Transportation Networks and the Geographic Concentration of Employment*

Dustin Frye†

August 10, 2023

KEYWORDS: interstate highways, local labor markets, spatial spillovers

JEL CLASSIFICATION: N72, N92, O18, R12, R23, R40

*The earlier title of this paper is "Transportation Networks and the Geographic Concentration of Industry". Thanks to Nate Baum-Snow, Taylor Jaworski, Carl Kitchens, Murat Iyigun, Lee Alston, Jonathan Hughes, Carol Shiue, Ann Carlos, Edward Kosack, Zachary Ward, Steven M. Smith, and Gisella Kagy, and conference/seminar participants at Vassar College, the US Census Bureau, Haverford College, Columbia University and Barnard College, Northwestern University, University of Wisconsin-Madison, Queens University, the Urban Economics Association Annual Meeting, the Economic History Association Annual Meeting, the All UC Graduate Student Dissertation Conference (UC Davis), and the CSMGEP Dissertation Session AEA Annual Conference for helpful comments and suggestions. I would like to thank Richard Weingroff for generously sharing the PR-511 reports and answering my many questions about it. I would also like to thank Courtney Geiss, Juan Felipe Laso, Gordon Lin, and Nina Zachariah for excellent research assistance. Appendix materials are available through the research page of my personal website. All errors are my own.

†Department of Agricultural and Applied Economics, University of Wisconsin - Madison, 427 Lorch St, Madison, WI 53706 (e-mail: dustin.frye@wisc.edu; website: http://dustinfrye.github.io/)

Abstract

This paper examines the effect of expanding transportation networks on spatial industrial growth across the United States from 1953 to 2016. I use a new methodological approach that applies network theory combined with a historic military map to address the two forms of endogeneity present in expanding transportation networks: route placement and construction timing. I find that Interstate counties experienced significant growth in employment and the number of establishments relative to non-Interstate counties. Growth rates are highest within two decades of receiving an Interstate. Results also reveal positive spillovers occurred in later decades among adjacent counties along the metropolitan periphery.

1 Introduction

Numerous policies focus on stimulating economic development in communities. Evaluating these policies depends on exogenous variation across locations, but when policymakers apply selective resource distribution over time it complicates the evaluation by introducing an additional source of endogeneity. For example, the construction of the Interstate Highway System in the United States, a major place-based policy, involved selective route determination by Congress and allocation of construction funds over several decades by state governments. This paper introduces a novel methodology that applies network theory to address endogeneity along both dimensions. I apply this new approach to estimate the impact of Interstates on employment and establishment growth in the entire U.S. from 1956 to 2016.

Interstate construction introduced over 40,000 miles of limited access highways, lowering travel costs and improving travel times. By the end of the twentieth century, Interstates had reshaped cities by altering the location choices of workers and firms (Baum-Snow, 2007, 2020; Duranton and Turner, 2012), encouraged trade by connecting regions and international markets (Duranton et al., 2014; Jaworski et al., 2020; Michaels, 2008), and raised aggregate welfare (Allen and Arkolakis, 2014, 2019). Despite our growing knowledge, limitations in data and empirical approaches have restricted our ability to assess spatial spillovers from Interstates and their long-term impact on industrial growth.

I use a reduced form analysis and instrumental variable approach to estimate the impact of an Interstate's presence on changes in employment and establishments, addressing both timing and location endogeneity. Non-random placement of Interstates has been highlighted in previous literature, which shows that Interstates were often directed to struggling metropolitan areas to encourage economic growth (Duranton and Turner, 2012; Redding and Turner, 2015). In line with previous literature, I use a proposed but never constructed historic highway plan as an instrument for eventual Interstate locations. Once routes were established, state and local officials were in charge of determining when particular segments would be constructed, providing another means of encouraging local development. To address endogenous timing, I implement the Newman-Girvan algorithm on the historic map to predict the timing of Interstate construction. The algorithm prioritizes segment construction based on their importance for network connectivity. I use this priority ranking with a simple social planner's problem to predict the construction year for each proposed Interstate segment from the historic plan.

To address concerns about centrality in a network potentially influencing economic growth independently of Interstates, I apply a correction discussed in Borusyak and Hull (2021) within the estimation strategy. Additionally, I support the instrument's validity through an event study analysis, which confirms a null result for a pre-trends test in the reduced form and highlights the endogeneity issue in the OLS approach.

This analysis relies on a newly constructed county-year panel dataset covering 1953-2016. The dataset is compiled from County Business Pattern (CBP) data, providing employment and establishment counts for all counties in the lower-48 states, as well as the number of firms across eight employment size categories. Detailed information on Interstate construction timing is obtained from the National Archives.

In addition to the novel empirical methodology, this paper bridges the existing empirical literature, which has often focused exclusively on metropolitan areas or rural areas, by considering all counties in the lower-48 states.[1] Including this set of "missing middle" counties, which were not previously categorized as metropolitan or rural in existing analyses, leads to comprehensive and broadly applicable findings. Results show that counties with Interstate highways experienced higher year-over-year employment and establishment growth compared to counties without Interstates. Between 1956 and 2016, findings suggest that Interstates contributed to an increase of 2,400 to 4,200 jobs for the median Interstate county. Establishment growth was similar in magnitude, with Inter-

[1]For example, seminal work by Michaels (2007) analyzes ≈1000 rural counties, and Duranton and Turner (2012) analyze 227 Metropolitan Statistical Areas which is ≈ 900 urban counties. This paper uses ≈3,100 counties, which indicates a large number of counties (≈1000) have been excluded from prior analysis.

state contributing to between 20 and 40 percent of annual growth. However, growth was concentrated among larger establishments with over 50 employees and came at the expense of smaller firms with fewer than 20 employees, suggesting that establishments are able to scale up.

This paper's inclusion of all counties in the lower 48-states allows for insights into the spatial spillovers from Interstates. Findings indicate Interstates generate spatial spillovers into the non-Interstate metropolitan periphery. This growth aligns with decentralized suburban job growth and strengthens the connection between Interstates and agglomeration spillovers at the labor market level (Baum-Snow, 2020). I find no evidence of spillovers into other adjacent areas, suggesting that the negative spillover effects on Interstate adjacent counties documented in Chandra and Thompson (2000) are limited to shifts between Interstate adjacent rural areas.

By using a time-varying instrument and annual outcome data, I analyze three distinct eras: the initial expansion (1956-1975), the completion (1976-1995), and the post-construction era (1996-2016). This approach reveals new dynamic patterns across eras and locations. Employment and establishment growth in Interstate counties is highest during the initial expansion era, while the non-Interstate metropolitan periphery experiences stronger growth in later decades. The results indicate significant localized industrial growth due to the Interstates, despite their initial disconnection during the expansion era.

This paper provides a solution for multidimensional endogeneity that is broadly applicable to settings where the allocation of resources can be patterned according to a network, where the most direct applications include other forms of infrastructure such as rail, airports, or electricity. This analysis enables a comparison between my instrument, which explicitly addresses the endogenous timing of construction, and instruments that do not. Results show that using an instrument solely for location endogeneity underestimates the impact of transportation on growth, especially during the period of Interstate expansion. This finding highlights the potential for similar types of endogeneity in other contexts including other place-based policies with bureaucratic discretion over funding allocation.

2 Data and Descriptive Evidence

2.1 Dataset Construction

My analysis examines Interstate highways' impact on industry dynamics using Census Bureau's County Business Patterns (CBP) data from 1953 to 2016, alongside historic transportation network information. I compiled CBP data from three sources: pre-1964 records were hand collected from archival documents, 1964-1970 data came from ICPSR (Ody and Hubbard, 2011), and subsequent years are published by the US Census Bureau.[2] The County Business Patterns consistently tabulate total employment, establishments, and the number of establishments in six employment size categories, which allows me to consider changes in the share of establishments within each of the size categories.[3] I construct annual information on the location and timing of the construction of the Interstate Highway System, depicted in Appendix Figure B.1.1a, by combining contemporary route locations with the PR-511 collection at the National Archives.

To account for factors that are correlated with changing industrial dynamics and the expansion of Interstate highways, I include data covering population, market size, geography, military installations, and alternative methods of transportation. Appendix A.3 provides a detailed description of the variables and how they are constructed.

[2]For more detail regarding CBP dataset construction and Interstate highway maps see Appendix A.

[3]Employment size groups typically include eight categories: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, and above 500 employees. The thresholds of smaller categories change across eras, so I combine them into 1-19 employees.

2.2 Growth Over Time and Evidence of Selection

From the three sources above, I construct a panel dataset from 1953–2016 to evaluate the effects of Interstate highways on employment and establishment growth across counties. Figure 1 plots the mean changes in both outcomes relative to 1953 by eventual Interstate status. The employment trends in panel (a) and establishments trends in panel (b) reveal both Interstate and non-Interstate counties grew over the period, with Interstate county growth outpacing non-Interstate growth. Appendix Table C.1.1 reports summary statistics, revealing average annual employment growth of 2.2 percent in Interstate counties compared to 1.7 percent in non-Interstate counties, and establishment growth of 1.5 percent in Interstate counties versus 0.9 percent in non-Interstate counties.

Panels (c) and (d) of Figure 1 plot the change in employment and establishments relative to the initial construction year among Interstate counties. Each figure plots the growth relationship by the decade of construction. Both panels reveal that by 2016, counties that constructed their Interstates earlier experienced more relative growth. The figures also suggest that growth was similar among counties that were built prior to the mid-1970s, but was much slower among those counties that were built in later decades.

