THE JOURNAL OF TRANSPORT AND LAND USE http://jtlu.org
VOL. 4 NO. 2 [Summer 2011] pp. 5–18 doi: 10.5198/jtlu.v4i2.188

JTLU

The impact of access to rail transportation on agricultural improvement

The American Midwest as a test case, 1850–1860

Jeremy Atack
Vanderbilt University and National Bureau of Economic Research[a]

Robert A. Margo
Boston University and National Bureau of Economic Research[b]

Abstract: During the 1850s, the amount of farmland in the United States increased by 40 million hectares (100 million acres), or more than one-third. Moreover, almost 20 million hectares, an area almost equal to that of the states of Indiana and Ohio combined, were converted from their raw, natural state into productive farmland. The time and expense of transforming this land into a productive agricultural resource represented a significant fraction of domestic capital formation at the time and was an important contributor to American economic growth. Even more impressive, however, was the fact that almost half of these total net additions to cropland occurred in just seven Midwestern states, which constituted somewhat less than one-eighth of the land area of the country at that time. Using a new GIS-based transportation database linked to county-level census data, we estimate that at least a quarter (and possibly two-thirds or more) of this increase in cultivable land can be linked directly to the coming of the railroad to the Midwest. Farmers responded to the shrinking transportation wedge, which raised agricultural revenue productivity, by rapidly expanding the area under cultivation and these changes, in turn, drove an increase in farm and land values.

Keywords: Railroads; Development; Midwest; Nineteenth century; Networks; Agriculture; Density

1 Introduction

Growth accounting suggests that capital formation accounted for approximately one-third of the increase in per-capita GNP in the United States between 1800 and 1840, and almost one-half of the increase between 1840 and 1900 (Engerman *et al.* 2000b). While some of this growth in the capital stock came from foreign sources, much was created through domestic savings and investment. According to Engerman *et al.* (2000b), for example, the domestic savings rate in the 1850s was between 17 and 20 percent depending on one's assumptions about prices and home manufacturing.

During the same period, a considerable portion of domestic investment took the form of land clearing and associated improvements to make the land suitable for agriculture. While the federal censuses prior to 1850 did not collect information on the amount of land in farms or improved acreage, we do know from other sources that, during the 1840s, the federal government sold off over 19 million acres (7.7 million hectares) of public land despite the bursting of the 1830s speculative land bubble (Carter *et al.* 2006, Table Cf79). Much of this land went (eventually) to farmers and, by 1850, the Census reported over 293 million acres (118 million hectares) of land in farms, of which about 38.5 percent had been improved (United States Census Office and de Bow 1854, Table 183).

Public land sales during the 1850s increased more than two-and-a-half fold. This represents an area 80 percent larger than the state of Ohio or 40 percent larger than Illinois(Carter *et al.* 2006, Table Cf79). Although it took time for the new land to be cleared and enter agricultural production, improved acreage grew by almost 20 million hectares during the decade and land in farms expanded by well over 40 million hectares. As a result, by 1860, a little over 40 percent of all land in farms had been improved (United States Census Office and Kennedy 1864, 184).

[a] jeremy.atack@vanderbilt.edu
[b] margora@bu.edu

Such additions to the stock of improved farmland may have accounted for more than half of all capital formation before 1820. Even as late as 1860, Engerman *et al.* (2000b) estimates that 22 percent of annual capital formation came from additions to the stock of improved acreage. However, while the quantitative significance of these improvements to farm land has been widely recognized (Engerman *et al.* 2000a; Primack 1962, 1977), the quantitative significance of the causal factors underlying such investments is less clear. Using a newly developed dataset, this paper sheds light on one such causal factor: the coming of the railroad, which lowered the costs of internal transportation and thus expanded commerce. In our analysis, we focus on the decade of the 1850s in the American Midwest since it was there that almost half of the increase in improved farmland during the decade occurred and was also where much of the railroad construction took place, especially between 1853 and 1856. Indeed, in 1856, almost 40 percent of the nation's new railroad construction occurred in a single Midwestern state, Illinois. Our new data consist of county-level information on the spread of the railroad network in seven Midwestern states (Illinois, Indiana, Iowa, Michigan, Missouri, Ohio, and Wisconsin) during the 1850s, which has been linked to existing county-level data on farm outcomes at the benchmark census dates. The transportation data, collected as part of a larger project, derive from the geographic information systems (GIS) processing of digitized nineteenth-century railroad maps and related sources as described below.

We measure the causal impact of the railroad on improved acreage using a difference-in-differences (DID) analysis.[1] This compares the change in farm outcomes between 1850 and 1860 in a treatment group of counties—identified as those that gained access to the railroad during the period—versus a control group that did not. We have excluded from consideration those counties that already had rail links in 1850 so that our treatment group consists of counties that gained rail access during the 1850s versus those counties that did not have access until after 1860. Our analysis is complicated by the fact that some counties' boundaries were redrawn during the period, while others did not even exist until the 1840s. Consequently, we have also restricted our analysis to the 278 counties in the Midwest that already existed in 1840 and did not change their boundaries during the 1850s.[2] The location of these counties is shown in Figure 1.

We have supplemented our analysis of the share of improved farm land with two robustness checks. First, because improved land was more valuable than unimproved land, we expect to find increases in the value of farm acreage associated with the coming of the railroad. We measure this as the average value of an hectare of land in farms in the county using census estimates of the "value of farms." This value includes the value of the land as well as improvements and structures, and our results confirm that the coming of the railroad did, indeed, have a treatment effect on farm values. Moreover, when we repeat this DID analysis of farm values but control for the percentage of land that was improved, we are only able to explain some of the increase in farm values, suggesting that the railroad's effects on land value were general—that is a pure location effect.

A second robustness check addresses a key assumption in our DID analysis: that (conditional on various observable determinants), the coming of the railroad to a county was a random event. If this assumption were false, the estimated treatment effect of the railroad on farm improvements would be biased. In particular, if county residents anticipated getting a railroad it would make economic sense to "build-in-advance"—that is, to improve acreage in advance of the railroad. If this were the case, then our DID treatment effect would be biased downwards and our estimates would undervalue the effect of the railroad.[3]

To address this possible bias, we compare our DID results with results of an analysis in which an instrument variable (IV) was used. This instrument is derived from various federal government transportation surveys conducted in the 1820s and early 1830s—that is to say, right at the start of the railroad age and well in advance of when we observe the possible effects of the railroad on these Midwestern counties. These results are consistent with our initial DID findings that the coming of the railroad had a relatively large effect, such that it accounted for at least a quarter of the increased fraction of improved land; this finding suggests that our DID estimates of the impact of the railroad may indeed be biased downwards.

[1] The difference in differences approach was largely pioneered by David Card (see Card and Krueger 1994) and is now widely used (see, for example, the mini courses given by Wooldridge (2007) to various organizations including the NBER, and Angrist and Pischke 2009). For criticisms of the methodology, see, for example, the special issue of the *Journal of Business and Economic Statistics*, especially Meyer (1995).

[2] This restriction does not substantively change the results but it makes our analysis cleaner.

[3] This, for example, is a key insight in the work of Levinson (2008) and Xie and Levinson (2010).

