Appendices for:

Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies

By Richard Hornbeck and Martin Rotemberg

A Data Appendix

A.1 County-Industry Manufacturing Data

We have digitized manufacturing data, by county and industry, for 1860, 1870, and 1880 from the original published tabulations of the Census of Manufactures (United States Census Bureau, 1860*b*, 1870, 1880). In 1860, the Census of Manufactures also collected information for enterprises outside of manufacturing (fisheries and mining) that we drop from our analysis for consistency.

The county-industry data report many industries in each decade, with some small variations, which we group together for our analysis. We homogenized industry names from each county to the list of industry names from US-industry tabulations in each decade: 331 names in 1880, 412 names in 1870, and 639 names in 1860. We then grouped these industries into 193 categories that were more consistent across decades, and further grouped these industries into 31 categories. Our estimates are not sensitive to these industry groupings (Table 2), but our goal was to balance industry-level details against statistical noise and to maintain comparability across decades and geographic areas.

Starting in 1870, the county-by-industry data do not list some "neighborhood industries" such as blacksmithing (Atack and Margo, 2019) or additional industries with less than $10,000 of revenue in total. We define a residual industry to capture the difference between county-level data and the summed county-by-industry data, and include this residual industry in our analysis. This residual "industry" includes less than 5% of manufacturing revenue in 1870 and 1880. For our county-industry results, the most relevant reason for a "residual" industry was that small producers of local products, such as many grist mills, were not included in the county-industry tabulations. We also created an "other" industry, representing less than 1% of revenue, reflecting named but small industries not otherwise classified.

These manufacturing data were collected by Census enumerators, who visited each manufacturing establishment to solicit responses. The Census then published aggregated statistics, including county-by-industry cells that contain only one manufacturing establishment (in 1860, 1870, 1880). For multi-industry establishments, such as grist & lumber mills, the Census would "[separate] the two parts of the business and [assign] each to its appropriate place in the Statistics of Industries" (United States Census Bureau, 1860*b*). We often refer to "firms" for convenience, though note that the Census enumeration is at the establishment level and activity is recorded where it takes place, not at headquarters, so this refers to single-establishment "firms."

The 1860 Census instructions to enumerators discuss the data collection guidelines in useful detail, which we quote below, and there is similar language in the instructions for other decades. Prior to 1850, there are greater concerns about the comprehensiveness of the data collection and the Census data collection was professionalized in 1850 (Atack and Bateman, 1999).

Our main variables of interest, from the manufacturing data, are:

Manufacturing Revenue (R). Total value of products, by county and industry from 1860, 1870, and 1880. These products were valued at the factory gate, excluding transportation costs to customers: "In stating the value of the products, the value of the articles *at the place of manufacture* is to be given, exclusive of the cost of transportation to any market" (emphasis original, United States Census Bureau, 1860*a*).

Manufacturing Materials Expenditure (E^M). Total value of materials, by county and industry from 1860, 1870, and 1880. These materials were valued at the factory gate, including transportation costs from suppliers: "this value is always to represent the cost of the article *at the place where it is used*" (emphasis original, United States Census Bureau, 1860*a*). Materials included fuel and "the articles used for the production of a manufacture," which the instructions noted might be manufactured by another establishment. Unused materials (on June 1) were to be excluded.

Manufacturing Labor Expenditure (E^L). Total amount paid in wages during the year, by county and industry from 1860, 1870, and 1880. Reported wages were intended to reflect total labor costs, including boarding costs paid in kind and the proprietor's own labor. From the Census instructions: "In all cases when the employer boards the hands, the usual charge of board is to be added to the wages, so that *cost of labor* is always to mean the amount paid, whether in money or partly in money and partly in board..." (emphasis original) and to be included was "the individual labor of a producer, working on his own account" (United States Census Bureau, 1860*a*). The measurement of labor costs raises some challenges, particularly in the treatment of owner-operator labor (Weeks, 1886), and Appendix B shows the robustness of our results to inflating measured labor costs to account for potential under-measurement of owners' labor.

Manufacturing Capital Expenditure (E^K). We impute annual capital expenditure by multiplying the reported total value of capital invested, in each county and industry (1860, 1870, 1880), by a state-specific mortgage interest rate that varies between 5.5% and 11.4%, with an average value of 8% (Fogel, 1964).[12] The establishment's capital value was directed to include "capital invested in real and personal estate in the business" (United States Census Bureau, 1860*a*). The measurement of capital is challenging, particularly in distinguishing

[12]The mortgage interest rates are similar to the antebellum returns to equity collected by Bodenhorn and Rockoff (1992). They are also around the implied interest rate currently used by the BLS to convert capital stocks to the flow value of capital services, when only considering assets that existed in the 19th century such as buildings, land, and steam equipment (Cunningham et al., 2021).

between nominal and resale values and potential non-reporting of rented land and equipment.[13] Appendix B reports that our estimates are not sensitive to alternative approaches to adjusting for measurement error in capital, in part because the annual cost of capital is substantially smaller than labor and materials expenditures and because the estimated percent impacts on capital expenditures are similar to the estimated percent impacts on labor and material expenditures.

Manufacturing Establishment Counts. The number of establishments in each county and industry (1860, 1870, 1880) with at least $500 in annual sales. The Census enumerators were instructed to survey every manufacturing establishment, except "household manufactures or small mechanical operations where the annual productions do not exceed five hundred dollars" (United States Census Bureau, 1860*a*). When multiple establishments were owned by the same party, and operated jointly, then Census enumerators were instructed to obtain separate details on the operations of each establishment. If this were impossible, particularly when one establishment manufactured the materials for the other establishment, then enumerators were instructed to "return the last manufacture, giving the raw materials for the first, and capital, fuel, and cost of labor, with the number of hands, in both" (United States Census Bureau, 1860*a*).

Civil War Related Industries. We coded two sets of industries as being "Civil War related." Our strict classification includes: artificial limbs and surgical appliances; awnings and tents; coffins; cutlery, edge tools, and axes; drugs; chemicals and medicines; explosives and fireworks; flags and banners; gun- and lock-smithing; gunpowder; lead; military goods; and ship and boat building. Our broad classification adds: bronze; canning and preserving; carriage and wagon materials; carriages and wagons; clothing (general); cooperage; gloves and mittens; and hats and caps.

A.2 Main Outcomes

The table below is a reference for the formulas used in calculating county productivity and its components. We use an upper bar to denote averages over the sample period. County-level values of revenue and input expenditures in each year reflect a sum of county-industry values in that year.

[13]Rented-in capital has only been irregularly collected in the modern Annual Survey of Manufactures, but Cunningham et al. 2021 report that it is a small share of total capital in years when measured.

Component	Formula	Notes
Revenue	R_{ct}	Gate value of revenue in the Census.
Capital	E_{ct}^{K}	Book value of capital in the Census, multiplied by interest rate.
Labor	E_{ct}^{L}	Wage bill in the Census.
Materials	E_{ct}^{M}	Gate value of materials in the Census.
s_{ct}^{k}	$\frac{E_{ct}^{k}}{R_{ct}}$	Revenue share of input k in county c in year t, with s_c^k representing the average across years.
α_{ct}^{k}	$\sum_i \frac{R_{cit}}{\sum_j R_{cjt}} \frac{\sum_c E_{cit}^{k}}{\sum_c \sum_\ell E_{cit}^{\ell}}$	County-level revenue share weighted sum of input's national industry cost share
ν_c	$\frac{1}{1-\left(\frac{1}{C}\sum_c \sum_k \overline{s_c^k}\right)}$	Used for re-scaling percent growth in county revenue into percent growth in county productivity.
Productivity	$\nu_c \left[R_{ct} - \sum_k s_c^k \ln E_{ct}^{k}\right]$	
TFPR	$\nu_c \left[R_{ct} - \sum_k \overline{\alpha_c^k} \ln E_{ct}^{k}\right]$	
Allocative Efficiency (AE)	$\nu_c \left[\left(\alpha_c^k - s_c^k\right) \ln E_{ct}^{k}\right]$	
Productivity Robustness: County Scalar	$\nu_c \left[P_{ct}Q_{ct} - \sum_k \overline{s_c^k} \ln W_{ct}^{k} X_{ct}^{k}\right]$	$\nu_c = \frac{1}{1-\left(\sum_k \overline{s_c^k}\right)}$ Drop counties with negative scalar values and top 1% of values
Productivity Robustness: Median Scalar	$\nu_c \left[P_{ct}Q_{ct} - \sum_k \widetilde{s_c^k} \ln W_{ct}^{k} X_{ct}^{k}\right]$	$\nu_c = \frac{1}{1-\left(\frac{1}{C}\sum_c \sum_k \widetilde{s^k}\right)}$ $\widetilde{s_c^k}$ is the median revenue share for k in county c, $\widetilde{s^k}$ is its national median
Productivity Robustness: 1860 Scalar	$\nu_c \left[P_{ct}Q_{ct} - \sum_k s_{c1860}^{k} \ln W_{ct}^{k} X_{ct}^{k}\right]$	$\nu_c = \frac{1}{1-\left(\frac{1}{C}\sum_c \sum_k s_{c1860}^{k}\right)}$ We also use α_{c1860}^{k} for decomposing into TFPR and AE

A.3 Other County-Level Data

For some specifications using manufacturing data from 1890 and 1900, when county-industry tabulations are unavailable, we use the corresponding county-level data (Haines, 2010). For 1850, the only values aggregated and published at the county level were manufacturing revenue and capital. Other county-level data are from the United States Census of Population and Census of Agriculture (Haines, 2010).

Population is defined as the reported total population in each county. In Appendix B, we inflate these population data due to potential undercounting in the Census that is estimated to vary by region and year: undercounting in the South by 7.6% in 1860, 8.8% in 1870, and 5.2% in 1880, and undercounting in the North by 5.6% in 1860, 6.0% in 1870, and 4.4% in 1880 (Hacker, 2013).

Agricultural land value is defined as the total value of land in farms, including the value of farm buildings and improvements. We follow Donaldson and Hornbeck (2016) in deflating these reported data, using Fogel's state-level estimates of the value of agricultural land only (Fogel, 1964, pp. 82-83).

We adjust county-level data to maintain consistent county definitions in each decade. We adjust data from each decade to reflect county boundaries in 1890 following the procedure outlined by Hornbeck (2010). Using historical United States county boundary files (from NHGIS), county borders in each decade are intersected with county borders in 1890. When counties in another decade fall within more than one 1890 county, data for each piece are calculated by multiplying that decade's county data by the share of its area in the 1890 county. For each other decade, each 1890 county is then assigned the sum of all pieces falling within its area. This procedure assumes that data are evenly distributed across county area, though for most counties in each decade there is little overlap with a second 1890 county. In three instances, we combine separately reported cities into a neighboring county for consistency: Baltimore City is combined into Baltimore County; St. Louis City is combined into St. Louis County; and Washington DC is combined into Montgomery County.

The regression sample is 1,802 counties that report county-industry manufacturing data in 1860, 1870, and 1880 (see Figure 2). The counterfactual sample is 2,722 counties with positive population and positive agricultural or manufacturing revenue in 1890 (see Figure 4).

B Results Appendix

B.1 Robustness: Regional Shocks

In this section, we explore whether counties experiencing relative growth in market access might otherwise have experienced relative growth in county productivity. We estimate similar pre-trends in counties prior to their relative growth in market access; control for time-varying effects of county characteristics in 1860, including counties' 1860 input wedges and input gaps; and adjust for potential differential effects of the Civil War. For example, controlling for the share of counties' 1860 revenue in each industry, interacted with year, adjusts for potential relative changes in industry output prices or other industry-specific shocks to differentially impact counties' growth.

An expanding national railroad network affected different counties' market access from 1860 to 1870 and from 1870 to 1880, and over each decade there were similar effects of market access on county productivity.[14] Splitting our baseline analysis by decade pair (1860 and 1870, 1870 and 1880, 1860 and 1880), increases in county market access lead to substantial increases in county productivity that are more often driven by increases in county AE than by increases in county TFPR (Appendix Table 12, rows 2, 3, 4). For the 1870 to 1880 period, there is more indication of productivity gains driven by TFPR growth. Pooling the post-Civil War period from 1870 to 1900, however, changes in county productivity are driven by county AE growth and in pooled models this difference for the 1870-1880 period is not statistically significant. Galiani, Jaramillo and Uribe-Castro (2022) find that manufacturing productivity growth in Canada from the opening of the Panama Canal was driven by AE growth rather than TFPR growth, and future analysis can further explore this tendency across time periods and places.

We also estimate little serial correlation in county market access, regressing changes in log market access from 1870 to 1880 on changes in log market access from 1860 to 1870. Controlling for state fixed effects and latitude/longitude the point estimate (on a one percent increase in market access from 1860 to 1870) is -0.02, with a standard error of 0.04. The estimate is 0.002 (0.04) when additionally controlling for contemporaneous and future growth in whether a county has any railroad and the length of its railroads.

We also also find that growth in county market access was not associated with differential pre-trends in county manufacturing activity. Estimating the effects of county market access and counties' future market access, controlling for contemporaneous railroads and future railroads, row 5 of Appendix Table 12 reports that contemporaneous county market access has significant effects on county productivity (driven by growth in county AE) but future

[14] These separate estimates also avoid potential issues with interpreting two-way fixed effects models with multiple time periods (De Chaisemartin and d'Haultfoeuille, 2020).

market access does not predict county productivity growth (i.e., these outcomes were changing similarly prior to growth in counties' market access).[15] From estimating this specification on an extended sample back to 1850, using the available county-level data on revenue from 1850, we also estimate no significant effect of future market access on manufacturing revenue growth from 1850 to 1860. Similarly, we estimate no significant effect on revenue growth (or productivity growth) from 1860 to 1870 from market access growth from 1870 to 1880.

Rows 6–17 control for county characteristics in 1860, interacted with year, to flexibly allow for those characteristics to have time-varying influences on county productivity. Rows 6 and 7 control for 1860 market access without railroads and actual 1860 market access. Rows 8 and 9 control for the share of counties' 1860 revenue in each industry, interacted with year, given the potential for relative changes in industry output prices and input prices or other industry-specific shocks to differentially impact counties' growth. Row 10 controls for county banking activity in 1860, interacted with year, including the presence of any bank and total bank deposits per capita. Rows 11 to 18 control for other county characteristics in 1860: counties' input-specific gaps in 1860 (the difference between the output elasticities and the revenue shares); the 1860 production function elasticities; 1860 input revenue shares; 1860 input wedges (the ratio of the output elasticities to the revenue shares); the 1860 HHIs of manufacturing industry revenue and employment; whether a county was on the "frontier" in 1860;[16] and jointly controlling for the gaps, elasticities, wedges, HHIs, and frontier status.[17]

As the Civil War occurred within our sample period, and had substantially different implications for different areas of the country, we explore the sensitivity of our results to adjusting for differential impacts of the Civil War and the abolition of slavery. First, counties that were initially concentrated in industries that produced more war-related goods may have changed differently over this period even in the absence of changes in the railroad network. For row 19, before calculating county productivity, we drop all industries strictly related to war production and in row 20 we drop industries more broadly related to war production (as defined in the Section A). Alternatively, in rows 21 and 22, we instead control for the 1860 share of revenue in war-related industries under each definition. These adjustments do not have large effects on the estimated coefficients. The Civil War itself may have had a

[15]This specification controls for contemporaneous and future values for whether a county has any railroad and the length of its railroads, and the estimates are similar with additional cubic polynomial controls for contemporaneous and future railroad length in the county and nearby areas.

[16]We follow the definition from Bazzi, Fiszbein and Gebresilasse (2020): counties with between two and six people per square mile in 1860 and that are within 100km of the boundary where population density fell below two people per square mile in 1860.

[17]The share of counties' 1860 revenue in each industry, the HHIs, and frontier status are moderately predictive of 1860 county gaps (a within R-squared of 0.15, after conditioning on state fixed effects and latitude/longitude, which mostly reflects the influence of the 31 industry shares).

direct effect on outcomes, and row 23 controls for whether a county had a Civil War battle, the number of battles (cubic polynomial), and the number of casualties (cubic polynomial), all interacted with year fixed effects.[18] Row 24 instead drops all counties with battles with over 500 casualties, while row 25 drops all counties with any noted battle. Row 26 drops counties on the border of the Union and the Confederacy. Row 27 drops Confederate states, row 28 includes only slave states, row 29 drops slave states, and row 30 drops the Southern region. The estimates are stable across these sample changes. When excluding areas from the regression sample, we continue to include them in the measurement of other counties' market access.

Our baseline empirical specification estimates the impacts of market access, controlling for county fixed effects and state-by-year fixed effects, such that the identifying variation is within-state relative changes in counties' market access. In rows 31 and 32, we report similar estimates when also controlling for region-by-year fixed effects (20 regions) or subregion-by-year fixed effects (106 subregions) that further restrict the analysis to relative changes in county market access within nearby economic regions that cut across state and county boundaries. We assign each county the share of its area in each region or subregion, and control for those shares interacted with year fixed effects.

Our baseline specification also adjusts for the general westward expansion of the United States, even within states, by controlling for year-interacted cubic polynomial functions of counties' latitude and longitude. As alternative functional forms, we find similar estimates when controlling for fifth-order polynomials (row 33) or linear functions of counties' latitude and longitude (row 34). Row 35 controls for state-specific linear functions of counties' latitude and longitude. We also estimate similar impacts of market access when excluding from our sample the Plains and West Coast regions of the United States in row 36. The West Coast sample states are California, Oregon, and Washington. The Plains sample states are Kansas, Nebraska, and Texas. Rows 37–39 omit the Northeast, East-North-Central, and West-North-Central regions.

Our main analysis adjusts the statistical inference for spatial correlation within states over time, reporting standard errors that are clustered by state. If we instead adjust for more gradual spatial correlation across counties, assuming that spatial correlation declines linearly up to an assumed distance cutoff and is zero thereafter (Conley, 1999; Hsiang, 2010; Baylis, 2020), we estimate smaller standard errors for our baseline specification. The standard errors are 5-7% lower for distance cutoffs of 200 miles or 300 miles, 12-19% lower for distance cutoffs between 400 miles and 700 miles, and 19-30% lower for distance cutoffs between 800 miles

[18]Our battle data includes all battlefields identified as significant by the Civil War Sites Advisory Commission's Report on the Nation's Civil War Battlefields.

and 1000 miles.

B.2 Robustness: Measurement of Productivity

This section discusses alternative methods of estimating county productivity and its decomposition into county TFPR and county AE. This includes adjustments for measurement error in inputs, alternative approaches for calculating production function elasticities, an analysis of home manufacturing in the Census of Agriculture, and the expansion of manufacturing activity into new counties.

