added to the cost in the examined county to produce the total cost from the seed county. Repeating this procedure until there are no further counties to examine in the queue yields the minimum transit costs from the original seed county to all other counties. While the adjacency rules in this procedure are not as precise as using network analysis, it is less computationally intensive, and the data that are used need not have the level of precision necessary for Network Analyst to produce results. However, for ports, the full network analysis approach is more straightforward (since there are no intermediate counties). All port counties are assumed to be adjacent to all other

[10] Travel along a railroad cost 0.63 cents per ton-mile, a canal 0.50 cents a ton-mile, river or other waterway 0.49 cents per ton-mile, wagon or overland 23.1 cents per ton-mile, and changing mode of transportation 50 cents per ton (e.g. unloading from a rail car onto a river barge). Also following Donaldson and Hornbeck (2013) I take θ to be 3.8.

port counties, with a distance computed by using modern waterway network data from the National Transportation Atlas Databases. Other county adjacency is computed by modifying a tool written by Chieko Maene.

5 Results and Discussion

5.1 Percent of County Near Transportation

In creating a connection between cheap access to transportation and innovation Sokoloff (1988) presented three cross sections, for 1805-1811, 1823-1829, and 1830-1836, all of which show a positive relationship between the logarithm of patents per million people in a county and its being located on a navigable river or canal (Sokoloff, 1988, pg. 838, table 5). As a starting point for examining the effect of transportation access on patenting rates, table 4 similarly looks at relationship between the percent of a county with in five miles of transportation and the logarithm of patents per million people for a cross section of the data from 1810, 1840 and 1870. The specification is:

$$\log\left(\frac{Patents_{is}}{Population_{is}/_{OneMillion}}\right) = \alpha + \beta\, TransportationAccess_{is} + \zeta \mathbf{X_{is}} + \gamma_s + \varepsilon \qquad (5)$$

where $\log(\frac{Patents_{is}}{Population_{is}/_{OneMillion}})$ is the measure of patenting at the county (i) level, $TransportationAccess_{is}$ is the measure of transportation access, $\mathbf{X_{is}}$ are county characteristics, and γ_s are state dummies used for each year. This specification is very similar, but not identical, to the one used in Sokoloff (1988). The notable differences are that in Sokoloff (1988) the left hand side considers patents issued in the bins noted above, annualized, rather than the two year increments used here, and that a number of variables that were indicators in Sokoloff (1988) are continuous variables here: the percent of a county that is urban (more than 2,500 people in an incorporated area), the percent of county that is metropolitan (more than 25,000 people in an incorporated area), and the transportation variables. Note that because the left hand size is logarithmic, and a single patent is a large observation, only counties with some patenting are included in this regression. The coefficient, β, on access to transportation is positive in all years, but it is small and imprecisely estimated for 1810 and 1840, and grows over time. Including the counties that do not patent as zeros increases the size and precision of this estimate, but not the qualitative relationship between the years. Also notable is that, as in Sokoloff (1988), being on the ocean has a negative relationship with patenting, that the importance of metropolitan population decreasing over time, while, unlike in Sokoloff (1988), the importance of urban population increases. Thus table 4 suggests a positive relationship between patenting and new transpiration access in all years, but only a weak one in the early part of the century; older transportation access, in the form of ports, may be negatively related to patenting.

In order to understand how this relationship between transportation access and patenting evolves over time, figures 9, 10, 11, and 12 show the mean patenting per 10,000 people as a function of the number of years until the railroad or a canal arrived. In all figures the point estimates are those on the dummy variables for the number of years to the arrival of the railroad in a county from a regression of:

$$\frac{Patents_{it}}{Population_{it}/_{10,000}} = \alpha + \beta \mathbf{YearstoArrivalDummies_{it}} + \gamma_i + \delta_t + \varepsilon \qquad (6)$$

where $\frac{Patents_{it}}{Population_{it}/_{10,000}}$ is the measure of patenting at the county level, $\mathbf{YearstoArrivalDummies_{it}}$ are dummy variables for the number of year until a county, i, receives a railroad (as above), γ_i are county fixed effects, and δ_t are year fixed effects; standard errors are clustered at the county level.

Figure 9 examines all counties in the North East, while figure 10 only includes the counties that eventual received a canal. In the first the arrival of a canal is not associated with a change in the predicted number of patents per capita, however when the sample is restricted to counties in which a canal would be built, the arrival of a canal was indeed assorted with a gradual increase in patenting. The counties that are excluded from this analysis are both some of the most developed (places with enough river transport that canals are not needed, such as New York or Suffolk) and most peripheral (counties in rural Maine). Thus, the canal may have promoted development and growth relative to a group that didn't otherwise have good transportation access.

In figure 11 the presence of the railroad is measured at ten-year intervals. That is, if the railroad arrived in a county in 1854, it would be measured as not present in 1850 and present in 1860, and 1850 would be coded at ten years until railroad arrives (or -10), despite the fact that in reality it was only four years until arrival, and 1860 would be coded as the date of arrival, zero years, despite the fact that in reality it was six years ago. In contrast figure 12 shows yearly data, but for Mid-western states only. In both figures 11 and 12 an increase in patents per 10,000 people once the railroad arrives in a county, the point estimates remain flat before the railroad arrives and then increase steadily afterward. Both of these pictures include all counties, suggesting that the effect of the railroad was widespread. While in the 1830 the small amount of rail seemed to be a substitute for water transport, by 1840 a large number of the most developed places that never received a canal had a railroad (including New York and Suffolk)—perhaps avoiding the cost of transfer modes. However, as figure 12 only considers the mid-west, these old port cities are not in the sample. These figures suggests that the arrival of inexpensive transportation encourages the growth of innovative activity. This effect might be seen for many reasons. Perhaps the arrival of transportation encourages general economic development, or encourages the growth of dense areas more conducive to innovation, or perhaps it brings with it greater connection to the formal institutions of intellectual property protection.

To further understand the process of transportation spread figure 13 examine the number of patents per 10,000 people in counties over time by the year that these counties received a canal or the railroad. This plots the mean of patents per 10,000 people in each railroad arrival category in each year. Note that the counties that received the railroad the earliest also had the highest level of patenting per capita (and also over all, as they tended to have larger populations). This pattern is one reason why in crossection the presence of transportation would be correlated with the number of patents per capita.

The increase shown in figures 9-12 suggests that individual counties increase the number of patents per capita, and thus it is not mere corssectional variation or population growth. To understand this better I turn to fixed effects regressions:

$$Patents_{it} = \alpha + \beta TransportationAccess_{it} + \zeta \mathbf{X_{it}} + \gamma_i + \delta_t + \varepsilon \tag{7}$$

where $Patents_{it}$ is the measure of patenting, which will either be $\frac{Patents_{it}}{Population_{it}/_{10,000}}$ or $\log(\frac{Patents_{it}}{Population_{it}/_{OneMillion}})$; $TransportationAccess_{it}$ is the measure of transportation access which will generally be the percent of the county's area within 1.5 or 5 miles of transportation (as above this is a number between zero and one, rather than between zero and a hundred); $\mathbf{X_{it}}$ are county characteristics; γ_i are county fixed effects, and δ_t are year fixed effects.

Table 5 column (1) shows the results of a regression of patents per 10,000 people on a dummy for the presence of a railroad in a county (and year and county fixed effects). The increase seen in figure 13 is reflected here. The other columns of table 5 use a slightly more nuanced measure of access to transportation, the percent of a county within a distance (1.5, 5, or 15 miles) of any transportation (port, canal, river or railroad) or of a railroad. Note that although the coefficients on 1.5 miles of transportation, and on rail in particular, are larger than the ones associated with in 5.0 miles of transportation, the elasticity estimates are not dissimilar. In contrast, the percent of a county within 15 miles of transportation or rail has a much smaller elasticity. These estimates suggest a positive relationship between transportation access and patenting that is stronger for the newer modes, rail rather than ports, and that the advantage of transportation access comes not just from advantages a farmer would get from occasionally bringing crops to the rail depot, but from having transportation close enough to be a reasonable round-trip in a day. This suggest that whatever is causing this increase it is not simply from the lowered cost the occasional shipments of goods. It is striking that the estimates only considering the railroad and considering all transportation don't vary much, though it is true that the vast majority of the variation in transportation access comes from the spread of the railroad.

Similarly, table 6 shows fixed effects regressions of the logarithm of patents per million people on measures of transportation access, thus looking at something more like specification used in table 4. Columns (1), (3) and (5) omit counties that do not patent, while columns (2), (4) and (6) include them as zeros (the minimum observed of log patents per million people is 2.5). This suggests that the intensive effect—the effect of increasing the number of patents in counties that patent is about 3 to 5 times smaller than the the effect considering those counties that move from not patenting to patenting. Given that for most counties the movement of one patent (weather from 0 to 1 or 5 to 6) represents a large change in patenting, so the extensive-margin effect is important. These estimates suggest that for a county that is patenting, a 1% increase in the area of the county that is within 1.5 miles of some form of transport corresponds to a to a 1.5% increase in patenting. In 1850 the counties that were patenting had a mean of 16% (0.16) of their area within 1.5 miles of transportation, with a standard deviation of 0.14, this has a large mass at values below 30% and a thin tail afterward, and 121 patents per million people, with a standard deviation of 105 and a distributions looks to be roughly Pareto. This means that a one standard deviation increase in area of a county with in 1.5 miles of transportation corresponds to about a 0.2 standard deviation increase in patenting—which given the zero bound and size of the tail is a substantial increase.

Table 7 shows a fixed effects regression of patents per ten-thousand people on measures of transportation plus county controls. Included are the year interacted with logarithm of the number of workers in manufacturing over number of workers in agriculture in 1820 and 1840 (so these values are attached to every year); the year interacted with percent of the county that is literate in 1840, the percent of the county that attended school in 1840 and the percent of farm land that was improved in 1850 (so, again, these values are attached to every year); the year interacted with the percent of county that is urban (defined as an incorporated area of more than 2,500 people) in that year and every one previous to the one in question; the year interacted with the percent of county that is metropolitan (defined as an incorporated area of more than 25,000 people) in that year and every one previous to the one in question; similarly the values of percent of the county that is literate in 1850 and 1870, the percent of the county that attended school in 1850 and 1870, the percent of farm land that was improved in 1860-1900; and year interacted with a full set of pretends for the number of patents per ten-thousand people and the transpiration measure. The other measures of transportation used in table 5 were tried, and are similar to the reported measures. Note that coefficient on transportation access falls significantly, but remains positive.

Given how much smaller the estimates become with the inclusion of county level controls, there is significant reason to worry that something is biasing the above estimate. To address this, I propose an instrument for the spread of the railroad across the country (this instrument is also found in Atack et al. (2012)).[11] After the independence of the U.S., the populace began

[11]This instrument owes its inspiration to Gutberlet (2014), which uses lines between pre-existing cities in Germany to predict the spread of the railroad, and Michaels (2008), which uses the orientation of a line between the nearest city and a county centroid to predict the presence of an interstate highway.

colonizing the interior of the continent, moving west from the eastern coast. While it is not literally the case that people simply walked due west from major cities, this was a popular story, enough that it drove people's understanding of westward space in the U.S. I take advantage of this westward orientation to instrument for where the railroad was eventually built. Figure 14 shows the 14 port cities with a customs houses or a public warehouses that had been built by the US government in 1826 (HRD, 1826), which were the largest ports and lines running due west and north from those locations. Note the way those strait westward lines correspond with where population had settled by the 1860s, heavily in upstate New York, but not proceeding onward to Michigan, across the southern mid-west (Ohio, Indian, and Illinois), and in a belt in the South. The one line running north of significance captures those that settled along the Mississippi. The variable is defined at the county level: a county takes on a one if one of these lines intersects it, and a zero otherwise.[12] Thus tables 8 and 9 show a fixed effects regressions as in equation 11 and the first stage and two stage least squares estimation given by:

$$TransportationAccess_{it} = \alpha + \beta\, OnCardinalDirectionLine_{it} + \zeta \mathbf{X_{it}} + \gamma_i + \delta_t + \varepsilon \qquad (8)$$

and

$$Patents_{it} = \alpha + \beta\, \widehat{TransportationAccess}_{it} + \zeta \mathbf{X_{it}} + \gamma_i + \delta_t + \varepsilon \qquad (9)$$

where $Patents_{it}$ is the measure of patenting, which will either be $\frac{Patents_{it}}{Population_{it}/10{,}000}$ or $\log(\frac{Patents_{it}}{Population_{it}/OneMillion})$; $TransportationAccess_{it}$ is the measure of transportation access which will generally be the percent of the county's area within 1.5 or 5 miles of a railroad; $OnCardinalDirectionLine_{it}$ is the indicator for weather or not one of the cardinal direc-

[12] Alternate possibilities to interments for presence of a railroad are: The lines between end points in congressional railroad surveys as used for the Midwest in Atack et al. (2010), however, this only covers the Midwest. A measure of roughness, however, as much of the wealth outside the industrial northeast was from agriculture, and most of the production was agricultural this instrument is correlated with many important features. It also lacks variation in much of the Midwest.

tion lines described above intersects a county; $\mathbf{X_{it}}$ are county characteristics; γ_i are county fixed effects, and δ_t are year fixed effects. Both tables 8 and 9 suggest significant downward bias the basic difference in differences estimations, piratically when controlling for county characteristics, of the effect of the percent of a county within 1.5 or 5.0 miles of a railroad on patents per 10,000 people. Only using the westward lines leads to basically the same result. Using the estimates from table 8 with controls, 32% of the increase in patenting between 1850 and 1860, and 17% of the increase in patenting between 1850 and 1870 can be explained by the spread of the rail; the estimates from table 9 with controls allow for the increase in the railroad network to explain 67% the increase in patenting between 1850 and 1860, and 32% of the increase in patenting between 1850 and 1870.

The arrival of the railroad in counties across the U.S. was assorted with increases in patenting above their population growth. In particular, the percent of the counties land area that is very close to the new transportation connection, in the range of 5 miles away, is associated with this increase even when controlling for the composition of the labor force, the literacy rate, and most importantly, the urbanization rate. IV estimates are much larger than simple fixed effects estimates. This suggests that being very near enough to a railroad to reasonably make a round trip in a day effects the patenting rate above and beyond the urbanization it brings (Atack et al., 2010). Patenting is encouraged because of the access to the larger world, ether from increased information flows or the possibility of larger returns due to selling the resulting improved products (ether directly, in the case of new products, or indirectly in the case of process innovation or in the case of sale of the intellectual property itself).

5.2 Market Access

The other measure of transportation proposed above its an approximation of market access. The percent of a county that is near transportation relies on the large cost differential between overland transportation and railroad or water transport. It doesn't take into account where in the larger that connection is made, so a port on Long Island Sound (with easy access to New York City) is treated as the same as a port on the Great Lakes. Approximate market access takes this into account, while the density of the network in any one location is downplayed. Thus, it reflects potential long distance rather than local connections.

Table 10, examines market access in three cross-sections, as table 4 does for the percent of a county within 5 miles of transportation. That is, this table looks at the relationship between the logarithm of market access as described in section 4.2 and the logarithm of patents per million people for a cross section of the data from 1810, 1840 and 1870. The specification is:

$$\log(\frac{Patents_{is}}{Population_{is}/_{OneMillion}}) = \alpha + \beta MA_{is} + \zeta \mathbf{X_{is}} + \gamma_s + \varepsilon \tag{10}$$

where $\log(\frac{Patents_{is}}{Population_{is}/_{OneMillion}})$ is the measure of patenting at the county (i) level, $Transportation_{is}$ is the measure of transportation access, $\mathbf{X_{is}}$ are county characteristics, and γ_s are state dummies used for each year. The coefficient, β, on log market access is negative and imprecisely estimated in 1810 and 1840 (note that the mean logarithm of market access is negative for both of those years). However, including counties that do not patent as zeros changes the sign, size, and precision of this estimate, with this inclusion the estimate it is positive in all years and, as here, grows over time. Supporting table 10, figures 16, 17, 18 show the crossselectional relationship between patents per 10,000 people and the logarithm of market access, making it clear that the relationship isn't strong until 1870.

Also notable in table 10 is that like in table 4 the importance of metropolitan population decreases over time, while the importance of urban population increases. This suggests that for counties that patent there isn't a strong relationship between market access and patenting until the later half of the nineteenth century, but counties that patent have more market access than counties that do not patent. One might worry that the approximation of only computing access to larger places might have more of a bias in years before access to transportation became wide spread—as one might expect the correlation between precedent of the county within 1.5 miles of transportation and the logarithm of market access is much stronger later in the nineteen century. Thus, it seems like local connections matter more than long distance ones, and the positive coefficient in 1870 reflects the density of the railroad by that year.