2 Nineteenth-century land clearing

While the first step to becoming a farmer was to acquire title to land, either through private purchase or directly from the federal government under the terms of the prevailing land legislation, the cost of purchasing raw land was generally minor compared with the cost of clearing it and readying it for cultivation. For example, according to contemporary estimates by Seaman (1852, 484, 656), imply the cost of clearing and fencing land averaged between about \$27 (in 1850) and \$30 (in 1840 when more of the land was wooded) per hectare, while Lebergott's data (Lebergott 1985, Table 3) suggest the median cost of improving a hectare of land in the Midwest was between about \$34 and \$62 in 1850. In contrast, the average price of raw land bought using military land warrants was less than a dollar per acre, or somewhat more than \$2 per hectare (Swierenga 1968, Table 6.1; Oberly 1990). These cost differentials reflect the fact that where the land was forested, trees had to be cut down and burned or hauled away. Even then, the more laborious and labor intensive task of stump removal remained. In areas where the land was naturally clear of trees, as on the Midwestern prairies, clearing expenses remained high as heavy, specialized equipment and crews were needed to break the prairie sod (Stellingwerff and Swierenga 2004). Even when the land had been cleared, it still had to be fenced to keep livestock in and predators out. Many of these activities were expensive, arduous, and time-consuming even when they could be done by the farmer and the family using only those tools and implements that came readily to hand on a nineteenth-century farm. As a result, farmers rarely cleared all the land they needed or wanted at once. According to Primack (1962, 1977), farmers could only clear two to four hectares a year. Thus, clearing the approximately 525 000 hectares of land (much of it forested) that was cleared annually in the Midwest in the 1850s must have occupied the time of about one-sixth of the Midwestern labor force at the time. Needless to say, with this kind of investment in cash and time, cleared land was highly valuable compared with virgin stock and accounted for much of the resale price of farm land.

3 The impact of transportation improvements on land values

The economic logic linking transportation costs to land improvements can be illustrated using a von Thünen model of agricultural land rents first proposed in 1826 (von Thünen and Hall 1966). To fix ideas, suppose that each farmer has a farm of size one hectare (of improved land), located at distance t from a "central place" where trade with the hinterland (i.e., the farmer) takes place. The farm supplies labor z at an opportunity cost of w. The cost of transporting output q per unit of distance is c and the price of output (at the central place) is p.

In a competitive market these assumptions imply that the rental price $R(t)$ of a farm at distance t from the central place is

$$R(t) = q(p - w \cdot z - c \cdot t) \tag{1}$$

Note that if $t = 0$ the farm is located on the boundary of the central place and $R(t)$ is at a maximum as determined by the yield of the land (q) given the labor input (z), output price, and the opportunity cost of the farmer's time. Conversely, there will exist a t^* that defines the distance from the central place beyond which it is not profitable to farm, $t^* = (p - w \cdot z)/c$. Economists sometimes refer to this as the extensive margin for market production.

It is straightforward to show that $\mathrm{d}R(t)/\mathrm{d}c = -q \cdot c < 0$ and $\mathrm{d}t^*/\mathrm{d}c = (p - w \cdot z)/c^2 > 0$; that is to say, a reduction in transport costs "flattens" the relationship between land rents and distance to the central place and, therefore, increases the distance t^* from the central place at which farming is still just profitable. A fall in transport costs therefore makes it economical to bring more land into production. However, before this land can be brought into production it has to be improved—that is, cleared and fenced—hence the link between transport costs and land improvements.

Of course, other factors also determine t^* and therefore affect the amount of land in production. A rise in p, for example, or a reduction in w, will also increase t^*. Moreover, in an explicitly dynamic setup, the permanent versus transitory nature of changes in these parameters will influence the likelihood of improving land and bringing it into production, as will expectations about future values. Railroad building was (in general) a permanent change in the environment, and thus could be expected to generate a greater response in land improvements than, for example, a shift in output prices, which might prove to be transitory. One reason for this was that much of the railroad investment was not only fixed but also "sunk," in that it was embedded in location-specific improvements. A railroad, once built, could not be relocated easily, and the benefit of that investment could only be fully captured by another transportation medium—investments in grading, filling, and cuttings, for example. According to the 1880 Census (United States Census Office *et al.* 1991), over 80 percent of railroad investment went into construction costs, of which only one or two percent represented the cost of the land itself; the rest was spent on surveying, grading, removing or bridging obstacles, and laying the track. While the railroad ties, ballast, and rails might be reused elsewhere and the land itself could be resold, the grading, cuttings, embankments, bridges, and drainage ditches had few alternative uses other than serving the needs of a transportation system—especially in the nineteenth century.[4] Indeed, railroads (and interstate highways) today still follow many of the routes blazed by the nineteenth-century railroads.[5]

If there were any adjustment costs involved in making improvements to their land—as would seem likely—then it would be rational for farmers to improve some land in advance of construction if they anticipated receiving rail access in the near term—that is, farmers are expected to be forward-looking rather than merely reactive.[6] We will return to this point later in our instrumental variable analysis, as it relates to a possible downward bias in our difference-in-differences estimation.

4 The data

We use a newly constructed geographic information system (GIS) database on the transportation infrastructure in the mid-nineteenth century Midwestern United States, which we have linked to an updated version of the well-known ICPSR county-level census database. This GIS-based methodology has several advantages over earlier approaches, which have generally involved matching historical transportation maps to county boundary maps by hand (and eye).[7] Our GIS database was created from geo-referenced digitized historical maps showing the transportation system—canals, navigable waterways, and railroads—at different benchmark dates.[8] We have not yet tried to render information on roads and trails.

The canal data were taken from Poor (1970) and Goodrich (1961) as well as from digitized maps accessed via the Library of Congress "American Memory" web site, including the maps prepared by Williams (1851) and Disturnell (Burr 1850). Information on navigable rivers was taken from U.S. Army Corps of Engineers GIS data (Vanderbilt University Engineering Center for Transportation Operations and Research 1999), supplemented by information from Hunter (1949) and contemporary newspaper accounts regarding steamboat service on specific rivers. Since very little change occurred in the canal and river systems from the 1840s until the Corps of Engineers embarked on its ambitious navigation plans beginning with the establishment of the Office of Western River Improvements in 1866, we have measured access to water-borne transportation as of a single benchmark date (1850) rather than by change over time.

The railroad portion of our GIS database is based on digitized state-level maps for 1911 from *The Century Atlas* (Whitney and Smith 1911) since these maps appear to be accurately drawn and the rail network was largely completed by that time but had not yet begun to shrink through closures.[9] This 1911 railroad network was then traced back through time to earlier dates using digitized transportation maps from the Library of Congress and other sources such as the maps created by Taylor and Neu (1956). Our implicit assumption in this approach is that most railroad investment was literally sunk in location-specific grading and other immoveable features, a presumption that is strongly supported by the available data.

[4] Today, of course, they find recreational uses in the "rails to trails" movement.

[5] And, of course, thanks to GIS, these can be quite precisely located. One might thus be tempted to begin with modern railroad maps and force the nineteenth century locations to conform to these. On average, this is probably a good strategy even though it would misrepresent earlier track locations that were changed through realignments. There are two downsides to this strategy—the potential for copyright infringement since these GIS SHP files are proprietary commercial ventures and the large number of individual segments which comprise the modern rail network. For example, the 177 000 mile rail network from Bureau of Transportation Statistics and distributed ESRI with ArcGIS contains over 131 000 segments. See United States Department of Transportation (1998).

[6] For a similar point, see, for example, Levinson (2008).