One concern is that mismeasurement of inputs, particularly capital expenditures, may generate spurious measurement of misallocation (Hulten, 1991; Rotemberg and White, 2021). Also, when capital investment reflects forward-looking investment decisions, then apparent market distortions can reflect dynamically efficient input decisions (Solomon, 1970; Fisher and McGowan, 1983; Fisher, 1987; Caplin and Leahy, 2010; Asker, Collard-Wexler and De Loecker, 2014). We estimate that growth in county allocative efficiency is predominately driven by growth in materials inputs (Table 3), and capital is a small share of total input expenditures and responds similarly to other input expenditures, such that our estimates are generally not sensitive to the measurement of capital inputs.

Appendix Table 13, row 2, reports similar impacts of county market access on county productivity, and its decomposition into county TFPR growth and county AE growth, when assuming zero misallocation in capital (such that the county's capital revenue share is equal to its cost share). Row 3 reports similar estimates when assuming that capital misallocation is equal to materials misallocation (such that the county's capital revenue share is adjusted so the ratio of its cost share to revenue share is equal to the ratio of that county's materials cost share to materials revenue share). In row 4, we additionally replace the labor wedge with the materials wedge and estimate similar percent effects of market access.

A related concern is that capital may be systematically under-measured (United States Census Bureau, 1880); indeed, Appendix Table 1 shows that national manufacturing capital expenditures are 15% of total labor costs and capital costs, which is below typical values of roughly one-third. In Appendix Table 13, rows 5 and 6, we double and triple the baseline measured values of capital and find similar effects of market access on productivity and allocative efficiency. This adjustment also addresses concerns that the flow rate of capital services should be larger than what the mortgage rate would imply. Estimated depreciation rates for equipment in this historical era are around 6% (Davis and Gallman, 1968), which is at the lower end of modern values estimated by the BEA because various high-depreciation capital inputs did not exist at the time (e.g., internal combustion engines and computers). Row 7 reports estimates when imputing annual capital expenditures using an average national interest rate of 8% instead of state-specific interest rates (from Fogel, 1964).

Given that the measurement of reallocation gains can be sensitive to the upper and lower tails (Rotemberg and White, 2021), in rows 8–13 we test the sensitivity of our results to lowering dispersion in input distortions. In rows 8–10, we shrink dispersion in the capital wedge by 5 percent, 10 percent, and 25 percent. To do so, we replace the observed wedge for capital with the weighted average of the observed capital wedge and its national median, where the weights on the county's observed values are respectively 95 percent, 90 percent, and 75 percent. We then impute consistent values for capital (so the relationship between imputed capital and observed revenue is consistent with the imputed wedge). While this adjustment mechanically lowers the potential gains from reallocation in exercises like those in Hsieh and Klenow (2009), our regression estimates are stable. In rows 11–13, we shrink dispersion in all of the input wedges, which also has little effect on our regression estimates. For the aggregate counterfactual estimates, replacing the observed wedges with their shrunken counterparts affects the welfare losses from removing the railroads by less than 1 percentage point.

Our main estimates hold fixed counties' revenue shares and production function elasticities (and therefore the wedges), using the observed average from 1860 to 1880 (as in Petrin and Levinsohn 2012's Törnqvist-Divisia approximation). In row 14, we instead hold these values fixed at their 1860 levels (and correspondingly use those values to calculate the scaling factor). In row 15, in addition to using these 1860 values in the measurement of county productivity in each decade, based on county revenue and county input expenditures in each decade, we hold county populations fixed at their 1860 values for calculating market access (as in Appendix Table 14, row 23).

In rows 16 and 17, we use alternative scaling factors in the definition of county productivity. The percent impact of market access on county productivity reflects a scaled percent impact of market access on revenue and input expenditures, as discussed in Section V.D, and our baseline estimates define this scaling factor as the average ratio of county revenue to county productivity over the 1860-1880 period. Alternatively, in row 16, we define this scaling factor based on the median ratio of county revenue to county productivity. In row 17, we use county-specific scaling factors instead of the national average (dropping counties with negative values and the top 1% of values, as these scaling factors become undefined as productivity approaches zero). Rather than using time invariant shares and scaling factor, we estimate a similar 23.8% increase in the log of county revenue minus total county input expenditures in each decade.

As an alternative approach to dealing with extreme values, we show similar impacts of market access when excluding sample counties with the largest and smallest changes in productivity from 1860 to 1880: row 18 excludes the top and bottom 1% of counties, and row 19 excludes the top and bottom 5% of counties. This latter specification suggests more county productivity growth through increases in county TFPR, though excluding values based on changes in the outcome variable is prone to introduce biases.

Labor inputs may also be under-reported in the Census of Manufactures, which would over-state establishments' productivity. Census enumerators ask establishments for their labor costs and these were intended to include in-kind boarding costs and labor supplied by establishment owners working on their own account, but there has been debate about whether establishment owner labor is fully reflected in these costs (Weeks, 1886; Atack, 1977; Sokoloff, 1984; Margo, 2014). Row 20 reports similar impacts of county market access, though, when inflating labor costs to reflect potentially omitted labor costs. For this specification, we add to county-by-industry labor costs the number of establishments multiplied by the average wage in that county and industry.

Our measurement of county productivity does not depend on an assumed production function, but we do assume Cobb-Douglas production with constant returns to scale for our decomposition of county productivity into TFPR and AE. Lafortune et al. (2021) estimate returns to scale around 0.95 for the late 19th century, which is at the lower end of scale estimates. David (1969) emphasizes the potential for increasing returns to scale in his critique of Fogel (1964), which Fogel (1979) disputes but considers scale coefficients around 1.05 to 1.10.

In Appendix Table 12, row 21, we assume decreasing returns to scale and re-scale the production function elasticities to add up to 0.95. In row 22, we correspondingly assume the returns to scale are 1.05. Assuming decreasing returns to scale increases the impact on county productivity through TFPR growth (column 2) and reduces the impact through AE growth (column 3), while mechanically leaving unchanged the impact on county productivity (column 1). Our baseline estimates could understate TFPR growth, which is exactly offset by sufficiently strong decreasing returns to scale, but this would also need to then be true across industry groups (Table 4).

We also show the sensitivity of our results to alternative methods for calculating the production function elasticities. These adjustments have no effect on the measurement of county productivity (in column 1) and, in practice, have little substantive effect on the estimated impacts through county TFPR (column 2) and county AE (column 3) because market access had similar percent effects on each input. In row 23, we calculate county-level elasticities averaging industry-level cost shares with weights equal to an industry's share of total expenditure in that county, rather than an industry's share of total revenue in that county (which could over-weight the influence of high-markup industries). In row 24, we calculate production function elasticities only using the "most efficient" counties for calculating each industry's cost shares (specifically, those counties whose total input expenditure gap is within one standard deviation of zero). In row 25, instead of using national values for calculating industry cost shares, we leave out each specific county-industry when calculating its production function elasticity. In row 26, we instead use only other counties located in the same state to calculate industry-level cost shares, and in row 27 we leave out that county in the calculation of state-level industry cost shares. In row 28, we use each observed county-industry cost share as our measure of production function elasticities, which imposes a constant wedge across inputs within the county.

If input frictions are consistently higher for one input than the others, then the measured cost shares will understate the output elasticity for that input relative to the others. To assess the sensitivity of our estimates, row 29 reports estimates when increasing the cost share for labor by 5 percentage points and proportionally decreasing the cost shares for materials and capital. Similarly, we increase the cost share for materials by 5 percentage points (in row 30) or for capital by 5 percentage points (in row 31).

We measure manufacturing productivity using data from the Census of Manufactures, though there may be additional manufacturing activity not included in the Census of Manufactures. The Census of Manufactures reports a butter and cheese industry with only two establishments in 1860 and $13 thousand of output, but reports 1195 establishments in 1870 with $16.5 million of output.[19] If we exclude the butter and cheese industry in each decade, though, we estimate similar impacts of market access (row 32).

The Census of Manufactures would potentially have missed some manufacturing establishments, perhaps smaller establishments (United States Census Bureau, 1870). To the extent that the Census coverage varies over time and geographic areas, this would be partly corrected for by our inclusion of state-by-year fixed effects and year-interacted controls for county latitude and longitude. The remaining concern is that changes in county market access might be systematically correlated with changes in Census data coverage. While the general concern is that smaller establishments are more likely to be missed by Census enumerators, we report in Table 5 that changes in market access are not associated with changes in average establishment size.

We can use data from the Census of Agriculture on the value of home manufactures to expand our analysis of manufacturing beyond the Census of Manufactures. For this analysis, we assume that home manufactures are not already included in the Census of

[19] In 1880, the Census of Manufactures reports 3250 establishments with $30.4 million of output. There were also large technological changes in dairy manufacturing during the time period (Boberg-Fazlic and Sharp, 2020). The Census of Agriculture reports quantities of butter and cheese produced, but does not report data on their values or associated inputs.

Manufactures. Spot checks of the Census of Agriculture manuscripts show the values of home manufactures tend to be substantially less than the $500 threshold used by the Census of Manufactures. The national value of home manufactures is only 1.2% of total manufacturing revenue in 1860 and 0.4% of total manufacturing revenue in 1870, summing output values from the Census of Agriculture and the Census of Manufacturing for our sample counties, though home manufactures are substantively important in some counties. Among our sample counties, the median revenue shares of home manufacturing are 3.6% in 1860 and 1.0% in 1870 and the average revenue shares of home manufacturing are 11.8% in 1860 and 6.5% in 1870. The Census of Agriculture stopped asking about home manufactures in 1880, and we assume for our analysis here that there was zero home manufacturing in 1880. The estimated impact of county market access on the value of manufacturing output is 0.192 (0.049), using only data from the Census of Manufactures, and the effect is 0.160 (0.043) when adding the value of home manufactures to data from the Census of Manufactures. Data on home manufactures also allow us to expand the balanced sample of counties to include 149 additional counties that do not report manufacturing revenue in at least one decade of the Census of Manufactures, but do report home manufactures in that decade. The estimated impact of county market access on county manufacturing revenue increases to 0.270 (0.041) when including these 149 additional counties and adding the value of home manufactures to data from the Census of Manufactures.

Indeed, the United States itself expanded substantially from 1860 to 1880. Our baseline estimates focus on a balanced sample of counties from 1860 to 1880, which includes 91% of population and 99% of manufacturing revenue in 1880. When focusing on this balanced panel, however, the analysis does not include impacts on the extensive margin of manufacturing growth in newly created counties. Over this period, from 1860 to 1880, we estimate that a one standard deviation increase in market access leads to a 4 percentage point increase in the probability that a county reports any manufacturing activity in the Census of Manufactures. We cannot estimate what happened to manufacturing productivity in these counties, which is not measured in the earlier periods, but increases in county market access are leading to growth on the extensive margin along with our estimated productivity effects on the intensive margin.

B.3 Robustness: Measurement of Market Access

In this section, we explore how the estimated effects of county market access on county productivity (and TFPR and AE) depend on the measurement of county market access. This includes: calculating county-to-county transportation costs under alternative assumptions; using different values for the trade elasticity θ or the average price per ton of transported goods $\overline{P}$; and adjusting for the influence of international trade and mismeasurement of population in the 1870 Census.

In Appendix Table 14, rows 2 and 3, we show that the estimated impacts on county productivity are not sensitive to omitting counties with the largest and smallest changes in market access. The estimated impact on county productivity through TFPR growth becomes moderately larger when omitting more counties (in row 3), but the estimates are not statistically different.

Measured changes in county market access reflect changes in county-to-county transportation costs, which are based on assumed rates for transporting goods using railroads, waterways, and wagons. Rows 4 and 5 report similar impacts of market access on county productivity when decreasing the costs of waterways and wagons to the lowest values considered by Fogel (1964). Row 6 reports similar estimates when removing the costs of transshipment within the waterway network. Rows 7–9 report similar estimates when increasing the cost of railroad transportation to reflect potential congestion, fragmented track ownership, or differences in gauges that would require transshipment of goods or more indirect routes within the railroad network (see, e.g., Gross (2020) on North-South gauge differences).

In calculating counties' market access, the "iceberg trade costs" (τ_{od}) reflect the measured county-to-county transportation costs (t_{od}) scaled by the average price per ton of transported goods ($\tau_{od} = 1+t_{od}/\overline{P}$). Our baseline estimates use an average price of 38.7 that we estimate, but rows 10 and 11 report similar estimated effects of market access if we instead use 20 or 50 dollars per ton. We continue to focus on the estimated impact of a one standard deviation greater increase in county market access, as market access itself is re-scaled when changing the transportation cost parameters.

The impact of market access is also similar when using alternative values of θ, the trade elasticity. Smaller values of θ compress the distribution of changes in market access, just as for larger average prices, but increase the effect of gaining access such that there remains a similar effect from a one standard deviation greater increase in market access. Rows 12 to 14 show similar results using the extremes of the 95% confidence interval around our baseline estimate of θ (1.95 to 3.925) and a larger value of θ (8.22) from Donaldson and Hornbeck (2016).

Our baseline measure of county market access reflects counties' access to all other counties' population, though we estimate similar impacts from a one standard deviation increase in modified definitions of county market access. To incorporate the influence of access to international markets, we inflate the population in 11 counties with major international ports to reflect the value of imports and exports in each year divided by GDP per capita (row 15). The Census of Population is known to have undercounted population, especially in the

South in 1870, and so in row 16 we inflate counties' population by year and region based on the estimated degree of undercounting by Hacker (2013). Our baseline measure of market access considers counties' access to population, but we also replace counties' population with counties' wealth as an alternative proxy for counties' market size (row 17). The estimates are also not sensitive to including a county's own population in its market access (row 18), which we omit in our baseline measure to avoid regressing county manufacturing activity on its own population (along with other counties' population).

Rows 19–22 report similar estimates when restricting the measurement of county market access to include only access to counties beyond 5 miles, 50 miles, 100 miles, or 200 miles, such that changes in county market access only reflect more-distant economic influences. One goal of excluding nearby counties in this "donut specification" is to adjust for spatially correlated growth, whereby a regional shock would increase economic activity in county A and its nearby counties and also increase county A's market access (which is a function of nearby counties' population). Omitting nearby counties from market access does not matter much in our context, though, because the spatial distribution of economic activity in the US is sufficiently concentrated at this time that county A's market access is generally more dependent on its cost of getting to major cities through the railroad and waterway network than the cost of getting to nearby counties.

Row 23 reports similar estimates when calculating market access in each period holding county populations fixed at 1860 levels, such that changes in county market access are only due to changes in county-to-county transportation costs. Our preferred specifications use actual market access, rather than measuring market access with fixed populations, because railroads also potentially affect market access through changes in the population distribution. Actual market access is highly correlated with population-fixed market access, however, and so our estimates are effectively unchanged whether we use actual market-access, market access with fixed populations, or if we instrument the former with the latter.

C Theory Appendix

In this section, we provide some additional details on the model from Section V. These details relate to deriving the log-linear relationship between market access and productivity, and our estimation of counterfactuals. To more clearly express input prices, rather than only denoting input k as costing W_o^k in county o, we also refer to the labor wage w_o, capital interest rate r_o, land rental rate q_o, and materials price index W_o^M.

C.1 Market Access and Productivity

As described in equation 8, trade flows follow a gravity equation:

$$\text{Exports}_{od} = \kappa_1 A_o \left(\prod_k \left(\left(1 + \psi_o^k\right) W_o^k \right)^{\alpha_o^k} \right)^{-\theta} \tau_{od}^{-\theta} Y_d P_d^{\theta}. \tag{13}$$

Consumer market access (CMA) in county d is an inverse transformation of the goods price index (Redding and Venables, 2004; Donaldson and Hornbeck, 2016):

$$CMA_d = P_d^{-\theta} = \kappa_1 \sum_o \tau_{od}^{-\theta} A_o \left(\prod_k \left(\left(1 + \psi_o^k\right) W_o^k \right)^{\alpha_o^k} \right)^{-\theta}. \tag{14}$$

Consumer market access is higher in county d when it has access to cheaper goods: when there are lower costs of transporting goods from counties with higher technical efficiency and lower "effective costs." Input frictions in county o lower consumer market access in county d because county d is not able to fully benefit from low marginal costs in county o.

Firm market access (FMA) in county o is a sum over firms' access to all destination counties, adjusting for those destination counties' access to other sources of goods:

$$FMA_o = \sum_d \tau_{od}^{-\theta} Y_d CMA_d^{-1}. \tag{15}$$

Firm market access is higher in county o when it has access to more product demand: when there are lower costs of transporting goods to counties with higher consumption, which have less access to other sources of goods (CMA_d). We can also represent consumer market access in county d as a sum over consumers' access to all origin counties:

$$CMA_d = \sum_o \tau_{od}^{-\theta} Y_o FMA_o^{-1}. \tag{16}$$

Similar to equation 15, consumer market access is higher in county d when it has access to more product supply: when there are lower costs of transporting goods from counties with higher production, which have less access to other destinations for goods (FMA_o). Indeed, equations 15 and 16 imply that a county's firm market access and consumer market access are exactly proportional: $FMA_o = \rho CMA_o$, where $\rho > 0$. This result depends on symmetric trade costs ($\tau_{od} = \tau_{do}$), as constructed in Section I.B. We therefore use a single measure of "market access" (MA), which reflects the ideas underlying both firm market access and consumer market access: $MA_o \equiv FMA_o = \rho CMA_o$. Given that workers receive a fixed share of revenue $\left(\frac{\alpha_d^L}{\left(1+\psi_d^L\right)}\right)$, in equation 9 we express market access in county o as a function of market access in all other counties d.

Goods markets clear in general equilibrium, where demand in each county is equal to supply. Production in each county is then equal to the sum of exports to all destinations (including itself). Summing equation 13 over all counties and taking logs gives:[20]

$$\text{(17)} \qquad \ln\left(Y_o\right) = \kappa_1 + \varkappa_{1o} + \left(\alpha_o^M + \alpha_o^L\right)\ln\left({W_o^M}^{-\theta}\right) - \theta\alpha_o^T \ln q_o + \ln\left(\sum_d \left(\frac{P_d}{\tau_{od}}\right)^\theta Y_d\right).$$

Replacing $q_o = \frac{\alpha_o^T Y_o}{T_o}$, plugging in $FMA_o = \rho(W_o^M)^{-\theta}$, and combining terms gives:[21]

$$\text{(18)} \qquad \ln Y_o = \kappa_1 + \varkappa_{2o} + \left(\frac{\alpha_o^M + \alpha_o^L + 1}{1+\theta\alpha_o^T}\right)\ln\left({W_o^M}^{-\theta}\right).$$

Once we have solved for output in each county, input quantities and input expenditures in each county follow directly from our assumption that within-county revenue shares are constant.