To understand how this relationship changes with changes in market access rather then time, table 11 looks at a fixed effects regression:

$$Patents_{it} = \alpha + \beta \log MA_{it} + \zeta \mathbf{X_{it}} + \gamma_i + \delta_t + \varepsilon \tag{11}$$

where $Patents_{it}$ is the measure of patenting, which will either be $\frac{Patents_{it}}{Population_{it}/_{10,000}}$ or $\log(\frac{Patents_{it}}{Population_{it}/_{OneMillion}})$; MA_{it} is an approximation of market access; $\mathbf{X_{it}}$ are county characteristics; γ_i are county fixed effects, and δ_t are year fixed effects. Note that including the counties that do not patent greatly increases the number of observations, as well as the size of β. While the levels regression still has a positive point estimate for the intensive margin only, the regression on the logarithm has a negative point estimate, as does table 6 column (1). Including controls actually reduces the size of this negative estimate, as it does for the negative estimate using the same specification and percent of the county within 1.5 miles of transportation (see table 7 column (2) and table 12 columns (3) and (4)). Thus one might assume that the effect of increased market access matters much more to the extensive margin, affecting whether a county does or does not engage in patenting. Indeed, the point estimate in a regression where the measure of patenting is a dummy for whether or not a patent was observed in the county is positive and precisely estimated. Market access is a stronger predictor of weather or not a county patents, than it is of the level of patenting among counties that patent.

Table 12 show a fixed effects regression with county-level controls. Included are: the year interacted with logarithm of the number of workers in manufacturing over number of workers in agriculture in 1820 and 1840 (so that these values are attached to every year); the year interacted with percent of the county that is literate in 1840; the year interacted with the percent of the county that attended school in 1840; the year interacted with the percent of farm land that was improved in 1850; the year interacted with the percent of county that is urban (defined as an incorporated area of more than 2,500 people) in that year and every one previous to the year in question; the year interacted with the percent of county that is metropolitan (defined as an incorporated area of more than 25,000 people) in that year and every one previous; similarly, the year interacted with the percent of the county that was literate in 1850 and 1870; the year interacted with the percent of the county that attended school in 1850 and 1870; the year interacted with the percent of farm land that was improved in 1860-1900; and year interacted with a full set of pretends for the number of patents per ten-thousand people. Columns (2), (4) and (6) also include a full set of pretends in Market Access, which decreases the point estimate. As before the estimate is positive if one includes counties that don't patent, but flat otherwise.

Figure 19 give some clues to this decrease. This figure shows the relationship between the first difference of patents per 10,000 people and the first difference of the logarithm of market access. The overall pattern is flat. However, the U-shape in the middle suggests that there may be multiple regimes, which will be explored below. Figure 15 helps illustrate these regimes. It plots the first difference of market access on a map, showing which counties are losing relative market access in 1870 and which counties are improving their market access in relative terms.[13] All counties are gaining access to larger populations as transport spreads; however, some places are growing faster than average, or having transportation built at a

[13]This is in contrast to absolute market access, which grows along with overall population growth, and is what Donaldson and Hornbeck (2013) calculates.

rate faster than average. Note that the places losing relative market access are the better developed locations. In those counties, regional-level transportation access is not increasing greatly because the infrastructure is already well developed.

Splitting the U-shape in the middle figure 19 into two parts helps examine the two regimes. Thus figure 20 looks at the right side of this U and figure 21 looks at the left side of it. Figure 20 shows the negative side of the U by plotting the first difference of patents per 10,000 people and the logarithm of negative one times the first difference of market access, so that on the negative side of the graph are the counties that are losing very little market access, and on the positive side are counties that are losing much more market access, which are the more developed counties discussed above. While this corrolation between losing market access and more patents is the opposite of what one might expect, it is explain by the fact that the eastern most places in the U.S. are the ones losing market access. Figure 21 shows the positive side of the U by plotting the first difference of patents per 10,000 people and the logarithm of the first difference of market access. This shows the expected relationship: places that gain more market access see a bigger increase in patents per capita.

Tables 13 and 14 both examine these two sides of the U separately, that is, by looking at counties that lose separately from counties that gain relative market access. Table 13 regresses the first difference of patents per 10,000 people on the first difference of market access, split into two variables: one that take on the values of market access only if it negative and zero otherwise, and a similar variable taking on positive values of market access. Note the negative value on negative market access is as one expects from figure 20 (places that lose greater relative market access have larger increases in patenting) and the generally expected positive estimate on positive market access (places that gain more market access have larger increases in patenting). Column (2) includes the logarithm of total population effects. Note that the point estimate on negative market access is approximately zero, while the coefficient on positive market access remains positive and significant. This suggests that being more developed does explain the relationship between losing a larger amount of relative market access and a larger increase in patenting, but does not explain the relationship between gaining more relative market access and a larger increase in patenting. Table 14 similarly splits these two regimes apart, but does so by separately regressing first difference of patents per 10,000 people on the logarithm of the first difference of market access or on the logarithm of negative one times the first difference of market access for the years 1810, 1840 and 1870. As expected, 1870 shows the same pattern seen in table 13, 1810 has a broadly similar pattern. However, the coefficient on all terms is negative in 1840. This could be explained by the change in patent law from a pure registration system to an examined system in 1836, meaning that on average the number of patents per capita actually decreased between 1830 and 1840. This decrease is not systematically related to total population or market access. Increases in market access are correlated with larger increases in patenting, even controlling for own population.

Increases in a county's market access were assorted with increases in patenting above their population growth. This is driven by the less developed counties in the mid-west, the effect is much stronger on the extensive margin than on the intensive margin—it seems the connection is instrumental in helping counties move from no patenting to some level of patenting, but not as important in determining the level of patenting thereafter.

6 Conclusion

This paper has expanded upon the work of Sokoloff (1988), testing his hypothesis that increased market access due to the transportation revolution was a key driver of innovation for the United States in the nineteenth century. Using a large dataset of geo-located patents cross-referenced with a map of the expansion of transportation infrastructure, it shows that the arrival of rail transport has a large and positive effect on patenting behavior. These gains are realized slowly over time, suggesting that market access causes a trend change in the overall rate of development rather than a sudden innovative shock. This expansion in patenting is largely due to locations registering their first patents. Results from a regression that instruments for the arrival of the railroad with cardinal detection lines from the most important ports in 1826 suggest that 30-70% of the increase in patenting between 1850 and 1860 was caused by the spread of the railroad in this period, and 15-30% of the increase between 1850 and 1870.

The main hypothesis of Sokoloff (1988) falls in line with the intuition gained from the Bustos (2011) and Lileeva and Trefler (2010) extensions to the Melitz (2003) model, which suggest that, upon gaining access to larger markets, it is worthwhile for the firms with the lowest marginal costs of production to invest in productivity upgrades. The increase in units sold allows them to recoup a larger investment.

However, a transportation link might increase patenting for many reason. Transportation brings information both on innovations elsewhere and on the patent system itself. Transportation helps encourage general economic development, which in turn leads to more resources available for inventive activity. Because transportation encourages urbanization, it encourages patenting in many ways. Urbanization allows for the better transmission of information, ether about innovations or the patent system, as much learning is social. To this end urbanization encourages the creation of civil society, spaces where learning can take place. In this period prevailing cultural attitudes were such that scientific and technical societies were popular. Urban areas allow for greater specialization, which might encourage patenting by giving people in those areas better access to the bureaucracy of patenting (e.g. lawyers, machinists, draftspeople), or by encouraging innovation directly. Urban areas also allow for better access to formal credit markets, lifting liquidity constraints for potential innovators. Finally, in an urban area, secrecy may not effectively protect inventions, leading to patenting.[14]

Several of these more subtle reasons, beyond the Bustos (2011) and Lileeva and Trefler (2010) trade-driven, investment enabled by integration with a larger market mechanism that has been explored in this paper, for why patenting might increase bear further exploration. I will explore several of the hypotheses presented above in future work, focusing on two different sets of explanations for increase: the increase in flow of information, both from better connections to the wider world and the development of civil society; and the increased connectivity to the bureaucracy of the patent system.

The increase caused by access to transport is not a sudden jump, but rather a slow build up, whatever process is occurring is one of gradual development. The effects are more extensive later in the nineteenth century, a time where steam (rather than water or animal) power was used in the most productive factories. Steam power allowed for greater set of factory locations to feasible. Market access increases most with the first link to a larger network, whereas the percent of the county within 5 miles of transportation varies more with the density of connections within a county. It is suggestive that the latter is predicts the increase in patenting better than the former; if the basic trade story were driving these results, the reverse would be true. Transportation was important for increasing patenting in peripheral areas, it did so by increasing the overall process of urbanization and development in those ares.

[14]It is not a priori obvious that a transportation link will lead to more innovation in peripheral areas. It might also have increased the importance of being in the center of the network, or lead to human capital flight by providing and easier way to migrate to urban areas where high human capital is better rewarded.

References

(1826). H.R. Doc. No. 93, 19th Cong., 1st Sess. Technical report. (On page 26.)

Amiti, M. and Konings, J. (2004). Trade Liberalization , Intermediate Inputs , and Productivity : Evidence from Indonesia. *American Economic Review*. (On pages 3, 9.)

Atack, J. (2013). On the use of geographic information systems in economic history: The american transportation revolution revisited. *The Journal of Economic History*, 73:313–338. (On pages 5, 16, 19, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75.)

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (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. (On pages 13, 26, 27.)

Atack, J., Bateman, F., and Margo, R. A. (2008). Steam power, establishment size, and labor productivity growth in nineteenth century American manufacturing. *Explorations in Economic History*, 45(2):185–198. (On page 7.)

Atack, J., Haines, M., and Margo, R. A. (2011). Railroads and the rise of the factory: Evidence for the united states, 1850-70. pages 162–179. (On pages 2, 7.)

Atack, J., Margo, R., and Perlman, E. (2012). The impact of railroads on school enrollment in nineteenth century America. (On page 25.)

Becker, S. O. and Egger, P. H. (2013). Endogenous product versus process innovation and a firm's propensity to export. *Empirical Economics*, pages 1–26. (On pages 3, 9.)

Bogart, D. (2005). Turnpike trusts and the transportation revolution in 18th century england. *Explorations in Economic History*, 42(4):479 – 508. (On page 13.)

Bracha, O. (2004). The Commodification of Patents 1600–1836: How Patents Became Rights and Why We Should Care. *Loyola of Los Angeles Law Review*, 38:177. (On pages 11, 14.)

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *The American Economic Review*, 101(February):304–340. (On pages 3, 4, 7, 9, 10, 33, 34.)

Deloecker, J. (2007). Do exports generate higher productivity? Evidence from Slovenia. *Journal of International Economics*, 73(1):69–98. (On pages 3, 9.)

Dhingra, S. (2013). Trading away wide brands for cheap brands. *American Economic Review*, 103(6):2554–84. (On page 10.)

Donaldson, D. and Hornbeck, R. (2013). Railroads and American economic growth: A "market access" approach. (On pages 2, 5, 18, 19, 30.)

Fenichel, A. H. (1966). *Growth and Diffusion of Power in Manufacturing, 1838-1919*, pages 443–478. NBER. (On page 7.)

Fernandes, A. M. (2007). Trade policy, trade volumes and plant-level productivity in Colombian manufacturing industries. *Journal of International Economics*, 71(1):52–71. (On pages 3, 9.)

Feyrer, J. (2009). Distance , Trade , and Income – The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment. (On page 8.)

Fogel, R. W. (1964). *Railroads and American Economic Growth: Essays in Econometric History*. Johns Hopkins University Press, Baltimore, MD. (On pages 2, 17, 19.)

Foster, L., Haltiwanger, J., and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *American Economic Review*, 98(1):394–425. (On pages 3, 9.)

Gutberlet, T. (2014). Cheap Coal versus Market Access: The Role of Natural Resources and Demand in Germany's Industrialization. (On pages 7, 25.)

Haines, M. R. (2010). Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 [Computer file]. (On pages 16, 19, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75.)

Hornbeck, R. (2010). Barbed Wire: Property Rights and Agricultural Development. *The Quarterly Journal of Economics*, 125(2):767–810. (On pages 16, 19, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75.)

Hrdy, C. A. (2013). State Patent Laws in the Age of Laissez-Faire. *Berkeley Technology Law Journal*, 28(1). (On pages 11, 37.)

Khan, B. Z. (2005). *The Democratization of Invention: Patents and Copyrights in American Economic Development*. Cambridge Univ Press, Cambridge, MA. (On page 11.)

Lamoreaux, N. R. and Sokoloff, K. L. (2005). The Decline of the Independent Inventor: A Schumpterian Story? (On page 12.)

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity. . . for some plants. *The Quarterly Journal of Economics*, 125(3):1051–1099. (On pages 3, 4, 7, 9, 10, 33, 34.)

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725. (On pages 3, 4, 7, 9, 10, 11, 33.)

Michaels, G. (2008). The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System. *Review of Economics and Statistics*, 90(4):683–701. (On page 25.)

Minnesota Population Center (2011). National Historical Geographic Information System. No citations.

Moser, P. (2004). Determinants of Innovation Evidence from 19th Century World Fairs. *The Journal of Economic History*, 64(2):548–552. (On pages 2, 12.)

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from Chilean plants. *The Review of Economics Studies*, 69. (On pages 3, 9.)

Sokoloff, K. L. (1988). Inventive Activity in Early Industrial America: Evidence From Patent Records, 1790–1846. *The Journal of Economic History*, 48(04):813–850. (On pages 1, 2, 3, 4, 5, 9, 16, 20, 21, 32, 33.)

Topalova, P. and Khandelwal, A. (2011). Trade liberalization and firm productivity: The case of india. *Review of Economics and Statistics*, 93(3):995–1009. (On pages 3, 9.)

Trajtenberg, M. (1990). A Penny for Your Quotes: Patent Citations and the Value of Innovations. *The RAND Journal of Economics*, 21(1):172. (On page 2.)

Van Biesebroeck, J. (2005). Exporting raises productivity in sub-Saharan African manufacturing firms. *Journal of International Economics*, 67(2):373–391. (On pages 3, 9.)

A Appendices

A.1 Notable Developments in Patent Law

The U.S. federal patents were introduced in 1790. Previously the colonies individually granted patents (Hrdy, 2013) in a manner more typical of the royal monopolies they emulated. When federal patents were introduced, inventors were given the opportunity to apply for federal patents on things they had previously patented at the states level; the last state patent was issued in 1798. In 1790 federal patent applications were examined by the Secretary of State, the Secretary of War, and the Attorney General to see "if they deem the invention or discovery sufficiently useful and important," however, in practice this meant Thomas Jefferson examined patents. Patents were granted for a maximum of 14 years, the board was to determine the grant length of each patent, the fee was between $4 and $5, and a specification and drawing were required while a model was encouraged. There was no official channel to appeal decisions. The board of Secretaries (and, as noted above, in particular Thomas Jefferson) had trouble balancing their many other duties and examination of patent applications. The largest subsequent changes in patent law occurred in 1793, when examination was removed, and 1836 when examination was reintroduced, however several other term and fee changes occurred throughout the nineteenth century.

In 1793, the requirement that an invention be "sufficiently useful and important" was removed and patents were no longer examined for novelty. The fee was increased to $30, and aliens were not allowed to obtain patents. In 1800, this requirement was changed so the aliens that have resided in the county for 2 years and declare an intention of becoming a citizen may receive patents.

The next major change to patent law occurred in 1836, with the patent office established as a distinct bureau who is charged with examining patent application (the patent office consisted of a commissioner, a chief clerk, an examiner, a machinist, two draftsmen, an inferior clerk, and a messenger). Patent office employees are forbidden to acquire any interest in a patent, and a library of scientific works for use by employees of the patent office is created. In addition to the grant length of 14 years, the option to apply for a 7 year extension made available. The fee remained $30 for citizens, but foreigners are allowed to hold patents with fees of $500 for British subjects, $300 for all others. Applicants must file a specification, a drawing, and a model, and any appeals made be made to three "disinterested persons" appointed by the Secretary of State. That same year a fire in a patent office destroyed many of the previously issued patents.

In 1839 the patent office is charged with collecting statistics on agriculture, this continues until the department of agriculture is created in 1862. In 1842 design patents become available.