[7] See, for example, Craig *et al.* (1998) who visually compared historical maps to county boundaries (which generally did not appear on the historical maps). For a discussion of the hand-matching procedure used by Craig and the pitfalls that can arise, see footnote #2 of Atack *et al.* (2010).

[8] Geo-referencing refers to the process of fixing specific points with known geographic coordinates between the digitized image—which was invariably drawn and printed with error (and which may also be subject to parallax error as a result of the digitization process)—and the geographically accurately rendered base boundary files (a "shapefile" in ESRI's parlance). Algorithms within the GIS software then distribute the error (the difference) between the historical images and the boundary file across the space between fixed points. In essence, the process treats the historical image as if it were printed on a sheet of rubber which is then stretched over the boundary file with pins holding it in place at fixed reference points between the two. Once done, it is then possible to "trace" features from the historical image onto the geographically coordinated boundary file. The resulting files can then be manipulated and used for computations using the GIS software. Accurate historical county boundary files are freely available from the National Historical Geographical Information System at the University of Minnesota (http://www.nhgis.org/) along with a wide variety of historical United States census data.

[9] These digitized images were purchased from Goldbug.com but the original source may be found in many libraries around the country, including Vanderbilt's.

For the Midwest, which is our focus here, we have supplemented these more general sources with detailed information for the seven states in the region. Especially detailed information about the spreading rail network in five of these states (Indiana, Illinois, Michigan, Ohio and Wisconsin) exists in the form of a series of crudely drawn maps and a data appendix prepared more than ninety years ago at the University of Wisconsin by Professor Frederick L. Paxson (Paxson 1914) and his students, who used extant contemporary travel guides published between 1848 and 1860 as their principal sources of information.[10] Since then, however, many of these travel guides (which were cheaply printed ephemera) seem to have been lost or to have disintegrated. This is unfortunate, as there is little doubt that they are, or rather would have been, the single most valuable resource for our research. These guides served the needs of the traveling public, providing up-to-date route maps and timetables. Consequently, they were published at fairly frequent and regular intervals by a number of different companies and, presumably, competition among them should have guaranteed that only the best (that is, the most accurate, complete and useful) would have continued publication for an extended period of time.[11] Our data for Iowa and Missouri are taken from digitized contemporary state maps from the online David Rumsey Map Collection (http://www.davidrumsey.com/) for 1856 (Colton 1856a,b), 1857 (Colton 1857) and 1859 (Mitchell 1859a,b) supplemented by national maps from the Library of Congress collection for 1858 (Sage 1858) and 1860 (Colton 1860). While GIS software is capable of generating a wide variety of different measures of transportation access, at the present time we use a simple county-level binary variable, ACCESS. This takes a value of 1 if a railroad crossed or formed a county's border. Our choice of this measure reflects the still-preliminary nature of our GIS database. While we are certain that it can create the binary "0-1" access variable accurately, we are not yet convinced that the original maps and our "tracings" of them are sufficiently precise to generate more refined measures of access, such as the miles of track in the county. Another measure of "access"—proximity to a railroad stations or depots—can only be measured beginning in the mid-1850s when maps begin to record this important information. For example, the earliest digitized map in the Library of Congress collection to mention stations and depots explicitly dates from 1856 (Ensign, Bridgman & Fanning 1856). Paradoxically, it shows many more named places along the right of way than the map from Dinsmore's travel guide for same year, which might have been expected to show prospective travelers their station destinations. Nor does it appear that the Ensign, Bridgman and Fanning map was overly sanguine in its reporting of railroad stations. Another map published two years later (Sage 1858) shows most of the same stations as Ensign, Bridgman and Fanning plus additional ones along the new rights of way as well as a few others scattered between some of the 1856 stations and depots. Beginning in the 1870s, Rand-McNally (Morgan 1969) and other publishers began to produce "commercial" maps for the convenience of shippers, showing destinations to which freight could be consigned; these maps made the reporting of depots and stations more reliable and systematic. As a practical matter, however, trains could potentially stop and load almost anywhere (albeit inconveniently)—much as they do in parts of Alaska even today—and stations were built only when freight and passenger volume had passed some critical economic threshold.

Despite our measure of access to transportation being crude and limited, it is still the best current measure of the historical spread of the United States railroad system at the county level for this period.[12] These transportation access data have then been linked with the county-level Haines-ICPSR census data using county FIPS codes, which are common to both

[10] These travel guides first appeared in the 1840s and include *Disturnell's Guide* (Disturnell 1847), *Doggett's Gazetteer* (Doggett 1848), *Appletons' Guide* (Appleton and Company 1848), *Dinsmore's Guide* (Cobb 1850), *Lloyd's Guide* (Lloyd 1857), *Travelers' Guide* (National Railway Publication Company 1868) and *The Rand-McNally Guide* (Rand-McNally and Company 1871). Some of these were published monthly; others, semi-annually or annually. Each typically went through many editions. All of the guides that we have physically handled are fragile, especially the multi-page fold-out maps, and not sturdy enough for scanning or copying, although a few guides have been digitized and are available on-line. See, for example, the June 1870 copy of the Travelers' Official Railway Guide at http://cprr.org/Museum/Books/I_ACCEPT_the_User_Agreement/Travellers_Guide_6-1870.pdf from the Central Pacific Railroad Museum. There are also at least two different editions of *Appleton's Guide* on Google Books such as http://books.google.com/books?vid=UOM39015016751375 as well as a number of other guides. See http://www.lib.utexas.edu/maps/map_sites/hist_guide_sites.html.

[11] However, as we note below, these sources do not seem to have been infallible—certainly at any moment of time.

[12] One obvious problem is that the access variable does not capture within-county variation in access and counties differed in size. In our econometric analysis we weight by county land area but we are unable at the present time to adjust for cases in which the railroad lay close to a county boundary but where the adjacent counties do not have rail access. County boundaries, in other words, are arbitrary ways of delineating local economies; the average farmer in a geographically small county in our control group that happened to be adjacent to a large county in the treatment group might have had better access to the railroad than the average person in the large county. If this were common it would produce a downward bias in our estimated treatment effects. Although we do not believe this bias to be important empirically we admit that the question is open.

databases. In 1840, the seven states and territorial areas in our study had 391 "counties" (these include, for example, Clayton "county" in the Iowa Territory, which then comprised most of what we now know as the state of Minnesota) (Thorndale and Dollarhide 1987). By 1860, however, all of the area under consideration had been organized into states and was divided into 623 counties. Because of boundary changes, we have restricted our analysis to those counties that were: (1) present beginning in 1840; (2) had the same county boundaries (as determined by square mileage in 1850 and 1860); and (3) did not have rail access as of 1850. Taken together, these restrictions produce a balanced panel of 278 Midwestern counties. The distribution of these counties is shown in Table 1, and their locations are mapped in Figure 1.[13] Qualitatively, our results are unchanged if we relax the requirement that boundaries remain unchanged.

We have divided this panel into a treatment group and a control group. The treatment group consists of those counties that did not have rail access in 1850 but gained it at some point between 1850 and 1860. The control group is the set of counties that did not have rail access in either 1850 or 1860. The 1840 data for these same counties then allows us to control for pre-1850s trends in certain variables in our analysis (see below).

By construction, none of the counties in our sample had rail access in 1850. By 1860, however, 195 of the 278 counties, representing almost 72 percent of the land area in the panel, had gained direct access to at least one railroad (See Table 2 and Figure 2). The share of farm land that was improved rose by 13 percentage points between 1850 and 1860 across all counties in the panel, while the logarithm of per-hectare value of farms, which we use in our robustness checks (see below), increased by 0.796 log points or 123 percent ($= \exp[0.796] - 1) \times 100$).