For estimating changes in productivity, what matters is changes in real output and real inputs. Because markups and other distortions are constant in this environment, we can convert from nominal output to real output using the changes in marginal costs. The price of capital is independent of market access. The price of land (q_o) is endogenous to market access:

$$\text{(19)} \qquad \frac{d\ln q_o}{d\ln MA_o} = \frac{\alpha_o^M + \alpha_o^L + 1}{1+\theta\alpha_o^T}.$$

The local prices for labor (w_o) and materials (W_o^M) are log-linear in market access, as

[20] $\kappa_1 = \left(-\frac{\theta}{1-\sigma}\right)\ln\left(\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right)$ and $\varkappa_{1o} = \ln\left(A_o\right) - \theta\alpha_o^L\ln\left(\left(1+\psi_o^L\right)\bar{U}\right) - \theta\alpha_o^K\ln\left(\left(1+\psi_o^K\right)r\right) - \theta\alpha_o^M\ln\left(1+\psi_o^M\right)$

[21] $\varkappa_{2o} = \frac{\varkappa_{1o}+\ln\rho-\theta\alpha_o^T\ln\frac{\alpha_o^T}{T_o}}{1+\theta\alpha_o^T}$

described by equation 14:

$$\frac{d \ln w_o}{d \ln MA_o} = \frac{d \ln W_o^M}{d \ln MA_o} = -\frac{1}{\theta}. \tag{20}$$

We can give more structure to the impact of market access on productivity by substituting into equation 10 the impacts of log market access on log real inputs, and converting from the gap to the wedge:

$$\begin{aligned} \frac{d \ln Pr_o}{d \ln MA_o} = \quad & \frac{R_o}{Pr_o} \left[\left(\alpha_o^L - s_o^L\right) \left(\frac{1}{\theta} + \frac{\alpha_o^M + \alpha_o^L + 1}{1 + \theta \alpha_o^T} \right) \right. && \text{(Labor)} \\ & + \left(\alpha_o^M - s_o^M\right) \left(\frac{1}{\theta} + \frac{\alpha_o^M + \alpha_o^L + 1}{1 + \theta \alpha_o^T} \right) && \text{(Materials)} \\ & \left. + \left(\alpha_o^K - s_o^K\right) \left(\frac{\alpha_o^M + \alpha_o^L + 1}{1 + \theta \alpha_o^T} \right) \right] && \text{(Capital)}, \end{aligned} \tag{21}$$

where the final term in parenthesis in each line corresponds to the elasticity of the corresponding input to market access.

C.2 Market Access and County Productivity

In this section, we describe the decomposition of Equation 2 in more detail. For considering why log county productivity increases with log county market access, it is useful to re-write the impact of log market access on log productivity in county c as a function of the impacts of log market access on log revenue (R_c) and log expenditures on k inputs (E_c^k):

$$\begin{aligned} \frac{\partial \ln Pr_c}{\partial \ln MA_c} &\equiv \frac{\partial \ln (R_c - \sum_k E_c^k)}{\partial \ln MA_c} \\ &= \frac{1}{Pr_c} \left[\frac{\partial R_c}{\partial \ln MA_c} - \sum_k \frac{\partial E_c^k}{\partial \ln MA_c} \right] \\ &= \frac{1}{Pr_c} \left[R_c \frac{\partial \ln R_c}{\partial \ln MA_c} - \sum_k E_c^k \frac{\partial \ln E_c^k}{\partial \ln MA_c} \right] \end{aligned} \tag{22}$$

$$\frac{\partial \ln Pr_c}{\partial \ln MA_c} = \frac{R_c}{Pr_c} \left[\frac{\partial \ln R_c}{\partial \ln MA_c} - \sum_k s_c^k \frac{\partial \ln E_c^k}{\partial \ln MA_c} \right], \tag{23}$$

where $s_c^k = \frac{E_c^k}{R_c}$ or the revenue share of input k.

Equation 23 can be further decomposed using estimates of production function elasticities (α_c^k). We add and subtract the growth in "expected output" caused by the changes in input expenditures from changes in market access: the sum over the growth rate of each input multiplied by its respective output elasticity (α_c^k). Rearranging terms gives Equation 3.

C.3 Counterfactual Estimation and Uniqueness

Given our model, we rationalize the observed distribution of population by estimating each county's technical efficiency (A_o) and quantity of fixed factors (T_o), as well as the national utility $\bar{U}$. As in Donaldson and Hornbeck (2016), we do not separately identify A_o, T_o, and $\bar{U}$, but only their combined value is needed for estimating the counterfactuals. In particular, we describe $A_i T_i^{\theta\alpha_i^T} \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)}$ as a measure of local "amenities," which combines both information on how productive a place is $(A_i T_i^{\theta\alpha_i^T})$ as well as a function of national utility $\bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)}$. In this section, we describe how we jointly estimate amenities, θ (the trade elasticity), and $\overline{P}$ (the average price per ton).

First, suppose that we have values for θ and $\overline{P}$. Prices in any location can be expressed as a function of prices and population in all locations:

$$P_d^{-\theta} = \sum_o \frac{\tau_{od}^{-\theta} P_o^{\frac{\left(1+\psi_o^L\right)L_o}{\alpha_o^L}}}{\sum_i \tau_{oi}^{-\theta} P_i^{1+\theta\frac{\left(1+\psi_i^L\right)L_i}{\alpha_i^L}}}. \tag{24}$$

This equation matches equation 15 in the appendix to Donaldson and Hornbeck (2016), with two changes: we allow for market distortions $\left(\psi_o^L\right)$, and we allow variation across counties in the production function elasticity for labor $\left(\alpha_o^L\right)$. As a result, many of the steps in our derivation match those in Donaldson and Hornbeck (2016), which in turn rely on results from Allen and Arkolakis (2014). Thus, we focus on describing where our new assumptions require a new approach. For example, the first step is the same: there is a steady state solution for prices that can be identified using the Fujimoto-Krause algorithm. Due to the structure of the equation, the solution is only unique up to proportionality. Normalizing the model prices so that $P_{\text{New York City}}^{-\theta}$ is equal to 1, we define γ as the constant of proportionality between the model-implied prices and the actual (unobserved) nominal prices in the data. That is,

$$\begin{aligned}
(\gamma^{-\theta})P_d^{-\theta} &= \sum_o \frac{\tau_{od}^{-\theta}(\gamma) P_o^{\frac{(1+\psi_{lo})L_o}{\alpha_{lo}}}}{\sum_i \tau_{oi}^{-\theta}(\gamma^{1+\theta}) P_i^{1+\theta\frac{(1+\psi_{li})L_i}{\alpha_{li}}}} \\
P_d^{-\theta} &= \sum_o \frac{\tau_{od}^{-\theta}(\gamma^{1+\theta}) P_o^{\frac{(1+\psi_{lo})L_o}{\alpha_{lo}}}}{\sum_i \tau_{oi}^{-\theta}(\gamma^{1+\theta}) P_i^{1+\theta\frac{(1+\psi_{li})L_i}{\alpha_{li}}}} \\
P_d^{-\theta} &= \sum_o \frac{\tau_{od}^{-\theta} P_o^{\frac{(1+\psi_{lo})L_o}{\alpha_{lo}}}}{\sum_i \tau_{oi}^{-\theta} P_i^{1+\theta\frac{(1+\psi_{li})L_i}{\alpha_{li}}}}.
\end{aligned}$$

Note that conditional on the allocation of labor, we can solve for relative prices without solving for amenities. However, the parameter γ is unknown, and for that we do need to solve for amenities, as well as θ and $\overline{P}$.

We first describe how we solve for amenities and γ assuming that we have already solved for θ and $\overline{P}$. We then describe how we solve for θ and $\overline{P}$ given γ.

We can rewrite equation 8 as

(25)

$$A_o T_o^{\theta\alpha_o^T} \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)} = \frac{P_o^{1+\theta\left(\alpha_o^L+\alpha_o^T+\alpha_o^M\right)} \left(\frac{(1+\psi_o^L)L_o}{\alpha_o^L}\right)^{1+\theta\alpha_o^T} \left(\sum_i \tau_{oi}^{-\theta} P_i^{1+\theta} \frac{(1+\psi_i^L)L_i}{\alpha_i^L}\right)^{-1}}{r_o^{-\theta\alpha_o^K} \alpha_o^{T-\theta\alpha_o^T} \left[\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right]^{-\frac{\theta}{1-\sigma}} (1+\psi_o^L)^{-\theta\alpha_o^L} (1+\psi_o^K)^{-\theta\alpha_o^K} (1+\psi_o^M)^{-\theta\alpha_o^M}}.$$

Amenities are invariant to the price scaling parameter γ, and so are determined only as a function of relative prices (conditional on θ and $\overline{P}$). This invariance result is due to our assumption that nominal interest rates are proportional to the New York City price index.

We now turn to solving for γ. To condense notation, define C_i such that:

(26)

$$C_o \equiv A_o T_o^{\theta\alpha_o^T} \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)} r_o^{-\theta\alpha_o^K} \alpha_o^{T-\theta\alpha_o^T} \left[\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right]^{-\frac{\theta}{1-\sigma}} \left(1+\psi_o^L\right)^{-\theta\alpha_o^L} \left(1+\psi_o^K\right)^{-\theta\alpha_o^K} \left(1+\psi_o^M\right)^{-\theta\alpha_o^M}.$$

Unlike amenities, C_i is a function of γ through the interest rate. The total value of exports from county o to county d can be rewritten as:

$$\text{Exports}_{od} = C_o \bar{U}^{\theta\left(\alpha_o^L+\alpha_o^T\right)} \left(\frac{(1+\psi_o^l)L_o}{\alpha_o^l}\right)^{-\theta\alpha_o^t} P_o^{-\theta(\alpha_o^l+\alpha_o^t+\alpha_o^m)} \tau_{od}^{-\theta} Y_d P_d^{\theta} \tag{27}$$

Leveraging Cobb-Douglas to set $Y_o = \frac{\bar{U} P_o (1+\psi_o^l) L_o}{\alpha_o^l}$, and rewriting prices in terms of γ:

$$\gamma^{-\theta\alpha_o^k} C_o \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)} \left(\frac{(1+\psi_o^l)L_o}{\alpha_o^l}\right)^{-\theta\alpha_o^t} (\gamma^{-\theta(\alpha_o^l+\alpha_o^t+\alpha_o^m)}) P_o^{-\theta(\alpha_o^l+\alpha_o^t+\alpha_o^m)} \tau_{od}^{-\theta} \frac{\bar{U}(1+\psi_o^l)L_o}{\alpha_o^l} (\gamma^{1+\theta}) P_d^{1+\theta}$$

$$= \gamma^{-\theta\alpha_o^k} C_o \left(\frac{(1+\psi_o^l)L_o}{\alpha_o^l}\right)^{-\theta\alpha_o^t} P_o^{-\theta(\alpha_o^l+\alpha_o^t+\alpha_o^m)} \tau_{od}^{-\theta} \frac{\bar{U}(1+\psi_o^l)L_o}{\alpha_o^l} P_d^{1+\theta} * \gamma^{1+\theta\alpha_o^k}$$

$$= \gamma \text{Exports}_{od}$$

Just as the model only solves for relative prices, it also solves for relative nominal trade flows: multiplying all of the prices by γ correspondingly increases the model dollar trade flows by a factor of γ without affecting quantities.

Summing across destination counties gives county level revenue: $\gamma Y_o = \sum_d \gamma \text{Exports}_{od}$.

We then solve for the γ that minimizes the mean absolute error between model predicted revenue γY_o and the data for county output $\hat{Y}_o$, given the assumed θ and $\overline{P}$:

$$\gamma\left(\theta, \overline{P}\right) = \arg\min_{\gamma} \sum_{o} \left|\gamma Y_o - \hat{Y}_o\right| \tag{28}$$

The estimated $\gamma\left(\theta, \overline{P}\right)$ allows us to match the pattern of nominal output across counties, given a value of the price per ton and the dispersion of productivity.

Having solved for amenities and γ as a function of θ and $\overline{P}$, we now turn to estimation of θ and $\overline{P}$ (as a function of γ). For any given $\overline{P}$ and θ, we can solve equation 9 to calculate market access in each decade. As a result, we can generate model-predicted values for the change in land values for each county over time, using Equation 19:

$$\ln(W_o^t) = \varkappa_o + \left(\frac{\alpha_o^m + \alpha_o^l + 1}{1 + \theta\alpha_o^t}\right) \ln(MA_o(\theta)) \tag{29}$$

For any $\overline{P}$, there is a corresponding θ that minimizes the residual sum of squared differences between actual and predicted land value changes from 1860-1900.[22] Our use of land values draws on the assumption that land markets are integrated across sectors within counties. We also use land values for this estimation of θ, as in Donaldson and Hornbeck (2016), because land values capture the net present value of market access and respond more immediately than population. Appendix Figure 6 shows the fit for alternative values of θ (for the optimal price, $\overline{P} = 38.7$). For the optimal value of θ, 3.05, we estimate that a one-standard deviation greater increase in market access increases land values by 28.6% with a standard error of 3.7%.

We now have for every $\overline{P}$, a corresponding θ and therefore a $\gamma\left(\theta\left(\overline{P}\right), \overline{P}\right)$. Therefore, we also have a model implied value for nominal shipments for every county pair, $\text{Exports}_{od}(\theta\left(\overline{P}\right), \overline{P})$. For every county pair od, the trade costs described in section I.B allow us to see if the least cost path from o to d uses the railroad network or not. The use of a railroad on a given path is determined within the network database, and is independent of other parameters in the model. As a result, we can use the model to generate a value for

[22] Adding input frictions to the model does not affect the impact of market access on county land value, which is the main estimated impact in Donaldson and Hornbeck (2016). We obtain the same predicted impact of market access on land value as Donaldson and Hornbeck (2016), when replacing our county-specific sum of the labor share and materials share ($\alpha_o^M + \alpha_o^L$) with their average labor share of value-added (α^L). As in Donaldson and Hornbeck (2016), we weight by initial land values.

total predicted railroad shipments for any price per ton $\overline{P}$:

(30)

$$\text{Shipments}\left(\overline{P}\right) = \sum_{od} \gamma(\theta\left(\overline{P}\right), \overline{P}) * \text{Exports}_{od}(\theta\left(\overline{P}\right), \overline{P}) * \mathbb{1}\{\tau_{od} \text{ uses the railroad network}\}.$$

We pick the $\overline{P}$ that most closely matches the actual value of railroad shipments from Adams (1895). Appendix Figure 7 shows the fit for a range of prices, given the optimal θ.

To estimate counterfactuals, we hold fixed θ, $\overline{P}$, and $A_i T_i^{\theta\alpha_i^T}$. We vary the transportation network, which potentially will endogenously affect national utility, national population, or both. We now turn to describing how we estimate the counterfactuals.

First, note that we can rewrite equation 8 as

$$C_o \sum_i \tau_{oi}^{-\theta} P_i^{1+\theta} \frac{\left(1+\psi_i^L\right) L_i}{\alpha_i^L} = P_o^{1+\theta\left(\alpha_o^L+\alpha_o^T+\alpha_o^M\right)} \left(\frac{\left(1+\psi_o^L\right) L_o}{\alpha_o^L}\right)^{1+\theta\alpha_o^T}, \tag{31}$$

and re-write equation 14 as

$$P_o^{-\theta} = \sum_i C_i P_i^{-\theta\left(\alpha_i^L+\alpha_i^T+\alpha_i^M\right)} \left(\frac{\left(1+\psi_i^L\right) L_i}{\alpha_i^L}\right)^{-\theta\alpha_i^T} \tau_{oi}^{-\theta}. \tag{32}$$

The model implies the relative allocation of labor is affected by the aggregate amount of labor. To solve this system above, we define:

$$\phi_o \equiv \frac{\left(P_o^{-\theta} C_o\right)}{\left(\frac{(1+\psi_o^L)L_o}{\alpha_o^L}\right)^{1+\theta\alpha_o^T} P_o^{1+\theta(\alpha_o^L+\alpha_o^T+\alpha_o^M)}}. \tag{33}$$

We now have the equations:

$$1 = \sum_i \frac{C_o}{P_o^{1+\theta(\alpha_o^L+\alpha_o^T+\alpha_o^M)} \left(\frac{(1+\psi_o^L)L_o}{\alpha_o^L}\right)^{1+\theta\alpha_o^T} \bar{U}^{\theta(\alpha_o^L+\alpha_o^T)}} \tau_{oi}^{-\theta} P_i^{1+\theta} \left(\frac{1+\psi_i^L}{\alpha_i^L}\right) L_o L_i \tag{34}$$

and

$$\phi_o = \sum_i \frac{C_o}{P_o^{1+\theta(\alpha_o^L+\alpha_o^T+\alpha_o^M)} \left(\frac{(1+\psi_o^L)L_o}{\alpha_o^L}\right)^{1+\theta\alpha_o^T} \bar{U}^{\theta(\alpha_o^L+\alpha_o^T)}} \tau_{oi}^{-\theta} P_i^{1+\theta} \left(\frac{1+\psi_i^L}{\alpha_i^L}\right) L_o L_i \phi_i. \tag{35}$$

To show uniqueness of the counterfactual solutions, following Allen and Arkolakis (2014), we use the Perron-Frobenius theorem. The Perron-Frobenius theorem says that for any matrix M, with all its elements positive, the equation $\hat{\phi}\lambda = M\hat{\phi}$ has a unique eigenvalue λ and a unique (up to proportionality) eigenvector $\hat{\phi}$ with all its elements positive. We define the matrix A with elements $[A_{od}] = \left[\frac{C_o}{P_o^{1+\theta\left(\alpha_o^L+\alpha_o^T+\alpha_o^M\right)} L_o^{1+\theta\alpha_o^T}} \tau_{od}^{-\theta} P_d^{1+\theta}\left(\frac{1+\psi_d^L}{\alpha_d^L}\right) L_d\right]$. For any possible values of C_o, P_o, L_o, and τ_{od}, the matrix A has all of its elements positive. Perron-Frobenius therefore implies that, for any values of P_o, L_o, τ_{od}, the matrix A will have only one unique (up to proportionality) eigenvector with all elements positive. In other words, when our system of equations 34 and 35 holds, we have two eigenvectors. The two eigenvectors are ϕ_o and 1, which therefore must be proportional since the eigenvector is unique (up to proportionality) and so $\phi_o = \phi^{-1} * 1$ for some constant ϕ^{-1}. Thus, we know that

$$P_o^{1+\theta\left(1+\alpha_o^T+\alpha_o^L+\alpha_o^M\right)} = \phi L_o^{-\left(1+\theta\alpha_o^T\right)} C_o. \tag{36}$$

We then solve for P_o in equation 36, and plug into equation 31 to get:

$$\begin{aligned} C_o \sum_i \tau_{oi}^{-\theta}\left(\phi \bar{U}^{-\theta\left(\alpha_i^L+\alpha_i^T\right)}\left(\frac{2}{\left(1+\psi_i^L\right) L_o L_i}\alpha_i^L\right)^{-\left(1+\theta\alpha_i^T\right)} C_i\right)^{\frac{1+\theta}{1+\theta\left(1+\alpha_i^T+\alpha_i^L+\alpha_i^M\right)}} \frac{\left(1+\psi_i^L\right) L_o L_i}{\alpha_i^L} = \\ \left(\phi \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)}\left(\frac{\left(1+\psi_o^L\right) L_o}{\alpha_o^L}\right)^{-\left(1+\theta\alpha_o^T\right)} C_o\right)^{\frac{1+\theta\left(\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}{1+\theta\left(1+\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}} \\ \times\left(\frac{\left(1+\psi_o^L\right) L_o}{\alpha_o^L}\right)^{1+\theta\alpha_o^T} \bar{U}^{\theta\left(\alpha_o^L+\alpha_o^T\right)} \end{aligned} \tag{37}$$

Rearranging and combining like terms, we get:

$$\begin{aligned} \left(\phi \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)}\right)^{-\frac{1+\theta\left(\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}{1+\theta\left(1+\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}} \bar{U}^{-\theta\left(\alpha_o^L+\alpha_o^T\right)} C_o^{\frac{\theta}{1+\theta\left(1+\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}} \\ \times \sum_i \tau_{oi}^{-\theta}\left(\phi \bar{U}^{-\theta\left(\alpha_i^L+\alpha_i^T\right)} C_i\right)^{\frac{1+\theta}{1+\theta\left(1+\alpha_i^T+\alpha_i^L+\alpha_i^M\right)}}\left(\frac{\left(1+\psi_i^L\right) L_o L_i}{\alpha_i^L}\right)^{1-\frac{\left(1+\theta\alpha_i^T\right)(1+\theta)}{1+\theta\left(1+\alpha_i^T+\alpha_i^L+\alpha_i^M\right)}} \\ = \left(\frac{\left(1+\psi_o^L\right) L_o}{\alpha_o^L}\right)^{\left(1+\theta\alpha_o^T\right)\left(1-\frac{1+\theta\left(\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}{1+\theta\left(1+\alpha_o^L+\alpha_o^T+\alpha_o^M\right)}\right)}. \end{aligned} \tag{38}$$

There is a unique ϕ that solves equation 38. To find that solution, we grid-search over the parameter space. For each initial guess, we use the Fujimoto-Krause algorithm to solve for the distribution of population (L_i) and we pick the parameters for which equation 38 holds.