This faster growth among highway counties is coupled with significantly higher levels of initial employment. Appendix Table C.1.2 presents summary statistics of the pre-Interstate county characteristics by eventual highway type. Not surprisingly, areas that built Interstate highways are considerably different from those that did not. The table shows highway counties were more populated, had a higher share of urban population, and had better access to alternative forms of transportation. These differences reinforce the selection concerns regarding Interstate locations.

2.3 Defining Spatial Adjacency

I assess the importance of adjacency by leveraging the national coverage of the data and by considering spatial spillovers within labor markets, which are defined by commuting zone boundaries according to Economic Research Service (2019). Figure 2 illustrates the spatial adjacencies of interest. The figure plots four commuting zones around Houston, TX. Each commuting zone is outlined by a dark bold line, and the interior of each commuting zone includes the county boundaries. The dashed line marks the eventual path of the Interstate highway and the light gray shaded counties are all Interstate counties. The dark gray shaded counties include both Interstates and the Houston MSA boundaries in 1960.

The lettered counties define the spatial adjacencies of interest. Counties labeled *A* and *B* are defined as Interstate adjacent within the commuting zone, sharing the same local labor market as Interstate treated counties.[4] These are different than the counties labeled with the letter *C*, which are a collection of non-Interstate counties, where none of the counties in the commuting zone were connected to the Interstate system. Finally, I distinguish commuting zone adjacent county types *A* and *B* because type *A* is in a commuting zone that includes an MSA. I define these as non-Interstate counties in the metropolitan periphery. I incorporate these spatial adjacencies directly into the analysis by including binary adjacency treatments in some estimation specifications.

3 Empirical Strategy

3.1 Estimating Equation

To estimate the effect of the Interstate Highway System on year over year employment and establishment growth, I exploit spatial and temporal variation in the location of In-

[4] The average adjacent county is 46 kms from an Interstate, whereas the average CZ adjacent county is 36 kms away.

terstate highways using the following first-difference specification:

$$ln(Y_{ct}) - ln(Y_{ct-1}) = \beta \cdot IHS_{ct} + \delta_{st} + \theta_{mt} + \gamma_{dt} + X'\rho_{ct} + \epsilon_{ct} \quad (1)$$

where Y_{ct} is the outcome of interest in county c at time t. The variable IHS_{ct} is an indicator that is equal to one if an Interstate highway intersects county c at year t. The coefficient of interest is β, which estimates the average effect of the Interstate highway system. The specification includes state $\times$ year fixed effects, δ_{st}, so the treatment effect of an Interstate highway is identified using variation within a state in a year. It includes market size $\times$ year fixed effects, θ_{mt}, to flexibly account for ways that metropolitan and rural areas grow over time. It also includes CBP adjustment $\times$ year fixed effects, γ_{dt}, to account for minor differences in CBP reporting, d, each year. Standard errors are two-way clustered by county and commuting zone $\times$ year to account for serial correlation and spatial correlation in the error term (Kelly, 2019, 2020).[5]

Although the first-difference specification accounts for fixed county level characteristics and includes a wide range of flexible time varying fixed effects, a concern remains that time varying county characteristics might be correlated with both Interstate highways and industrial growth. To account for this, I include a set of controls, $X'\rho_{ct}$, that flexibly account for preexisting differences. The controls, which are described in detail in Appendix A.3 and are presented in Appendix Table C.1.2, include baseline employment, establishments, and population measures, as well as additional measures to account for urbanization, transportation infrastructure, and geographic characteristics. This specification identifies the effect of an Interstate highway for Interstate counties relative to non-Interstate counties while allowing for subsequent endogenous policy decisions (Kline and Moretti, 2014). This model does not allow me to separately identify the effect of Interstates on growth and relocation, but rather relative differences between the two types of

[5]OLS results are robust to using spatially weighted standard errors proposed by Conley (2010) and Hsiang (2010).

locations.[6]

3.2 Addressing Highway Endogeneity

The history of highway construction indicates that route placement, construction timing, and funding of highways was an intensely political process (Rose, 1990). Politicians, lobbyists, and heads of industry all contributed to the current locations of Interstate highways and state and local officials were in charge of allocating resources for construction (Kaszynski, 2000; Lewis, 1997). If these outside contributors viewed highway construction as a place-based economic development policy, they may have been more likely to add segments of Interstate or reroute planned segments to reach less developed counties or start construction earlier to promote more growth. Therefore both location choice and timing of construction are potentially endogenous. To account for this, I construct an instrument that predicts both the location of an Interstate and the year of construction.

To address endogeneity concerns regarding Interstate locations, I use the 1921 national network plan developed by the War Department under the supervision of General John J. Pershing as an instrumental variable to predict eventual Interstate locations. This plan is commonly referred to as the Pershing Map and was designed to prioritize the military needs of the early-1920s (Michaels et al., 2019). Proposed Interstate locations, depicted in Figure B.1.1b, are based on the digitized Pershing Map from the Bureau of Public Roads collection at the National Archives.[7]

I address the endogenous timing of Interstate construction using an application from network theory to predict the optimal timing of Interstate construction. I implement the Newman-Girvan Algorithm to determine a construction timing priority for each segment

[6] For recent examples of quantitative spatial models across different forms of transit infrastructure see Donaldson and Hornbeck (2016); Baum-Snow et al. (2017); Jaworski and Kitchens (2019); Jaworski et al. (2020); Jedwab and Storeygard (2022); Fajgelbaum and Schaal (2020).

[7] The commonly used instrument in the U.S. Interstate literature is the National Interregional Highway Committee from 1947. However, in Section 3.3.4, I discuss its limitations when applied to a national setting with extensive margin treatment.

of the proposed Interstate networks.[8] In order to apply the algorithm to the Pershing Map, I decompose the proposed routes into a mathematical network of nodes and edges, where each node occurs at the intersection of two edges or at the end of an edge. Each edge is weighted by its length in kilometers. The Newman-Girvan Algorithm calculates edge-betweenness for every edge by finding the shortest paths between all node pairs and counting the number of times each edge is used for these trips. Edges with higher betweenness values are more important for connecting the network.

Figure 3 presents two stylized highway graphs to illustrate calculating edge-betweenness. The first panel presents a simplified highway network with ten cities (nodes) connected by thirteen Interstates (edges). The approximate mileage between each node is printed along each edge. Consider an example trip between New York and El Paso. The shortest path between these two nodes passes through Cheyenne and covers 2,460 miles, which is slightly shorter than passing through Jacksonville (2,540 miles). So the edge-betweenness value would increase along the edges from New York to Cheyenne and from Cheyenne to El Paso. This process gets repeated between every pair of nodes. The second panel presents the resulting edge-betweenness calculation for this network, where the betweenness score is presented both as the value on the edge and illustrated by the thickness of the edge, where thicker edges have higher betweenness scores.

To predict a construction year for each Interstate segment, my procedure sequentially builds the network edges with the highest betweenness value subject to an annual construction budget. I derive the annual budget appropriation based on estimated construction costs of the entire network equally divided over a fixed construction time horizon. I calculate the total network construction cost by estimating construction costs for each segment based on weighted average costs of the urban and rural mileage for that segment. I use construction cost estimates for urban and rural cost per mile from a 1955

[8]This algorithm was originally used to identify important connections in biological and social networks (Girvan and Newman, 2002).

Congressional highway proposal.[9]

To determine the annual construction constraint, I divide the total network construction cost over a twenty five period, approximating actual Interstate construction for the Pershing Military Plan. I then rank proposed network edges based on betweenness scores and build them in that order until the construction spending matches the annual constraint. Unbuilt edges carry over to the next year, and the process repeats, providing a construction year for each edge and creating an Interstate instrument predicting both location and year of construction. Appendix Figure B.1.2 illustrates how the proposed construction horizon compares to the timing of construction for the Interstate Highway System. The figures show that the twenty five year construction horizon closely matches the actual construction horizon.

3.3 Instrument Validity

For my proposed instrument to be valid it must be correlated with the endogenous variable and also only impact the outcomes of interest via its impact on the endogenous variable. In this section, I discuss the first-stage relationship, how the different components of the exclusion restriction are plausibly satisfied, and present three empirical falsification tests that support using the Pershing Map as a valid instrument.

3.3.1 First-Stage

To test whether the proposed instrument, that is a network of roads with the associated year of construction, sufficiently predicts whether a county will have an Interstate highway at time t, I estimate the following first-stage regression:

$$IHS_{ct} = \alpha \cdot Plan_{ct} + \lambda_{st} + \tau_{mt} + \sigma_{dt} + X'\mu_{ct} + \upsilon_{ct} \quad (2)$$

[9]Estimates derived from House Document 120, submitted to the 84th Congress during the first session. Urban mileage had an estimated cost of \$2,431,818 per mile, while rural costs are significantly lower at \$378,787 per mile, both in 1955 dollars.

The variable $Plan_{ct}$ is an indicator for whether a county c is predicted to have an Interstate from the proposed network in year t. The specification includes the full set of fixed effects and controls from equation 1. First-stage coefficients and standard errors are included at the bottom of each set of results in the main specifications. Kleibergen-Paap F-statistics are reported with every specification (Stock and Yogo, 2005). The first stage is consistently strong with Kleibergen-Paap F-statistics ranging from 30 - 160.