Table 2 also shows the sample averages for percent improved and the logarithm of farm values per hectare conditional on their treatment status. In the treatment counties, the percentage of land that was improved rose by 15.6 percentage points from 1850 to 1860 or almost double the 8.4 percentage point increase in the control counties. The difference between these—7.2 percentage points—is the difference-in-differences estimate of the treatment effect of gaining rail access with no other control variables in the regression (see below). This is equal to 18 percent of the mean proportion of land that was improved in 1850. As such, the effect is economically large. The equivalent difference-in-differences estimate of the treatment effect of rail access on per-acre farm values is 0.072, or about 7.5 percent. Both estimates are statistically significant at the five-percent level.

In an ideal difference-in-differences analysis, the treatment and control counties would be sufficiently "similar" or "matched" that treatment could be (plausibly) argued to be randomly assigned. However, it is immediately evident from Table 2 that this was not the case with rail access. Counties that gained rail access in the 1850s already had a greater fraction of their farmland improved and this was also more valuable land by 1850—that is, before gaining rail access.

The observations that the treatment counties already had more of their farm acreage improved and more valuable farms (in part, because the land was more improved, see below) are manifestations of the same general issue: railroads did not arrive randomly. In particular, railroad promoters and investors in the Midwest paid close attention to agricultural development because the profitability of the railroad depended upon it (Fishlow 1965; Gates 1934). This is illustrated in Table 3, which shows the correlates, agricultural and otherwise, of gaining rail access in the 1850s using a simple linear probability model. In column 1 of Table 3, we have included

Table 1: Distribution of balanced panel of sample counties.

State	Number of counties	Percentage of total
Illinois	71	25.5
Indiana	66	23.7
Iowa	17	6.1
Michigan	12	4.3
Missouri	41	14.7
Ohio	57	20.5
Wisconsin	14	5.0
Total	278	100.0

Notes: To be included in the sample, counties must: (1) be present in all three census years (1840, 1850 and 1860); (2) have fixed county boundaries; and (3) either not have rail access in 1850 but have gained rail access by 1860 (treatment group) or not have rail access in either 1850 or 1860 (control group).

[13] By "balanced" we mean that the same counties appear in 1850 and 1860; no new counties enter the sample during the 1850s. Balancing ensures that county fixed effects are "differenced away" when we compute the change in economic outcomes from 1850 to 1860; this would not be the case if new counties entered the sample in the 1850s. We restrict our basic analysis to counties with fixed land area because the ICPSR census data are not adjusted for changes in land area over time. Results are qualitatively the same if we do not impose the restriction that the county not have rail access by 1850 and if we do not impose the requirement that county boundaries be the same in 1850 and 1860.

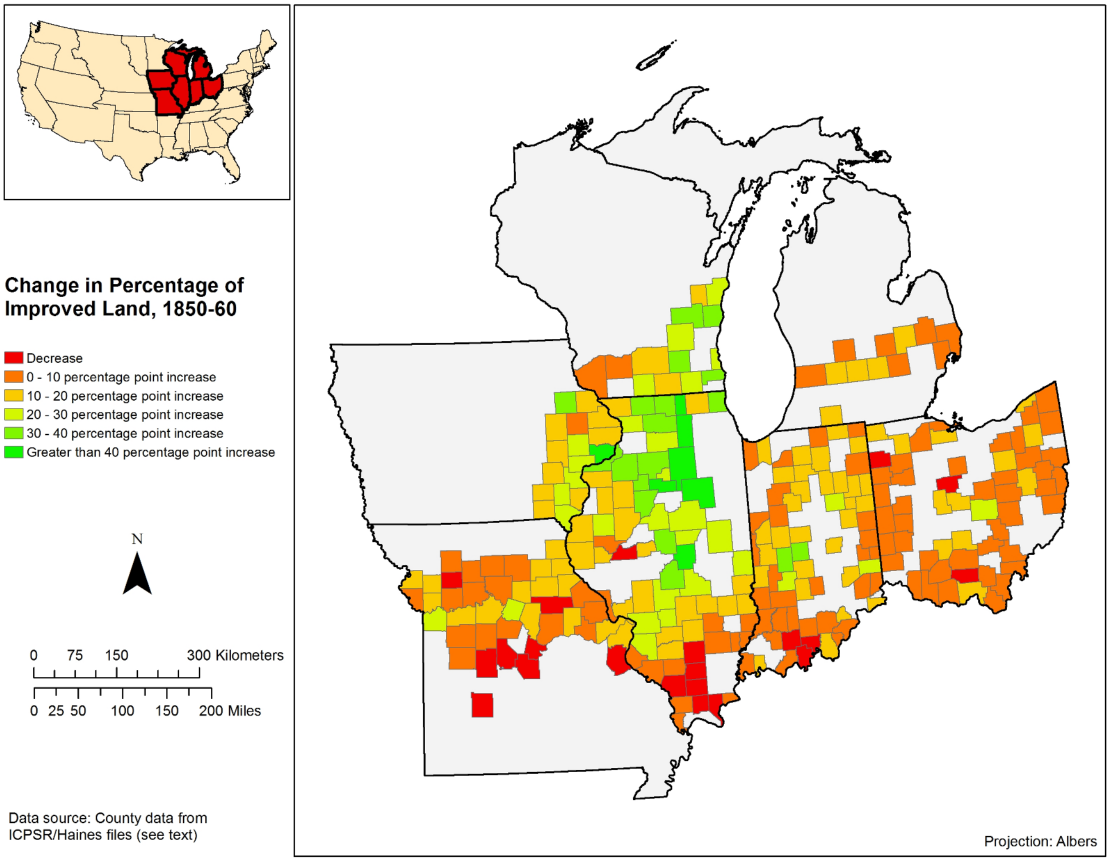

Figure 1: Percentage change in improved farm land 1850–1860 in sample panel of counties with stable boundaries.

Table 2: Sample statistics of key variables in aggregate and in control and treatment counties.

	1850	1860	Δ (1860 – 1850)
Rail = 1	0	0.717	0.717
% Improved Land in Farms	0.390	0.526	0.136
Treatment Counties	0.411	0.567	0.156
Control Counties	0.337	0.421	0.084
(Treatment) – (Control)	0.074	0.146	0.072
Ln (Per Hectare Farm Value)	3.024	3.820	0.796
Treatment Counties	3.116	3.933	0.817
Control Counties	2.792	3.537	0.745
(Treatment) – (Control)	0.324	0.396	0.072

Source: See Table 1.

There are 278 counties in the sample. Treatment = 1 if county gained rail access between 1850 and 1860. Control = 1 (Treatment = 0) if the county does not have rail access before the Civil War. There are 195 treatment counties and 83 control counties. Observations are weighted by surface area prior to calculating sample means.