In our counterfactuals, we estimate how much production inputs would have changed in each county given a different vector of costs τ_{od} (e.g., without the railroads) and given a value of population (e.g., holding utility constant or holding population constant), as in Donaldson and Hornbeck (2016) and Fajgelbaum and Redding (2018). Note that relative population levels are not independent of total population levels, unlike in Donaldson and Hornbeck (2016), because the production function elasticities vary over space but quantitatively this effect is relatively small. The counterfactual impact on national aggregate productivity is then given by the Domar-weighted sum of these counterfactual changes in county production inputs multiplied by the county-level gap for that input (equation 11).

Appendix References

Adams, Henry C. 1895. *Report on transportation business in the United States at the eleventh census: 1890.* Vol. 1, US GPO.

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics*, 129(3): 1085–1140.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic inputs and resource (mis) allocation." *Journal of Political Economy*, 122(5): 1013–1063.

Atack, Jeremy. 1977. "Returns to scale in antebellum United States manufacturing." *Explorations in Economic History*, 14(4): 337.

Atack, Jeremy, and Fred Bateman. 1999. "Nineteenth-century US industrial development through the eyes of the Census of Manufactures." *Historical Methods*, 32(4): 177–188.

Atack, Jeremy, and Robert A Margo. 2019. "Gallman revisited: blacksmithing and American manufacturing, 1850–1870." *Cliometrica*, 13(1): 1–23.

Baylis, Patrick. 2020. "Temperature and temperament: Evidence from Twitter." *Journal of Public Economics*, 184: 104161.

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse. 2020. "Frontier culture: The roots and persistence of "rugged individualism" in the United States." *Econometrica*, 88(6): 2329–2368.

Boberg-Fazlic, Nina, and Paul Richard Sharp. 2020. "Immigrant Communities and Knowledge Spillovers: Danish-Americans and the Development of the Dairy Industry in the United States." *CEPR Discussion Paper No. DP13757.*

Bodenhorn, Howard, and Hugh Rockoff. 1992. "Regional interest rates in antebellum America." In *Strategic factors in nineteenth century American economic history: A volume to honor Robert W. Fogel.* 159–187. University of Chicago Press.

Caplin, Andrew, and John Leahy. 2010. "Economic Theory and the World of Practice: A Celebration of the (S, s) Model." *Journal of Economic perspectives*, 24(1): 183–202.

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1–45.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf. 2021. "Dispersion in dispersion: Measuring establishment-level differences in productivity."

David, Paul A. 1969. "Transport Innovation and Economic Growth: Professor Fogel on and off the Rails." *Economic History Review*, 22(3): 506–524.

Davis, Lance E, and Robert E Gallman. 1968. "The Share of Savings and Investment in Gross National Product During the 19th Century in the USA." 437–466, De Gruyter Mouton.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review*, 110(9): 2964–96.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *Quarterly Journal of Economics*, 131(2): 799–858.

Fajgelbaum, Pablo, and Stephen Redding. 2018. "Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914." *NBER WP*, 20217.

Fisher, Franklin M. 1987. "On the misuse of the profits-sales ratio to infer monopoly power." *The Rand Journal of Economics*, 384–396.

Fisher, Franklin M, and John J McGowan. 1983. "On the misuse of accounting rates of return to infer monopoly profits." *American Economic Review*, 73(1): 82–97.

Fogel, Robert W. 1964. *Railroads and American Economic Growth: Essays in Econometric History.* Johns Hopkins University Press.

Fogel, Robert W. 1979. "Notes on the Social Saving Controversy." *Journal of Economic History*, 39: 1–54.

Galiani, Sebastian, Luis F Jaramillo, and Mateo Uribe-Castro. 2022. "Free-Riding Yankees: Canada and the Panama Canal."

Gross, Daniel P. 2020. "Collusive Investments in Technological Compatibility: Lessons from US Railroads in the Late 19th Century." *Management Science*, 66(12): 5683–5700.

Hacker, J. David. 2013. "New estimates of census coverage in the United States, 1850–1930." *Social Science History*, 37(1): 71–101.

Haines, Michael R. 2010. "Historical, Demographic, Economic, and Social Data: The United States, 1790–2002." *ICPSR*, 2896.

Hornbeck, Richard. 2010. "Barbed wire: Property rights and agricultural development." *Quarterly Journal of Economics*, 125(2): 767–810.

Hsiang, Solomon M. 2010. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *PNAS*, 107(35): 15367–15372.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403–1448.

Hulten, Charles R. 1991. "The measurement of capital." In *Fifty years of economic measurement: The jubilee of the conference on research in income and wealth.* 119–158.

Lafortune, Jeanne, Ethan G Lewis, José Pablo Martínez, and José Tessada. 2021. "Changing Returns to Scale in Manufacturing 1880-1930: The Rise of (Skilled) Labor?"

Margo, Robert A. 2014. "Economies of Scale in Nineteenth-Century American Manufacturing Revisited: A Resolution of the Entrepreneurial Labor Input Problem." In *Enterprising America: Businesses, Banks, and Credit Markets in Historical Perspective*. 215–244. University of Chicago Press.

Petrin, Amil, and James Levinsohn. 2012. "Measuring aggregate productivity growth using plant-level data." *RAND Journal of Economics*, 43(4): 705–725.

Redding, Stephen, and Anthony Venables. 2004. "Economic geography and international inequality." *Journal of International Economics*, 62(1): 53–82.

Rotemberg, Martin, and T Kirk White. 2021. "Plant-to-Table(s and Figures): Processed Manufacturing Data and Measured Misallocation." *Working Paper*.

Sokoloff, Kenneth L. 1984. "Was the transition from the artisanal shop to the nonmechanized factory associated with gains in efficiency? Evidence from the US Manufacturing censuses of 1820 and 1850." *Explorations in Economic History*, 21(4): 351–382.

Solomon, Ezra. 1970. "Alternative rate or return concepts and their implications for utility regulation." *The Bell Journal of Economics and Management Science*, 65–81.

United States Census Bureau. 1860*a*. "Instructions to U.S. Marshals. Instructions to Assistants."

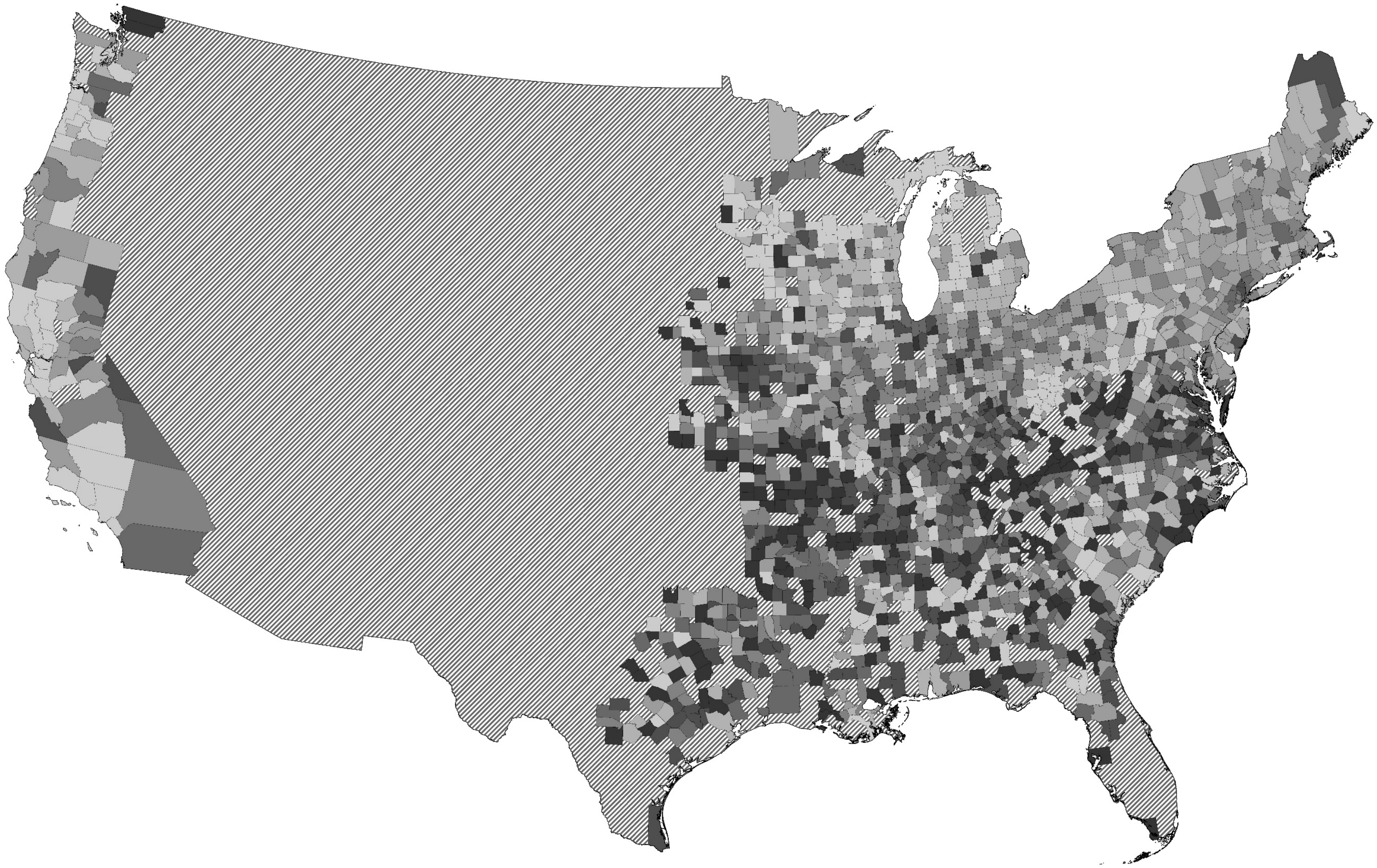
United States Census Bureau. 1860*b*. "Manufacturers of the United States in 1860."

United States Census Bureau. 1870. "Volume 3. The Statistics of Wealth and Industry of the United States."

United States Census Bureau. 1880. "Report on the Manufactures of the United States at the Tenth Census."

Weeks, Joseph Dame. 1886. *Report on the statistics of wages in manufacturing industries: with supplementary reports on the average retail prices of necessaries of life, and on trades societies, and strikes and lockouts.* Vol. 20, US GPO.

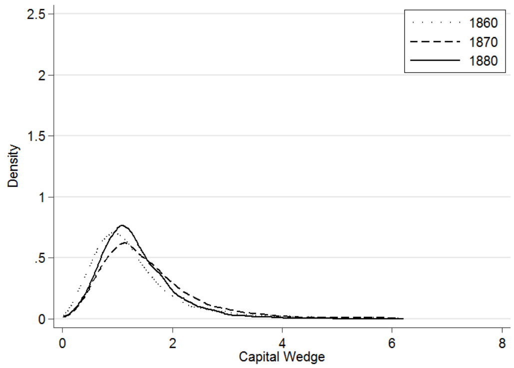
Appendix Figure 1. County-level Wedges in the Main Regression Sample, 1860-1880



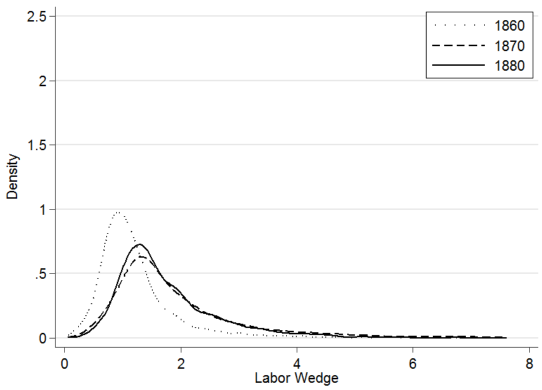
Notes: This map shows counties shaded according to their average wedge in the 1860-1880 period, averaged across capital, labor, and materials: darker shades denote larger wedges, and counties are divided into seven equal groups. This map includes the 1,802 sample counties in the regression analysis, which are all counties that report non-zero manufacturing activity from 1860, 1870, and 1880. The excluded geographic areas are cross-hashed. County boundaries correspond to county boundaries in 1890.

Appendix Figure 2. Cross-County Dispersion in Input Wedges, by Decade

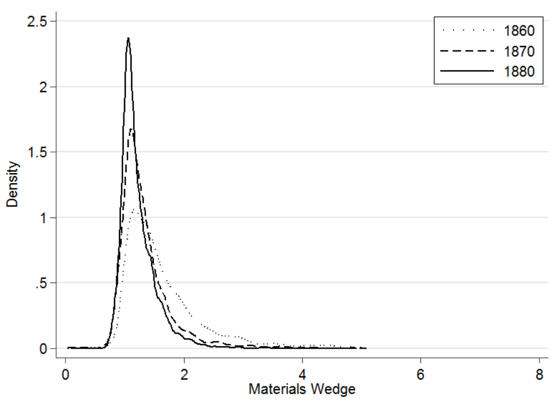
A. Cross-County Dispersion in Capital Wedges, by Decade



B. Cross-County Dispersion in Labor Wedges, by Decade



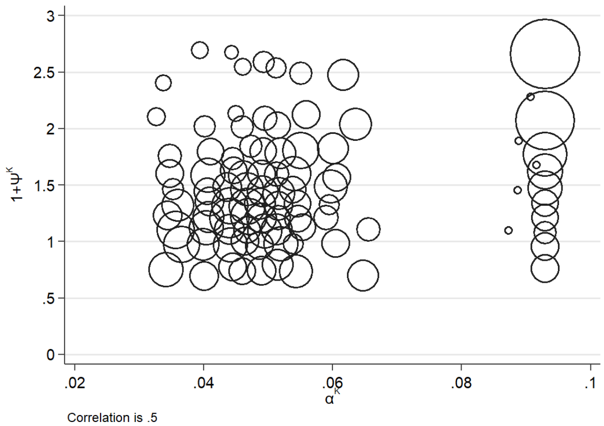
C. Cross-County Dispersion in Materials Wedges, by Decade



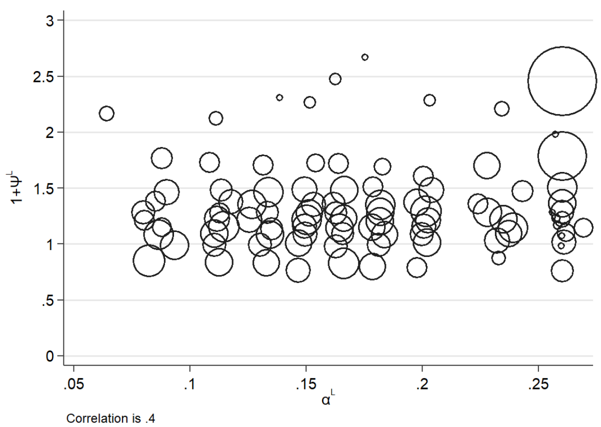
Notes: Each panel plots the cross-county dispersion in input wedges + 1, by decade, defined as the input's output elasticity divided by its revenue share.