3.3.2 Exclusion Restriction: Pershing Route Locations

There are several advantages in the design of the Pershing Map that support the argument for the exclusion restriction. One advantage is the strong military influence in creating the map and the lack of input from outside political and economic agents. The Pershing plan originated shortly after World War One and the legacy of the domestic war efforts are evident in the route decisions, which did not extend into southern Florida and emphasized coasts and borders (Swift, 2011). The system was designed with straight line connections, avoiding city centers, and creating a network-style graph akin to those evaluated by Banerjee et al. (2020) and Faber (2014). Straight line connections remove the possibility of local officials subtly manipulating the locations of the plan.

My empirical strategy continues to condition on historic population, economic conditions, transportation, geography, and spatial controls to account for county characteristics that may be correlated with employment growth and industrial development. Additionally, I construct four controls to explicitly account for military interests in the early 1920s. First, I include nodal fixed effects for counties with nodes in the Pershing Map. Second, I control for distance from county seats to the nearest node. Third, I calculate proximity to nearly 700 World War One posts, camps, and stations in the US in 1918 (Center of Military History, 1931). Finally, I account for distances to pre-World War One military conflicts. Each of these time invariant controls are interacted with year fixed effects to flexibly account for changes over time. Given the historic narrative, limitations of military planners to forecast the mid-twentieth century economic environment, and the rich set of covariates, it seems plausible that the Pershing Map is a suitable instrument for Interstate locations.

3.3.3 Exclusion Restriction: Centrality and Timing

Predicted construction timing relies on centrality in the proposed Interstate network. This approach abstracts from the endogenous state level decision-making process. Borusyak and Hull (2021) show that centrality partially drives the relationship between transportation expansion and regional economic growth. I address this potential confounding relationship by constructing a centrality fixed effect based on a 1947 Federal and State Highway map that is in the spirit of the correction proposed by Borusyak and Hull (2021).[10] With these centrality fixed effects in all specifications, the comparisons are identified from variation in Interstate treatment within similarly central counties. To address the conceptual issue differently, I add a continuous market potential control to all specifications based on Harris' approach (1954), using 1950 population and straight line distances between county seats.

3.3.4 Evidence in Support of the Exclusion Restriction

Effects Prior to Construction: An empirical concern is that influential route designers simply identified places that were poised for growth. To directly test whether Pershing routes affected employment prior to construction, I construct a county level panel dataset from 1930 to 1954 from Haines et al. (2010)[11] and estimate equation 1, using a time-invariant binary Interstate highway indicator, with the full set of controls and fixed effects outlined in Section 3.1 and the new covariates and fixed effects introduced in Section 3.3. The outcomes of interest are decade over decade changes in total county employment, bank deposits, and the number of firms engaged in manufacturing, retail sales, and

[10] Appendix B.2 describes the measure in detail.

[11] Coverage includes each decade from the Decennial Census and the 1954 County Data Book.

wholesale trade.

Panels (A)–(B) of Appendix Table B.3.1 report OLS results and IV results using the Pershing IV. OLS coefficients in Panel (A) are positive and statistically significant for several outcomes reflecting possible selection of faster growing counties or anticipatory growth based on expectations of highway development. Reassuringly, coefficient estimates using the Pershing locations as an instrument in Panel (B) are not statistically significant. Panel (C) replicates the IV results using the 1947 Interregional Highway Committee Plan (See Appendix Figure B.1.1c). The significant coefficients on employment and retail sales raise concerns that the 1947 Plan does not sufficiently address the endogeneity concerns when considering a national set of counties and measuring highway treatment through a binary indicator.[12]

Finally, as noted in Section 3.3.3, there is concern that centrality may have a direct effect on economic growth. I replace the binary Interstate indicator with a continuous Pershing centrality measure to assess if centrality predicts pre-Interstate economic growth. Panel (D) presents results and across all five outcomes, the centrality coefficients result in precisely estimated zeros, indicating that centrality is not associated with growth in these outcomes prior to Interstate construction.

Event Study: To enhance the instrument's credibility, it's important to verify that the predicted Interstate locations and timing are not correlated with pre-construction growth in the construction era. To test this, I leverage the recent advances in event-study methodologies by Goodman-Bacon (2021) and Schmidheiny and Siegloch (2020) in conjunction with the new county-year panel data. Specifically, I estimate the following linear panel model with dynamic highway treatment effects:

$$ln(Y_{ct}) = \sum_{m=-20}^{60} \beta_m hwy_{c,t+m} + \alpha_c + \delta_{st} + \theta_{mt} + \gamma_{dt} + X'\rho_{ct} + \epsilon_{ct} \tag{3}$$

[12]Prior work used urban rays (Agrawal et al., 2017; Baum-Snow, 2007), urban mileage (Duranton and Turner, 2011, 2012), or binary treatment in rural areas (Herzog, 2021; Michaels, 2008).

The estimated parameters of interest, β_m, separately identify binary highway treatment effects for each period beginning 20 years prior to construction and extending 60 years after construction. Following the conventions in Freyaldenhoven et al. (2021), I exclude the year prior to the Interstate opening, $t-1$, so each coefficient is interpreted as the effect of highways m years after the Interstate opens. The specification includes county fixed effects, α_c, and the full set of time varying fixed effects and controls from equation 1 and Section 3.3. Consistent with the prior specifications, standard errors are two-way clustered by county and commuting zone $\times$ year to account for serial correlation and spatial correlation in the error term.

Event study figures show there is no relationship between employment and establishment growth prior to the predicted construction of an Interstate. Figures B.3.1 and B.3.2 plot the coefficients, β_m, and 95% confidence intervals that result from estimating equation 3 where the outcomes of interest are the log of employment and log of establishments, respectively. I present sub-figures for actual and predicted Interstate locations using the instrument, both for population-weighted and unweighted outcomes. All figures display a positive effect of Interstates on the outcome after construction.[13] When using the instrument, estimated effects on the outcomes before the year of construction are not significantly different from zero, satisfying a necessary test of pre-trends and supporting the instrument's validity.

When the specification is run with population weights the endogeneity concern, first highlighted in Baum-Snow (2007) and Duranton and Turner (2012), that additional Interstate mileage was steered towards lower performing metropolitan areas, becomes more apparent. With population-weighted outcomes, actual Interstate construction shows a negative 'effect' before their arrival, indicating negative selection among higher population counties. However, this negative selection is not present when using the instrument.

[13]Confidence intervals widen near construction year endpoints due to fewer within-construction year observations, as seen in panels (e) and (f) of Figures B.3.1 and B.3.2. The reduction in pre-construction years is due to data gaps caused by the CBP being published every three years between 1953 and 1962.

The event study figures demonstrate that the instrument is not correlated with significant employment/establishment growth before predicted construction, even with population weighting.

4 Results

4.1 Interstates and Firm Dynamics

Table 1 reports OLS and IV results from estimating equation 1 after adding the new covariates and fixed effects described in Section 3.3 that address military motives and centrality concerns. Panel A shows total county employment, while Panel B displays total county establishments. The first two columns use a binary Interstate indicator, while columns three to eight decompose Interstate effects across space using spatial adjacency definitions from Figure 2. The lower section of each panel reports the Kleibergen-Paap F-Statistic, which indicates the instruments are sufficient predictors of highway status.

Interstates induced significant employment and establishment growth. Employment increased 0.4-0.7 percent faster yearly in Interstate counties relative to non-Interstate counties (columns 1 and 2 of Table 1). Given the average employment growth rate of 2.2 percent (see Appendix Table C.1.1), this suggests that between 18 and 31 percent of employment growth in those counties is attributable to Interstates. For the median Interstate county, between 1956 and 2016, employment increased by roughly 13,700 workers and the coefficient estimates suggest that Interstates were responsible for between 2,400 and 4,200 jobs. Similar results are seen for establishment growth, as the number of establishments grew 0.3–0.6 percent faster among Interstate counties (Panel B). With an average growth rate of 1.5 percent per year among Interstate counties, this indicates that between 22 and 40 percent of the establishment growth could be attributed to Interstates. Transit expansion may influence the size distribution of establishments if they promote economies of scale (You, 2021). I re-estimate equation 1 to focus on changes in the share of firms within each of the six employment size categories and find that Interstates promoted scale increases in employment, leading to fewer small firms and more growth among firms with 50-99 employees and 250-499 employees (Appendix Table C.1.3). The magnitudes of these estimates are meaningful relative to the average changes (Appendix Table C.1.4) and suggest that Interstates played a significant role in shaping firm location decisions and created a local environment that encouraged economies of scale.

Decomposing the effects of Interstates in space is important for understanding the spatial extent of transportation network benefits. I analyze the definitions of spatial adjacency, visualized in Figure 2, by adding binary adjacency interactions into equation 1. Columns (3)–(8) of Table 1 incorporate the three different adjacency interactions. Columns (3) and (4) examine spatial adjacency to treated highway counties. The specification includes binary interactions for Interstate and adjacent counties, with the comparison group being counties farther from the Interstate system. The specifications in columns (5) and (6) introduce a single binary interaction for adjacent counties within a commuting zone (types A and B from Figure 2). Columns (7) and (8) partition the commuting zone adjacency based on whether the commuting zone contains an MSA. This specification introduces two separate binary treatments for county types A and B. In columns (5)–(8), the relevant comparison set of counties are those in commuting zones with no Interstates.