The act of March 2, 1861 extended the patent grant length to 17 years, while removing the possibility of extension. The fee structure was changed so the total was $35 in two payments, $15 at application and $20 on grant, to any person who was a citizen of a county “not discriminating against the US,” by 1924 this had increased to $40, $20 at application and $20 at issue. A permanent board of appeals consisting of three examiners was created. In August of 1861 the Confederate Patent Office granted its first patent, that office would grant a total of 266 patents.

In 1870 the filing requirements are changed so that a model only need be provided if requested, until 1880 models were generally requested—on Sept. 24, 1877 a fire in the patent office destroyed many models. The commissioner was also given leave to print copies of patents for the public. In 1871 congress discontinued its reports on patents issued that year, but distributes individual patents: “for the first time printed patent specifications became available to the public at a nominal charge. Hitherto, in order to study patents, it had been necessary to consult original drawings and specifications in the Patent Office or have copies made at considerable expense.” Continuing this movement toward information distribution, in 1872 the patent office starts publishing weekly excepts from patents and law in the Official Gazette of the United States Patent Office.

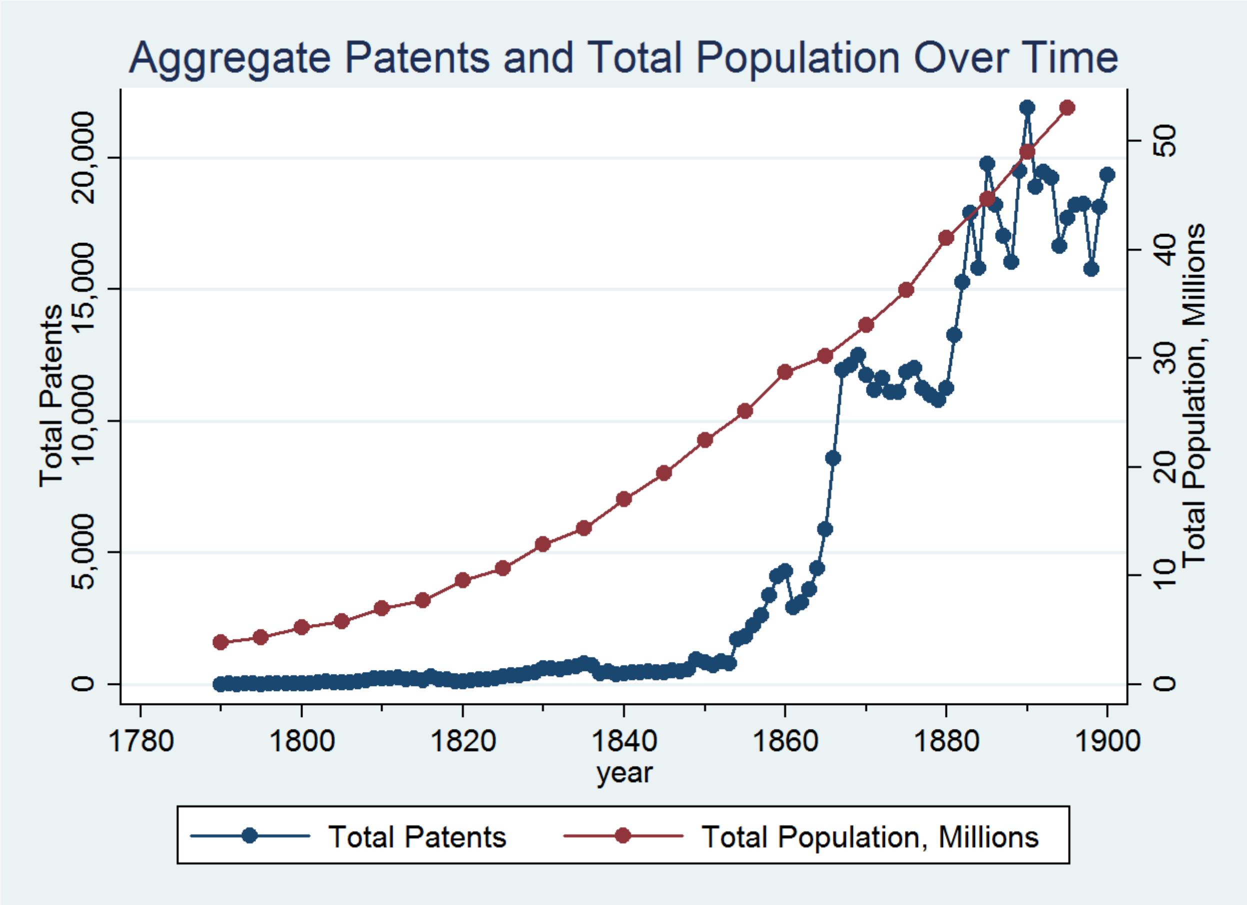
In 1887 the U.S. joins the International Convention for the Protection of Industrial Property that had been formed in Paris in 1883. That same year the question of the first inventor of the telephone draws wide public attention.

On June 10, 1898 a Classification Division is formed to reclassify all patents—thought the first classification had been published many years before in 1830. It consisting of sixteen categories, it was updated and expanded to 22 categories in 1847 and then several times there after.

B Figures and Tables

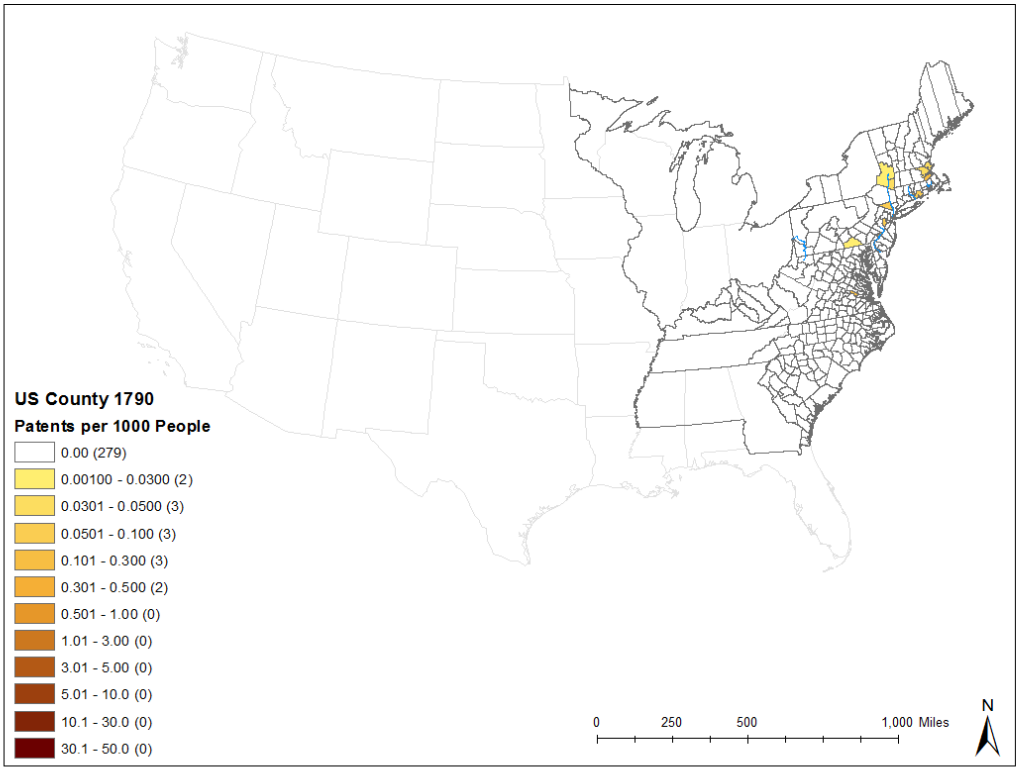
B.1 Figures

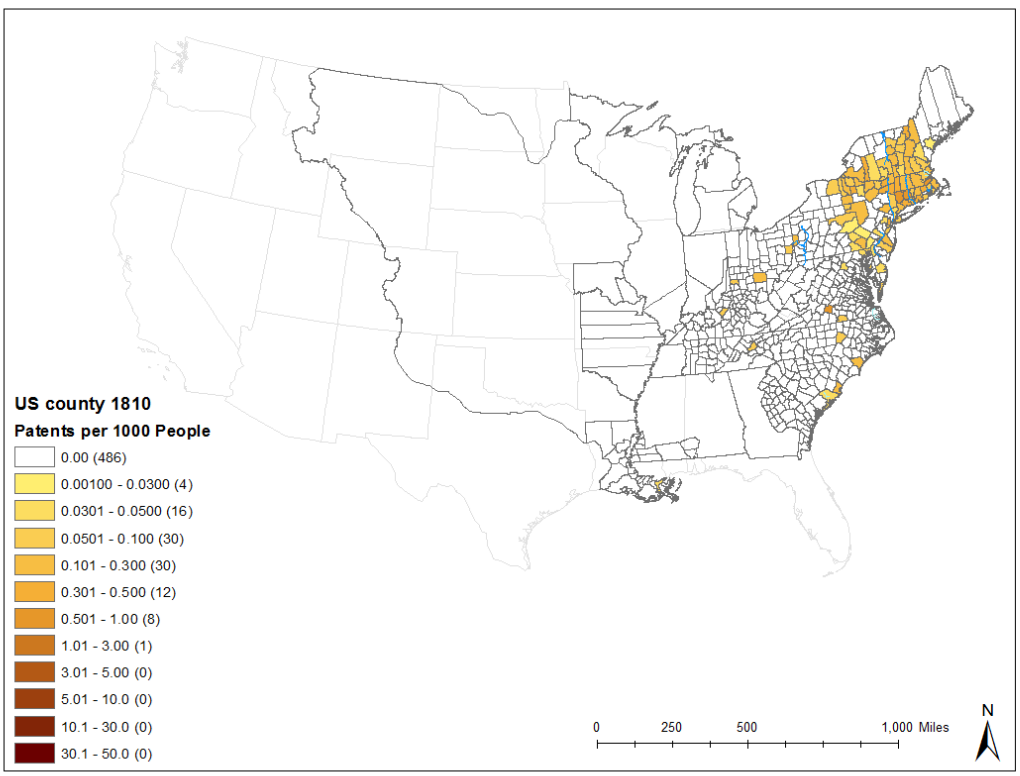
Figure 1: Number of Patents Issued Each Year and Total Population over Time



Back to page 16.

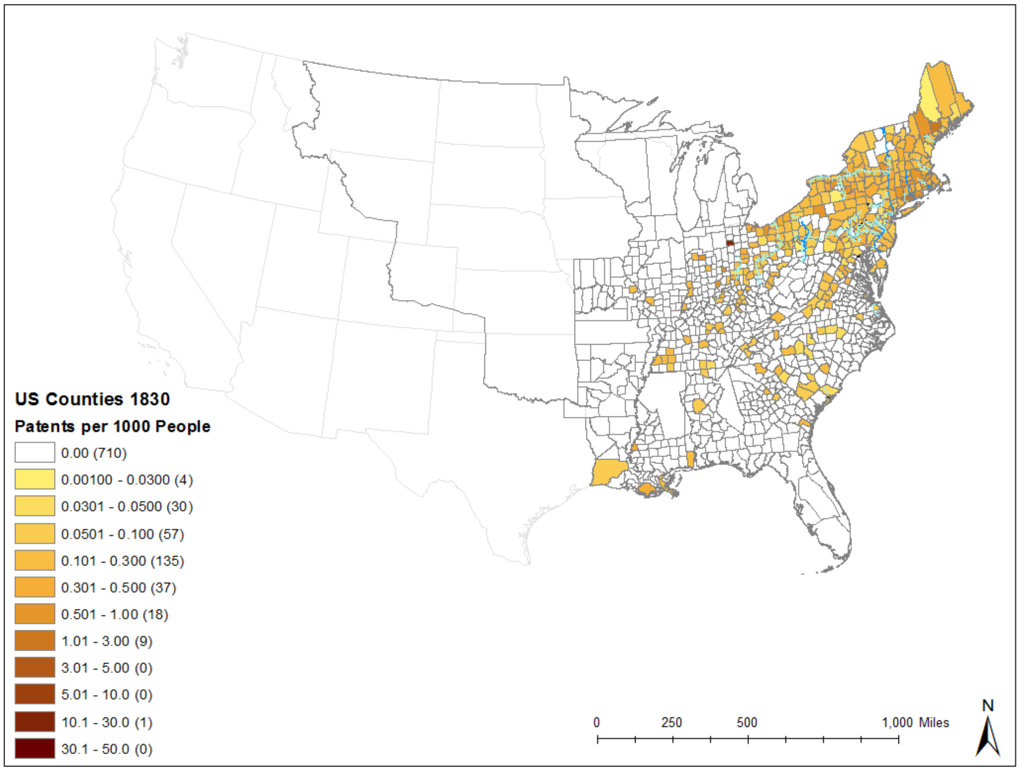
Figure 2: Patent per 1000 People with Transport, 1790 and 1810
Patent counted in five year increments.

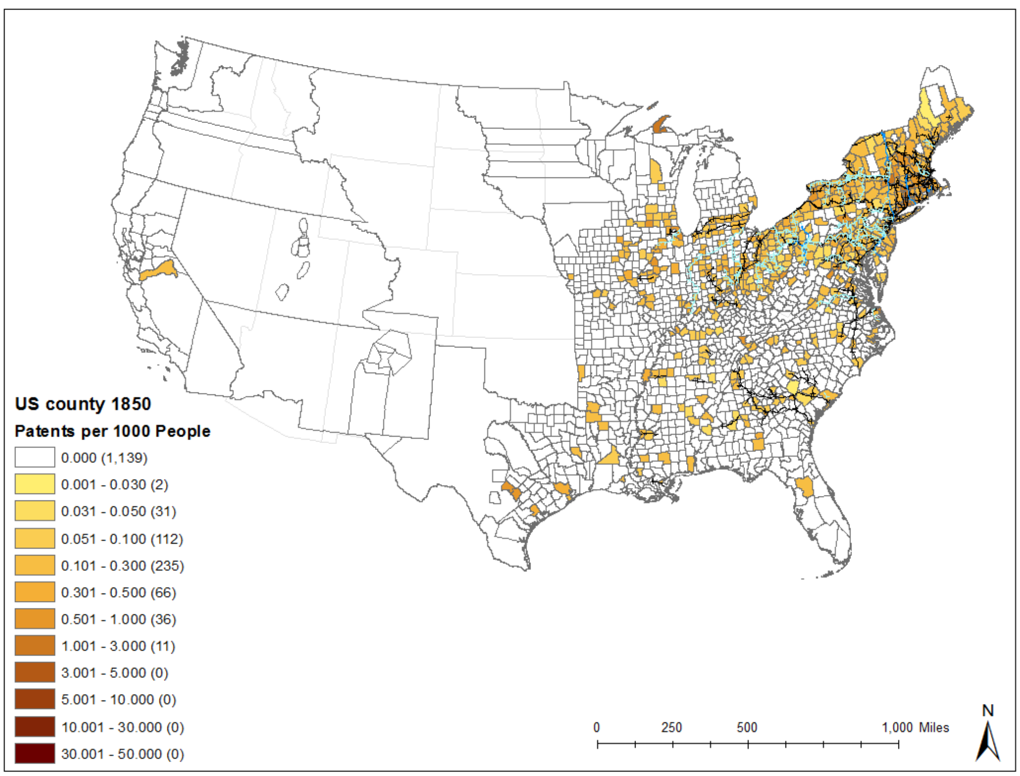




Back to page 16.