Appendix Figure 3. Correlations between Input Wedges and their Output Elasticities

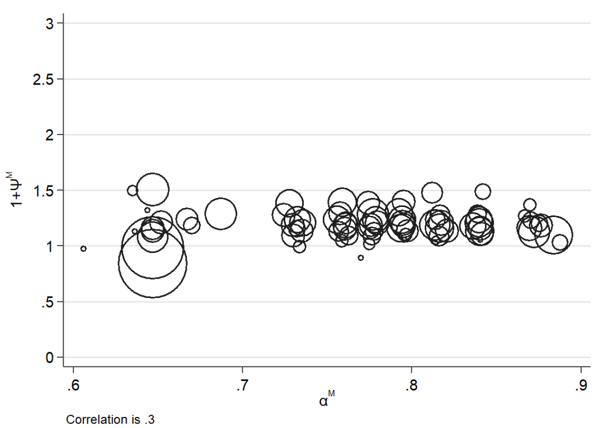
A. Capital Wedges and Production Function Elasticities



B. Labor Wedges and Production Function Elasticities



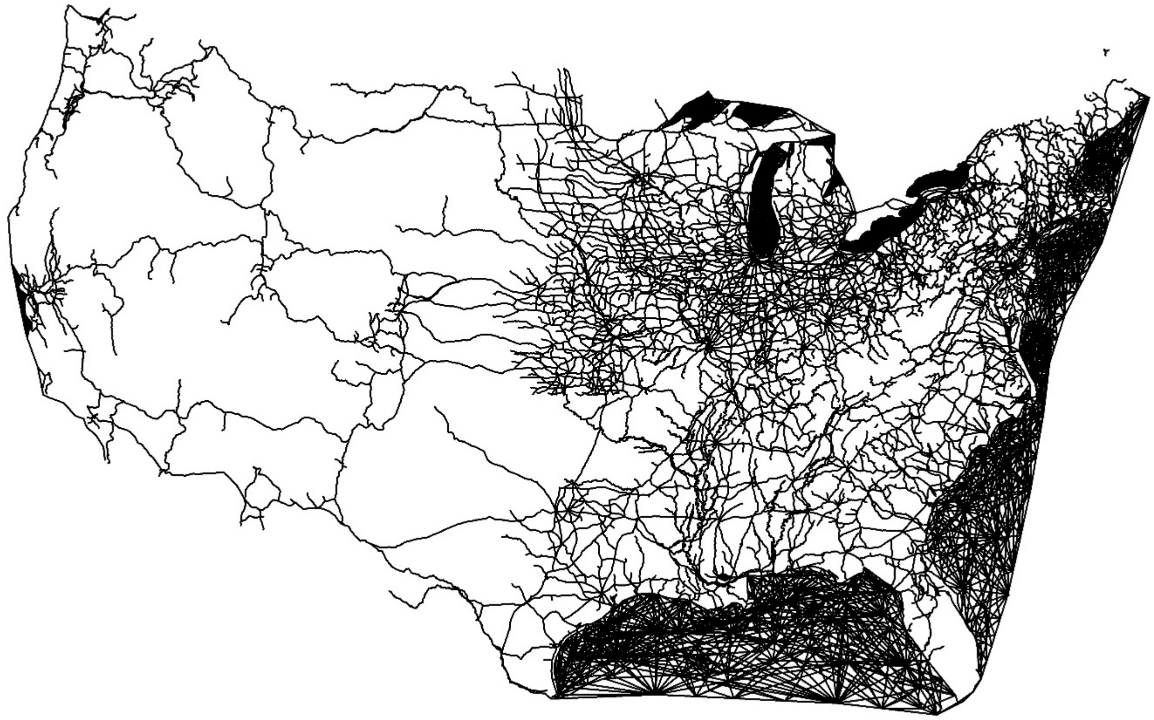
C. Materials Wedges and Production Function Elasticities



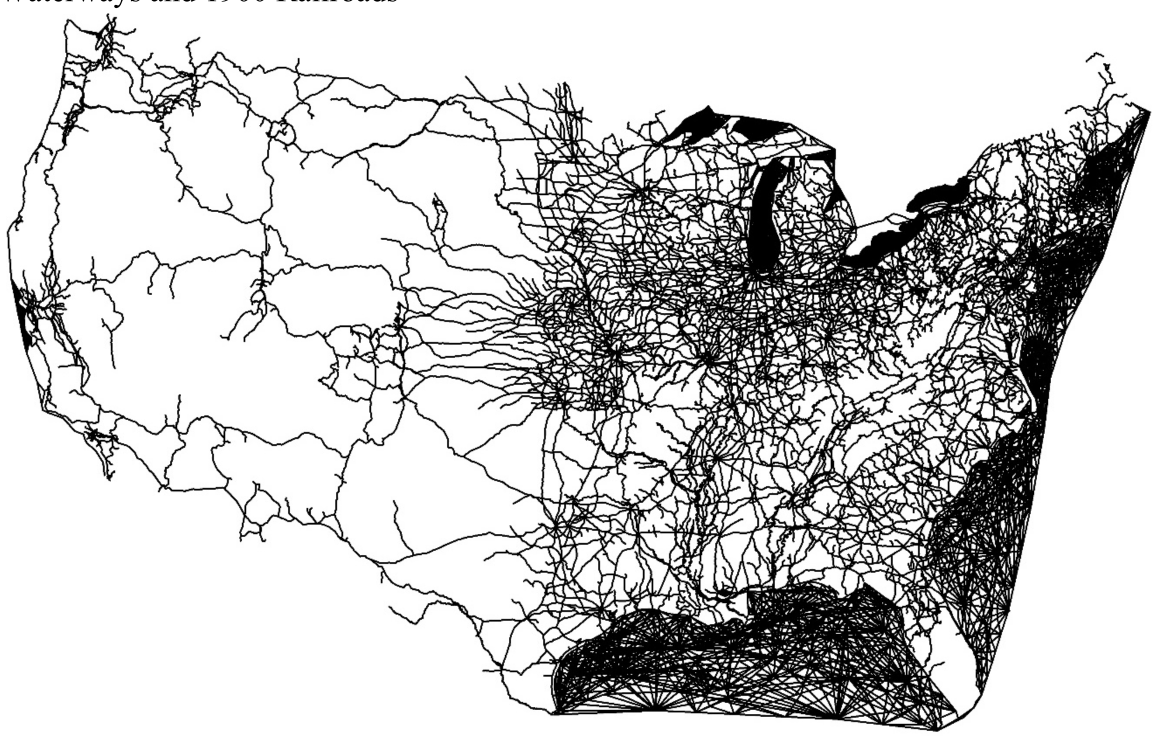
Notes: This figure plots the relationship between each input's wedge + 1 (output elasticity / revenue share) and its production function elasticity on the x-axis. To make the figure, we create 10 bins each for the wedge and the elasticity. Each bubble corresponds to one combination of bins, where the area corresponds to the number of counties in the group, and the location corresponds to the median values within the group. Each panel reports the correlation coefficient for the wedge and elasticity: 0.5 in Panel A, 0.4 in Panel B, and 0.3 in Panel C.

Appendix Figure 4. Waterways and Railroads, 1890 and 1900

A. Waterways and 1890 Railroads



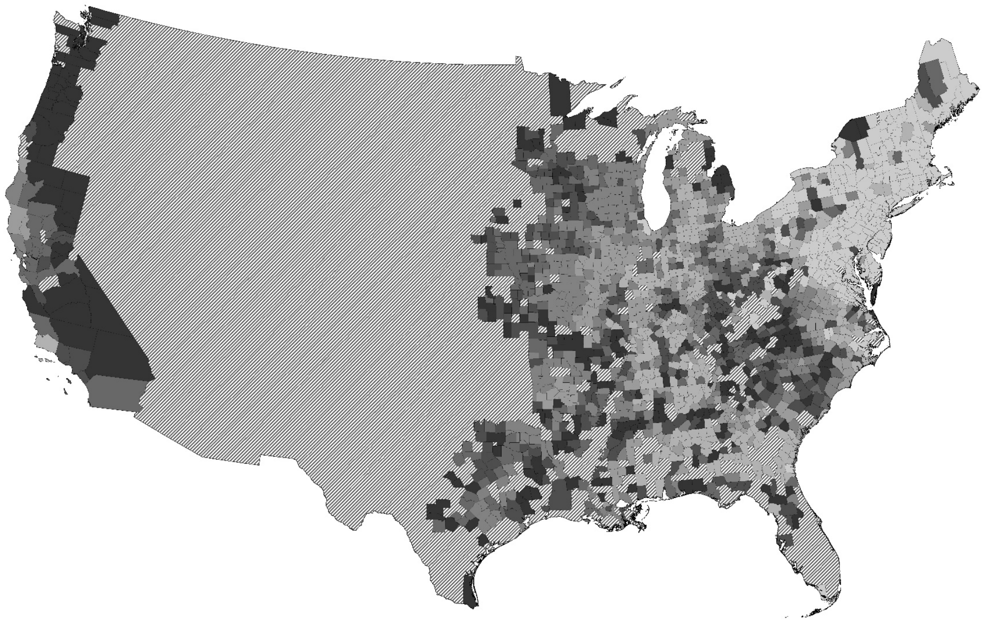
B. Waterways and 1900 Railroads



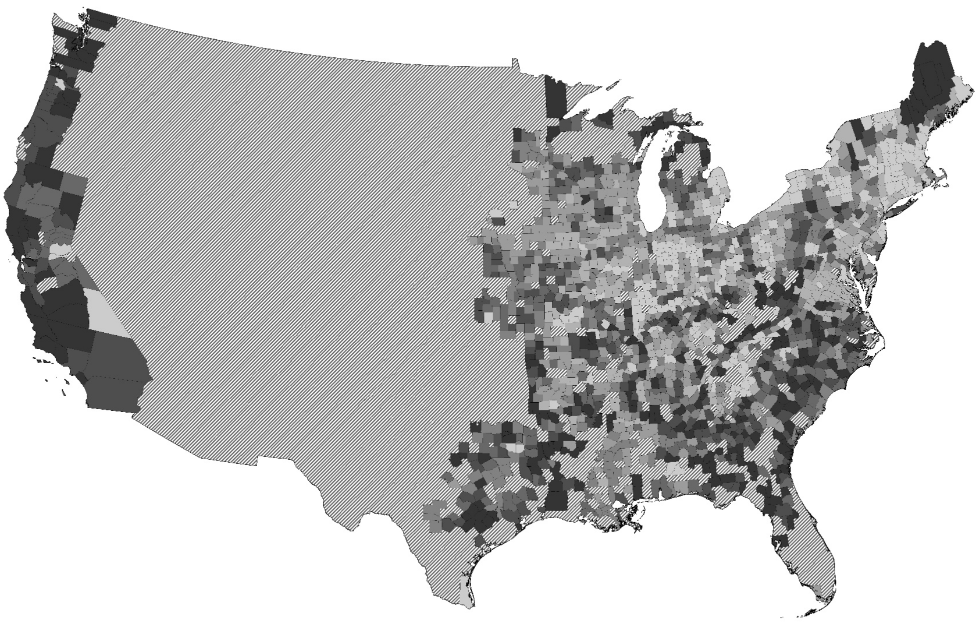
Notes: Similar to Figure 1, Panel A shows the railroads constructed by 1890, as well as the natural waterways (including navigable rivers, lakes, and oceans) and constructed canals. Panel B adds railroads constructed between 1890 and 1900.

Appendix Figure 5. Calculated Changes in Log Market Access, by County

A. From 1880 to 1890

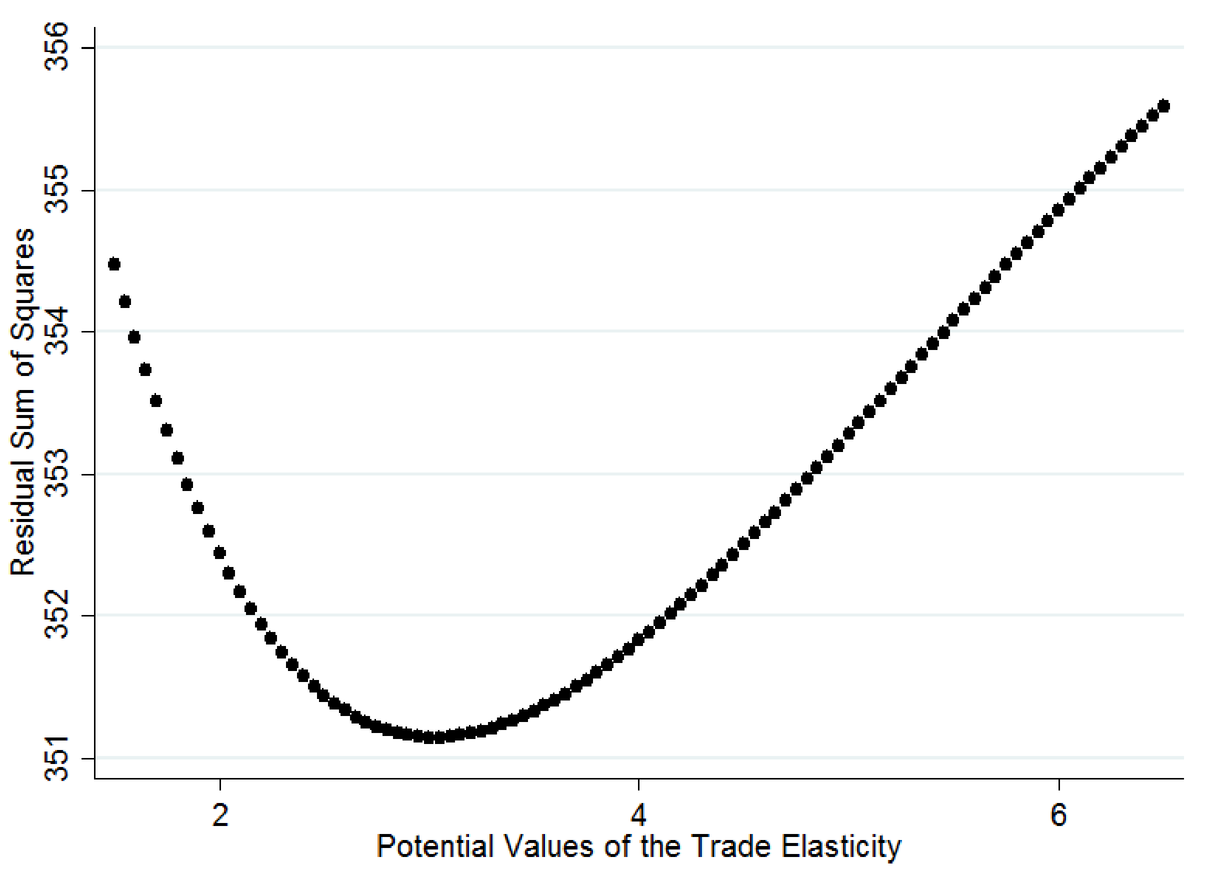


B. From 1890 to 1900



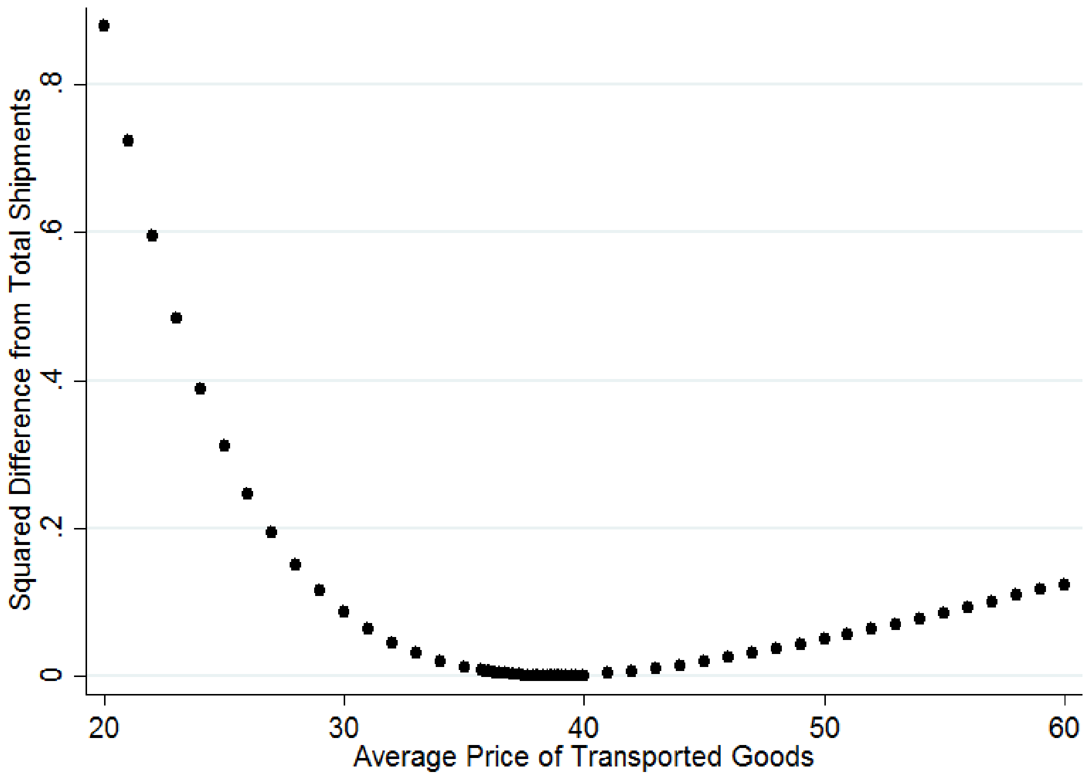
Notes: In each panel, counties are shaded according to their calculated change in market access from 1880 to 1890 (Panel A) and from 1890 to 1900 (Panel B). Counties are divided into seven groups (with an equal number of counties per group), and darker shades denote larger increases in market access. These maps include the 1,802 sample counties in the regression analysis, which are all counties that report non-zero manufacturing activity from 1860, 1870, and 1880. The excluded geographic areas are cross-hashed. County boundaries correspond to county boundaries in 1890.

Appendix Figure 6. Model-fit between Land Values and Market Access



Notes: This plot shows the residual sum of squares between the model-implied relationship between land values and market access and the corresponding relationship in the data, for different potential values of the trade elasticity θ.

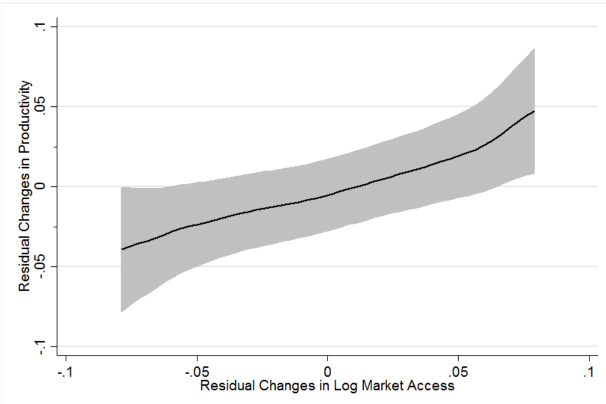
Appendix Figure 7. Model-fit for Total Railroad Shipments



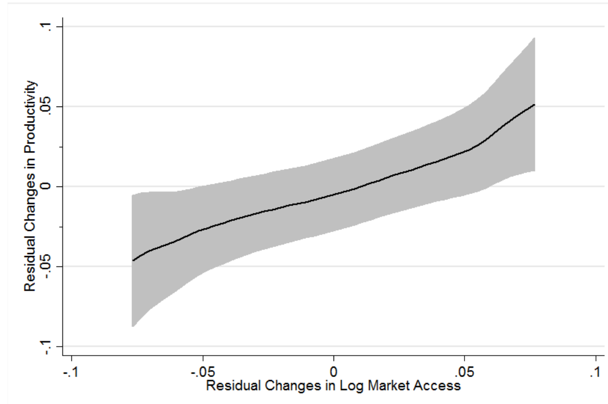
Notes: This plot shows the squared difference between actual reported railroad shipments and model-implied railroad shipments for each value of the average price of transported goods.