Results highlight fundamental differences in the spatial spillovers of Interstates, depending on which type of adjacency is being considered. Results from columns (3)–(4) show there is no change in either employment or establishments for adjacent counties once endogeneity has been addressed. The null results persist when the definition of adjacency is altered to be non-Interstate counties within an Interstate treated labor market. However, there are significant positive spillovers within the metropolitan periphery. That is, once we partition counties into those that are adjacent within a labor market that includes an MSA and those that are adjacent within a labor market that does not, columns (7)–(8), there are significant positive impacts on industrial growth for both Interstate counties and non-Interstate counties in the metropolitan periphery.

4.2 Time Path of Treatment

I leverage the time-varying instrumental variable to explore the time path of Interstate development by analyzing three 20 year eras: the expansion era (1956–1975), the completion era (1976–1995), and the post-construction era (1996–2016). The expansion era had a largely disconnected network because construction decisions were made at the state level.[14] Roughly 75 percent of the system was complete by 1975, and the original plan was fully finished in 1992, followed by additional lane mileage in post-construction expansion (Turner et al., 2020).

I empirically incorporate these eras by modifying equation 1 as follows,

$$ln(Y_{ct}) - ln(Y_{ct-1}) = \sum_{e} \beta_e \cdot IHS_{ct} \times I_e + \delta_{st} + \theta_{mt} + \gamma_{dt} + X'\rho_{ct} + \epsilon_{ct} \tag{4}$$

where I interact the original Interstate measure with era indicators, I_e, to estimate the effect of Interstates separately by era. The rest of the estimating equation follows directly from equation 1.

Results reveal dynamic patterns of growth over time, with the initial expansion period experiencing the largest impact. Table 2, columns (1) and (2), report the three era specific coefficients of interest for OLS and IV specifications. OLS estimates suggest that Interstates led to stable gains in employment and establishments, with similar coefficients across the three eras. IV estimates reveal a different pattern, where the largest gains occurred during the expansion era (1956–1975). For employment, these elevated growth differences decrease slightly during the completion era before tapering off during the post-construction era. Establishment growth diminishes during the second era, before recovering in the final two decades.

[14] Appendix Figure B.1.2c illustrates the disconnected system.

I extend the analysis in columns (3) and (4) to allow for spatial spillovers in the metropolitan periphery by including era interactions of the non-Interstate counties within Interstate treated labor markets that contained an MSA (type *A* from Figure 2). Results reveal that both employment and establishment growth in the metropolitan periphery was strongest during the completion and post-construction eras. This suggests that Interstates initially promote spatially concentrated growth, but over time that growth can spread to non-Interstate parts of metropolitan areas.

Next, I compare results found with the new methodology that addresses both timing and location endogeneity with an instrument that addresses only location endogeneity, revealing that only addressing location endogeneity underestimates early era growth gains from Interstates. Table 2, columns (5) and (6), show OLS and IV results if a time invariant Interstate indicator is used. Comparing columns (2) and (6), results show smaller magnitude estimates using a fixed highway status, with the most pronounced differences during the expansion era.

This analysis reveals dynamic endogeneity patterns across eras. Larger IV estimates in early eras (Table 2, columns 2 and 6) suggest policymakers targeted lower performing areas with both Interstate routes *and* early construction. These results confirm the location selection found by Duranton and Turner (2012) and document that state and local officials amplified the negative selection through their decisions to prioritize construction in these same urban areas, likely contributing to the 1960s highway revolts (Brinkman and Lin, 2022).

5 Bridging Prior Literature

This paper ties together the seminal works of Duranton and Turner (2012) and Chandra and Thompson (2000) by filling the geography gap. The findings are closely related to Duranton and Turner (2012), which estimated the impact of metropolitan Interstate mileage on employment growth during the 1980s and 1990s. While my study doesn't directly estimate employment elasticities, using Table 2 estimates, I find that introducing an Interstate induced 14 percent more employment growth over a similar era (1976-1995).[15] I find similar results using Interstate density estimates in Appendix Table C.1.5, where I estimate four percent more employment growth following a one standard deviation increase in Interstate density. These estimates are similar to Duranton and Turner (2012), but show that strong employment growth extended beyond metropolitan areas and was strongest during the initial expansion era (1956-1975).

The national dataset used in this paper allows for the identification of spatial patterns that prior literature has not been able to consider due to data limitations. In a comparison among rural areas, Chandra and Thompson (2000) find that Interstates led to earnings declines in adjacent rural counties relative to non-adjacent rural counties. My paper expands on that result, extending adjacency to rural and non-rural counties, and shows that in a broader national sample, there is weak evidence of negative spillovers. The results in Table 1 instead point to positive spillovers in adjacent areas within labor markets that included a major metropolitan area.

6 Conclusion

Global investments in infrastructure have increased in recent decades and these investments have important consequences for the spatial distribution of economic activity within countries and regions (OECD, 2021). Evaluating the effects of network based infrastructure over time requires a new empirical approach that addresses the endogenous placement and timing of construction. Analyzing the Interstate Highway System, this study addresses both dimensions of endogeneity and finds early and persistent growth among connected counties, with much of the employment growth concentrated among larger

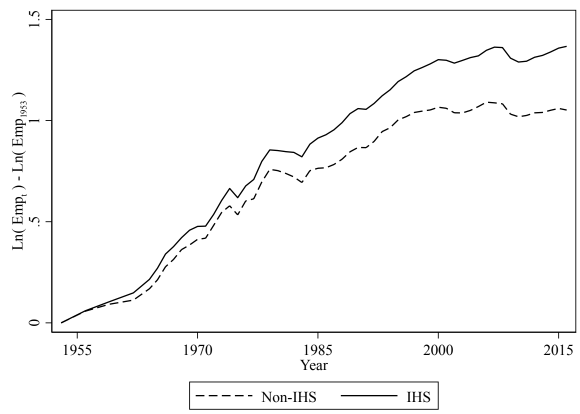
[15] Result is calculated by compounding the 0.008 percent growth from column (2) of Table 2 over 20 years.

firms. Interstates also induced positive spatial spillovers in employment and establishments among non-Interstate counties adjacent to metropolitan areas.

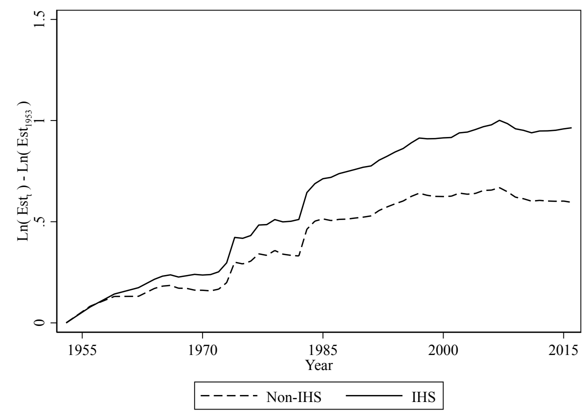
The methodology presented in this paper has the potential for broader applications beyond highways. In general, graph theory based algorithms could be valuable in scenarios involving interactions between places or individuals structured as a network, such as situations where firms are exchanging inputs and outputs or where governments are allocating project-based funding over time.

7 Figures

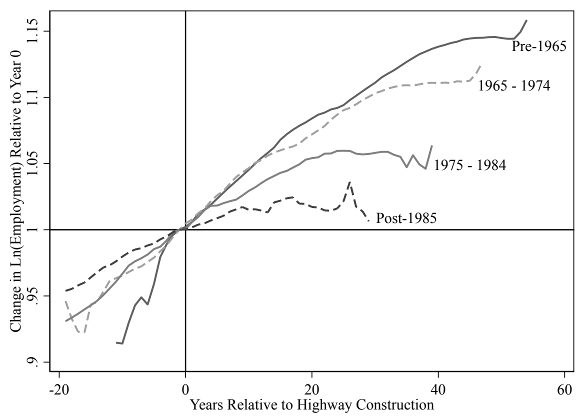
Figure 1: Illustrating Employment and Establishment Growth by Interstate Status and Construction Era



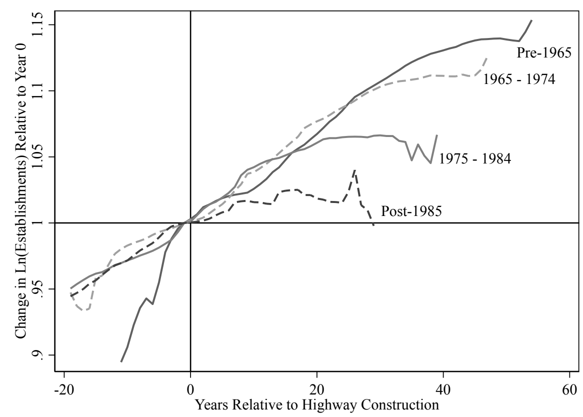
(a) Employment Growth by Interstate Status