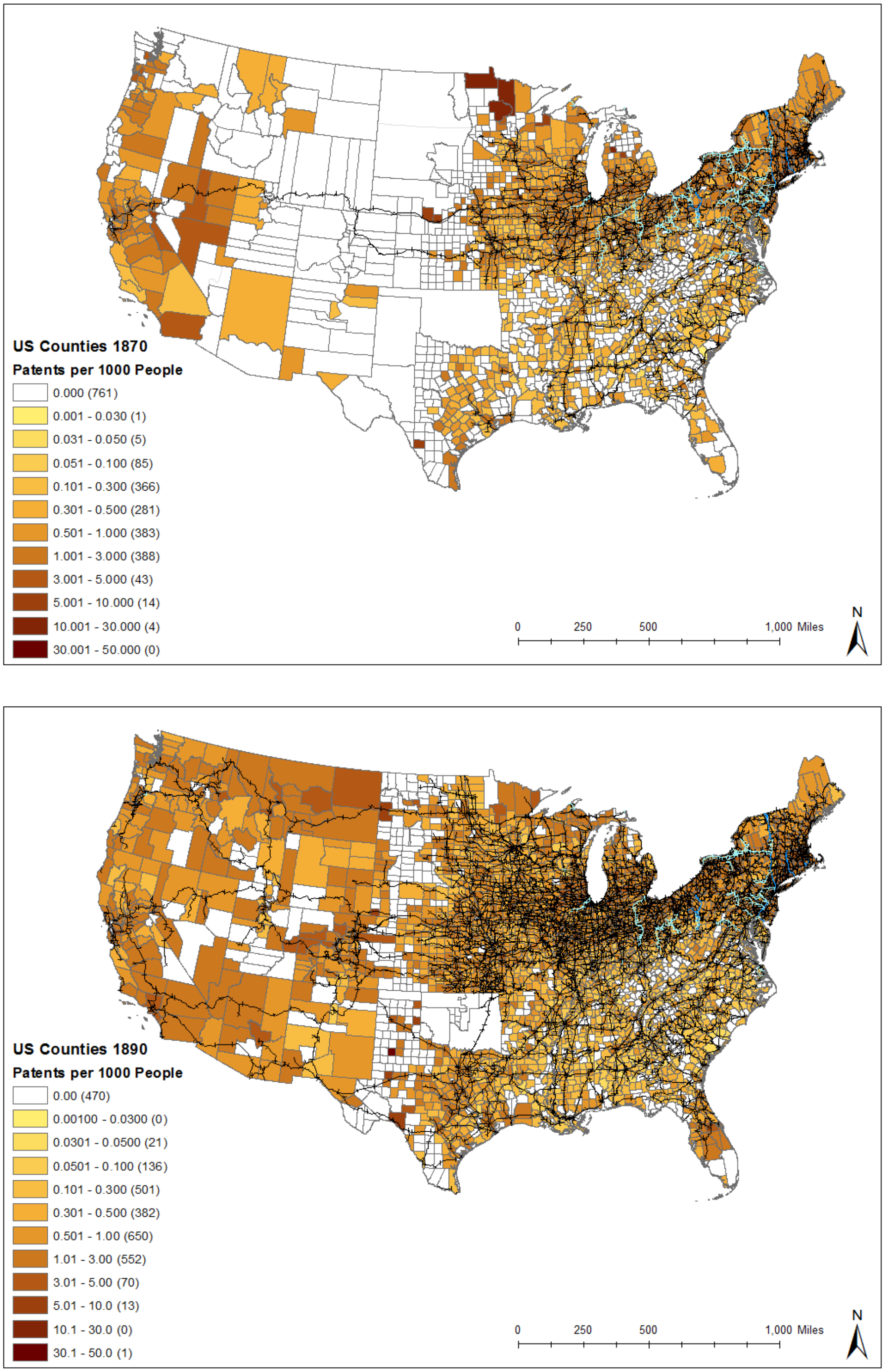
Figure 3: Patent per 1000 People with Transport, 1830 and 1850
Patent counted in five year increments.





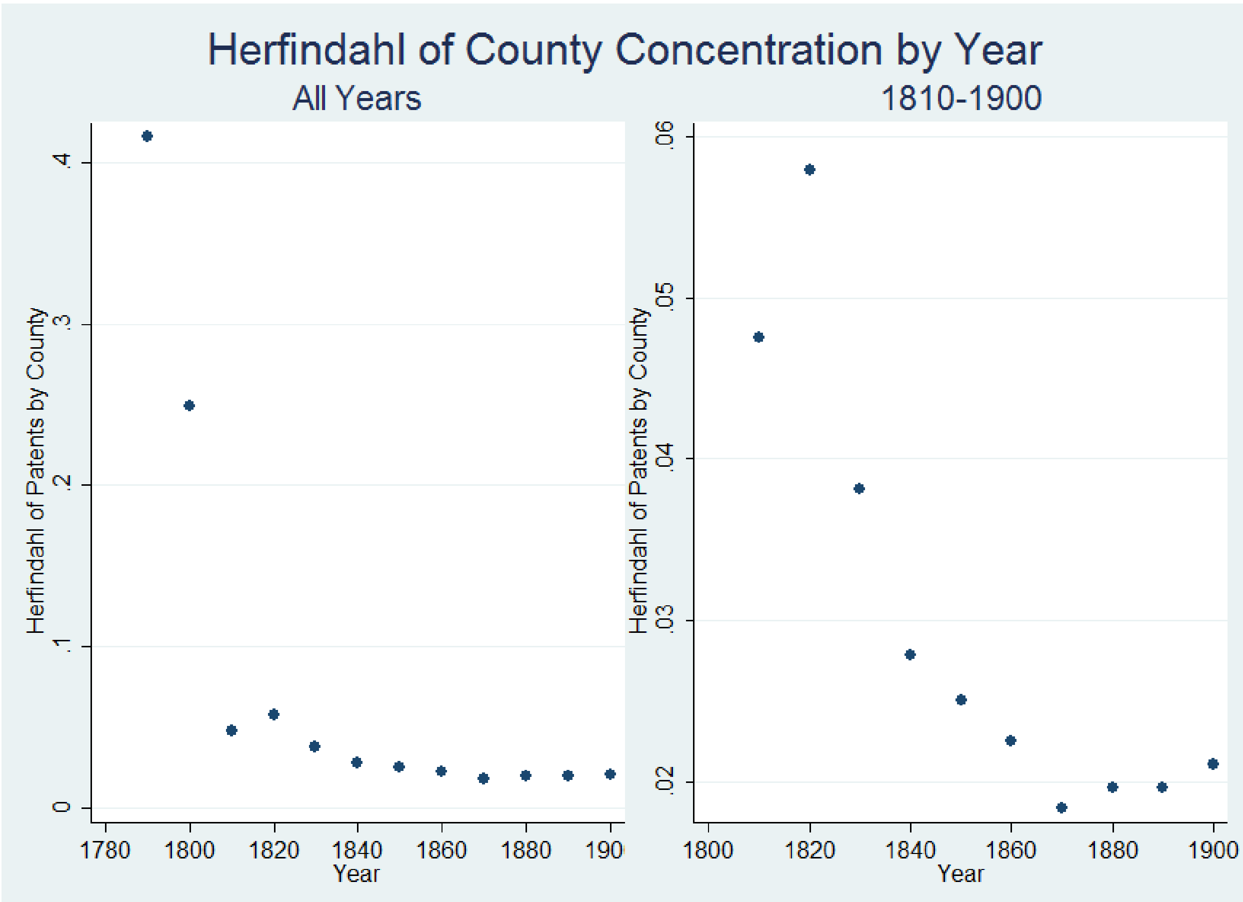
Back to page 16.

Figure 4: Patent per 1000 People with Transport, 1870 and 1890
Patent counted in five year increments.



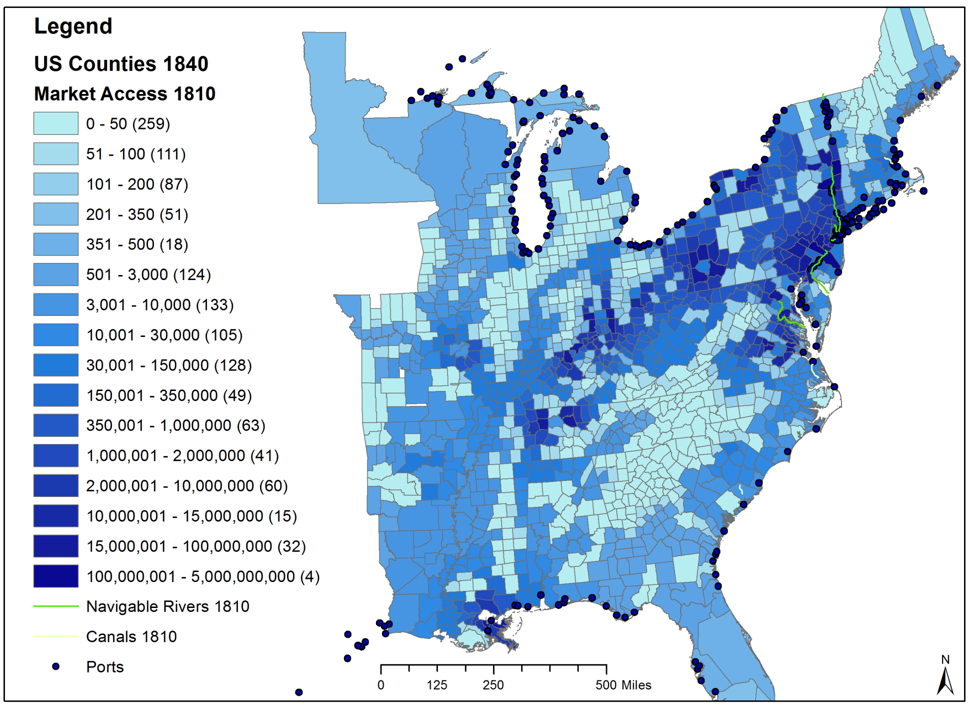
Back to page 16.

Figure 5: Concentration of Patents in Counties by Year, Herfindahl Index



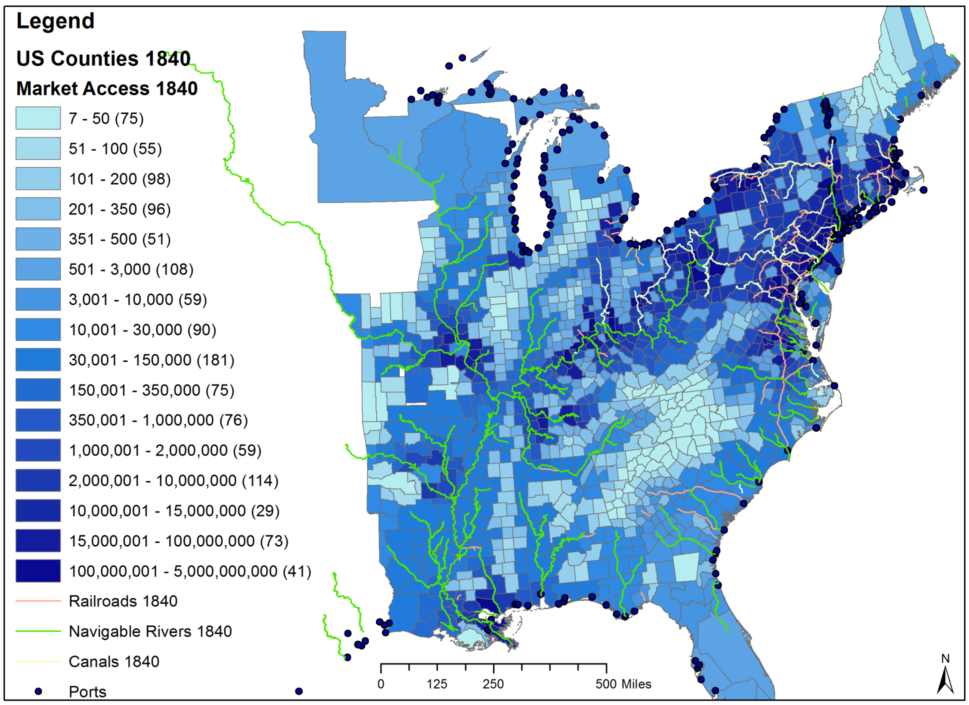
Back to page 16.

Figure 6: Computed Market Access in 1810



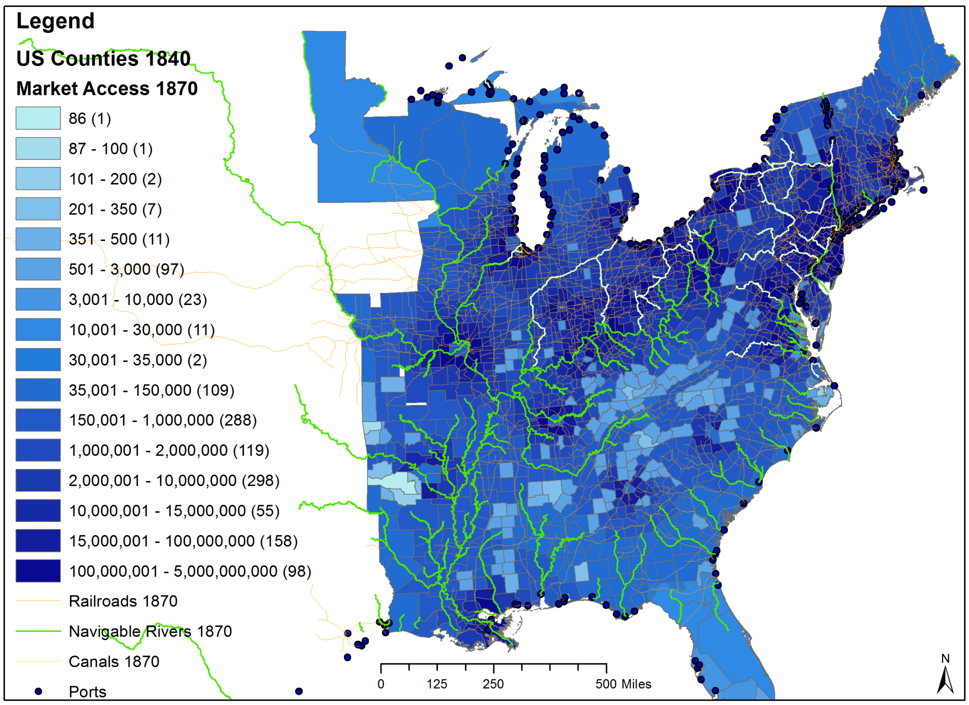
Back to page 19.

Figure 7: Computed Market Access in 1840



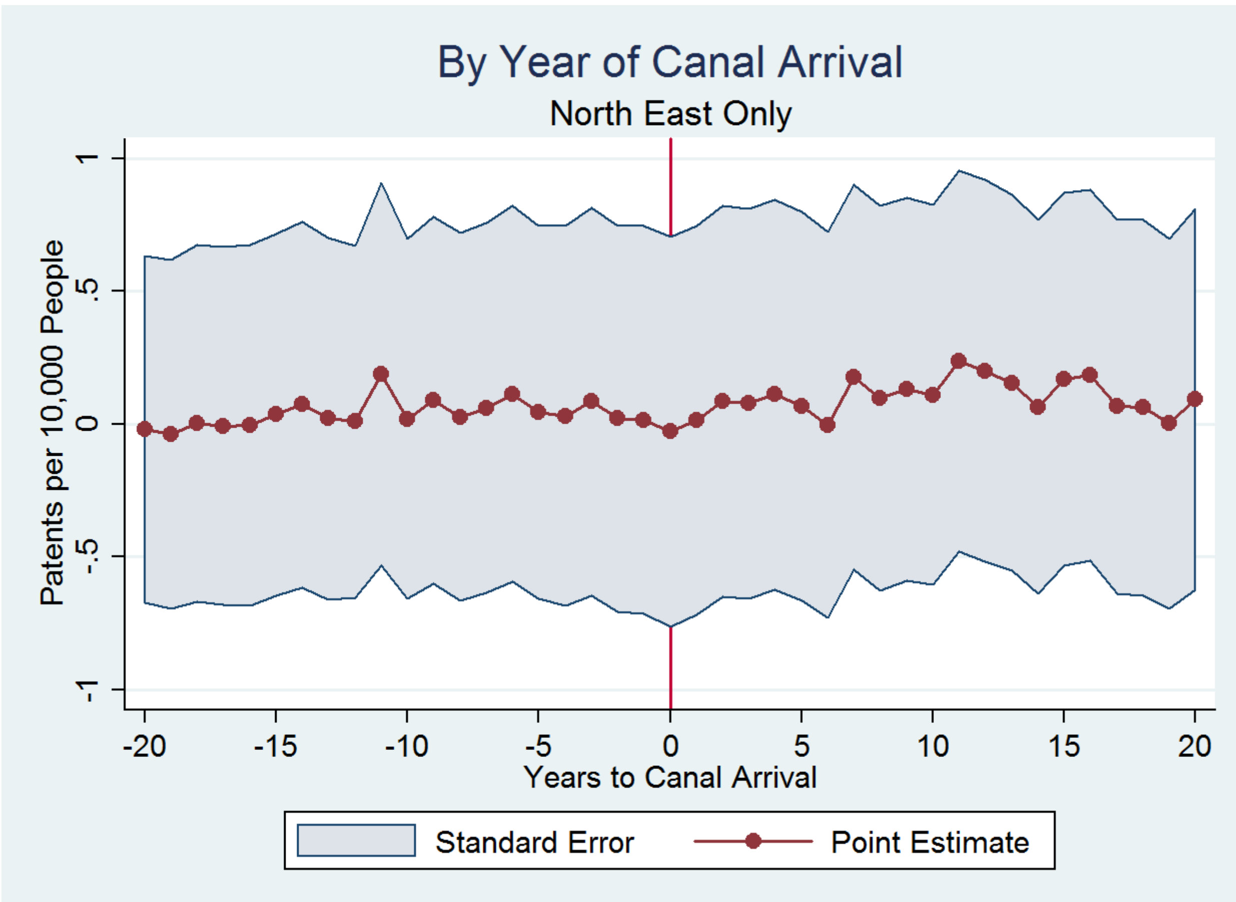
Back to page 19.