Appendix Figure 8. Local Polynomial Relationships Between County Productivity and Market Access, Using Approximated Market Access and Model-Defined Market Access

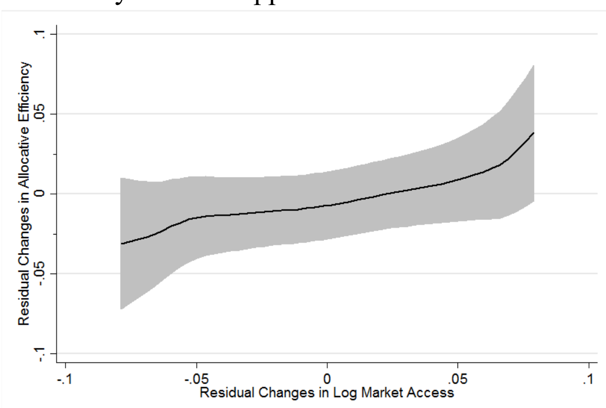
A. County Productivity and Approximated MA



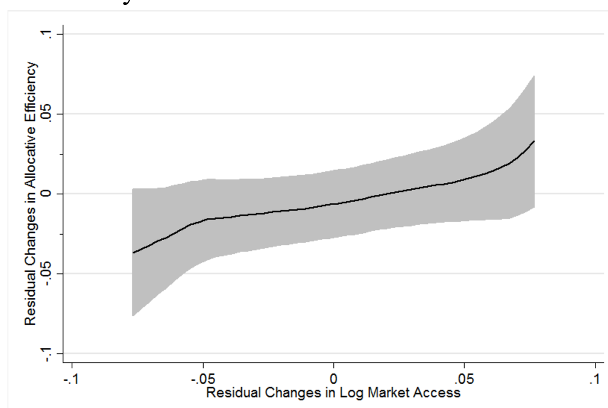
B. County Productivity and Model-Defined MA



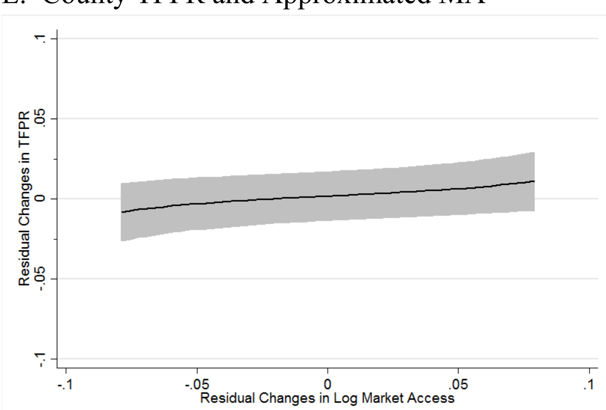
C. County AE and Approximated MA



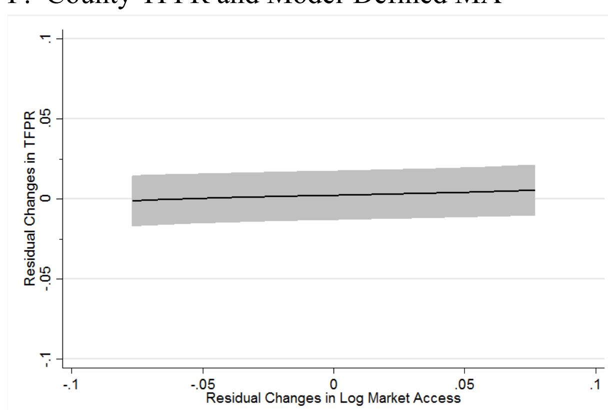
D. County AE and Model-Defined MA



E. County TFPR and Approximated MA

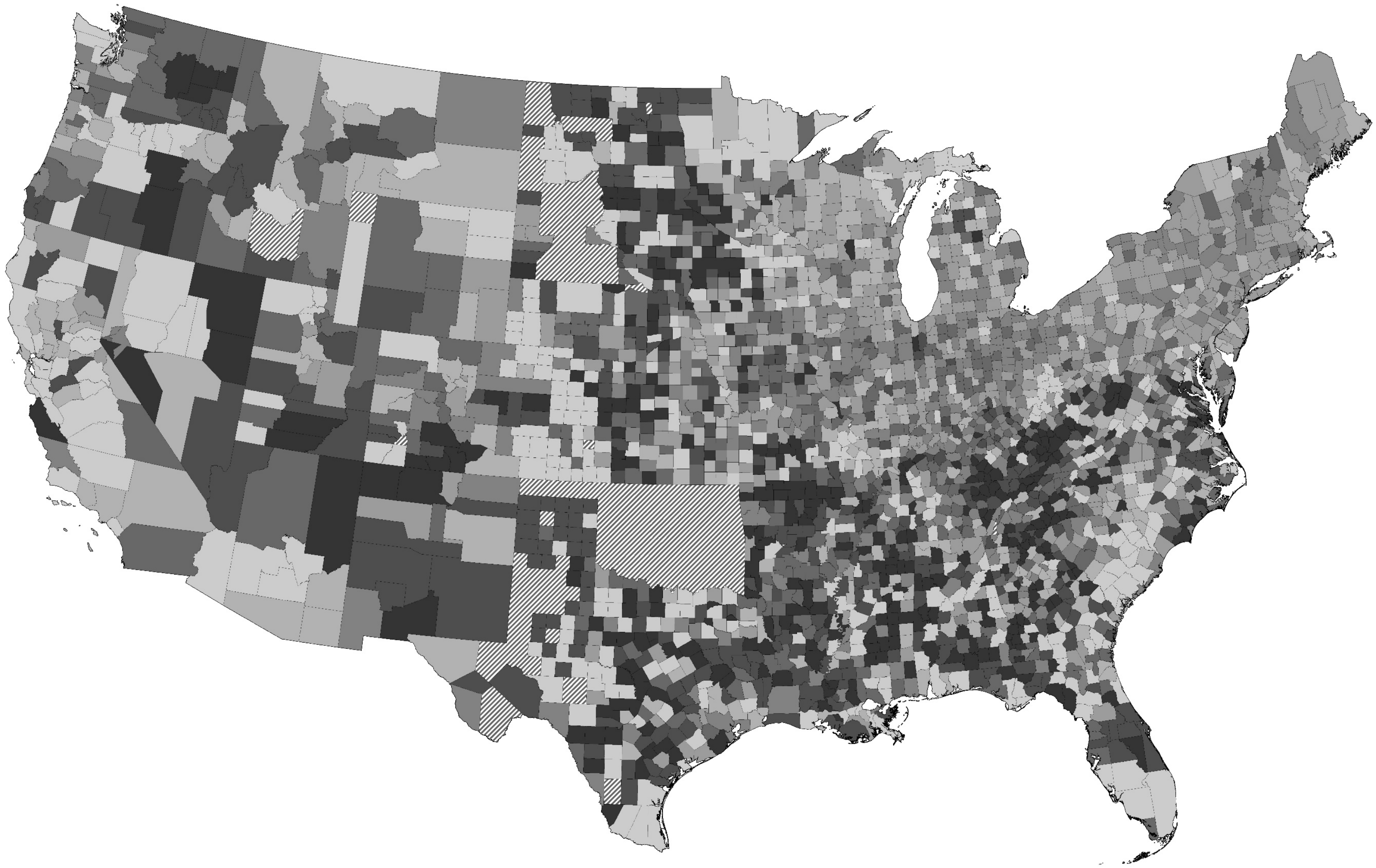


F. County TFPR and Model-Defined MA



Notes: Each panel plots the local polynomial relationship between residual productivity (y-axis) and residual market access (x-axis), where market access is based on our approximated measure (in Panels A, D, E) or based on our model-defined measure (in Panels B, C, and F). Residuals are calculated after controlling for county fixed effects, state-by-year fixed effects, and year-interacted cubic polynomial functions of county longitude and latitude (the controls in Equation 4). The local polynomial is based on an Epanechnikov kernel function with default bandwidth of 0.03. The shaded region reflects the 95% confidence interval.

Appendix Figure 9. County-level Gaps in the Counterfactual Analysis



Notes: This map shows counties shaded according to their estimated sum of gaps between the output elasticity for each input (materials, labor, capital) and the revenue share for that input: darker shades denote a larger sum of input gaps, and counties are divided into seven equal groups. This counterfactual sample includes all 2,722 counties that in 1890 report positive population and positive revenue (agriculture and/or manufacturing). The excluded geographic areas are cross-hashed. County boundaries correspond to county boundaries in 1890.

Appendix Table 1. Measured Manufacturing Output Elasticities by Decade and Region

		By Region:				
	National	Plains	West Coast	Midwest	Northeast	South
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Average Materials Elasticity						
1860	0.71	0.73	0.74	0.76	0.69	0.73
	[0.05]	[0.10]	[0.06]	[0.06]	[0.04]	[0.07]
1870	0.72	0.72	0.75	0.74	0.71	0.74
	[0.04]	[0.07]	[0.03]	[0.04]	[0.03]	[0.06]
1880	0.75	0.79	0.72	0.77	0.74	0.76
	[0.05]	[0.07]	[0.05]	[0.05]	[0.05]	[0.05]
Panel B. Average Labor Elasticity						
1860	0.25	0.22	0.22	0.20	0.26	0.23
	[0.05]	[0.09]	[0.05]	[0.06]	[0.04]	[0.07]
1870	0.24	0.24	0.21	0.22	0.25	0.22
	[0.04]	[0.06]	[0.03]	[0.04]	[0.03]	[0.05]
1880	0.21	0.16	0.23	0.18	0.22	0.19
	[0.05]	[0.06]	[0.04]	[0.05]	[0.04]	[0.05]
Panel C. Average Capital Elasticity						
1860	0.04	0.05	0.04	0.04	0.04	0.04
	[0.01]	[0.01]	[0.01]	[0.01]	[0.00]	[0.01]
1870	0.04	0.04	0.04	0.04	0.04	0.04
	[0.00]	[0.01]	[0.01]	[0.00]	[0.00]	[0.01]
1880	0.04	0.05	0.06	0.04	0.04	0.05
	[0.01]	[0.02]	[0.01]	[0.01]	[0.01]	[0.01]

Notes: This table reports measured output elasticities in the manufacturing sector, by decade. County output elasticities for each input in each decade are equal to: each national industry's expenditure on that input divided by total industry expenditure, multiplied by the share of county revenue in that industry. Column 1 reports these elasticities at the national level (reporting the unweighted average across counties), and columns 2 to 6 report these elasticities by region, which weight county elasticities by county revenue in that decade. Panels A to C report elasticities for materials, labor, and capital. Standard deviations are reported in brackets.

Appendix Table 2. Measured Manufacturing Wedges, by Decade and Region

		By Region:				
	National	Plains	West Coast	Midwest	Northeast	South
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Average of Input Wedges						
1860	1.36	1.80	1.30	1.30	1.37	1.33
	[0.21]	[1.02]	[0.20]	[0.27]	[0.14]	[0.28]
1870	1.31	1.38	1.14	1.29	1.30	1.49
	[0.20]	[0.40]	[0.11]	[0.18]	[0.12]	[0.59]
1880	1.24	1.27	1.24	1.31	1.21	1.31
	[0.18]	[0.28]	[0.12]	[0.18]	[0.14]	[0.33]
Panel B. Average Materials Wedge						
1860	1.28	1.91	1.63	1.28	1.27	1.34
	[0.23]	[1.30]	[0.45]	[0.19]	[0.10]	[0.47]
1870	1.25	1.44	1.42	1.29	1.23	1.31
	[0.21]	[0.27]	[0.21]	[0.34]	[0.11]	[0.30]
1880	1.17	1.21	1.17	1.15	1.18	1.22
	[0.09]	[0.11]	[0.08]	[0.08]	[0.08]	[0.15]
Panel C. Average Labor Wedge						
1860	1.35	1.62	0.91	1.32	1.37	1.35
	[0.27]	[1.07]	[0.17]	[0.38]	[0.17]	[0.38]
1870	1.32	1.43	1.10	1.39	1.27	1.73
	[0.45]	[1.04]	[0.19]	[0.33]	[0.17]	[1.58]
1880	1.23	1.19	1.21	1.29	1.20	1.35
	[0.23]	[0.53]	[0.15]	[0.22]	[0.14]	[0.66]
Panel D. Average Capital Wedge						
1860	1.43	1.87	1.37	1.31	1.48	1.31
	[0.40]	[1.54]	[0.47]	[0.47]	[0.31]	[0.56]
1870	1.35	1.28	0.90	1.20	1.41	1.43
	[0.34]	[0.51]	[0.24]	[0.27]	[0.29]	[0.64]
1880	1.32	1.42	1.35	1.49	1.24	1.36
	[0.44]	[0.53]	[0.30]	[0.45]	[0.41]	[0.57]

Notes: This table reports measured wedges in the manufacturing sector, by decade, where the we report one plus the wedge (i.e., that input's output elasticity divided by its revenue share). Column 1 reports these wedges at the national level, and columns 2 to 6 report these wedges by region, which weight county-level wedges by county revenue in that decade. Panel A reports the unweighted average of these wedges across inputs, and panels B to D report wedges for materials, labor, and capital. Standard deviations are reported in brackets.

Appendix Table 3. Measured Manufacturing Gaps, by Decade and Region

		By Region:				
	National	Plains	West Coast	Midwest	Northeast	South
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Sum of Average Input Gaps						
1860	0.22	0.33	0.23	0.21	0.22	0.21
	[0.06]	[0.18]	[0.09]	[0.06]	[0.04]	[0.09]
1870	0.19	0.24	0.22	0.21	0.19	0.22
	[0.07]	[0.09]	[0.10]	[0.08]	[0.06]	[0.08]
1880	0.15	0.16	0.15	0.15	0.15	0.17
	[0.06]	[0.06]	[0.03]	[0.04]	[0.07]	[0.06]
Panel B. Average Materials Gap						
1860	0.15	0.27	0.26	0.16	0.14	0.16
	[0.05]	[0.11]	[0.09]	[0.05]	[0.04]	[0.07]
1870	0.14	0.20	0.21	0.15	0.13	0.16
	[0.07]	[0.09]	[0.08]	[0.08]	[0.06]	[0.10]
1880	0.11	0.13	0.10	0.10	0.11	0.13
	[0.06]	[0.06]	[0.04]	[0.05]	[0.07]	[0.07]
Panel C. Average Labor Gap						
1860	0.06	0.05	-0.03	0.04	0.07	0.05
	[0.04]	[0.10]	[0.05]	[0.04]	[0.03]	[0.05]
1870	0.05	0.03	0.01	0.05	0.05	0.06
	[0.03]	[0.07]	[0.04]	[0.03]	[0.03]	[0.06]
1880	0.04	0.02	0.04	0.04	0.03	0.04
	[0.03]	[0.04]	[0.03]	[0.03]	[0.02]	[0.06]
Panel D. Average Capital Gap						
1860	0.01	0.01	0.00	0.01	0.01	0.00
	[0.01]	[0.03]	[0.06]	[0.01]	[0.01]	[0.03]
1870	0.01	0.00	-0.01	0.00	0.01	0.01
	[0.01]	[0.02]	[0.02]	[0.01]	[0.01]	[0.02]
1880	0.01	0.01	0.01	0.01	0.00	0.01
	[0.02]	[0.04]	[0.02]	[0.02]	[0.01]	[0.03]

Notes: This table reports measured gaps in the manufacturing sector, by decade, where the input gaps are equal to that input's output elasticity minus its revenue share. Column 1 reports these gaps at the national level, and columns 2 to 6 report these gaps by region, which weight county-level gaps by county revenue in that decade. Panel A reports the sum of these gaps across inputs, and panels B to D report gaps for materials, labor, and capital. Standard deviations are reported in brackets.

Appendix Table 4. Measured Manufacturing Revenue Shares, by Decade and Region

		By Region:				
	National	Plains	West Coast	Midwest	Northeast	South
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Average Materials Revenue Share						
1860	0.56	0.46	0.48	0.60	0.55	0.57
	[0.07]	[0.18]	[0.11]	[0.08]	[0.05]	[0.12]
1870	0.59	0.52	0.54	0.59	0.59	0.58
	[0.07]	[0.11]	[0.09]	[0.08]	[0.07]	[0.11]
1880	0.64	0.66	0.61	0.68	0.63	0.63
	[0.09]	[0.08]	[0.05]	[0.07]	[0.09]	[0.07]
Panel B. Average Labor Revenue Share						
1860	0.19	0.17	0.24	0.16	0.20	0.17
	[0.05]	[0.09]	[0.07]	[0.05]	[0.04]	[0.06]
1870	0.19	0.21	0.19	0.17	0.20	0.16
	[0.05]	[0.11]	[0.05]	[0.05]	[0.04]	[0.07]
1880	0.17	0.15	0.19	0.14	0.19	0.16
	[0.05]	[0.06]	[0.03]	[0.04]	[0.04]	[0.06]
Panel C. Average Capital Revenue Share						
1860	0.03	0.04	0.05	0.03	0.03	0.04
	[0.02]	[0.03]	[0.06]	[0.02]	[0.01]	[0.03]
1870	0.03	0.03	0.05	0.03	0.03	0.03
	[0.01]	[0.02]	[0.02]	[0.01]	[0.01]	[0.02]
1880	0.04	0.04	0.04	0.03	0.04	0.04
	[0.01]	[0.03]	[0.02]	[0.02]	[0.01]	[0.02]

Notes: This table reports measured revenue shares in the manufacturing sector, by decade, where the input revenue shares are equal to county expenditure on that input divided by county revenue. Column 1 reports these revenue shares at the national level, and columns 2 to 6 report these revenue shares by region, which weight county-level revenue shares by county revenue in that decade. Panels A to C report revenue shares for materials, labor, and capital. Standard deviations are reported in brackets.

Appendix Table 5. Impacts of County Bank Activity on County Manufacturing Wedges

	Log Bank Capital per Capita in County:			County has a Bank:		
	Total	State	National	Any	State	National
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Capital Wedge						
	-0.029	-0.020	-0.038	-0.014	0.064	0.053
	(0.014)	(0.013)	(0.018)	(0.029)	(0.058)	(0.062)
Panel B. Materials Wedge						
	0.005	0.010	-0.002	-0.016	-0.004	-0.006
	(0.006)	(0.007)	(0.007)	(0.024)	(0.012)	(0.014)
Panel C. Labor Wedge						
	-0.037	-0.009	-0.057	-0.007	0.110	0.128
	(0.013)	(0.011)	(0.016)	(0.070)	(0.082)	(0.090)

Notes: Each column reports the estimated relationship between county banks (total, state-chartered, national-chartered) and county manufacturing input wedges in each panel. Columns 1 - 3 report impacts of log bank capital per capita for all banks (column 1), all state-chartered banks (column 2), and all national-chartered banks (column 3). Columns 4 - 6 report impacts of whether a county has a bank (column 4), has a state-chartered bank (column 5), or has a national-chartered bank (column 6).