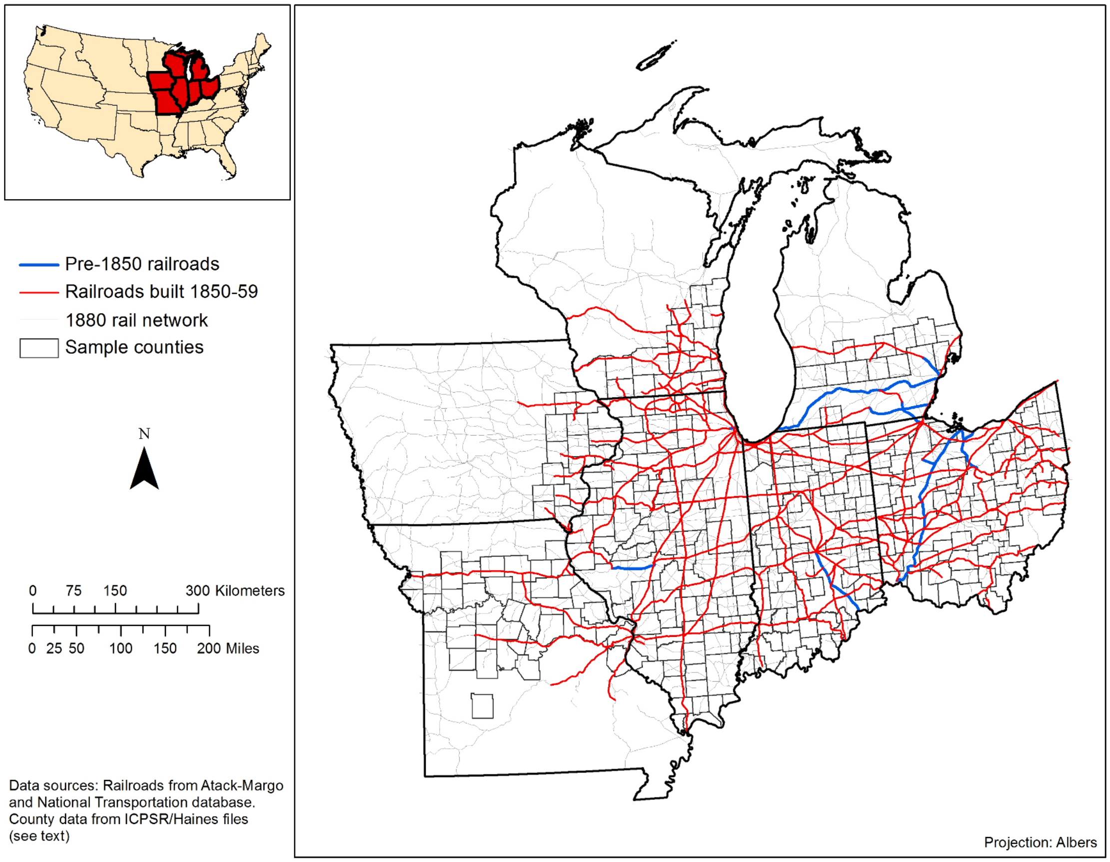

Figure 2: The Midwestern railroad network at benchmark dates relative to our panel of counties with stable boundaries.

just three covariates: the logarithm of the value of agricultural output per hectare (a measure of the land's "yield," in this case its revenue productivity); the percentage of total agricultural output in 1840 represented by wheat production; and the change in the wheat output percentage between 1840 and 1850.[14] Wheat is a crop of particular importance at this time as it was the primary "cash crop" of the period and, therefore, likely a key determinant of whether a rail line would be profitable. Each of these variables had a positive and significant effect on the likelihood of gaining rail access; for example, a 10 percentage point increase in the wheat share boosts the probability of gaining rail access by 5.3 percentage points.

However, the significant positive coefficients on the agricultural variables may also be due to other factors that are omitted from the regression. For example, an earlier paper (Atack *et al.* 2010) examined the relationship between gaining rail access and measures of urbanization, population density and our water transportation dummies using the same panel. Accordingly, in column 2 of Table 3 we have included measures of urbanization and population density in 1840 as well as changes in these variables between 1840 and 1850. We have also included dummy variables for access to different modes of water transportation (river, canals, and abutting the Great Lakes) while in column 3 we have also included dummy variables for each state. Adding additional controls reduces the magnitudes of the coefficients of the agricultural variables but the coefficients on the wheat "yield" variable and the "pre-trend" (1840–50) in the wheat share remain statistically significant. Moreover (not shown) several of these control variables are themselves statistically significant; for example, the

[14] The 1840 census reports the value of wheat production and the total value of agricultural output. For 1850 we use an estimate of the value of total agricultural output based on national prices multiplied by quantities; the wheat share is therefore the wheat output (in bushels) multiplied by wheat price divided by the estimated value of agricultural output. It is possible that our procedure for estimating agricultural values may overstate the growth in the percentage of wheat over the 1840s (because 1840 output was probably valued at local prices) but any such bias should be mitigated once we control for state fixed effects (since state-level variation arguably captures the most salient price variation). Our agricultural yield variable also uses the estimated value of agricultural output in its construction (the numerator).

Table 3: County-level agricultural correlates of gaining rail access in the 1850s: linear probability regression.

Urban and Pop. Density Variables Included?	No		Yes		Yes	
Water Transportation Dummies Included?	No		Yes		Yes	
State Dummies Included?	No		No		Yes	
	Coefficient	(Std. err.)	Coefficient	(Std. err.)	Coefficient	(Std. err.)
Constant	0.516*	(0.113)	0.627*	(0.123)	0.644	(0.15)
Log (Agricultural Yield in 1850)	0.325*	(0.077)	0.290*	(0.078)	0.280*	(0.092)
Percent Wheat in 1840	0.533*	(0.244)	0.349	(0.253)	0.356	(0.283)
Change in Percent Wheat, (1850) – (1840)	0.684*	(0.205)	0.652*	(0.272)	0.615*	(0.275)
Adjusted R^2	0.111		0.190		0.211	

Source: see text and Table 1.

Notes: Unit of observation is the county (N = 278). Dependent variable = 1 if county gains rail access by 1860 (treatment group), 0 otherwise (control group). Agricultural Yield: value of agricultural output/(improved + unimproved acres in farms). Percent wheat: value of wheat output relative to value of total agricultural output. Urban and Population Density variables: percent urban in 1840 is ln (population/square miles) in 1840; change in percent urban between 1840 and 1850 is change in ln (population/square miles) between 1840 and 1850. Water transportation: canal = 1 if canal existed within county boundary (or part of boundary), river = 1 if navigable river existed within county boundary or part of county boundary, Great Lakes = 1 if county abutted one of the Great Lakes. Prior to estimation, observations are weighted by the number of square miles in the county.

*: significant at five percent level.

presence of a canal was positively associated with gaining rail access, whereas having a navigable river reduced the statistical likelihood of gaining rail access.

Because gaining rail access clearly was not a random event, it is important, at the very least, that we control for the factors shown in Table 3 in our difference-in-differences analysis as they may have had independent effects on the change in percentage of improved acreage between 1850 and 1860. Column 1 of Table 4 reports our base DID estimate (that is, absent any controls for the other correlates of rail access) for the percentage of total farm acres that were improved. This is the same as the figure reported in Table 2. In column 2, we report the DID estimate of the treatment effect, taken from a regression in which we have included interactions between the full set of control variables in Table 3 and the dummy variable for the 1860 census year. Including these control variables reduces the DID treatment effect from 0.072 (base estimate) to 0.048, or 4.8 percentage points. However, although the addition of controls reduces the impact of getting rail access, the treatment effect remains significantly positive: counties that gained rail access in the 1850s increased improved acreage to a significantly greater extent than counties that remained without rail access before the Civil War even after controlling for other factors.