Figure 8: Computed Market Access in 1870



Back to page 19.

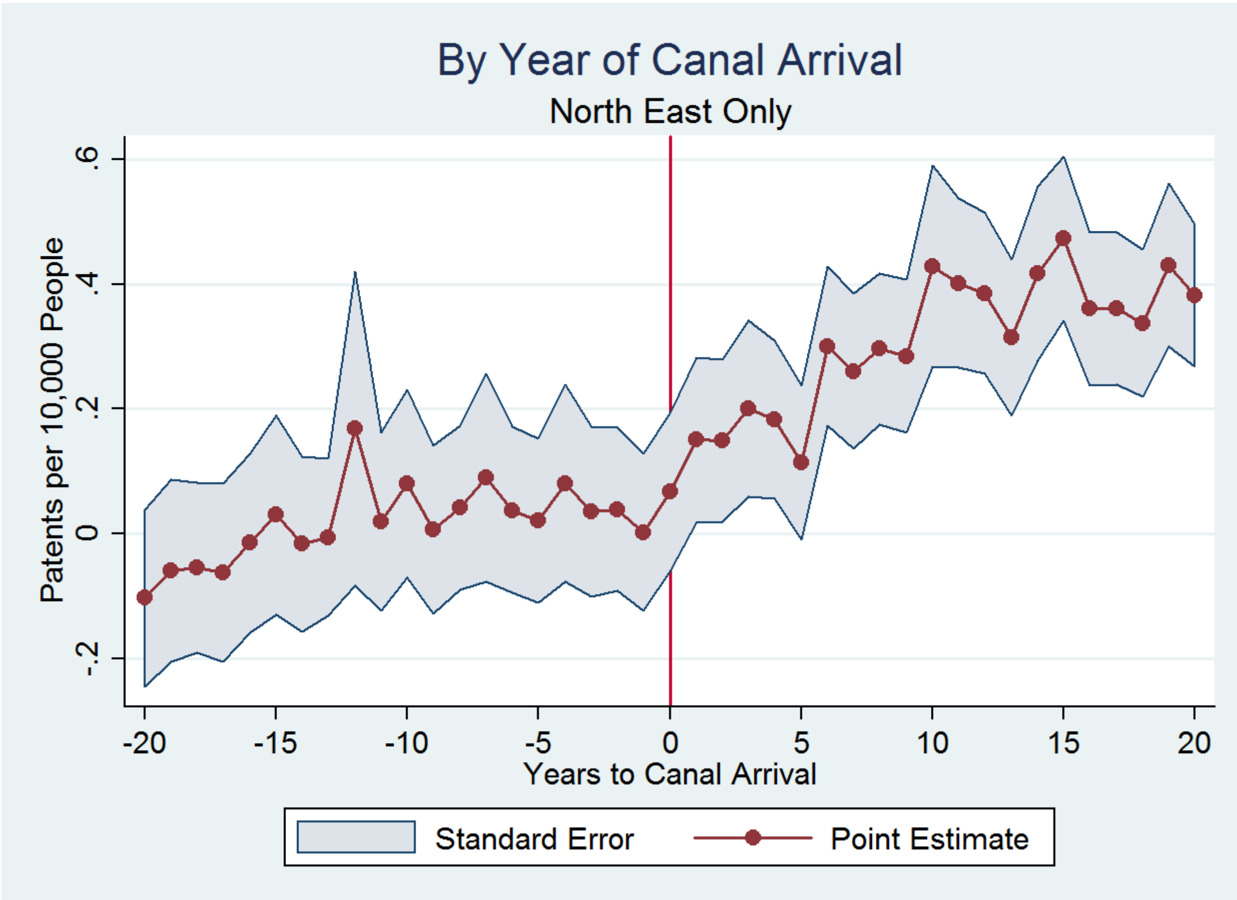
Figure 9: The Mean Patents per 10,000 People by the Years to Canal Arrival



The point estimates and standard errors come from a regression of patents per capita on the dummy variables for the number of years to the arrival of a canal in a county and year and county fixed effects; standard errors are clustered at the county level.

Back to page 21.

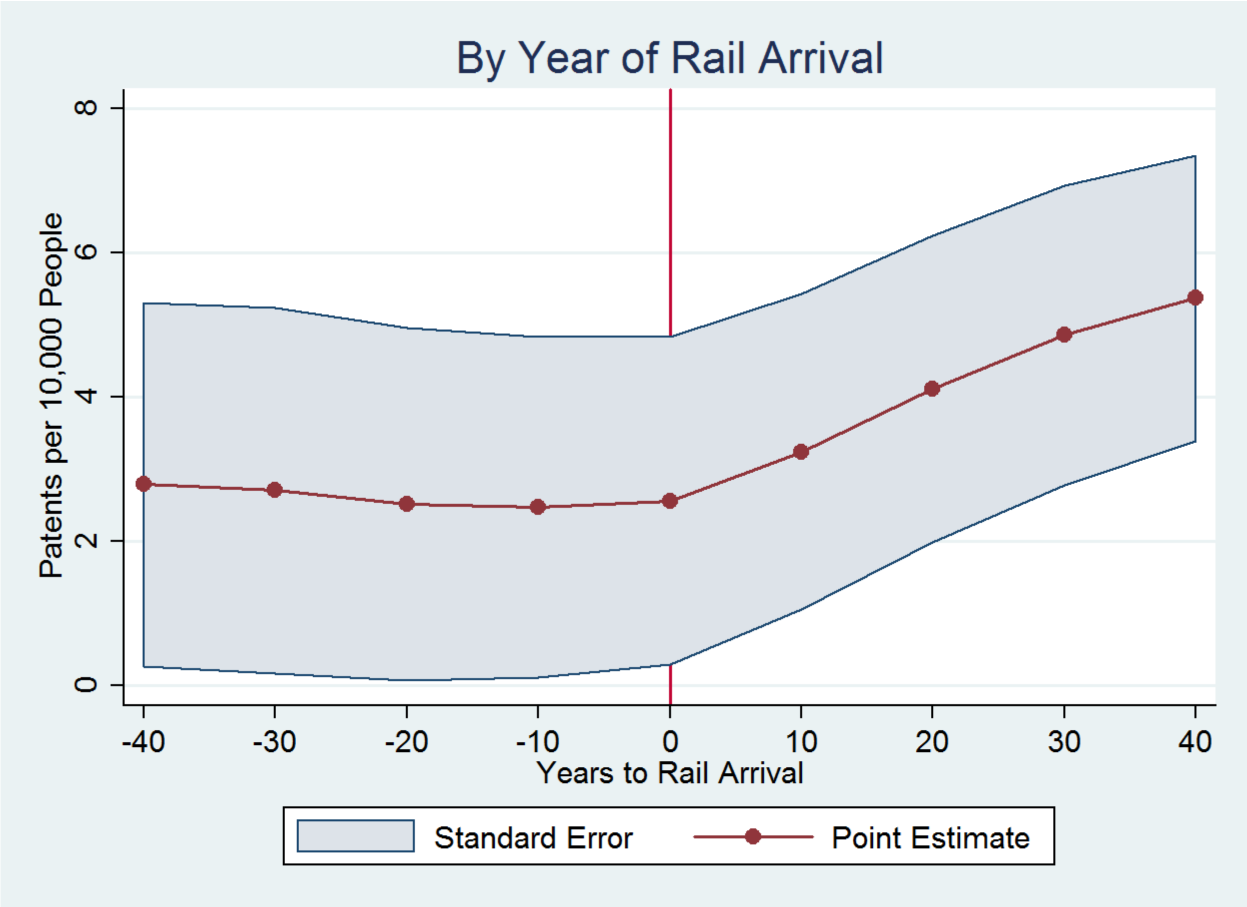
Figure 10: The Mean Patents per 10,000 People by the Years to Canal Arrival, Only Counties that Receive a Canal



The point estimates and standard errors come from a regression of patents per capita on the dummy variables for the number of years to the arrival of a canal in a county and year and county fixed effects; standard errors are clustered at the county level. Sample restricted to counties that receive a canal.

Back to page 21.

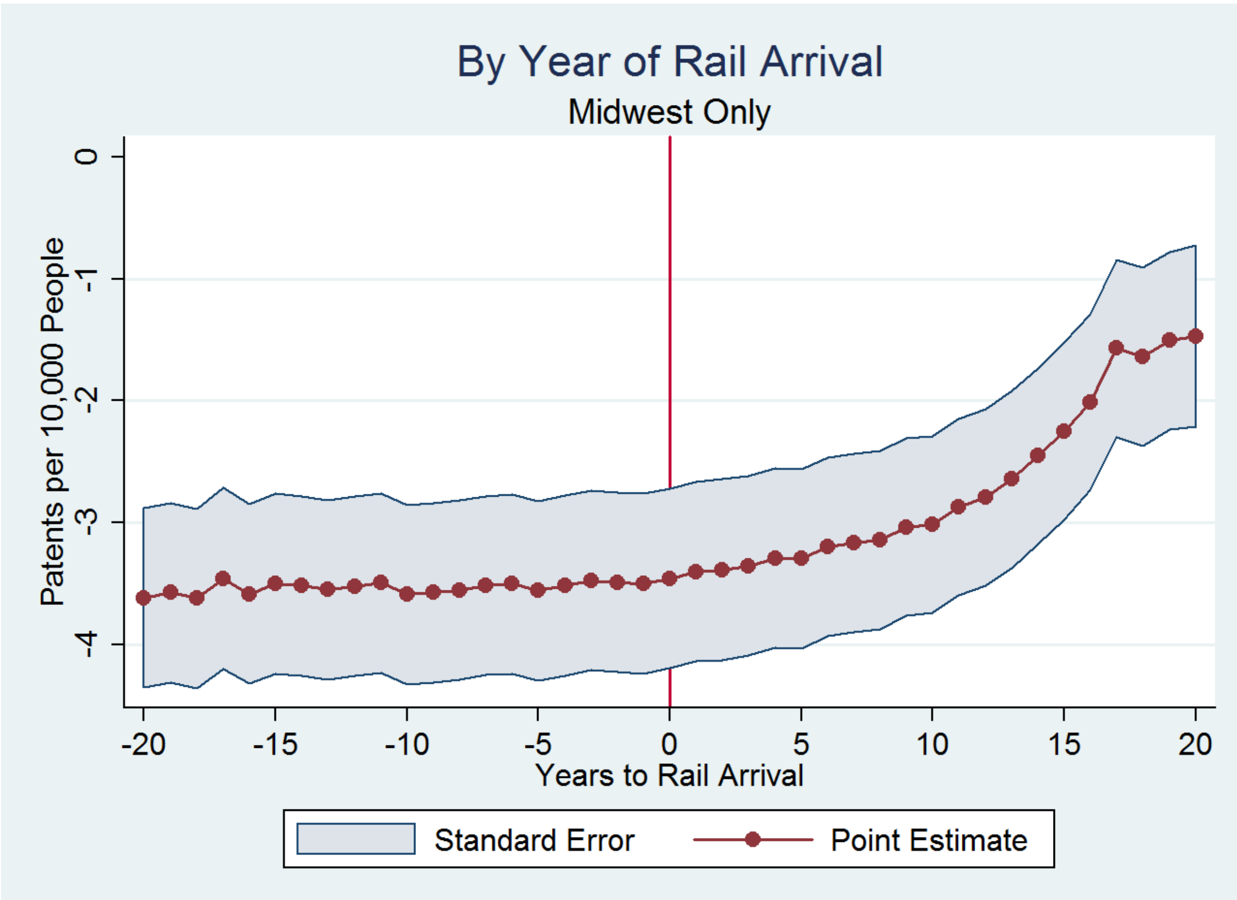
Figure 11: The Mean Patents per 10,000 People by the Years to Railroad Arrival



The point estimates and standard errors come from a regression of patents per capita on the dummy variables for the number of years to the arrival of the railroad in a county and year and county fixed effects; standard errors are clustered at the county level.

Back to page 21.

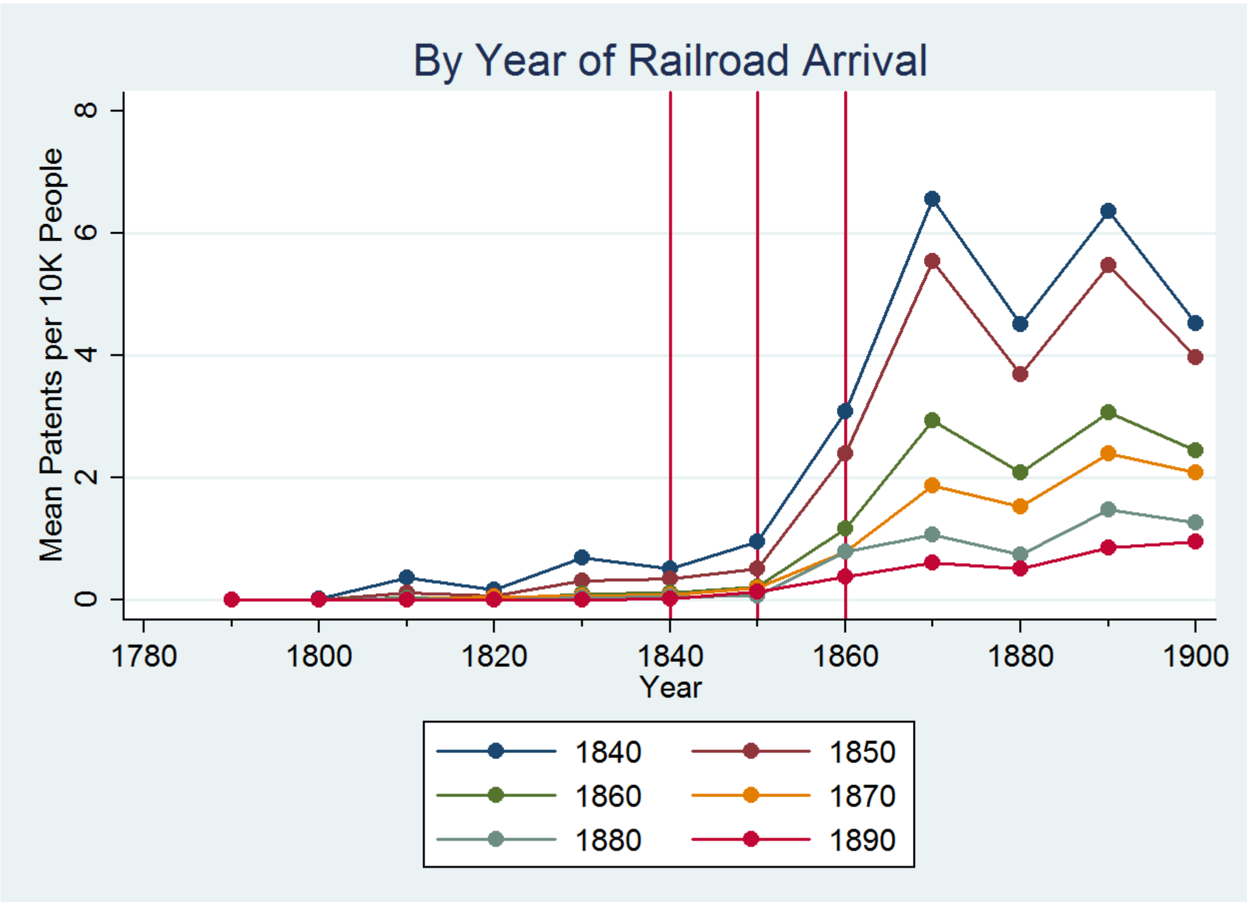
Figure 12: The Mean Patents per 10,000 People by the Years to Railroad Arrival



The point estimates and standard errors come from a regression of patents per capita on the dummy variables for the number of years to the arrival of the railroad in a county and year and county fixed effects; standard errors are clustered at the county level.