All regressions include county fixed effects, state-by-year fixed effects, and year-interacted cubic polymials in county latitude and longitude. The samples are drawn from our main balanced panel of 1,802 counties in 1860, 1870, and 1880, which are sometimes smaller due to missing bank data for some counties in some years. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880 from the full sample of 1,802 counties. Robust standard errors clustered by state are reported in parentheses.

Appendix Table 6. Aggregated Industry Group Revenue Shares and Cost Shares

	Revenue Share	Cost Shares: Materials	Labor	Capital
	(1)	(2)	(3)	(4)
Panel A. All Industries				
Total in 1860	1	0.72	0.24	0.04
Total in 1870	1	0.73	0.23	0.04
Total in 1880	1	0.76	0.20	0.04
Panel B. Aggregated Industry Groups				
Clothing, Textiles, Leather in 1860	0.32	0.69	0.27	0.04
Clothing, Textiles, Leather in 1870	0.23	0.76	0.21	0.03
Clothing, Textiles, Leather in 1880	0.22	0.73	0.23	0.04
Food and Beverage in 1860	0.25	0.90	0.07	0.03
Food and Beverage in 1870	0.21	0.89	0.08	0.03
Food and Beverage in 1880	0.29	0.91	0.06	0.02
Lumber and Wood Products in 1860	0.17	0.59	0.36	0.05
Lumber and Wood Products in 1870	0.16	0.64	0.31	0.05
Lumber and Wood Products in 1880	0.14	0.68	0.27	0.05
Metals and Metal Products in 1860	0.15	0.63	0.32	0.05
Metals and Metal Products in 1870	0.18	0.66	0.29	0.05
Metals and Metal Products in 1880	0.13	0.68	0.27	0.05
Other Industries in 1860	0.12	0.66	0.29	0.05
Other Industries in 1870	0.22	0.66	0.30	0.04
Other Industries in 1880	0.22	0.64	0.29	0.07

Notes: Panel A reports aggregate statistics on manufacturing in the United States, by decade, from summing the county-by-industry data (reporting the unweighted average across industries): annual expenditures on materials (column 2), labor (column 3), and capital (column 4) as a share of total annual expenditures. Panel B reports these statistics for aggregated industry groups, along with that industry group's share of total revenue (column 1). The "Clothing, Textiles, Leather" industry group contains: clothing; yarn, cloth, and other textiles; leather; leather products; boots and shoes. The "Food and Beverage" industry group contains: flour and grist mills; bread and bakery products; butter and cheese; tobacco; liquors and beverages. The "Lumber and Wood Products" industry group contains: lumber; wood products; cooperage; carriages and wagons; furniture; paper; printing and publishing; ship and boat building. The "Metals and Metal Products" industry group contains: iron and steel; iron and steel products; brass and other metal products; tin, copper, and sheet-iron ware; jewelry, pottery, and decorative work.

Appendix Table 7. Measured Manufacturing Wedges, by Decade and Industry

	All Industries	Clothing, Textiles, Leather	Food and Beverage	Lumber and Wood Products	Metals and Metal Products
	(1)	(2)	(3)	(4)	(5)
Panel A. Average Input Wedge					
1860	1.47	1.29	1.67	1.46	1.36
	[0.46]	[0.23]	[0.50]	[0.43]	[0.26]
1870	1.49	1.33	1.66	1.46	1.38
	[0.58]	[0.33]	[0.54]	[0.40]	[0.88]
1880	1.47	1.25	1.90	1.38	1.34
	[0.51]	[0.22]	[0.69]	[0.23]	[0.32]
Panel B. Average Materials Wedge					
1860	1.34	1.35	1.13	1.48	1.42
	[0.67]	[0.27]	[0.15]	[0.55]	[0.41]
1870	1.25	1.14	1.18	1.37	1.27
	[0.43]	[0.14]	[0.26]	[0.80]	[0.32]
1880	1.14	1.18	1.04	1.12	1.21
	[0.20]	[0.10]	[0.09]	[0.18]	[0.35]
Panel C. Average Labor Wedge					
1860	1.62	1.26	2.31	1.43	1.29
	[0.89]	[0.42]	[1.10]	[0.60]	[0.42]
1870	1.64	1.49	2.12	1.60	1.32
	[0.91]	[0.62]	[0.96]	[0.59]	[0.61]
1880	1.76	1.38	2.51	1.76	1.46
	[0.93]	[0.54]	[1.20]	[0.68]	[0.57]
Panel D. Average Capital Wedge					
1860	1.45	1.26	1.58	1.48	1.37
	[0.67]	[0.51]	[0.75]	[0.70]	[0.51]
1870	1.57	1.35	1.68	1.42	1.53
	[1.17]	[0.61]	[1.09]	[0.59]	[2.18]
1880	1.52	1.18	2.14	1.28	1.35
	[0.87]	[0.41]	[1.18]	[0.43]	[0.65]

Notes: This table reports measured wedges in the manufacturing sector, by decade, where the we report one plus the wedge (i.e., that input's output elasticity divided by its revenue share). Column 1 reports these wedges at the national level, and columns 2 to 5 report these wedges for consistent aggregated industry groups, which weight county-industry wedges by county-industry revenue in that decade. Panel A reports the unweighted average of these wedges across inputs, and panels B to D report wedges for materials, labor, and capital. Standard deviations are reported in brackets.

Appendix Table 8. Measured Manufacturing Gaps, by Decade and Industry

	All Industries	Clothing, Textiles, Leather	Food and Beverage	Lumber and Wood Products	Metals and Metal Products
	(1)	(2)	(3)	(4)	(5)
Panel A. Sum of Average Input Gaps					
1860	0.22	0.22	0.16	0.26	0.23
	[0.09]	[0.06]	[0.07]	[0.09]	[0.09]
1870	0.19	0.16	0.18	0.25	0.19
	[0.16]	[0.06]	[0.09]	[0.09]	[0.08]
1880	0.15	0.16	0.10	0.18	0.17
	[0.12]	[0.04]	[0.04]	[0.08]	[0.08]
Panel B. Average Materials Gap					
1860	0.15	0.18	0.09	0.17	0.17
	[0.09]	[0.06]	[0.08]	[0.09]	[0.10]
1870	0.11	0.08	0.11	0.12	0.12
	[0.14]	[0.07]	[0.09]	[0.10]	[0.10]
1880	0.07	0.10	0.03	0.05	0.09
	[0.12]	[0.05]	[0.07]	[0.11]	[0.10]
Panel C. Average Labor Gap					
1860	0.06	0.04	0.06	0.08	0.05
	[0.06]	[0.05]	[0.04]	[0.07]	[0.06]
1870	0.07	0.07	0.06	0.11	0.06
	[0.07]	[0.04]	[0.03]	[0.06]	[0.06]
1880	0.07	0.06	0.06	0.13	0.07
	[0.06]	[0.04]	[0.03]	[0.06]	[0.07]
Panel D. Average Capital Gap					
1860	0.01	0.00	0.01	0.01	0.01
	[0.02]	[0.01]	[0.01]	[0.02]	[0.04]
1870	0.01	0.00	0.01	0.01	0.01
	[0.02]	[0.01]	[0.01]	[0.01]	[0.02]
1880	0.00	0.00	0.01	0.01	0.01
	[0.03]	[0.01]	[0.01]	[0.02]	[0.02]

Notes: This table reports measured gaps in the manufacturing sector, by decade, where the input gaps are equal to that input's output elasticity minus its revenue share. Column 1 reports these gaps at the national level, and columns 2 to 5 report these gaps for consistent aggregated industry groups, which weight county-industry gaps by county-industry revenue in that decade. Panel A reports the sum of these gaps across inputs, and panels B to D report gaps for materials, labor, and capital. Standard deviations are reported in brackets.

Appendix Table 9. Impacts of Market Access on Input Gaps and Wedges, by Region

		By Region:					Frontier
	National	Plains	West Coast	Midwest	Northeast	South	Counties
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Input Gaps							
Materials	0.012	0.020	0.094	0.019	0.027	0.021	0.019
	(0.006)	(0.008)	(0.009)	(0.009)	(0.025)	(0.017)	(0.013)
Labor	-0.001	0.000	0.047	-0.015	0.004	-0.014	-0.003
	(0.005)	(0.009)	(0.011)	(0.003)	(0.016)	(0.009)	(0.007)
Capital	0.001	0.000	0.006	0.004	0.000	0.002	0.001
	(0.002)	(0.002)	(0.008)	(0.004)	(0.002)	(0.002)	(0.004)
Panel B. Input Wedges							
Materials	-0.028	0.019	-0.024	0.031	0.021	0.133	-0.099
	(0.040)	(0.020)	(0.031)	(0.039)	(0.027)	(0.090)	(0.127)
Labor	-0.048	0.061	0.102	-0.165	-0.039	-0.478	0.004
	(0.068)	(0.045)	(0.074)	(0.053)	(0.148)	(0.110)	(0.065)
Capital	0.022	0.173	-0.096	-0.005	0.062	-0.006	0.146
	(0.036)	(0.114)	(0.192)	(0.064)	(0.046)	(0.113)	(0.081)

Notes: Panels A and B report the estimated impacts of market access on county-level input gaps and county-level input wedges in manufacturing. Column 1 reports estimates at the national level, as in Columns 2 and 3 of Table 3. Columns 2 to 6 report estimates from separate regressions for each region, and Column 7 reports estimates in the sample of "Frontier Counties," defined following Bazzi et al. (2020) as counties with between two and six people per square mile in 1860 and that are within 100km of the boundary where population density fell below two people per square mile in 1860.

All regressions include county fixed effects, state-by-year fixed effects, and year-interacted cubic polymials in county latitude and longitude. The sample is our main balanced panel of 1,802 counties in 1860, 1870, and 1880. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880. Robust standard errors clustered by state are reported in parentheses.

Appendix Table 10. Impacts of Market Access on Input Gaps and Wedges, by Industry

		By Industry Group:			
	All Industries	Clothing, Textiles, Leather	Food and Beverage	Lumber and Wood Products	Metals and Metal Products
	(1)	(2)	(3)	(4)	(5)
Panel A. Input Gaps					
Materials	0.018	-0.001	-0.003	0.010	-0.021
	(0.014)	(0.008)	(0.004)	(0.004)	(0.015)
Labor	0.010	0.015	0.002	0.000	0.026
	(0.004)	(0.005)	(0.003)	(0.006)	(0.015)
Capital	-0.001	0.001	0.000	0.005	0.004
	(0.005)	(0.003)	(0.002)	(0.003)	(0.006)
Panel B. Input Wedges					
Materials	-0.004	-0.023	0.010	0.027	-0.072
	(0.019)	(0.035)	(0.016)	(0.040)	(0.096)
Labor	0.050	0.093	-0.010	0.006	0.420
	(0.029)	(0.078)	(0.078)	(0.033)	(0.259)
Capital	0.001	0.062	-0.068	0.071	-0.040
	(0.038)	(0.077)	(0.049)	(0.070)	(0.105)

Notes: Panels A and B report the estimated impacts of market access on county-industry input gaps and county-industry input wedges, where industry is defined using these four aggregated industry groups and other industries. Column 1 reports pooled estimates, from an unweighted regression, and Columns 2 to 5 allow the effect of market access to vary in each consistent aggregated industry group.

These are county-industry regressions, as in Table 4. The sample is drawn from our main balanced panel of 1,802 counties in 1860, 1870, and 1880, though each industry group is not reported in each county and decade. We omit county-industries that appear only once, but do not restrict the sample to county-industries that appear all three years. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880. Robust standard errors clustered by state are reported in parentheses.

Appendix Table 11. Impacts of Market Access on County Specialization

	Revenue Shares	Value-Added Shares	Surplus Shares	Employment Shares
	(1)	(2)	(3)	(4)
Panel A. Cross-Sector Specialization Index (Manufacturing vs. Agriculture)				
Log Market Access	-0.0103	0.0008	-0.0018	0.0016
	(0.0111)	(0.0062)	(0.0119)	(0.0053)
Number of Counties	1,774	1,774	1,713	1,687
County/Year Obs.	5,322	5,322	5,139	5,061
Panel B. Within-Manufacturing Specialization Index (Across Industries)				
Log Market Access	-0.0240	-0.0379	-0.0281	-0.0098
	(0.0100)	(0.0678)	(0.0403)	(0.0084)
Number of Counties	1,802	1,802	1,802	1,802
County/Year Obs.	5,406	5,406	5,406	5,406

Notes: For the indicated outcome variable, each column and panel reports the estimated impact of log market access from our baseline specification (as in column 1 of Table 1). In panel A, the outcome variables reflect a cross-sector specialization index: the share of county value in manufacturing minus its national share (squared) plus the share of county value in agriculture minus its national share (squared), where those values are based on revenue (column 1), value-added (column 2), surplus (column 3), and employment (column 4) as defined in Table 5. In panel B, the outcome variables reflect a within-manufacturing specialization index: the share of county manufacturing value in each industry minus that industry's national manufacturing share (squared and summed across each industry), where the values for manufacturing are as defined in panel A.

All regressions include county fixed effects, state-by-year fixed effects, and year-interacted cubic polymials in county latitude and longitude. The samples are drawn from our main balanced panel of 1,802 counties in 1860, 1870, and 1880, which are sometimes smaller due to missing data for some counties in some years. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880 in the full sample of 1,802 counties. Robust standard errors clustered by state are reported in parentheses.

Appendix Table 12. Impacts of Market Access, Robustness (Regional Shocks)

		Estimated Impact of Market Access on:		
		Productivity	TFPR	AE
		(1)	(2)	(3)
1.	Baseline Specification	0.204	0.036	0.168
		(0.051)	(0.025)	(0.051)
2.	Only 1860 and 1870	0.188	0.020	0.168
		(0.078)	(0.038)	(0.079)
3.	Only 1870 and 1880	0.201	0.123	0.078
		(0.102)	(0.059)	(0.113)
4.	Only 1860 and 1880	0.215	0.006	0.209
		(0.067)	(0.034)	(0.057)
5.	Current Market Access, Controlling for Market Access	0.177	0.043	0.134
	10 Years in the Future	(0.045)	(0.025)	(0.048)
	Market Access 10 Years in the Future, Controlling for	0.035	0.012	0.023
	Current Market Access	(0.085)	(0.066)	(0.069)
6.	Controlling for 1860 Waterway Access	0.203	0.033	0.170
		(0.060)	(0.048)	(0.073)
7.	Controlling for 1860 Market Access	0.255	0.059	0.196
		(0.089)	(0.050)	(0.082)
8.	Controls for Industry Shares	0.198	0.024	0.174
		(0.056)	(0.026)	(0.058)
9.	Controls for Detailed Industry Shares	0.161	0.037	0.124
		(0.066)	(0.023)	(0.074)
10.	Controls for Banks in 1860	0.203	0.034	0.168
		(0.052)	(0.026)	(0.053)
11.	Controls for County Gaps in 1860	0.182	0.003	0.179
		(0.045)	(0.025)	(0.059)
12.	Controls for County Elasticities in 1860	0.217	0.039	0.178
		(0.052)	(0.025)	(0.055)
13.	Controls for County Elasticities and Gaps in 1860	0.194	0.004	0.190
		(0.049)	(0.024)	(0.063)
14.	Controls for County Revenue Shares in 1860	0.195	0.002	0.194
		(0.047)	(0.023)	(0.060)
15.	Controls for County Wedges in 1860	0.207	0.066	0.141
		(0.055)	(0.023)	(0.056)
16.	Controls for County HHIs in 1860	0.190	0.046	0.144
		(0.053)	(0.026)	(0.054)
17.	Controls for Frontier in 1860	0.165	0.040	0.126
		(0.049)	(0.026)	(0.046)

18.	Controls for Gaps, Elasticities, Wedges, HHIs, and Frontier in 1860	0.162 (0.053)	0.035 (0.021)	0.127 (0.065)
19.	Excludes Civil War related industries (strict)	0.209 (0.050)	0.037 (0.026)	0.172 (0.052)
20.	Excludes Civil War related industries (broad)	0.221 (0.050)	0.047 (0.027)	0.174 (0.052)
21.	Controls for share of War Related Industries (strict)	0.205 (0.051)	0.035 (0.026)	0.170 (0.052)
22.	Controls for share of War Related Industries (broad)	0.223 (0.048)	0.035 (0.025)	0.188 (0.050)
23.	Controls for Civil War battles and casualties	0.203 (0.052)	0.037 (0.026)	0.167 (0.051)
24.	Drop Counties with battles, > 500 casualties	0.216 (0.054)	0.037 (0.025)	0.178 (0.053)
25.	Drop Counties with Civil War battles	0.199 (0.050)	0.033 (0.027)	0.166 (0.054)
26.	Drop Counties on Civil War Border	0.234 (0.056)	0.040 (0.027)	0.194 (0.054)
27.	Drop Confederate states	0.210 (0.066)	0.010 (0.037)	0.200 (0.063)
28.	Only Slave states	0.240 (0.080)	0.016 (0.050)	0.224 (0.085)
29.	Drop Slave states	0.171 (0.076)	-0.002 (0.044)	0.173 (0.077)
30.	Drop Southern region	0.198 (0.063)	0.030 (0.028)	0.168 (0.061)
31.	Fixed Effects for 20 "resource regions"	0.229 (0.054)	0.045 (0.029)	0.184 (0.052)
32.	Fixed Effects for 106 "resource subregions"	0.186 (0.067)	0.039 (0.049)	0.147 (0.056)
33.	Fifth Order Polynomial	0.207 (0.055)	0.026 (0.029)	0.180 (0.059)
34.	First Order Polynomial	0.176 (0.047)	0.023 (0.026)	0.154 (0.042)
35.	Control for State-Year-Geographic Coordinates	0.158 (0.043)	0.021 (0.030)	0.137 (0.039)
36.	Drop Western region (Plains and West Coast)	0.204 (0.079)	0.016 (0.035)	0.189 (0.077)
37.	Drop Northeast region	0.214	0.041	0.173

	(0.054)	(0.025)	(0.054)
38. Drop East-North-Central	0.178	0.036	0.142
	(0.053)	(0.026)	(0.056)
39. Drop West-North-Central	0.196	0.050	0.146
	(0.050)	(0.035)	(0.054)

Notes: Row 1 reports our baseline estimates, from Table 2, for the impacts of market accesss on county productivity (column 1) and the impacts through changes in county TFPR or revenue total factor productivity (column 2) and county AE or allocative efficiency (column 3). All regressions include county fixed effects, state-by-year fixed effects, and year-interacted cubic polymials in county latitude and longitude, unless otherwise noted. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880. The sample is a balanced panel of 1,802 counties (1860, 1870, 1880). Robust standard errors clustered by state are reported in parentheses.