(b) Establishment Growth by Interstate Status



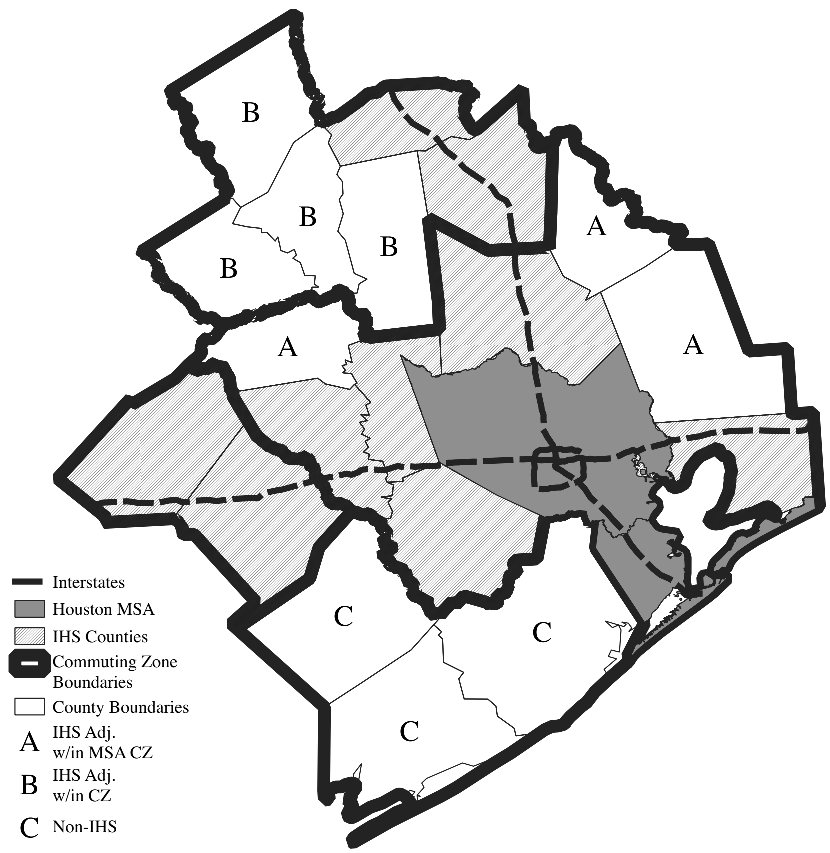
(c) Employment Growth by Interstate Construction Era



(d) Establishment Growth by Interstate Construction Era

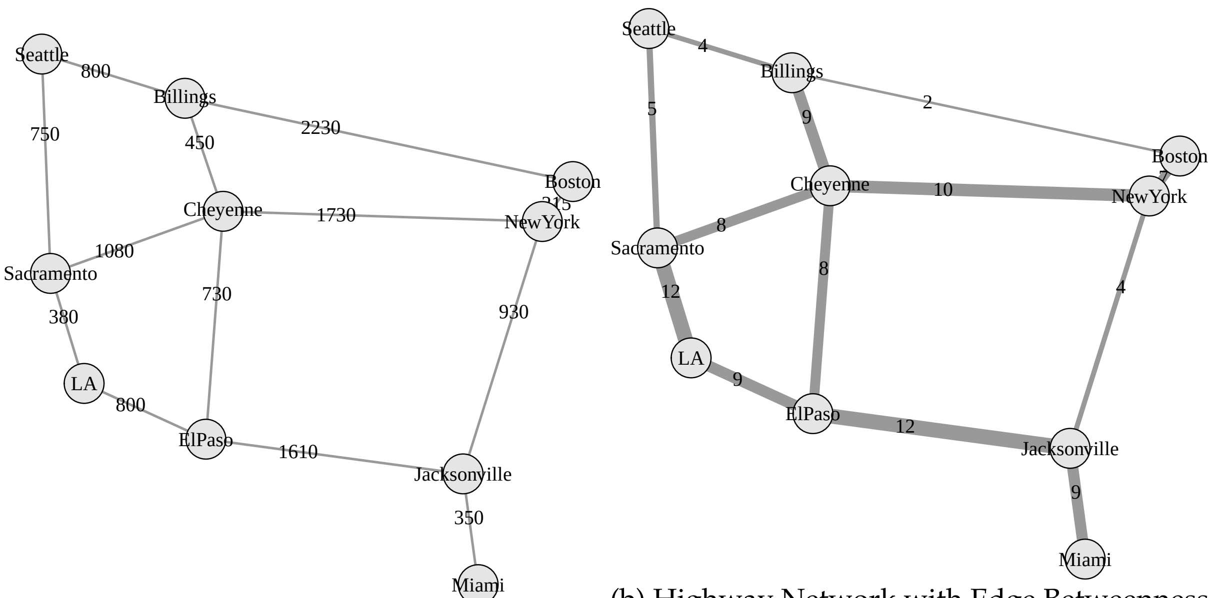
Notes: Employment and establishment data from 1953–2016 County Business Patterns annual reports. Highway designation based on highway status in 2016. Construction date information from PR-511 collection. Top row compares interstate and non interstate counties over time relative to their 1953 level of employment (a) or establishments (b). The bottom row looks compares growth in interstate counties over time based on the specific decade the interstates were constructed.

Figure 2: Example Spatial Structure



Notes: Figure presents the spatial structure around Houston, TX, which illustrates the typical adjacency structure. The small polygons indicate county boundaries, with the bold outlines mapping the commuting zones. Dark shaded counties indicate the Houston MSA boundaries in 1960. Interstate are recorded with the dashed line and interstate counties are shaded in light gray or dark gray. The lettered counties illustrate the three types of county adjacency identified in Section 2.3.

Figure 3: Illustrating Edge Betweenness



(a) Simple Highway Network with Distances

(b) Highway Network with Edge Betweenness Values

Notes: Figure presents stylized highway map illustrating the edge betweenness in a small network. Figure 3a presents the ten cities with approximate distances between cities listed along each edge. Figure 3b presents the edge betweenness calculation along each edge and adjusts the line width to reflect higher betweenness values.

8 Tables

Table 1: Interstate Highways and Growth in Employment and Establishments Across Locations

Panel A: Employment								
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Interstate (0/1)	0.004***	0.007***	0.003***	0.007**	0.004***	0.006*	0.005***	0.007**
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
IHS Adj			-0.001*	-0.000				
			(0.001)	(0.003)				
IHS Adj in CZ					-0.001*	0.001	-0.002***	-0.001
					(0.001)	(0.002)	(0.001)	(0.002)
IHS Adj in MCZ							0.002***	0.003***
							(0.001)	(0.001)
KP F-Stat		117.74		44.50		46.35		30.52
Panel B: Establishments								
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Interstate (0/1)	0.003***	0.006***	0.003***	0.006**	0.004***	0.005**	0.004***	0.006**
	(0.0004)	(0.002)	(0.001)	(0.002)	(0.000)	(0.003)	(0.000)	(0.003)
IHS Adj			-0.001*	-0.001				
			(0.000)	(0.002)				
IHS Adj in CZ					-0.001	0.001	-0.001***	-0.002
					(0.000)	(0.002)	(0.000)	(0.002)
IHS Adj in MCZ							0.003***	0.004***
							(0.000)	(0.001)
KP F-Stat		117.74		44.50		46.35		30.52

Notes: Every specification reports results from estimating equation 1, where the outcome of interest is the year over year change in either employment or establishments. Columns (1)–(2) report results with a binary interstate highway indicator. Columns (3)–(4) report results with a binary interstate highway indicator, plus an indicator for counties directly adjacent to Interstates. Columns (5)–(6) report results with both the binary indicator and an indicator for all broadly adjacent counties within an Interstate treated commuting zone (Type *A* & *B* in Figure 2). Columns (7)–(8) partition the broad adjacency to distinguish adjacent commuting zones that include an MSA (*Adj in MCZ*). These are type *A* in Figure 2. Every specification includes state × year fixed effects, along with the full set of controls outlined in Appendix A. Employment and establishment data are from 1956–2016 County Business Patterns annual reports. Panel A has 171,940 observations and Panel B has 171,938 observations. Every model covers 3,071 counties and includes 2,688 state × year fixed effects. Standard errors are two-way clustered by county and commuting zone × year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2: Interstates and Growth During Three Construction Phases

Panel A: Employment

	Time Varying Interstates				Fixed Interstates	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
IHS × I(56-75)	0.004***	0.013**	0.004***	0.013*	0.003***	0.009**
	(0.001)	(0.007)	(0.001)	(0.007)	(0.001)	(0.005)
IHS Adj. MSA CZ × I(56-75)			-0.001	0.002		
			(0.001)	(0.002)		
IHS × I(76-95)	0.003***	0.008**	0.003***	0.007*	0.003***	0.008**
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.004)
IHS Adj. MSA CZ × I(76-95)			0.002**	0.004***		
			(0.001)	(0.001)		
IHS × I(96-16)	0.005***	0.004	0.005***	0.003	0.005***	0.003
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
IHS Adj. MSA CZ × I(96-16)			0.002**	0.003***		
			(0.001)	(0.001)		
KP F-Stat		35.23		16.97		32.17

Panel B: Establishments

	Time Varying Interstates				Fixed Interstates	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
IHS × I(56-75)	0.003***	0.010**	0.004***	0.009**	0.003***	0.005*
	(0.001)	(0.004)	(0.001)	(0.004)	(0.001)	(0.003)
IHS Adj. MSA CZ × I(56-75)			-0.000	0.002*		
			(0.001)	(0.001)		
IHS × I(76-95)	0.003***	0.003	0.003***	0.002	0.003***	0.004
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
IHS Adj. MSA CZ × I(76-95)			0.003***	0.005***		
			(0.001)	(0.001)		
IHS × I(96-16)	0.004***	0.007***	0.003***	0.006**	0.004***	0.007***
	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)	(0.002)
IHS Adj. MSA CZ × I(96-16)			0.003***	0.003***		
			(0.001)	(0.001)		
KP F-Stat		35.23		16.97		32.17

Notes: Every specification reports results from estimating equation 1, where the outcome of interest is the year over year change in either employment or establishments. Columns (1)–(2) report results with a binary interstate highway measure interacted with mutually exclusive era indicators. Columns (3)–(4) extend the prior specifications to include era indicator interactions for the adjacent commuting zone counties that include an MSA. Columns (5)–(6) replace the time varying interstate indicator with a time invariant interstate indicator. For other details see Table 1.