We have made two robustness checks on this finding. First, we estimate the treatment effect of gaining rail access on the

Table 4: Difference-in-differences estimates of the impact of gaining rail access on the percentage of farm land that was improved: 278 Midwestern counties, 1850–60.

Controls?	No		Yes	
	Coeff.	(Std. err.)	Coeff.	(Std. err.)
Rail Access = 1	0.072*	(0.015)	0.048*	(0.015)

The unit of observation is the county. The dependent variable is the percentage of farm area in the county that was improved as reported in the census of agriculture. The coefficient of Rail Access = 1 is the coefficient of a dummy variable for the treatment counties (Treatment = 1) interacted with a dummy variable for the year 1860. In column 1, there are no control variables other than fixed effects for counties and Year = 1860. In column 2, the regression also includes interaction terms between the control variables in Table 3, column 3 (see the notes to Table 3) and Year = 1860. Prior to estimation, observations are weighted by surface area of the county. Standard errors shown in parentheses are clustered at the county level.

*: significant at the five percent level.

logarithm of the value of farms per hectare of farmland. If the logic underlying the von Thünen model is correct, then the impact of gaining rail access should also be reflected in higher farm values per hectare. We have already shown that this was the case with the simple DID calculations in Table 2, but those estimates do not control for additional factors associated with gaining rail access. Moreover, the size of the treatment effect should be smaller still if we also control for the change in the percent improved (since improved land itself was more valuable)—again, something that is not reflected in the DID estimate in Table 2.

Table 5 shows the difference-in-differences estimates for farm values. In contrast to the results for improved farmland, the effect of gaining rail access is larger once we add the additional controls. Note that, when we also add the change in the percentage of improved area to the regression, the treatment effect is reduced in size (as the von Thünen framework would predict), but controlling for the percentage improved does not eliminate the effect of gaining rail access on land values. This suggests that transportation improvements raised the value of unimproved land as well. This, of course, generally should be the case since most unimproved land has an option value—the farmer could always choose to improve it in the future.

Our second check on the robustness of our results is made using an instrumental variable. Although we can control for the various correlates in Table 3 in our DID analysis, it is still possible that the analysis is invalid because the treatment—gaining rail access—is correlated with the error term even after controlling for observable determinants of gaining rail access. This problem is of particular concern here because the 1840 census reported no acreage statistics making it impossible for us to control for the pre-1850 trend in the percentage of land that was improved. This will likely bias the treatment effect towards zero.

To understand why this is so, consider the following thought experiment. Suppose that farmers in a county anticipated gaining rail access in the near future—a reasonable supposition for Midwestern farmers at the start of the 1850s.[15] If the costs of improving farmland were increasing at the margin which was almost certainly the case since the most readily improvable land would be improved first, it would make economic sense to incur some of the costs prior to gaining rail access. If this were the case, counties that gained rail access in the 1850s would have experienced above-average growth in improved farmland somewhat earlier, and this would be reflected in a higher average percentage of improved land in 1850—as we observed was the case (recall Table 2). Conversely, some of the control counties might have anticipated (correctly, as it happens) that they would only gain rail access after 1860. By this same logic, they would have experienced above-average growth in improved land in the 1850s. The bias can be "corrected" to the extent that our control variables capture this process, but we cannot be sure we have captured all of it. Consequently, the percentage of improved land will grow too slowly in the treatment counties in the 1850s and too fast in the control counties, causing the DID estimate of the treatment effect to be biased downwards.

The appropriate correction for this sort of bias is to estimate the relationship using an instrumental variable—a variable that predicts gaining rail access in the 1850s when we control for other factors but which is otherwise uncorrelated with the outcomes we are examining. Using a variable which isolates plausibly exogenous variation in rail access, as we propose here, is similar to what would have been the case if rail access had been randomly assigned. We then predict rail access in 1860 using the instrumental variable and examine the effect of that predicted access on the percentage of land that was improved in 1860.

The historical narrative of internal improvements in America, particularly that for canal construction, assigns an important role to government in promoting these advances (Goodrich 1961). One such important source of government assistance was the assignment of the Army Corps of Engineers to conduct surveys for potential transportation routes. Beginning in 1824, the President was granted authority to survey routes for "such roads and canals as he may deem of national importance, in a commercial or military point of view, or necessary for the transportation of the public mail" (United States Congress 1824). Although railroads were not mentioned in the original act (hardly surprising since it predates even the Stockton to Darlington Railway in England, the first steam railroad in the world), it was not long before surveys conducted under this legislation also considered them. For exam-

[15] For example, based on new rail construction in 1848 and 1849, farmers in five counties in the balanced panel (Boone, Dearborn and Washington Counties in Indiana, Genesee County in Michigan and Henry County in Ohio) might reasonably have expected that they too would have a railroad in their county within a year or so. In fact, if a dummy variable (ANTICIPATE = 1) is included in the rail access regression in Table 3, the coefficient is positive ($\beta = 0.314$), indicating that such counties did gain rail access in the 1850s. However, the standard error of the coefficient is relatively large (s.e. = 0.195); more importantly, including this additional dummy variable has virtually no effect on the other coefficients in Table 3. Further, if ANTICIPATE is included in the DID specifications (Table 4 or Table 5, full set of controls) the estimated treatment effects of rail access are unchanged. Presumably this is because the number of counties (five) is so small in the 1840s, most of the rail construction was taking place much further east.

Table 5: Difference-in-differences estimates of rail access on Ln (value of farms/farm hectares): 278 Midwestern counties, 1850–1860.

Controls?	No		Yes		Yes	
	Coefficient	(Std. err.)	Coefficient	(Std. err.)	Coefficient	(Std. err.)
Rail Access = 1	0.072*	(0.033)	0.139*	(0.034)	0.107*	(0.032)
% Hectares Improved in Farms					0.657*	(0.164)

The dependent variable is the natural logarithm of the value of farms as reported by the census (this includes the value of land plus structures, fencing, ditching, etc.) divided by total farm acres (improved + unimproved). See notes to Tables 3 and 4 for controls in columns 2 and 3. All regressions include fixed effects for counties and year = 1860. Standard errors are clustered at the county level.

*: significant at the five percent level.

ple, in 1825 a survey to "ascertain the practicability of uniting the headwaters of the Kenawha [sic] with the James river and Roanoke river" expressly mentions railroads. Railroad routes soon came to dominate the surveys with perhaps as many as 61 such surveys being made before the law was repealed by Andrew Jackson's administration effective in 1838 (Haney 1908, 277).

Our instrumental variable is derived from the government surveys reported in American State Papers and compiled by Haney (Haney 1908, 283). For each survey, we have identified the counties at the start and endpoint of the proposed line. For example, an 1831 railway survey plotted a route from Portage Summit on the Ohio Canal (near Akron) to the Hudson River (we used Albany as the terminus as the city was not otherwise specified) (Haney 1908, 286). In some cases we inferred both endpoints, as in the case of an 1832 survey for a route between "the Mad River and Lake Erie" in Ohio (Haney 1908, 286). We used Springfield and Sandusky as the termini of this projected railroad. Having identified the starting and ending counties, we then drew a straight line between the center of the "start" and "end" counties. Counties that lay along this straight line received a value of one, while those that did not were coded as zero. That is, if a railroad were built, our instrument presumes that it would be built in a straight line as the shortest distance between the two points.[16]

These U.S. Army Corps of Engineers' surveys provided valuable information to the general public and prospective railroad promoters alike regarding topography and other factors that would affect potential construction costs. Since the costs of these surveys were borne by the public purse, their existence should have raised the likelihood that a railroad would eventually be built by lowering its private costs. Indeed, George Rogers Taylor (Taylor 1951, 95) even argued that "[a]s trained engineers were still very scarce...the government rendered a uniquely valuable service by making its experts available for such surveys." Moreover, as Haney (Haney 1908, 284) observed, "it is of some significance that in most cases the routes of these government surveys were early taken by railways...in the great majority of cases these early surveys have been closely followed" [emphasis added]. Indeed, many of the interstate highways today follow these same routes. Presumably, therefore, the surveys were found to be very useful and so it would seem that our "Congressional Survey" instrument should be well suited to predicting whether or not a county gained rail access.