Rows 2 to 39 report alternative estimates, which generally relate to controlling for other regional shocks. Rows 2, 3, and 4 report estimates when restricting the sample period to two decades only, focusing on changes over that period only. Row 5 reports the impacts of current market access along with the impacts of future market access (i.e., the 10-year pre-trend), controlling for contemporaneous and future values for whether a county has any railroad and the length of its railroads. Row 6 controls for counties' market access through waterways in 1860, interacted with decade. Row 7 controls for counties' market access in 1860, interacted with decade. Row 8 controls for counties' 1860 share of revenue in each of 31 industries, interacted with decade. Row 9 controls for counties 1860 share of revenue in each of 193 industries, interacted with decade. Row 10 controls for county banking activity in 1860, interacted with year, including the presence of any bank and total bank deposits per capita. Rows 11 to 17 control for other county characteristics in 1860, interacted with decade: county gaps; county elasticities; county elasticities and gaps; county revenue shares; county wedges; county employment concentration within manufacturing and across sectors; and whether a county was on the "frontier." Row 18 controls for all of those variables, interacted with decade. Rows 19 and 20 exclude from the data those industries most related to the Civil War or more broadly related to the Civil War (see Data Appendix for the list of industries). Rows 21 and 22 instead control for counties' revenue share in Civil War related production, interacted with decade. Row 23 controls for whether a county had a Civil War battle, the number of battles (cubic polynomial), and the number of casualties (cubic polynomial), all interacted with decade fixed effects. Row 24 excludes 99 counties with recorded Civil War battles that had more than 500 recorded casualties, and Row 25 excludes 177 counties with recorded Civil War battles. Row 26 drops 93 counties on the North-South border, Row 27 drops 745 counties in Confederate states, Row 28 includes only slave states, Row 29 drops 980 counties in slave states, and Row 30 drops 765 counties in the Southern region. Row 31 controls for region-by year fixed effects (20 regions), and row 32 controls for subregion-by-year fixed effects (106 subregions). Rows 33 and 34 modify the controls for county latitude and longitude to be a fifth-order polynomial or first-order polynomial, respectively. Row 35 controls for state-specific linear functions of counties' latitude and longitude. Row 36 excludes 201 counties in the Plains region and West Coast region of the sample. Rows 37 to 39 omit the Northeast, East-North-Central, and West-North-Central regions.

Appendix Table 13. Impacts of Market Access, Robustness (Measurement of Productivity)

		Estimated Impact of Market Access on:		
		Productivity	TFPR	AE
		(1)	(2)	(3)
1.	Baseline Specification	0.204	0.036	0.168
		(0.051)	(0.025)	(0.051)
2.	Set capital wedges to zero	0.207	0.027	0.179
		(0.049)	(0.028)	(0.053)
3.	Use materials wedge for capital wedges	0.206	0.030	0.176
		(0.049)	(0.029)	(0.053)
4.	Use materials wedge for capital and labor wedges	0.200	0.051	0.149
		(0.073)	(0.065)	(0.066)
5.	Doubling firm capital costs	0.206	0.051	0.155
		(0.052)	(0.029)	(0.051)
6.	Tripling firm capital costs	0.207	0.082	0.125
		(0.061)	(0.038)	(0.060)
7.	Using National instead of State Interest Rates	0.209	0.036	0.173
		(0.050)	(0.025)	(0.051)
8.	Decrease dispersion of capital wedges by 5%	0.206	0.029	0.177
		(0.049)	(0.024)	(0.050)
9.	Decrease dispersion of capital wedges by 10%	0.205	0.026	0.179
		(0.049)	(0.023)	(0.050)
10.	Decrease dispersion of capital wedges by 25%	0.198	0.015	0.183
		(0.049)	(0.020)	(0.051)
11.	Decrease dispersion of all wedges by 5%	0.208	0.029	0.179
		(0.049)	(0.024)	(0.050)
12.	Decrease dispersion of all wedges by 10%	0.208	0.025	0.183
		(0.049)	(0.023)	(0.051)
13.	Decrease dispersion of all wedges by 25%	0.204	0.014	0.190
		(0.049)	(0.019)	(0.051)
14.	Using 1860 values for Wedges and Scaling Factor	0.205	0.045	0.161
		(0.067)	(0.022)	(0.068)
15.	Using 1860 values for Wedges and Scaling Factor,	0.205	0.044	0.161
	and 1860 population for calculating market access	(0.067)	(0.022)	(0.067)
16.	Using Median Scaling Factor	0.224	0.028	0.197
		(0.055)	(0.023)	(0.051)
17.	Using County-Specific Scaling Factors	0.223	0.040	0.183
		(0.055)	(0.024)	(0.052)
18.	Dropping top/bottom centile, change in Productivity	0.188	0.052	0.136
		(0.047)	(0.027)	(0.047)

19	Dropping top/bottom 5 centiles, change in Productivity	0.169	0.058	0.111
		(0.049)	(0.034)	(0.051)
20.	Inflating firm labor costs	0.184	0.039	0.145
		(0.053)	(0.031)	(0.049)
21.	Decreasing returns to scale (0.95)	0.204	0.081	0.123
		(0.051)	(0.027)	(0.042)
22.	Increasing returns to scale (1.05)	0.204	-0.014	0.218
		(0.051)	(0.030)	(0.064)
23.	Using elasticities weighted by costs instead of revenues	0.204	0.031	0.173
		(0.051)	(0.026)	(0.053)
24.	Using elasticities from most-efficient counties	0.204	0.035	0.170
		(0.051)	(0.028)	(0.052)
25.	Using national industry cost shares, omitting own county	0.204	0.033	0.171
		(0.051)	(0.025)	(0.053)
26.	Using state industry cost shares	0.204	0.042	0.162
		(0.051)	(0.026)	(0.052)
27.	Using state industry elasticities, omitting own county	0.204	0.038	0.167
		(0.051)	(0.029)	(0.054)
28.	Using local elasticities	0.204	0.045	0.159
		(0.051)	(0.021)	(0.049)
29.	Inflate labor cost share by 5 percentage points	0.204	0.030	0.174
		(0.051)	(0.022)	(0.056)
30.	Inflate materials cost share by 5 percentage points	0.204	0.035	0.170
		(0.051)	(0.032)	(0.052)
31.	Inflate capital cost share by 5 percentage points	0.204	0.041	0.163
		(0.051)	(0.022)	(0.053)
32.	Exclude butter and cheese industry	0.206	0.034	0.172
		(0.049)	(0.025)	(0.050)

Notes: Row 1 reports our baseline estimates, from Table 2, for the impacts of market accesss on county productivity (column 1) and the impacts through changes in county TFPR or revenue total factor productivity (column 2) and county AE or allocative efficiency (column 3). All regressions include county fixed effects, state-by-year fixed effects, and year-interacted cubic polymials in county latitude and longitude. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880. The sample is a balanced panel of 1,802 counties (1860, 1870, 1880). Robust standard errors clustered by state are reported in parentheses.

Rows 2 to 32 report alternative estimates, which generally relate to adjusting our measurement of productivity and its decomposition into TFPR and AE. Row 2 calculates county productivity and its components when assuming zero misallocation in capital, such that the county's capital revenue share is equal to its cost share. Row 3 calculates county productivity and its components when assuming that capital misallocation is equal to materials misallocation, such that the county's capital revenue share is adjusted so the ratio of its cost share to revenue share is equal to the ratio of that county's materials cost share to materials revenue share. Row 4 assumes that capital and labor misallocation are both equal to materials misallocation, as in Row 3, using counties' materials wedge to proxy for their capital and labor wedge (adjusting county revenue shares for capital and labor to equal their output elasticities divided by the materials wedge. Rows 5 and 6 inflate firm capital costs, doubling or tripling the assumed interest rate to calculate annual capital expenditures. Row 7 uses the average national interest rate to calculate capital expenditures, rather than state-specific interest rates. Rows 8 to 10 report estimates when lowering the difference between measured county capital wedges and the median county capital wedges by 5%, 10%, or 25%, and Rows 11 to 13 report estimates when lowering the cross-county dispersion in all input wedges. Row 14 fixes county input wedges at their 1860 levels, rather than counties' average wedges over the 1860-1880 sample period, and fixes the scaling factor at its 1860 level of 4.9 (the average of county revenue divided by county productivity) and Row 15 does the same while also calculating counties' market access holding counties' population fixed at 1860 levels. Row 16 uses the median county scaling factor (4.9), rather than the average county scaling factor (5.1). Row 17 uses county-specific scaling factors, dropping counties with negative values and the top 1% of values. Rows 18 and 19 drop counties with the largest and smallest changes in productivity from 1860 to 1880: row 18 excludes the top and bottom 1% of counties, and row 19 excludes the top and bottom 5% of counties. Row 20 inflates firm labor costs, adding to county-by-industry labor costs the number of establishments multiplied by the average wage in that county and industry. Rows 21 and 22 modify our baseline assumption of constant returns to scale, and re-scale the cost shares to add up to 0.95 (row 21) or 1.05 (row 22). Rows 23 to 28 adjust the measurement of county output elasticities: averaging industry-level cost shares with weights equal to an industry's share of total expenditure in that county (rather than revenue); averaging over industry-level cost shares from only the most efficient counties (those with gaps within one standard deviation of zero); calculating leave-out elasticities based on industry-level cost shares in other counties (omitting own industries); calculating state-specific industry-level cost shares; and calculating county-industry cost shares (which imposes a constant wedge across inputs). Rows 29 to 31 modify the relative cost shares for each factor, inflating by 5 percentage points the cost shares of labor (row 29), materials (row 30), and capital (row 31), and proportionally reducing the cost shares of the other factors. Row 32 excludes the butter and cheese industry from the analysis, for which coverage in the Census of Manufactures changes from 1860 to 1870.

Appendix Table 14. Impacts of Market Access, Robustness (Measurement of Market Access)

		Estimated Impact of Market Access on:		
		Productivity	TFPR	AE
		(1)	(2)	(3)
1.	Baseline Specification	0.204	0.036	0.168
		(0.051)	(0.025)	(0.051)
2.	Dropping top/bottom centile, change in market access	0.190	0.044	0.146
		(0.060)	(0.030)	(0.064)
3.	Dropping top/bottom 5 centiles, change in market access	0.226	0.101	0.124
		(0.081)	(0.041)	(0.087)
4.	Reduces the cost of water to 0.198 cents per ton mile	0.173	0.028	0.145
		(0.050)	(0.025)	(0.045)
5.	Reduces the cost of wagons to 14 cents per ton mile	0.220	0.037	0.182
		(0.055)	(0.028)	(0.057)
6.	No transshipment costs between waterways	0.202	0.033	0.169
		(0.051)	(0.025)	(0.051)
7.	Include transshipment between Northern and Southern RRs	0.205	0.036	0.169
		(0.052)	(0.026)	(0.051)
8.	Raise railroad cost to 0.735 cents per ton mile	0.201	0.034	0.167
		(0.052)	(0.025)	(0.050)
9.	Raise railroad cost to 0.878 cents per ton mile	0.193	0.032	0.162
		(0.052)	(0.025)	(0.049)
10.	Average price of goods, $\overline{P}$, set to 20	0.205	0.037	0.169
		(0.051)	(0.026)	(0.051)
11.	Average price of goods, $\overline{P}$, set to 50	0.203	0.035	0.168
		(0.052)	(0.025)	(0.051)
12.	Trade elasticity, Θ, set to 1.95	0.203	0.036	0.168
		(0.051)	(0.025)	(0.051)
13.	Trade elasticity, Θ, set to 3.925	0.205	0.036	0.169
		(0.051)	(0.025)	(0.051)
14.	Trade elasticity, Θ, set to 8.22	0.208	0.037	0.171
		(0.051)	(0.026)	(0.051)
15.	Include access to international markets	0.203	0.035	0.167
		(0.051)	(0.025)	(0.051)
16.	Adjustment for Census undercounting	0.204	0.036	0.168
		(0.051)	(0.025)	(0.051)

17. Measure access to county wealth	0.197	0.033	0.164
	(0.049)	(0.025)	(0.050)
18. Include access to own market	0.205	0.036	0.170
	(0.051)	(0.025)	(0.051)
19. Limit access to counties beyond 5 miles	0.203	0.036	0.167
	(0.051)	(0.025)	(0.051)
20. Limit access to counties beyond 50 miles	0.200	0.036	0.164
	(0.051)	(0.025)	(0.051)
21. Limit access to counties beyond 100 miles	0.196	0.034	0.161
	(0.050)	(0.025)	(0.050)
22. Limit access to counties beyond 200 miles	0.185	0.032	0.153
	(0.048)	(0.024)	(0.049)
23. Fix Population to 1860 Level	0.196	0.035	0.161
	(0.049)	(0.025)	(0.049)

Notes: Row 1 reports our baseline estimates, from Table 2, for the impacts of market accesss on county productivity (column 1) and the impacts through changes in county TFPR or revenue total factor productivity (column 2) and county AE or allocative efficiency (column 3). All regressions include county fixed effects, state-by-year fixed effects, and year-interacted cubic polymials in county latitude and longitude. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880. The sample is a balanced panel of 1,802 counties (1860, 1870, 1880). Robust standard errors clustered by state are reported in parentheses.

Rows 2 to 23 report alternative estimates, which generally relate to adjusting our measurement of market access. Rows 2 and 3 drop counties with the largest and smallest changes in market access from 1860 to 1880: row 2 excludes the top and bottom 1% of counties, and row 3 excludes the top and bottom 5% of counties. Row 4 reduces the cost of water transportation from 0.49 cents per ton mile to 0.198 cents per ton mile, and row 5 reduces the cost of wagon transportation from 23.1 cents per ton mile to 14 cents per ton mile. Row 6 removes transshipment costs (50 cents) when transfering goods within the waterway network. Row 7 adds transshipment costs between Northern and Southern railroads, and rows 8 and 9 raise to cost of railroad transportation (from 0.63 cents per ton mile to 0.735 cents or 0.878 cents) to reflect general congestion or indirect routes along the railroad network (as considered in Donaldson and Hornbeck 2016). Rows 10 and 11 replace our baseline estimated average price of transported goods (38.7) with alternative assumed values of 20 or 50. Rows 12 and 13 replace our baseline estimated trade elasticity (3.05) with alternative assumed values that reflect its estimated 95% confidence internal (1.95 to 3.925), and row 14 assumes a value of 8.22 from Donaldson and Hornbeck (2016). Row 15 adjusts our measurement of counties' market access to reflect access to international markets, inflating the population in counties with major international ports based on the value of imports and exports (scaled by GDP per capita). Row 16 adjusts counties' population for different under-enumeration rates in the Census of Population, by decade and region. Row 17 measures counties' market access based on their access to other counties' wealth, rather than other counties' population. Row 18 includes counties' own population in their market access, and Rows 19 to 23 measure counties' market access when excluding other counties within 5 miles, 50 miles, 100 miles, or 200 miles. Row 23 reports estimates using a measure of counties' market access in each decade that holds counties' population levels fixed at 1860 levels.

Appendix Table 15. Impacts of Lagged Market Access on Manufacturing Employment

	Log Manufacturing Employment	
	(1)	(2)
Log Market Access	0.201	0.238
	(0.078)	(0.078)
Lagged Log Market Access	0.069	0.053
	(0.038)	(0.044)
State-Year FE	Yes	No
Number of Counties	1,437	1,438
County-Year Obs.	7,185	7,190

Notes: Column 1 reports estimated impacts of market access on manufacturing employment, from estimating equation 4 with the additional inclusion of decade lagged log market access as a regressor. Column 2 reports estimates from this same specification, but replacing state-year fixed effects with year fixed effects.

The sample is the balanced panel of 1,438 counties with manufacturing output in 1850 through 1900, with data on manufacturing employment starting in 1860. Column 1 has one fewer county observation because one state has only one county in this sample.

In each column, we report the estimated impact of a one standard deviation greater change in market access from 1860 to 1900. Robust standard errors clustered by state and county are reported in parentheses.

Appendix Table 16. Impacts of Market Access on County Revenue and Productivity, Interacted with Initial Distortions

	Include Interaction	County Industry FE
	(1)	(2)
Panel A. Log Revenue		
Log Market Access	0.118	
	(0.036)	
Interaction	-0.0002	0.0005
	(0.0007)	(0.0008)
Panel B. Log Productivity		
Log Market Access	0.092	
	(0.042)	
Interaction	0.0396	0.0418
	(0.0048)	(0.0053)
Number of Counties	1,800	1,442
County-Year Obs.	16,685	15,272

Notes: this table reports estimates from regressions at the county-by-industry level, after aggregating the more-detailed industries to five industry groups: clothing, textiles, leather; food and beverage; lumber and wood products; metals and metal products; and other industries. We extend our baseline estimating equation 4 to include county-industry fixed effects and state-year-industry fixed effects. Column 2 additionally includes county-year fixed effects. The interaction is market access multiplied by the (log) of each county's elasticity (α^k) weighted wedges (ψ^k), as in Hsieh and Klenow (2009): $\Sigma\alpha^k \ln(1+\psi^k)$. The sample is a balanced panel of 1,800 counties with a panel of county/industries, as in Table 4. Column 2 has fewer counties since counties with only one industry group are absorbed by the county-industry fixed effect. We continue to report the estimated impact of a one standard deviation greater change in market access from 1860 to 1880. Robust standard errors clustered by state are reported in parentheses.

Appendix Table 17. Counterfactual Impacts on National Aggregate Productivity, with Varying Aggregate Population Declines

	Percent Decline in Population:				
	Fixed Worker Utility: 66% Decline	Restricted Worker Mobility: 51% Decline	Exclude Foreign Born and Children: 33% Decline	Exclude Foreign Born: 15% Decline	Fixed Population: 0% Decline
	(1)	(2)	(3)	(4)	(5)
Change in Aggregate Productivity	-26.7%	-20.1%	-14.2%	-9.0%	-5.5%
Change in Worker Utility	0.0%	-12.8%	-22.1%	-28.7%	-32.7%

Notes: Each column reports estimated changes in national aggregate productivity and worker utility, in a counterfactual without the 1890 railroad network, under alternative scenarios for the aggregate decline in US population. In all scenarios, population is allowed to relocate endogenously within the country. The counterfactual sample includes all 2,722 counties that in 1890 report positive population and positive revenue (agriculture and/or manufacturing).

Column 1 reports impacts for a 65.89% decline in aggregate population, which reflects the model-predicted decline holding fixed worker utility. Column 2 reports impacts for a 50.73% decline in aggregate population, which reflects 77% of the population decline in column 1 based on the contemporaneous response of manufacturing employment to market access (Appendix Table 15) as a share of the model-predicted response (Table 8, column 2). Column 3 reports impacts for a 33.07% decline in aggregate population, which reflects the removal of the foreign born population in 1890 and the native born children of white foreign born parents. Column 4 reports impacts for a 14.69% decline in aggregate population, which reflects the removal of the foreign born population in 1890. Column 5 reports impacts for a fixed aggregate population.