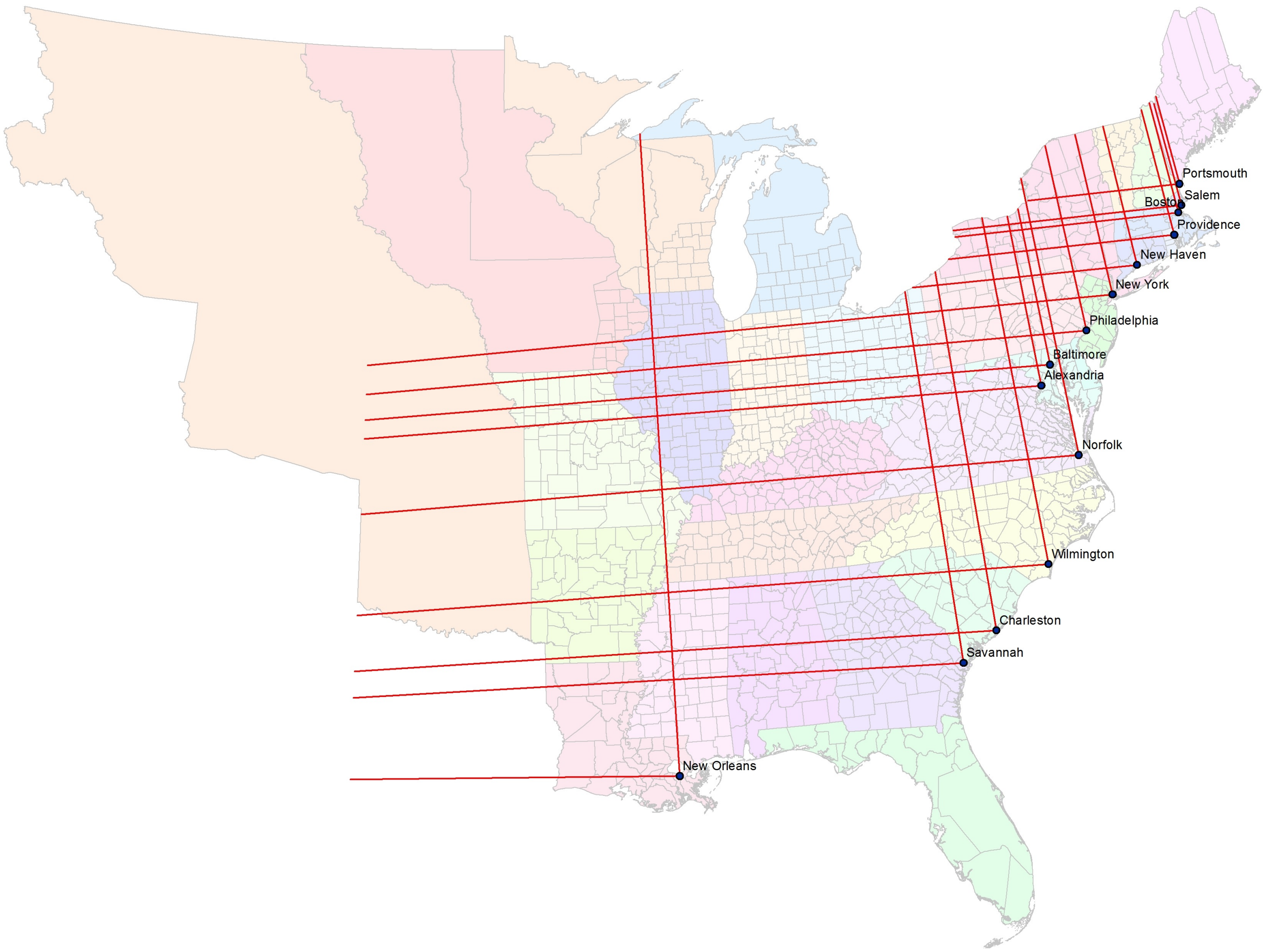
Back to page 21.

Figure 13: The Mean Patents per 10,000 People by the Year of Railroad Arrival



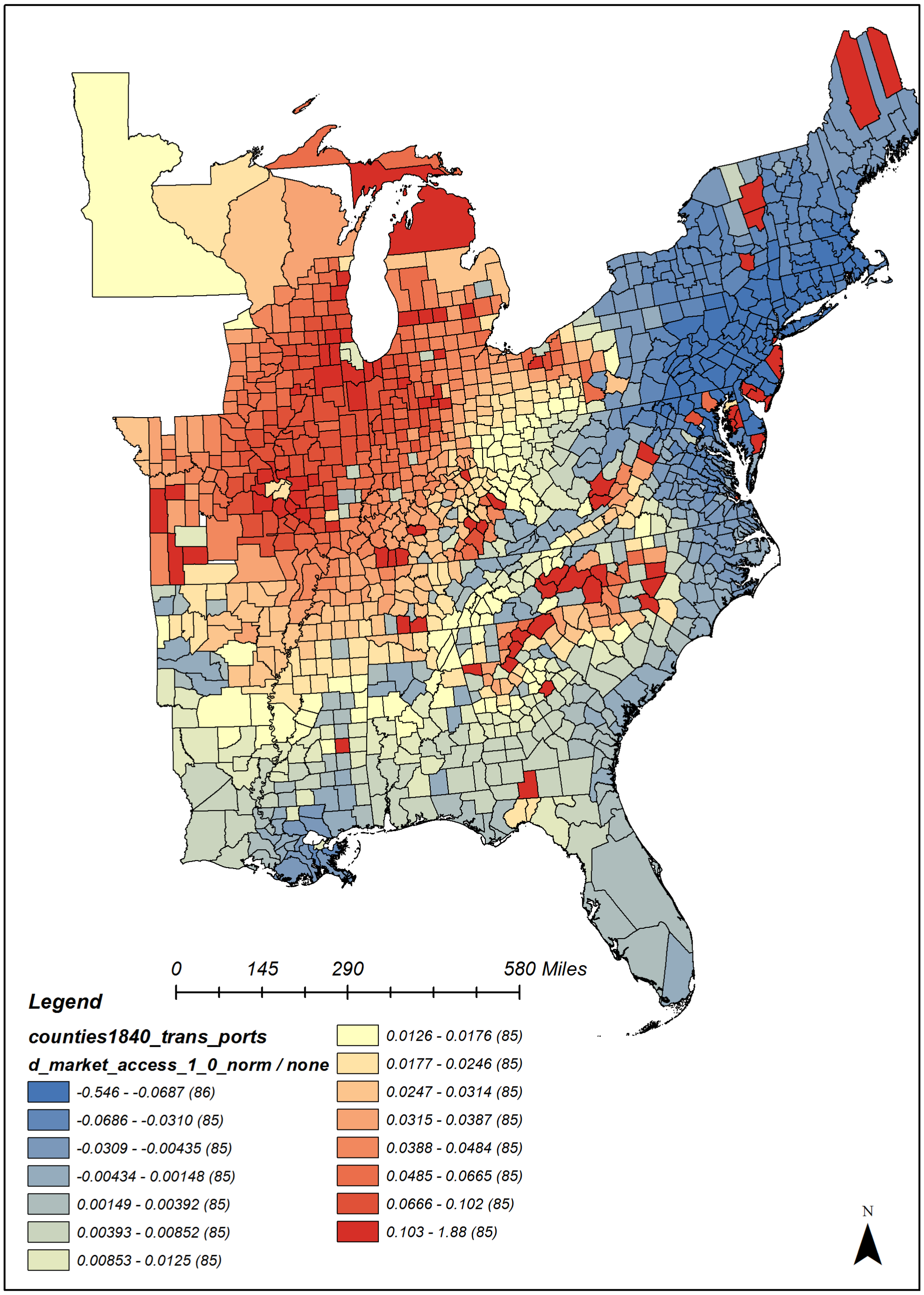
Back to page 23.

Figure 14: Port City Driven Instrument



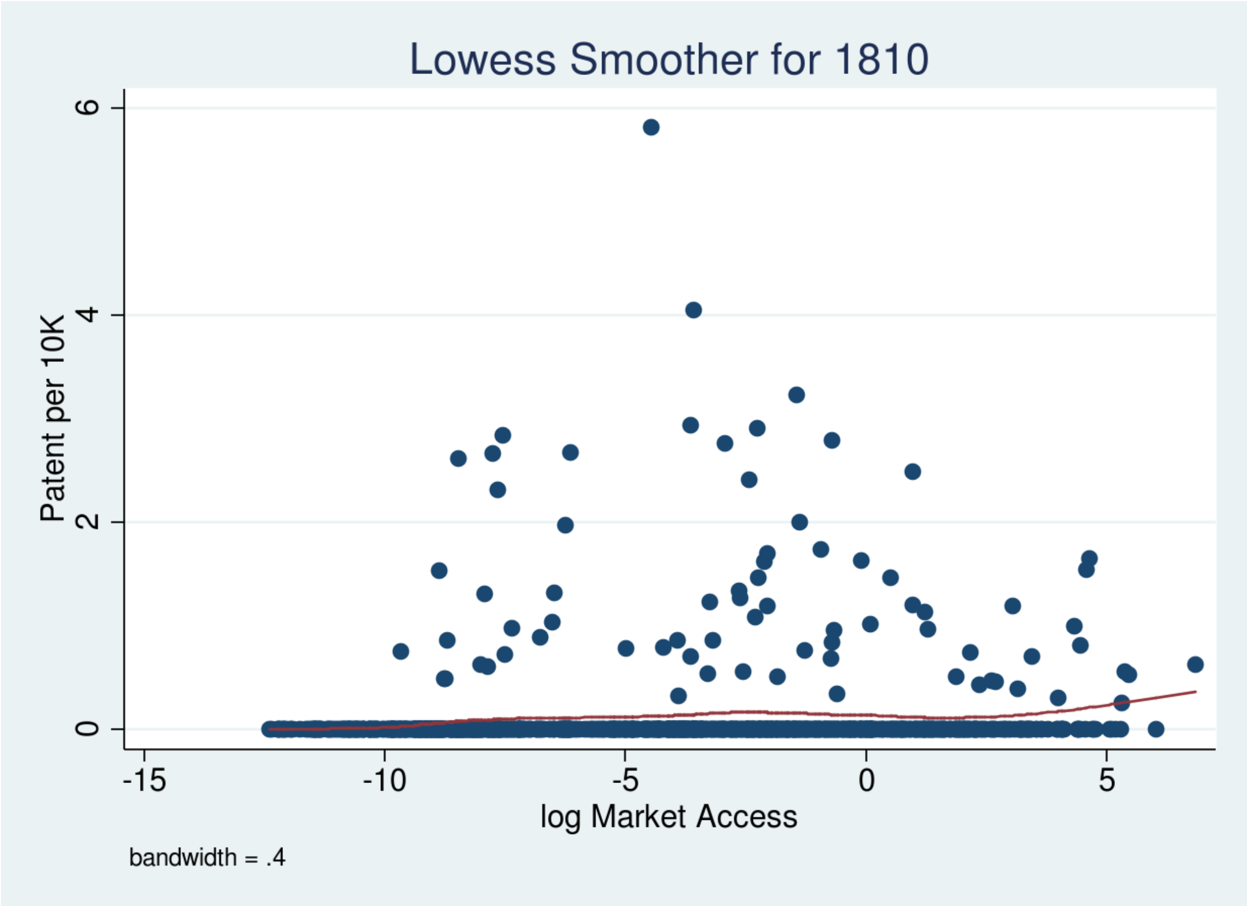
Back to page 26.

Figure 15: Map of First Differences in Market Access in 1870



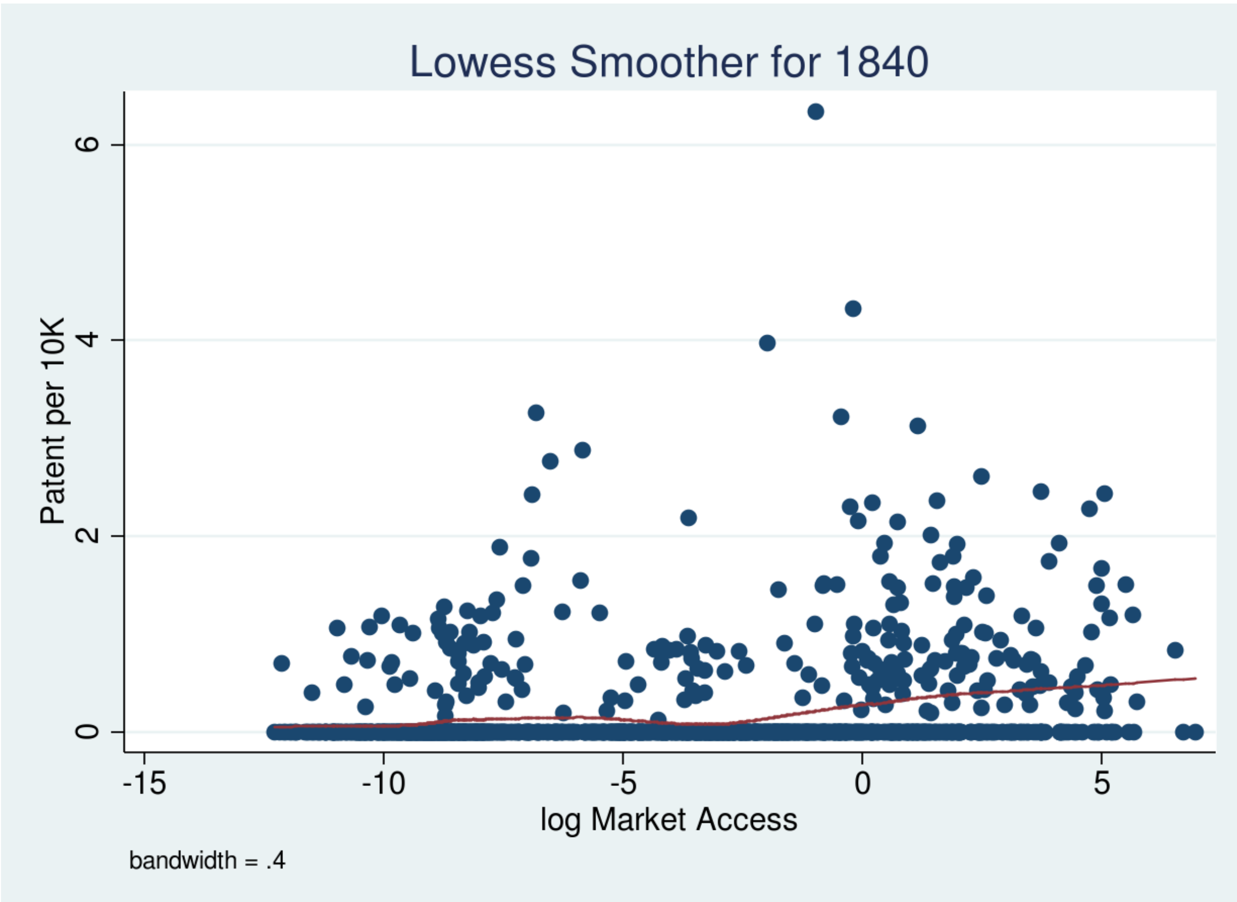
Back to page 30.