Table 6 reports our estimates using this instrumental variable. The Congressional Survey instrument does quite well in predicting treatment (gaining rail access in the 1850s) even when we control for all of the other variables included in the DID analysis (this is the "first stage" coefficient shown). We then use this first stage regression to predict the probability of gaining rail access, and use the predicted values of treatment in the second stage of the two-stage least squares (2SLS) regression of 1860 outcomes. The 2SLS coefficient is positive and much larger than our DID estimate (0.192 versus 0.048). The difference between the two is also statistically significant at the six percent level. The much larger IV coefficient strongly suggests that the anticipatory logic described earlier was present, leading to a downward bias in the estimated treatment effect of gaining rail access on improved acreage.

[16] Our use of a "straight-line" instrument is inspired in part by Banerjee *et al.* (2006) who construct a similar instrument for their study of the impact of rail access on wages in modern Chinas. Or course, many topographic features other than the shortest-distance criterion—grade, hills or mountains, and so on—influenced railroad building, but these topographic features likely affected density and urbanization directly, and thus are not candidates for instrumental variables (they fail the exclusion restriction).

Table 6: Instrumental variable estimates of the treatment effect of gaining rail access on the percentage of farm land that was improved: 278 Midwestern counties, 1850–60.

	First Stage, IV		Second Stage (2SLS)	
	Coefficient	(Std. err.)	Coefficient	(Std. err.)
Rail Access = 1			0.192*	(-0.073)
Congressional Survey = 1	0.316*	(-0.068)		
IV – DID = 0, Significance Level			0.06	
Sample	1860		1860	
Controls?	Yes		Yes	

The first stage shows the coefficient of the Congressional Survey instrumental variable from a linear probability regression of rail access. Controls are the same as in column 3 of Table 3. The second stage shows the coefficient of the predicted value of rail access from a regression of percent improved using the 1860 observations from the panel; control variables are the same as in column 1, above.

IV – DID = 0: significance level of a test of the difference between the second stage coefficient (0.192) and the DID coefficient from Table 4, column 2 (0.048).

5 Concluding remarks

We have used a novel data set on the antebellum transportation networks in the Midwest, derived from applying GIS software to digitized historical maps, to estimate the impact of gaining rail access on investments in improving land for agricultural activity. Such investment during the nineteenth century was a major component of United States capital formation at the time. Counties that gained early access to the railroad—in our case, prior to 1860 rather than later—experienced significantly greater increases in the percentage of farm acres that were improved. These counties also experienced above average increases in the value of their farms. This was due in part to the rise in the share of improved farmland, but it also reflected the fact farm land in general was more valuable because of its higher revenue product and locational advantage with the advent of the railroad.

We can use the estimated treatment effects to gauge the overall explanatory power of the railroad as a factor behind these agricultural improvements. Using the difference-in-differences estimate (0.048, Table 3, column 4), the predicted change in percentage of improved farmland between 1850 and 1860 due to gains in rail access is 0.034, or 3.4 percentage points ($0.048 \times 0.717 = 0.034$). This predicted increase is 25 percent ($0.034/0.136 = 0.25$) of the increase in the percentage of improved land over the 1850s. If, however, we were to use the much larger IV coefficient (0.192), the predicted increase in percentage of improved farmland is correspondingly larger: 0.138. Taken at face value, this would imply that the coming of the railroad accounts for all of the increase in percentage of improved farmland over the decade—indeed, that expansion would have been even more rapid but for some (unidentified) retarding factors. Although the IV coefficient is almost significantly different from the DID coefficient at conventional significance levels, it is not significantly different from, say, the average of the DID and IV coefficients. If we use this average ($[0.048+0.192]/2$), 0.12, then we can attribute slightly more than two-thirds (68 percent = $[0.12 \times 0.771]/0.136$) of the growth in improved farmland to the spread of the railroad.

In short, whatever else might have led Midwestern farmers to undertake the back-breaking labor of clearing their land for crops and livestock, it is likely that no other single factor was as important as the potential gains from trade deriving from the coming of the railroad.

About the authors

Jeremy Atack (jeremy.atack@vanderbilt.edu) is Professor of Economics and Professor of History at Vanderbilt University, and a Research Associate at the National Bureau of Economic Research (NBER).

Robert A. Margo (margora@bu.edu) is Professor of Economics at Boston University and a Research Associate at the National Bureau of Economic Research (NBER).

References

Angrist, J. D. and J. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Appleton and Company. 1848. *Appletons' Illustrated Railway and Steam Navigation Guide*. D. Appleton & Co.

Atack, J., F. Bateman, M. Haines, and R. A. Margo. 2010. Did railroads induce or follow economic growth? Urbanization and population growth in the American Midwest, 1850–60. *Social Science History*, 34(2):171–197. doi: 10.1215/01455532-2009-024.

Banerjee, A., E. Duflo, and N. Qian. 2006. The Railroad to Success: The Effect of Infrastructure on Economic Growth. Unpublished presentation, Brown University.

Burr, H. A. 1850. *Disturnell's New Map of the United States and Canada*. Disturnell.

Card, D. and A. B. Krueger. 1994. Minimum wages and employment: A case study of the Fast-Food industry in new jersey and pennsylvania. *The American Economic Review*, 84(4):772–793. ISSN 00028282. URL http://www.jstor.org/stable/2118030.

Carter, S. B., S. S. Gartner, M. R. Haines, A. L. Olmstead, R. Sutch, G. Wright, and L. P. Cain, eds. 2006. *Historical Statistics of the United States Millennial Edition Online*. New York: Cambridge University Press. URL http://hsus.cambridge.org/HSUSWeb/.

Cobb, C. 1850. *American Railway Guide, and Pocket Companion for the United States ... Together with a Complete Railway Map*. New York,: C. Dinsmore & Co.

Colton, G. 1856a. Map of Iowa. J.H. Colton & Co., New York. URL http://www.davidrumsey.com/luna/servlet/s/kc46t2.

Colton, G. 1857. Map of Iowa. J.H. Colton & Co., New York. URL http://www.davidrumsey.com/luna/servlet/s/g0346t.

Colton, G. W. 1856b. Map of Missouri. (with) vicinity of St. Louis. J.H. Colton & Co, New York. URL http://www.davidrumsey.com/luna/servlet/s/l13y9g.

Colton, J. H. 1860. Colton's New Railroad & County Map of the United States and the Canadas &c. J. H. Colton, New York. URL http://hdl.loc.gov/loc.gmd/g3701p.rr000410.

Craig, L. A., R. Palmquist, and T. Weiss. 1998. Transportation improvements and land values in the Antebellum United States: A hedonic approach. *Journal of Real Estate Finance and Economics*, 16(2):173–189.