References

Agrawal, A., A. Galasso, and A. Oettl (2017). Roads and innovation. *Review of Economics and Statistics 99*(3), 417–434.

Allen, T. and C. Arkolakis (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics 129*(3), 1085–1140.

Allen, T. and C. Arkolakis (2019). The welfare effects of transportation infrastructure improvements. Working Paper 25487, National Bureau of Economic Research.

Atack, J. (2016). *Historical Geographic Information Systems (GIS) database of U.S. Railroads for 1916*.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *The American Economic Review 103*(6), 2121–2168.

Banerjee, A., E. Duflo, and N. Qian (2020). On the road: Access to transportation infrastructure and economic growth in china. *Journal of Development Economics*, 102442.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics 122*(2), 775–805.

Baum-Snow, N. (2020). Urban transport expansions and changes in the spatial structure of us cities: Implications for productivity and welfare. *Review of Economics and Statistics 102*(5), 929–945.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang (2017). Roads, railroads, and decentralization of chinese cities. *Review of Economics and Statistics 99*(3), 435–448.

Borusyak, K. and P. Hull (2021). Non-random exposure to exogenous shocks: Theory and applications. Technical report, National Bureau of Economic Research.

Brinkman, J. and J. Lin (2022). Freeway revolts! the quality of life effects of highways. *Review of Economics and Statistics*, 1–45.

Center of Military History (1931). *Order of Battle of the United States Land Forces in the World War: Zone of the interior. pt. 1. Organization and activities of the War Department. pt. 2. Territorial departments, tactical divisions organized in 1918, posts, camps, and stations. pt. 3. Directory of troops*. CMH pub. Center of Military History, U.S. Army.

Chandra, A. and E. Thompson (2000). Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system. *Regional Science and Urban Economics*.

Conley, T. G. (2010). Spatial econometrics. In *Microeconometrics*, pp. 303–313. Springer.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics 131*(2), 799–858.

Duranton, G., P. M. Morrow, and M. A. Turner (2014). Roads and Trade: Evidence from the US. *The Review of Economic Studies 81*(2), 681–724.

Duranton, G. and M. A. Turner (2011). The fundamental law of road congestion: Evidence from us cities. *American Economic Review 101*(6), 2616–52.

Duranton, G. and M. A. Turner (2012). Urban Growth and Transportation. *The Review of Economic Studies 79*(4), 1407–1440.

Economic Research Service (2019). *Commuting Zones and Labor Market Areas*. U.S. Department of Agriculture.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from china's national trunk highway system*. *The Review of Economic Studies*, rdu010.

Fajgelbaum, P. D. and E. Schaal (2020). Optimal transport networks in spatial equilibrium. *Econometrica 88*(4), 1411–1452.

Freyaldenhoven, S., C. Hansen, J. P. Pérez, and J. M. Shapiro (2021). Visualization, identification, and estimation in the linear panel event-study design. Technical report, National Bureau of Economic Research.

Girvan, M. and M. E. J. Newman (2002). Community structure in social and biological networks. In *Proceedings of the National Academy of Science*, pp. 7821–7826.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277.

Haines, M. R. et al. (2010). Historical, demographic, economic, and social data: the united states, 1790–2002. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*.

Harris, C. D. (1954). The, market as a factor in the localization of industry in the united states. *Annals of the association of American geographers 44*(4), 315–348.

Herzog, I. (2021). National transportation networks, market access, and regional economic growth. *Journal of Urban Economics 122*, 103316.

Hines, F. K., D. L. Brown, and J. M. Zimmer (1975). *Social and Economic Characteristics of the Population in Metro and Nonmetro Counties, 1970*. Agricultural economic report. Economic Research Service, U.S. Department of Agriculture.

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

Hsiang, S. M. (2010). Temperatures and cyclones strongly associated with economic production in the caribbean and central america. *Proceedings of the National Academy of sciences 107*(35), 15367–15372.

Jaworski, T., C. Kitchens, and S. Nigai (2020). Highways and globalization. *International Economic Review*.

Jaworski, T. and C. T. Kitchens (2019). National policy for regional development: Historical evidence from appalachian highways. *Review of Economics and Statistics 101*(5), 777–790.

Jedwab, R. and A. Storeygard (2022). The average and heterogeneous effects of transportation investments: Evidence from sub-saharan africa 1960–2010. *Journal of the European Economic Association 20*(1), 1–38.

Kaszynski, W. (2000). *The American Highway: The History and Culture of Roads in the United States*. McFarland.

Kelly, M. (2019). The standard errors of persistence. Technical report.

Kelly, M. (2020). Understanding persistence. Technical report.

Kline, P. and E. Moretti (2014). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics 129*(1), 275–331.

Lewis, T. (1997). *Divided Highways: Building the Interstate Highways, Transforming American Life*. Viking.

Long, J. H. (1995). Atlas of historical county boundaries. *The Journal of American History 81*(4), 1859–1863.

Michaels, G. (2007). The long-term consequences of regional specialization.

Michaels, G. (2008). The effect of trade on the demand for skill: evidence from the Interstate Highway System. *The Review of Economics and Statistics 90*(4), 683–701.

Michaels, G., F. Rauch, and S. J. Redding (2019). Task specialization in us cities from 1880 to 2000. *Journal of the European Economic Association 17*(3), 754–798.

National Historical Geographic Information System (2011). *"National Historical Geographic Information System: Version 2.0"*. Minneapolis, MN: University of Minnesota.

Newman, M. E. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E 64*(1), 016132.

Newman, M. E. and M. Girvan (2004). Finding and evaluating community structure in networks. *Physical Review E 69*(2), 26113.

Ody, C. J. and T. N. Hubbard (2011). County Business Patterns, 1962, 1964-1970: U.S. Summary, State, and County Data. Inter-university Consortium for Political and Social Research [distributor].

OECD (2021). Infrastructure Investment (indicator. Technical report.

Redding, S. J. and M. A. Turner (2015). Transportation costs and the spatial organization of economic activity. In *Handbook of regional and urban economics*, Volume 5, pp. 1339–1398. Elsevier.

Rose, M. (1990). *Interstate: Express Highway Politics, 1939-1989*. Gaia future series. University of Tennessee Press.

Schmidheiny, K. and S. Siegloch (2020). On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization. *Journal of Applied Econometrics*.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg 1*.

Swift, E. (2011). *The Big Roads: The Untold Story of the Engineers, Visionaries, and Trailblazers Who Created the American Superhighway*. Houghton Mifflin Harcourt.

Turner, M., G. Duranton, and G. Nagpal (2020). Transportation infrastructure in the us. Technical report, National Bureau of Economic Research.

U.S. Census Bureau (1958). *Statistical Abstract of the United States*. Number 79. U.S. Government Printing Office.

You, W. (2021). The economics of speed: the electrification of the streetcar system and the decline of mom-and-pop stores in boston, 1885–1905. *American Economic Journal: Applied Economics 13*(4), 285–324.

Appendix A Data Appendix

Appendix A.1 County Business Patterns

In 1946, the United States Census Bureau began publishing industry-level employment and establishment counts for nearly every US county. From archival publications, I collected previously undigitized records from 1953 to 1964 using a combination of OCR scanning and hand collection.[16] Data from 1964 to 1970 are published for a limited number of industries on ICPSR (Ody and Hubbard, 2011). The remaining years are available from the US Census Bureau and the National Archives. The data contain information for total employment, the total number of establishments, and the number of establishments in different sized employment bins. Bin sizes vary across CBP reports. From 1953-1973, there were 8 bins: 0-3, 4-7, 8-19, 20-49, 50-99, 100-249, 250-499, 500+. From 1974-1997, there were 13 bins: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, 1000+, 1000-1499, 1500-2499, 2500-4999, 5000+. From 1998-2016, there were 13 bins, similar to those above. I aggregated the bins to the largest consistent bin size to be consistent across every wave of the CBP. Prior to 1997, CBPs were arranged according to the SIC classification system. From 1998 to the present, industries are classified by NAICS codes. I follow Autor et al. (2013) in unifying broad industry codes over time.

From the raw data I made the following changes for uniformity and completeness: First, for confidentiality purposes, the Census Bureau censored the county level employment data for some smaller industries. Similar to Duranton et al. (2014), I impute employment values using the distribution of establishment count data.[17] Second, prior to 1964, some counties were reported as county groups. This occurs in Georgia, Illinois, Kansas, Kentucky, Missouri, New Mexico, New York, North Carolina, South Dakota, Texas, and Virginia. It is most common in Georgia, Texas, and Virginia (ICs). There were fewer cases in the other states. To address this issue, I split the combined data by the employment shares in 1964 (the first year I observe split counties). For Yellowstone NP in MT, I use the share of 1950 employment from the US Census. Finally, I adjust for county boundary changes using 1950 boundary definitions following Hornbeck (2010) and consolidate independent cities into their surrounding counties similar to Jaworski and Kitchens (2019).[18]

Appendix A.2 Interstate Highway System Maps

I use several data sources to construct an annual county level panel dataset with Interstate Highway System location and mileage information from 1953 to 2016. The first

[16] Prior to 1962, published establishment and employment information was combined for some counties in eight states. I partition the data in these counties using weights from 1964, the first year every county is reported separately.