Figure 16: Lowess Smoothed Patents per Capita Vs. log Market Access in 1810



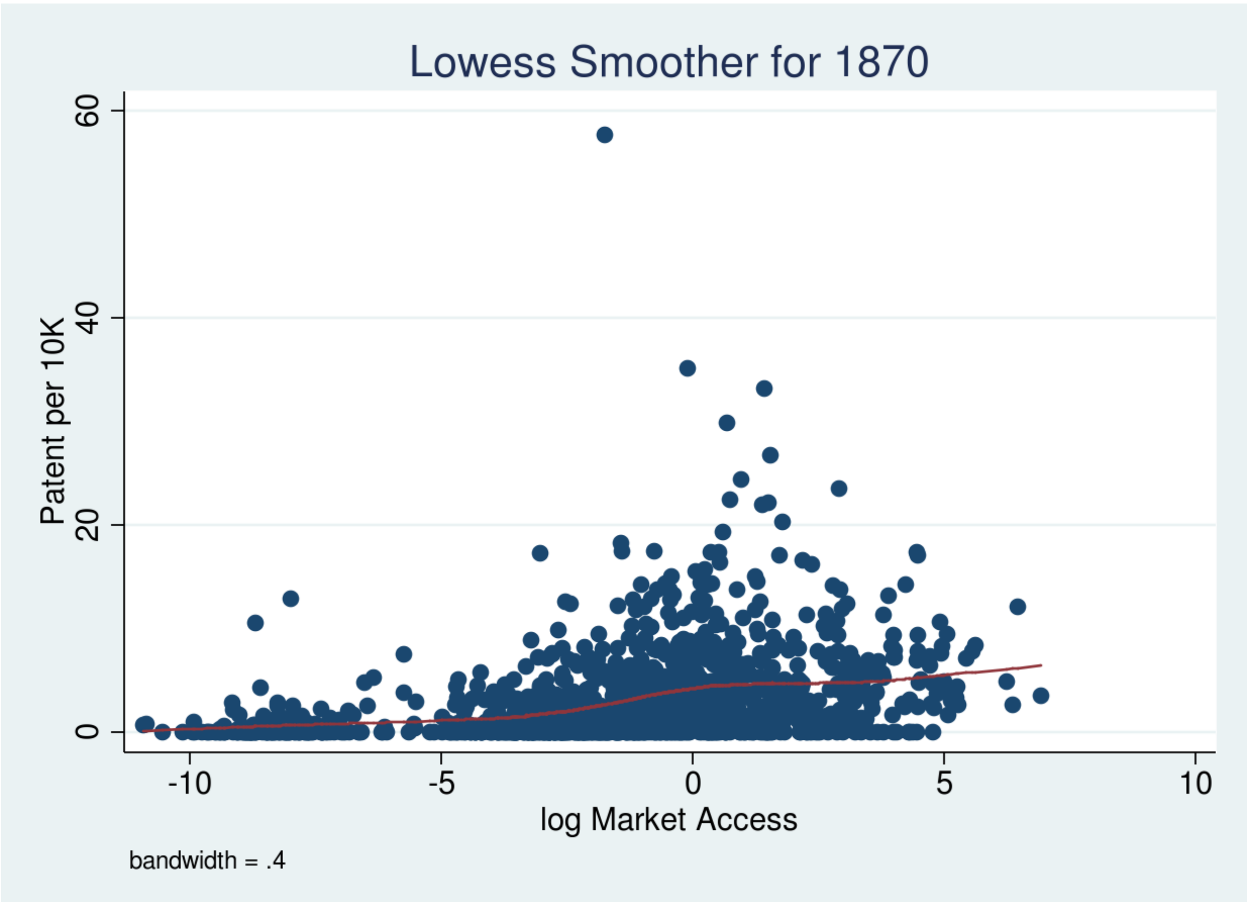
Back to page 28.

Figure 17: Lowess Smoothed Patents per Capita Vs. log Market Access in 1840



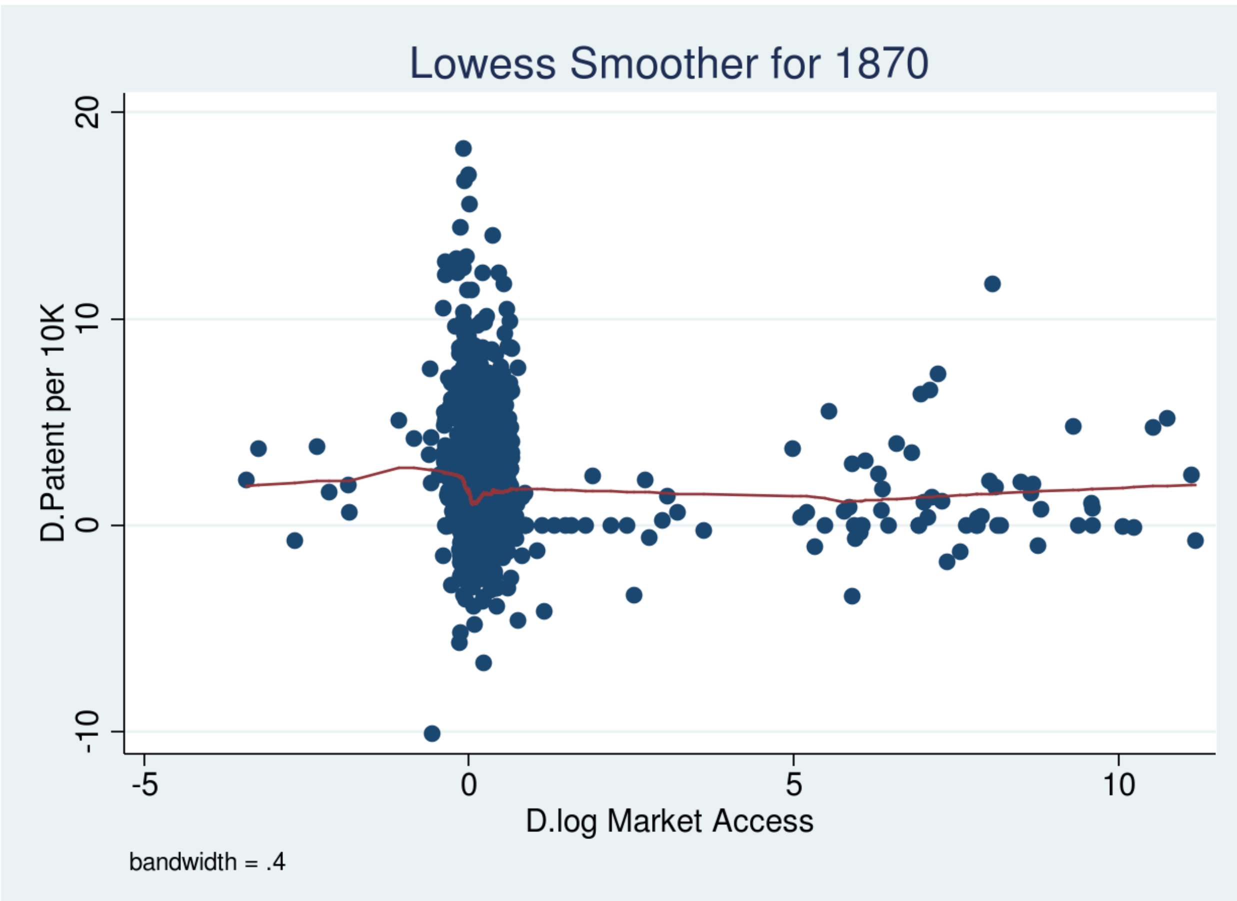
Back to page 28.

Figure 18: Lowess Smoothed Patents per Capita Vs. log Market Access in 1870



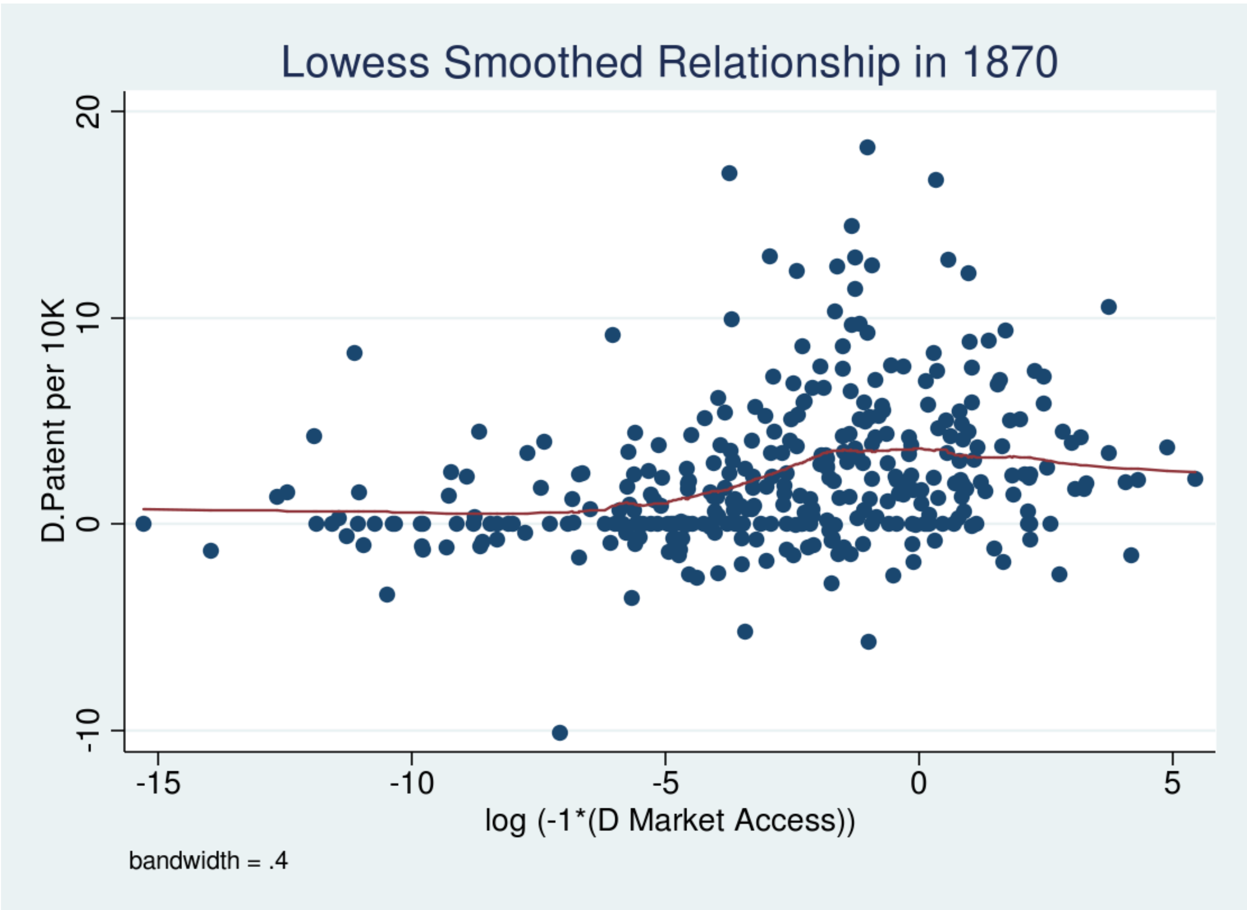
Back to page 28.

Figure 19: Lowess Smoothed First Differences Patents per Capita Vs. First Differences log Market Access in 1870



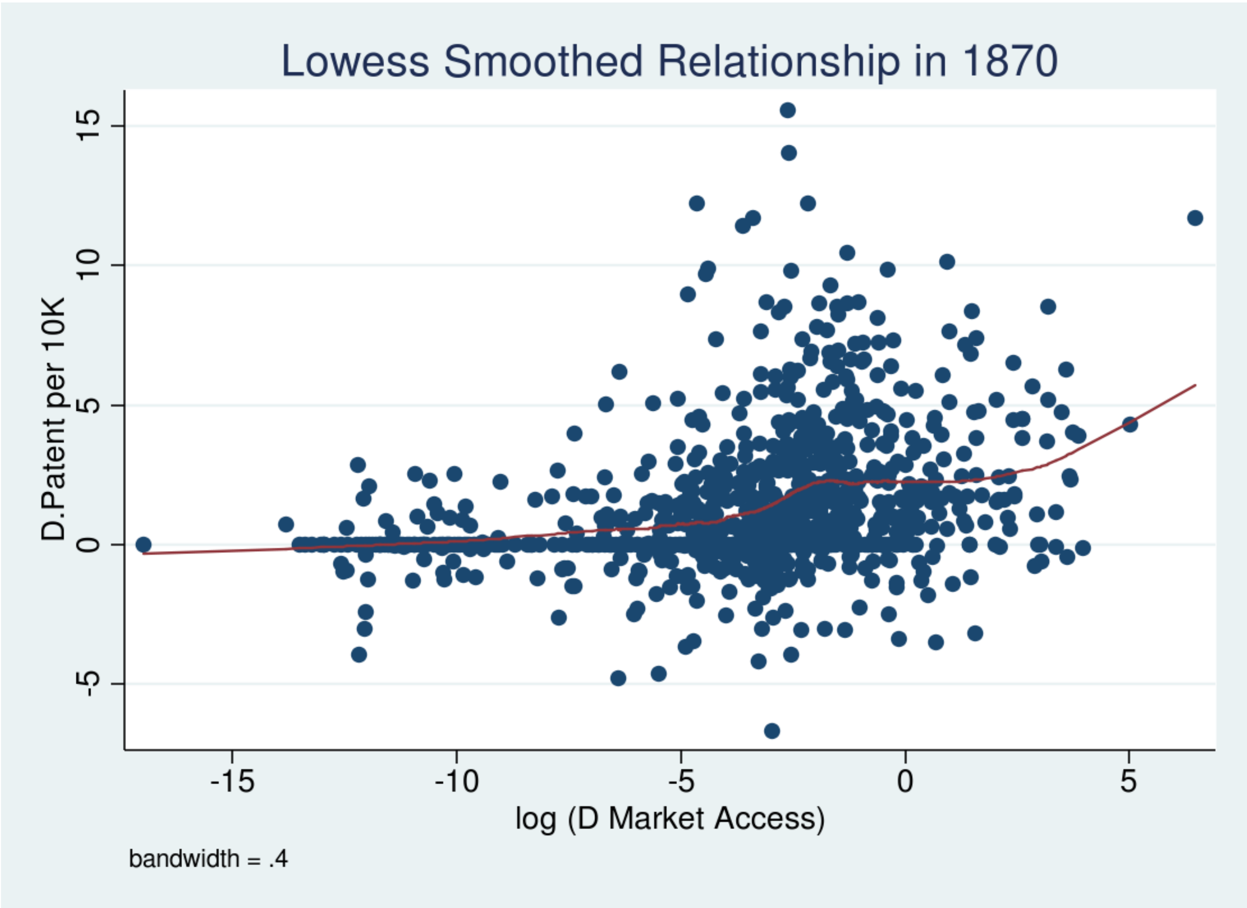
Back to page 30.

Figure 20: Lowess Smoothed First Differences Patents per Capita Vs. log Negative One times First Differences Market Access in 1870



Back to page 31.

Figure 21: Lowess Smoothed First Differences Patents per Capita Vs. log First Differences Market Access in 1870



Back to page 31.

B.2 Tables

Table 1: Means by Year 1790-1840

	1790			1800			1810		
VARIABLES	mean	p50	p95	mean	p50	p95	mean	p50	p95
Number of Patents	0.0188	0	0	0.00313	0	0	0.244	0	1
Total Population	664.5	569.9	1,569	7,684	4,616	25,643	7,274	3,848	26,205
Patent per 10K	2.967	0	0	0.00180	0	0	0.100	0	0.760
log Patent per Mil	8.967	8.263	11.56	3.316	3.475	3.802	4.656	4.595	5.683
% Urban, 2500+	0.0132	0	0.0223	0.0148	0	0.00185	0.0153	0	0
% Urban, 25K+	0.00276	0	0	0.00304	0	0	0.00327	0	0
% within 1.5 miles of transport	0.00306	0	0.00488	0.00362	0	0.00799	0.00524	0	0.0142
% within 5 miles of transport	0.0169	0	0.0681	0.0193	0	0.0999	0.0236	0	0.146
% within 15 miles of transport	0.0552	0	0.492	0.0666	0	0.598	0.0800	0	0.688
log Market Access (Over 3.8)	1.470	0.783	9.365	7.197	6.376	14.87	8.052	7.891	15.69

	1820			1830			1840		
VARIABLES	mean	p50	p95	mean	p50	p95	mean	p50	p95
Number of Patents	0.131	0	0	0.603	0	2	0.641	0	3
Total Population	8,710	4,904	31,017	10,585	6,441	36,354	13,363	8,745	41,080
Patent per 10K	0.0447	0	0.146	0.196	0	1.108	0.174	0	1.101
log Patent per Mil	4.144	4.104	5.610	4.588	4.471	5.965	4.366	4.319	5.493
Number First in Class							0.169	0	1
Number First in Subclass							2.236	1	7
Number NBER Subcategory							2.418	1	6
% Urban, 2500+	0.0149	0	0.0128	0.0194	0	0.117	0.0264	0	0.190
% Urban, 25K+	0.00416	0	0	0.00390	0	0	0.00519	0	0
% within 1.5 miles of transport	0.0254	0	0.142	0.0436	0	0.195	0.0725	0.0233	0.256
% within 1.5 miles of rail				0.000247	0	0	0.0164	0	0.120
% within 5 miles of transport	0.0865	0	0.476	0.141	0	0.616	0.214	0.127	0.688
% within 5 miles of rail				0.00140	0	0	0.0510	0	0.392
% within 15 miles of transport	0.238	0	1.000	0.364	0.121	1	0.501	0.550	1
% within 15 miles of rail				0.00682	0	0	0.141	0	0.976
log Market Access (Over 3.8)	8.614	8.687	16.45	9.102	9.254	16.78	10.15	10.41	17.73
Manufacturing Employment	315.4	61.82	1,535				640.1	176	2,544
% Literate							0.880	0.914	1

Back to page 16.