Disturnell, J. 1847. *Disturnell's guide through the middle, northern, and eastern states containing a description of the principal places, canal, railroad, and steamboat routes, tables of distances, etc.: Compiled from authentic sources*. J. Disturnell.

Doggett, J. 1848. *Doggett's Railroad Guide and Gazetteer for 1848 with Sectional Maps of the Great Routes of Travel*. New York: John Doggett Jr.

Engerman, S. L., R. E. Gallman, J. Atack, F. Bateman, and W. N. Parker. 2000a. Northern agriculture and the westward movement. In *The Cambridge Economic History of the United States: The Long Nineteenth Century*, volume 2, pp. 285–328. Cambridge, UK; New York, NY: Cambridge University Press.

Engerman, S. L., R. E. Gallman, and R. E. Gallman. 2000b. Economic growth and structural change in the long nineteenth century. In *The Cambridge Economic History of the United States*, volume 2. Cambridge, UK; New York: Cambridge University Press.

Ensign, Bridgman & Fanning. 1856. Railroad map of the United States, showing the depots & stations. URL http://hdl.loc.gov/loc.gmd/g3701p.

Fishlow, A. 1965. *American Railroads and the Transformation of the Antebellum Economy*. Cambridge, MA: Harvard University Press.

Gates, P. W. 1934. *The Illinois Central Railroad and Its Colonization Work*. Cambridge, MA: Harvard University Press.

Goodrich, C. 1961. *Canals and American Economic Development*. New York,: Columbia University Press.

Haney, L. H. 1908. *A Congressional History of Railways in the United States*. Madison, WI: Democrat Publishing Co.

Hunter, L. C. 1949. *Steamboats on the Western Rivers: An Economic and Technological History*. Cambridge, MA: Harvard University Press.

Lebergott, S. 1985. The demand for land: The United States, 1820–1860. *The Journal of Economic History*, 45(2):181–212. URL http://www.jstor.org/stable/2121686.

Levinson, D. 2008. Density and dispersion: The co-development of land use and rail in London. *Journal of Economic Geography*, 8(1):55–77. doi: 10.1093/jeg/lbm038.

Lloyd, E. 1857. *Lloyd's American Guide: Containing New Arranged Time Tables, So Simple and Correct that a Child Can Understand Them*. Philadelphia: E. Lloyd.

Meyer, B. D. 1995. Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13(2):151–161. ISSN 07350015. URL http://www.jstor.org/stable/1392369.

Mitchell, S. A. 1859a. A New Map of the State of Iowa. Charles Desilver, Philadelphia. URL http://www.

davidrumsey.com/luna/servlet/s/tpip6z.

Mitchell, S. A. 1859b. A New Map of the State of Missouri. Charles Desilver, Philadelphia. URL http://www.davidrumsey.com/luna/servlet/s/x11h77.

Morgan, D. L. 1969. *Rand McNally's Pioneer Atlas of the American West; Containing Facsimile Reproductions of Maps and Indexes from the 1876 First Edition*. Chicago: Rand McNally.

National Railway Publication Company. 1868. *Travelers' Official Railway Guide of the United States and Canada*. Ann Arbor, MI: National Railway Publication Co. University Microfilms.

Oberly, J. W. 1990. *Sixty Million Acres: American Veterans and the Public Lands before the Civil War*. Kent, OH: Kent State University Press.

Paxson, F. L. 1914. The railroads of the "Old Northwest" before the Civil War. *Transactions of the Wisconsin Academy of Sciences, Arts, and Letters*, 17(Part 1):247–274.

Poor, H. V. 1970. *History of the Railroads and Canals of the United States of America*. New York: A. M. Kelley.

Primack, M. L. 1962. Land clearing under nineteenth-century techniques: Some preliminary calculations. *The Journal of Economic History*, 22(4):484–497. URL http://www.jstor.org/stable/2116108.

Primack, M. L. 1977. *Farm Formed Capital in American Agriculture, 1850 to 1910*. New York: Arno Press.

Rand-McNally and Company. 1871. *The Rand-McNally Official Railway Guide and Hand Book*. Chicago: National Railway Publication Co.

Sage, J. 1858. J. Sage & Son's new & reliable rail road map comprising all the railroads of the United States and Canadas with their stations and distances, compiled from the most acurate statistics. J. Sage & Sons, Buffalo. URL http://hdl.loc.gov/loc.gmd/g3701p.rr000360.

Seaman, E. C. 1852. *Essays on the Progress of Nations, in Civilization, Productive Industry, Wealth, and Population*. New York: Scribner (Reprinted: Johnson Reprint Corp., 1967).

Stellingwerff, J. and R. P. Swierenga. 2004. *Iowa Letters: Dutch Immigrants on the American Frontier*. Grand Rapids, MI: Wiliam B. Eerdmans Co. No 47.

Swierenga, R. P. 1968. *Pioneers and Profits: Land Speculation on the Iowa Frontier*. Ames, IA: Iowa State University Press.

Taylor, G. R. 1951. *The Transportation Revolution 1815–1860*. Holt, Rinehart & Winston.

Taylor, G. R. and I. D. Neu. 1956. *The American Railroad Network, 1861–1890*. Cambridge, MA: Harvard University Press.

Thorndale, W. and W. Dollarhide. 1987. *Map Guide to the U.S. Federal Censuses, 1790–1920*. Baltimore: Genealogical Publishing Co.

United States Census Office and J. D. B. de Bow. 1854. *Statistical View of the United States*. Washington, D.C.: Beverly Tucker, Senate Printer.

United States Census Office and J. C. G. Kennedy. 1864. *Agriculture of the United States in 1860: Compiled from the Original Returns of the Eighth Census, under the Direction of the Secretary of the Interior*. Washington, D.C.: Government Printing Office (Reprinted by Norman Ross Publ., 1990). ISBN 0883544172.

United States Census Office, F. A. Walker, and C. W. Seaton. 1991. *Tenth Census of the United States, 1880*. New York, NY: Norman Ross Pub.

United States Congress. 1824. Laws of the United States, Vii, 239. Acts of 1823–24, Ch. 276.

United States Department of Transportation. 1998. North American Transportation Atlas Data. United States Department of Transportation, Bureau of Transportation Statistics. URL http://www.bts.gov/publications/north_american_transportation_atlas_data/.

Vanderbilt University Engineering Center for Transportation Operations and Research. 1999. National waterway network. U.S. Army Corps of Engineers Navigation Data Center.

von Thünen, J. H. and P. G. Hall. 1966. *Isolated State: An English Edition of Der Isolierte Staat*. Oxford and New York: Pergammon Press.

Whitney, W. D. and B. E. Smith. 1911. *The Century Dictionary and Cyclopedia, with a New Atlas of the World*. New York: The Century Co.

Williams, W. 1851. A new map of the United States. Upon which are delineated its vast works of internal communication. Lippincott, Grambo & Co., Philadelphia. URL http://hdl.loc.gov/loc.gmd/g3700.rr000180.

Wooldridge, J. 2007. Difference-in-differences estimation. Notes for Lecture delivered Tuesday, July 31st. National Bureau of Economic Research. URL www.nber.org/WNE/lect_10_diffindiffs.pdf.

Xie, F. and D. Levinson. 2010. How streetcars shaped suburbanization: A Granger-causality analysis of land use and transit in the Twin Cities. *Journal of Economic Geography*, 10(3):453–470. doi: 10.1093/jeg/lbp031.