[17] For each industry, I regress the county sectoral employment on the full set of eight establishment count groups and I use the resulting regression coefficients to impute the number of employees. The R^2 for each regression is between 0.945 and 0.999.

[18] County boundary locations from 1950 to 1990 are defined from Long (1995). For changes after 1990, I rely on the reported boundary changes from the US Census Bureau.

data source is current highway location information from NationalAtlas.gov (2016). To incorporate construction timing, I combine this file with highway construction information from several sources. My primary source is the PR-511 collection at the National Archives.[19] The PR-511 reports were not available prior to 1960, so I digitized annual Rand McNally highway maps from 1955 to 1959. For years after 2000, I relied on detailed interstate highway expansion information from the US Department of Transportation.

After combining these sources, I have annual information on the location and timing of the construction of the Interstate Highway System. I intersected this progress with a map of county locations in 1950, which allows me to know the year a county was connected to the Interstate Highway System.[20] For each county, I determine whether an interstate highway intersects that county and the year of arrival and the completed mileage constructed in each county in each year. Figure B.1.1a shows the current interstate highway locations.

Appendix A.3 Supplemental Data

To account for factors that are correlated with changing industrial dynamics and the expansion of interstate highways, I supplement the CBP and interstate location information with data covering population, market size, geography, military installations, and alternative methods of transportation. I use county level population data from the U.S. Census from 1910 to 1950, including the percent of the population living in urban areas in 1950 from National Historical Geographic Information System (2011). To account for differences in market size, I rely on Metropolitan Statistical Area (MSA) boundary definitions from National Historical Geographic Information System (2011) and metropolitan and rural classifications from Hines et al. (1975). To account for differences in the proximity to major metropolitan areas, I calculated the distance from each county seat to the centroid of the 1960 MSA boundary.

The historic narrative highlights that national defense and military interests played a role in the system design. To account for potential confounders caused by these interests I geo-located the coordinates to major military forts, naval bases, and airfields from the mid-1940s. With these locations, I calculated the distance from each county seat to the nearest military facility. To address geographic concerns I calculated the area, latitude, longitude, mean elevation, and ruggedness of each county. To account for potential spatial spillovers in Section **??**, I assign each county to a local labor market following the commuting zones definitions in Economic Research Service (2019).

I constructed several measures for alternative methods of transportation that existed prior to the construction of interstate highways, which could have influenced subsequent

[19]This series contains maps produced quarterly that show the progress of interstate highway construction. I digitized these maps and traced the annual construction progress of interstate highways in GIS. I denoted a segment of interstate highway completed once construction of that segment was finished and it was completely open to traffic. I used the fall quarter of each year when available. While I tried to be careful to accurately track annual construction progress it is possible that I classified counties as receiving interstate highways either before or after they actually did. This variation is likely to be random and corrected within the next year, which leads to short-term noise in the date of arrival.

[20]I adjust all of the county locations and data to be consistent with the 1950 county borders.

economic growth or the interstate construction decision. First, using two newly digitized historic maps of major highways from 1918 and 1947 and railroad route information from 1911 from Atack (2016), I calculate the distance from each 1950 county seat to the closest highway from each year and 1911 railroad. I also calculated the length of railroad track present in each county. Next, I collect location information for airports and ports in 1955 and 1956 from the Statistical Abstract of the United States (U.S. Census Bureau, 1958). I determine the latitude and longitude for each airport and port and calculate the distance from these locations to each 1950 county seat.

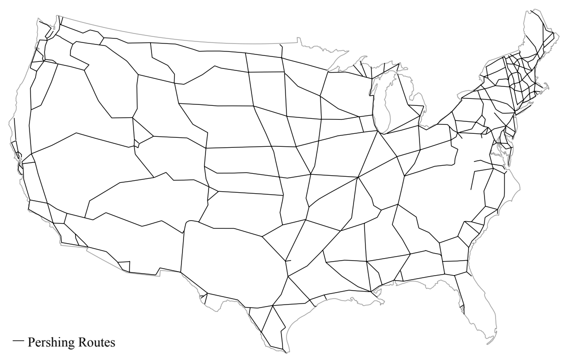
Every specification includes constructed geographic controls for the total area of the county, to account for the fact that geographically large counties are more likely to be traversed by the interstate system or the proposed IVs and are more likely to be located in growing western states. I also control for both latitude and longitude and their squares for each county seat.

Appendix B Exhibits Relating to IV

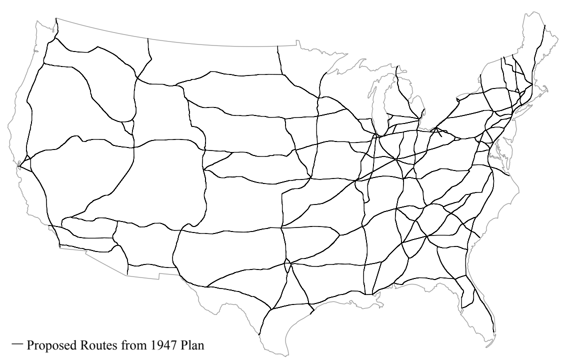
Appendix B.1 Maps and Interstate Construction

Figure B.1.1: Proposed and Constructed National Highway Network Plans

(a) Constructed Interstate Highway Locations



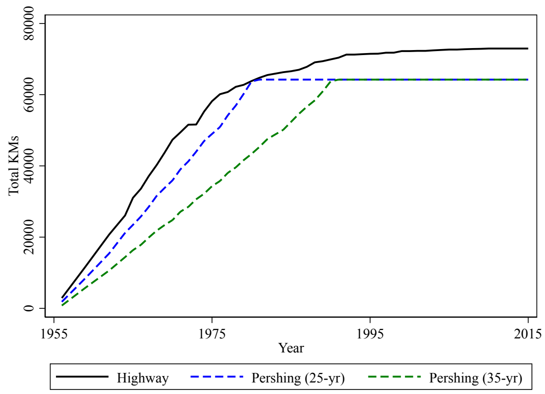
(b) Pershing Military Plan



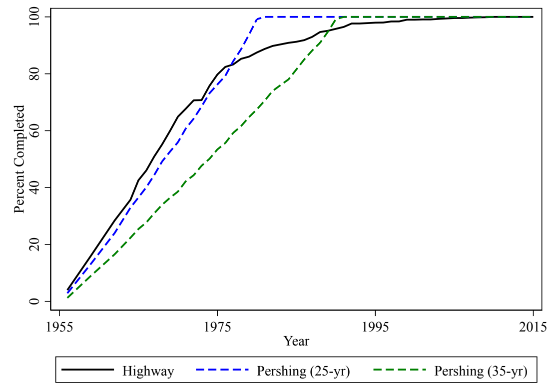
(c) 1947 Plan from Interregional Highway Committee

Notes: Panel (a) plots proposed Pershing routes digitized from the original map housed at the U.S. National Archives. Panel (b) plots the current Interstate Highway System locations from Federal Highway Administration. Panel (c) plots the proposed interstate system plan produced by the Interregional Highway Committee.

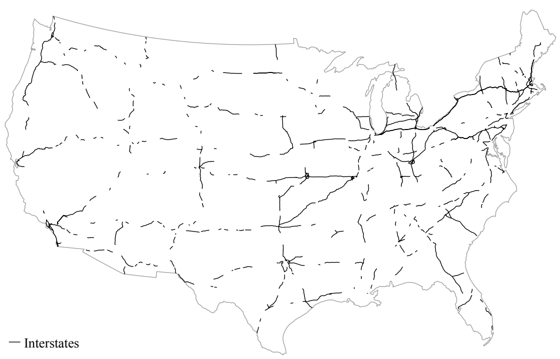
Figure B.1.2: Timing of Highway Construction and Predicted Construction



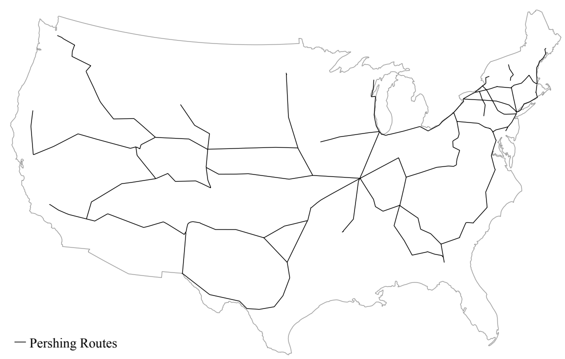
(a) Construction in KMs over 25 years



(b) Construction as a Share Completed over 25 years



(c) Interstate Highway Construction in 1965



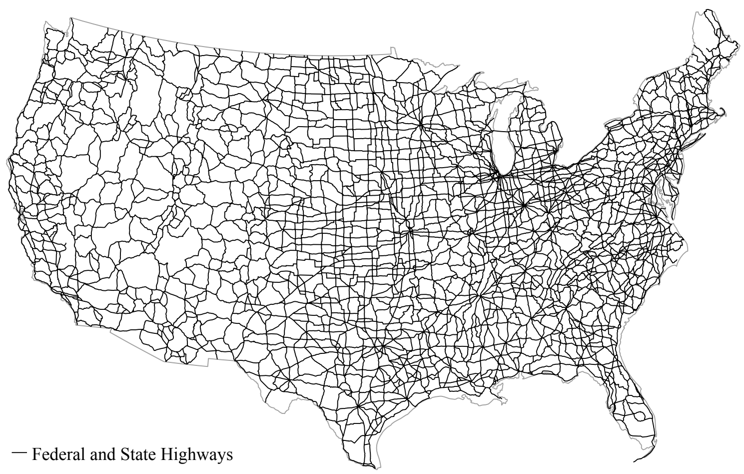
(d) Predicted Pershing Construction in 1965

Notes: Top figures plot the expansion of interstate highways over time (solid black line) with figure (a) presenting mileage and figure (b) presenting the share completed. Dashed lines plot alternative construction time horizons of the Pershing Map with blue denoting a 25-year construction horizon and green denoting a 35-year construction horizon. Figures (c) and (d) map completed and predicted segments of the Interstate Highway system in 1965. Sub-figure (c) maps the completed mileage in 1965 according to the PR-511 reports, and sub-figure (d) maps the predicted Pershing system mileage that would have been completed if construction had been prioritized according to centrality as defined by Girvan and Newman (2002) and the described budget allocation process.