Table 2: Means by Year 1850-1870

VARIABLES	1850 mean	1850 p50	1850 p95	1860 mean	1860 p50	1860 p95	1870 mean	1870 p50	1870 p95
Number of Patents	1.409	0	5	6.406	1	20	18.10	2	59
Total Population	17,807	12,178	48,638	23,409	16,046	61,435	27,733	18,011	66,225
Patent per 10K	0.315	0	1.930	1.446	0.618	5.283	3.040	1.545	11.40
log Patent per Mil	4.513	4.383	5.805	5.089	5.057	6.518	5.633	5.692	7.161
Number First in Class	0.0453	0	0	0.0147	0	0	0.00789	0	0
Number First in Subclass	2.502	1	10	2.779	1	8	2.911	1	10
Number NBER Subcategory	2.921	2	10	3.965	3	11	6.187	5	16
% Urban, 2500+	0.0428	0	0.306	0.0642	0	0.384	0.0921	0	0.481
% Urban, 25K+	0.0108	0	0	0.0150	0	0	0.0216	0	0
% Acres Improved	0.399	0.373	0.727	0.443	0.433	0.779	0.474	0.471	0.818
% within 1.5 miles of transport	0.0931	0.0630	0.291	0.154	0.139	0.394	0.189	0.174	0.437
% within 1.5 miles of rail	0.0361	0	0.198	0.103	0.0702	0.317	0.141	0.122	0.381
% within 5 miles of transport	0.269	0.214	0.754	0.422	0.438	0.887	0.495	0.524	0.935
% within 5 miles of rail	0.108	0	0.588	0.299	0.255	0.802	0.388	0.399	0.892
% within 15 miles of transport	0.599	0.726	1	0.795	0.985	1	0.855	1.000	1
% within 15 miles of rail	0.267	0	1	0.634	0.812	1	0.734	0.955	1
log Market Access (Over 3.8)	11.33	11.75	18.25	13.25	13.66	18.65	13.97	14.33	18.95
Manufacturing Employment	738.1	104	2,293	984.3	140	3,454	1,533	276	5,100
% Literate	0.868	0.895	0.989				0.887	0.914	0.984
% Born Out of State	0.226	0.178	0.589				0.189	0.156	0.471
% Foreign Born	0.0522	0.0131	0.236	0.0670	0.0234	0.276	0.0661	0.0229	0.264

Back to page 16.

Table 3: Means by Year 1880-1900

	1880			1890			1900		
VARIABLES	mean	p50	p95	mean	p50	p95	mean	p50	p95
Number of Patents	15.91	2	45	29.01	4	92	25.38	4	72
Total Population	34,788	22,958	82,325	41,701	25,363	100,852	49,693	28,281	120,367
Patent per 10K	2.165	1.185	7.652	3.205	1.775	11.42	2.484	1.583	7.954
log Patent per Mil	5.240	5.259	6.797	5.462	5.524	7.143	5.263	5.327	6.770
Number First in Class	0.00319	0	0	0.000951	0	0	0.00188	0	0
Number First in Subclass	1.594	0	5	1.796	0	6	1.269	0	4
Number NBER Subcategory	5.424	4	15	6.386	4	19	6.265	4	19
% Urban, 2500+	0.113	0	0.540	0.152	0.0464	0.639	0.179	0.103	0.683
% Urban, 25K+	0.0309	0	0.247	0.0414	0	0.453	0.0528	0	0.569
% Acres Improved	0.528	0.521	0.854	0.572	0.566	0.875	0.582	0.573	0.870
% within 1.5 miles of transport	0.256	0.238	0.555	0.287	0.269	0.577	0.366	0.361	0.646
% within 1.5 miles of rail	0.216	0.187	0.527	0.251	0.229	0.556	0.336	0.329	0.625
% within 5 miles of transport	0.610	0.651	0.988	0.667	0.721	0.995	0.786	0.845	1
% within 5 miles of rail	0.532	0.572	0.978	0.601	0.655	0.990	0.748	0.805	1.000
% within 15 miles of transport	0.912	1	1	0.946	1	1	0.987	1	1
% within 15 miles of rail	0.845	1.000	1	0.900	1	1	0.974	1	1
log Market Access (Over 3.8)	14.59	14.79	19.27	15.11	15.16	19.59	15.49	15.44	19.87
Manufacturing Employment	2,072	255.3	6,494	3,558	468.6	11,370	3,924	606	13,923
% Literate									
% Born Out of State	0.160	0.133	0.402						
% Foreign Born	0.0596	0.0191	0.245	0.0599	0.0181	0.261	0.0525	0.0153	0.234

Back to page 16.

Table 4: Cross-section by Year

VARIABLES	(1) 1810 log Patents per Mil People	(2) 1840 log Patents per Mil People	(3) 1870 log Patents per Mil People
% within 5 miles of transport	0.0580	0.330	0.718**
	(0.941)	(0.208)	(0.118)
log manufacturing over agriculture	-0.110	-0.169+	0.0333
	(0.0748)	(0.100)	(0.0277)
% Urban, 2500+	0.593	0.532	1.357**
	(0.531)	(0.471)	(0.191)
% Urban, 25K+	1.719*	0.715	0.292
	(0.660)	(0.568)	(0.184)
% within 5 miles of a port	-0.172	0.341	-0.674**
	(1.046)	(0.378)	(0.240)
Elasticity–Full Sample	0.002	0.024	0.071
	(0.030)	(0.015)	(0.012)
State Dummies	Yes	Yes	Yes
Counties	74	222	817
R-squared	0.509	0.267	0.570

Robust standard errors in parentheses, standard errors clustered by county.
** p<0.01, * p<0.05, + p<0.1

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 20.

Table 5: Fixed Effects: Patents per Capita vs. the Presence of a Transport or a Railroad

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Patents per 10K People	Patents per 10K People	Patents per 10K People	Patents per 10K People	Patents per 10K People	Patents per 10K People	Patents per 10K People
Rail Dummy	0.205+						
	(0.109)						
% within 1.5 miles of transport		4.805*					
		(2.448)					
% within 1.5 miles of rail			8.324**				
			(0.444)				
% within 5 miles of transport				2.454**			
				(0.554)			
% within 5 miles of rail					2.652**		
					(0.186)		
% within 15 miles of transport						0.322	
						(0.411)	
% within 15 miles of rail							0.332+
							(0.175)
Elasticity–Full Sample	0.072	0.495	0.699	0.633	0.553	0.141	0.118
	(0.039)	(0.252)	(0.037)	(0.143)	(0.039)	(0.180)	(0.062)
Years	1830-1880	1790-1900	1830-1900	1790-1900	1830-1900	1790-1900	1830-1880
County Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Counties	1250	1250	1250	1250	1250	1250	1250
Observations	7,432	13,256	9,932	13,256	9,932	13,256	7,432
R-squared	0.487	0.137	0.591	0.137	0.565	0.135	0.487

Robust standard errors in parentheses, standard errors clustered by county.
** $p<0.01$, * $p<0.05$, + $p<0.1$

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 23.

Table 6: Fixed Effects: log Patents per Capita vs. the Presence of a Transport or a Railroad

	(1)	(2)	(3)	(4)	(5)	(6)
	log Patents per Mil People	log Patents per Mil People Zeros	log Patents per Mil People	log Patents per Mil People Zeros	log Patents per Mil People	log Patents per Mil People Zeros
VARIABLES						
Rail Dummy	-0.239**	0.969**				
	(0.0505)	(0.0798)				
% within 1.5 miles of transport			1.452**	4.949**		
			(0.182)	(0.383)		
% within 5 miles of transport					0.359**	2.170**
					(0.0920)	(0.323)
Elasticity–Full Sample	-0.038	0.196	0.073	0.326	0.042	0.358
	(0.008)	(0.016)	(0.009)	(0.025)	(0.011)	(0.053)
Years	1830-1900	1830-1900	1790-1900	1790-1900	1790-1900	1790-1900
County Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Counties	1203	1251	1203	1251	1203	1251
Observations	5,320	10,008	5,463	15,012	5,463	15,012
R-squared	0.688	0.665	0.693	0.688	0.685	0.684

Robust standard errors in parentheses, standard errors clustered by county.
** $p<0.01$, * $p<0.05$, + $p<0.1$

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 24.

Table 7: Fixed Effects: Patents per Capita vs. the Presence of a Transport or a Railroad with Controls and Pretrends

VARIABLES	(1) Patents per 10K People	(2) log Patents per Mil People	(3) Patents per 10K People	(4) log Patents per Mil People
% within 1.5 miles of transport	0.667*	0.208		
	(0.264)	(0.202)		
% within 5 miles of transport			0.204*	0.162
			(0.0866)	(0.111)
Elasticity–Full Sample	0.028	0.01	0.023	0.054
	(0.212)	(2.761)	(0.01)	(209.468)
Years	1790-1900	1790-1900	1790-1900	1790-1900
Pretrends	Yes	Yes	Yes	Yes
County Controls	Yes	Yes	Yes	Yes
County Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Counties	1055	1013	1055	1013
Observations	11,645	4,728	11,645	4,728
R-squared	0.989	0.775	0.989	0.776

Robust standard errors in parentheses, standard errors clustered by county.
** p<0.01, * p<0.05, + p<0.1

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 25.

Table 8: IV 1.5 Miles

VARIABLES	(1) OLS Patents per 10K People	(2) First Stage Percent RR 1.5 Miles	(3) IV Patents per 10K People	(4) OLS Patents per 10K People	(5) First Stage Percent RR 1.5 Miles	(6) IV Patents per 10K People
Port IV		0.0514**			0.0128**	
		(0.00812)			(0.00358)	
% within 1.5 miles of rail	8.324**		15.67**	0.848*		8.439
	(0.444)		(2.691)	(0.363)		(5.738)
Years	1830-1900	1830-1900	1830-1900	1830-1900	1830-1900	1830-1900
Pretrends	No	No	No	Yes	Yes	Yes
County Controls	No	No	No	Yes	Yes	Yes
County Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Counties	1250	1250	1250	1055	1055	1055
Observations	9,932	9,932	9,932	8,432	8,432	8,432
R-squared	0.591	0.774	0.244	0.847	0.908	0.685
T - squared		40.01			12.78	
Wald Stat.			45.93			14.80
Non-Robust F			173.2			4.890e+08

Robust standard errors in parentheses, standard errors clustered by county.
** $p<0.01$, * $p<0.05$, + $p<0.1$

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 26.

Table 9: IV 5.0 Miles

VARIABLES	(1) OLS Patents per 10K People	(2) First Stage Percent RR 5.0 Miles	(3) IV Patents per 10K People	(4) OLS Patents per 10K People	(5) First Stage Percent RR 5.0 Miles	(6) IV Patents per 10K People
IV Port		0.0726**			0.0171*	
		(0.0140)			(0.00830)	
% within 5 miles of rail	2.652**		11.11**	0.266*		7.216
	(0.186)		(2.364)	(0.130)		(5.100)
Years	1830-1900	1830-1900	1830-1900	1830-1900	1830-1900	1830-1900
Pretrends	No	No	No	Yes	Yes	Yes
County Controls	No	No	No	Yes	Yes	Yes
County Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Counties	1250	1250	1250	1055	1055	1055
Observations	9,932	9,932	9,932	8,432	8,432	8,432
R-squared	0.565	0.790	-0.043	0.846	0.892	0.587
T - squared		26.90			4.24	
Wald Stat.			30.74			4.870
Non-Robust F			127.2			293570

Robust standard errors in parentheses, standard errors clustered by county.
** p<0.01, * p<0.05, + p<0.1

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 26.

Table 10: Cross-section by Year

VARIABLES	(1) 1810 log Patents per Mil People	(2) 1840 log Patents per Mil People	(3) 1870 log Patents per Mil People
log Market Access	-0.0432	-0.0101	0.0397**
	(0.0291)	(0.0123)	(0.0101)
log manufacturing over agriculture	-0.0567	-0.0502	0.0371+
	(0.0790)	(0.0667)	(0.0198)
% Urban, 2500+	0.542	0.577	1.775**
	(0.471)	(0.413)	(0.182)
% Urban, 25K+	1.442*	0.437	0.157
	(0.607)	(0.547)	(0.187)
% within 5 miles of a port	-0.130	0.402	-0.575*
	(0.487)	(0.358)	(0.231)
State Dummies	Yes	Yes	Yes
Counties	75	223	860
Observations	75	223	860
R-squared	0.554	0.254	0.553

Robust standard errors in parentheses, standard errors clustered by county.
** p<0.01, * p<0.05, + p<0.1

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 28.

Table 11: Fixed Effects: Patents per Capita vs. the Market Access

VARIABLES	(1) Patents per 10K People	(2) Patents per 10K People No Zeros	(3) log Patents per Mil People Zeros	(4) log Patents per Mil People
log Market Access	0.0427**	0.105	0.103**	-0.0236*
	(0.0156)	(0.193)	(0.00931)	(0.00932)
Years	1790-1900	1790-1900	1790-1900	1790-1900
County Dummies	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Counties	1250	1203	1250	1203
Observations	13,256	5,463	14,995	5,448
R-squared	0.135	0.375	0.671	0.676

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1
Standard errors clustered by county.

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 29.

Table 12: Fixed Effects: Patents per Capita vs. Market Access with Controls and Pretrends

VARIABLES	(1) Patents per 10K People	(2) Patents per 10K People	(3) log Patents per Mil People	(4) log Patents per Mil People	(5) log Patents per Mil People Zeros	(6) log Patents per Mil People Zeros
log Market Access	0.0128+	-0.00757	-0.000129	-0.0163	0.0300**	0.00876
	(0.00664)	(0.00879)	(0.00866)	(0.0116)	(0.00756)	(0.0119)
Years	1790-1900	1790-1900	1790-1900	1790-1900	1790-1900	1790-1900
Pretrends–Patents	Yes	Yes	Yes	Yes	Yes	Yes
Pretrends–Transport	No	Yes	No	Yes	No	Yes
County Controls	Yes	Yes	Yes	Yes	Yes	Yes
County Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Counties	1055	1055	1013	1013	1055	1055
Observations	11,645	11,645	4,728	4,728	12,656	12,656
R-squared	0.989	0.989	0.777	0.784	0.754	0.758

Robust standard errors in parentheses
** p<0.01, * p<0.05, + p<0.1
Standard errors clustered by county.

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 30.

Table 13: First Differences Patents per Capita Vs. First Differences Market Access, split Negative and Positive in 1870

VARIABLES	(1) 1870 First Differences Patents per 10K People	(2) 1870 First Differences Patents per 10K People
D Market Access–Neg Only	-0.0152+	0.00853
	(0.00892)	(0.0104)
D Market Access–Pos Only	0.0173**	0.0140**
	(0.00163)	(0.00108)
log Total Pop		2.497**
		(0.627)
Neg Elasticity–At Means	0.008	-0.005
	(0.005)	(0.006)
Pos Elasticity–At Means	0.015	0.012
	(0.002)	(0.001)
Counties	1273	1273
Observations	1,273	1,273
R-squared	0.006	0.181

Robust standard errors in parentheses, standard errors clustered by county.
** p<0.01, * p<0.05, + p<0.1

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 31.

Table 14: First Differences Patents per Capita Vs. log First Differences Market Access and log Negative One times First Differences Market Access

	(1)	(2)	(3)	(4)	(5)	(6)
	1810	1810	1840	1840	1870	1870
VARIABLES	First Differences Patents per 10K People	First Differences Patents per 10K People	First Differences Patents per 10K People	First Differences Patents per 10K People	First Differences Patents per 10K People	First Differences Patents per 10K People
log (D Market Access)		0.00184		-0.00424		0.158**
		(0.00297)		(0.00733)		(0.0198)
log (-1*(D Market Access))	-0.00981		-0.0223**		0.00979	
	(0.00710)		(0.00767)		(0.0693)	
log Total Pop	0.220**	0.106**	-0.0740	-0.0390+	3.585**	1.287**
	(0.0636)	(0.0195)	(0.0615)	(0.0221)	(1.362)	(0.147)
Counties	344	346	162	1052	347	926
Observations	344	346	162	1,052	347	926
R-squared	0.114	0.106	0.061	0.003	0.224	0.185

Robust standard errors in parentheses, standard errors clustered by county.
** $p<0.01$, * $p<0.05$, + $p<0.1$

Sources: Patent data as described in the text, U.S. Census Data is from Haines (2010) (county boundaries harmonized to 1840 as in Hornbeck (2010)), transportation data from Atack (2013).

Back to page 31.