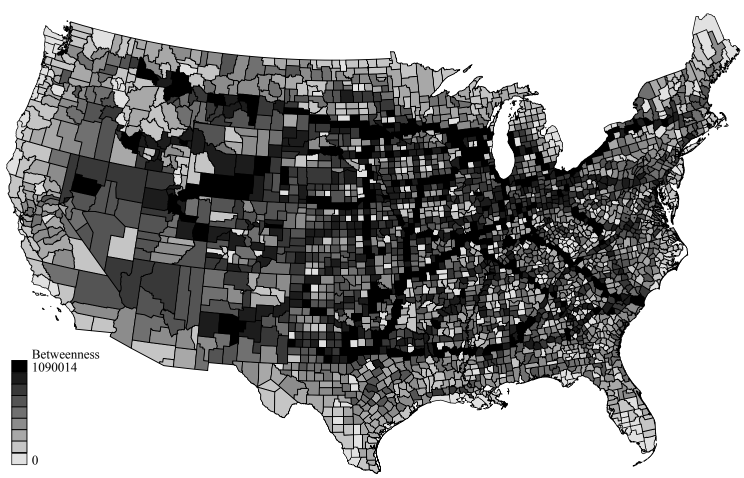
Appendix B.2 Pre-Interstate Centrality

Figure B.2.1 presents two figures related to constructing a centrality correction. Subfigure (a) maps the 1947 Rand McNally atlas of Federal and State Highways. I apply the Newman-Girvan Algorithm to the full pre-interstate highway network to calculate a betweenness centrality for every route in the network. I assign each county its maximum betweenness value, approximating a value of network importance to the county. I partition the distribution of these roughly 3,100 centrality values into 20 bins and assign each county to a fixed effect based on its bin. I then interact this set of fixed effects with year dummies to flexibly account for the role of pre-interstate network centrality over time.

Figure B.2.1: Mapping Pre-Interstate Highway Centrality



(a) Federal and State Highways in 1947



(b) Mapping Centrality from 1947 State and Federal Highways

Notes: Panel (a) plots the full set of federal and state highways in 1947. Panel (b) maps centrality values for this same network based on Newman (2001) and Newman and Girvan (2004) into nine categories.

Appendix B.3 Evidence in Support of the Exclusion Restriction

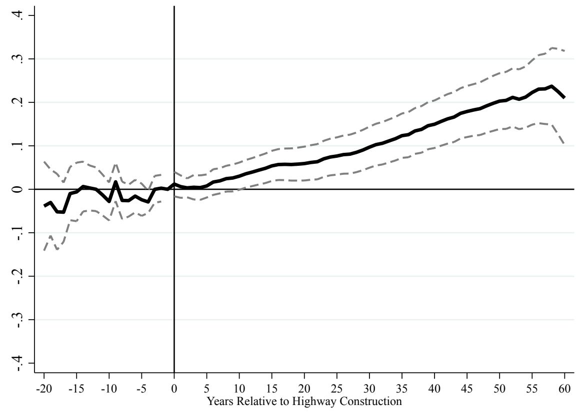
Table B.3.1: Effects of Interstates on Employment, Banking, and Industry Prior to Construction

	Employment	Bank Deposits	Manufacturing	Retail	Wholesale
	(1)	(2)	(3)	(4)	(5)
Panel A: OLS Estimates					
Interstate (0/1)	0.0022***	0.0011*	0.0001	0.0013***	0.0015*
	(0.0007)	(0.0007)	(0.0010)	(0.0005)	(0.0001)
Observations	12,342	11,666	10,176	12,389	11,887
Counties	3,100	2,932	2,556	3,098	3,013
State X Years	192	188	192	192	192
Panel B: Pershing IV Estimates					
Interstate (0/1)	-0.0049	-0.0034	0.0050	0.0000	-0.0050
	(0.0045)	(0.0036)	(0.0057)	(0.0026)	(0.0050)
Observations	12,342	11,666	10,176	12,389	11,887
Counties	3,100	2,932	2,556	3,098	3,013
State X Years	192	188	192	192	192
KP F-Statistic	85.50	74.90	64.61	84.63	80.28
Panel D: IV Specification Using 1947 Plan					
Interstate (0/1)	0.00200***	0.00075	-0.00119	0.00126*	0.00090
	(0.00050)	(0.00099)	(0.00144)	(0.00068)	(0.00130)
Observations	12,342	11,666	10,176	12,389	11,887
Counties	3,100	2,932	2,556	3,098	3,013
State X Years	192	188	192	192	192
KP F-Statistic	1,460.39	1,410.04	1,176.95	1,438.99	1,415.98
Panel C: Pershing Centrality					
Centrality Score	0.0002	0.0002	-0.0001	0.0000	-0.0000
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)
Observations	5,188	4,967	4,584	5,198	5,101
Counties	1,300	1,247	1,148	1,300	1,284
State X Years	192	188	192	192	192

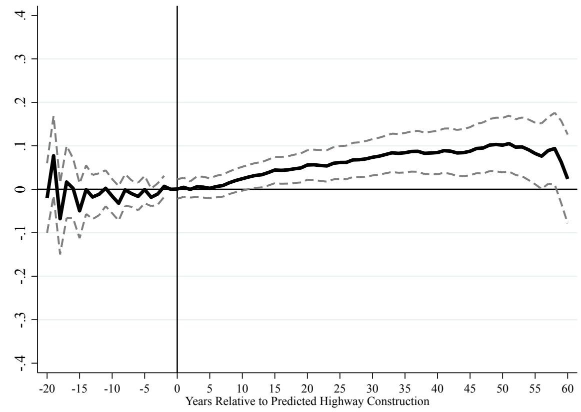
Notes: Panel A reports results from estimating a modified equation 1, which regresses a time-invariant binary interstate highway indicator on average changes in the log outcome. Panel B reports IV estimates using the time-invariant Pershing locations. Panel C reports IV estimates replacing the Pershing locations with the 1947 Interregional Highway Plan. Panel D replaces the binary interstate treatment with a continuous measure of the maximum centrality score associated with constructing the Pershing plan in the county. Every specification includes state $\times$ year fixed effects, along with the full set of controls. Outcome data from 1930–1950 decadal censuses and 1954 County Business Patterns reported at the county level. Employment reflects the total county employment; Bank Deposits are the total inflation-adjusted value of deposits; Manufacturing, Retail Trade, and Wholesale Trade all reflect the number of establishments. Each measure is used to calculate the average log change from the prior period. Standard errors are two-way clustered by county and commuting zone $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Figure B.3.1: Employment Changes as an Event Study with Interstate Highways and Proposed Pershing Highways

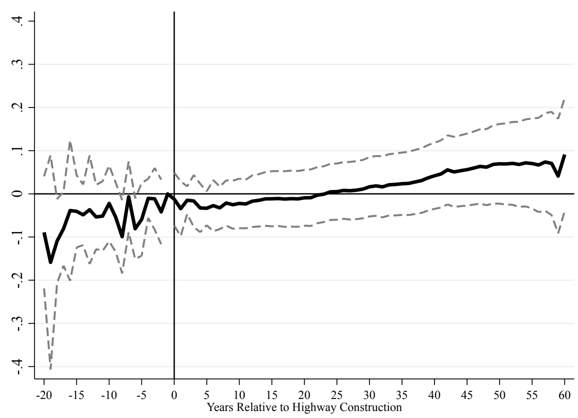
Interstate Highways | **Pershing Routes**



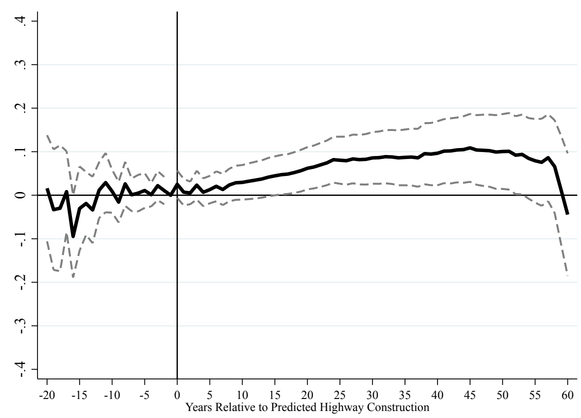
(a) Unweighted Employment



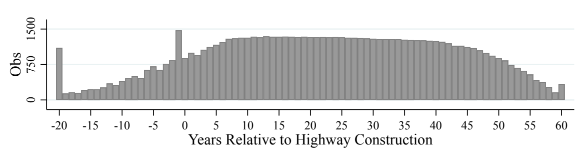
(b) Unweighted Employment



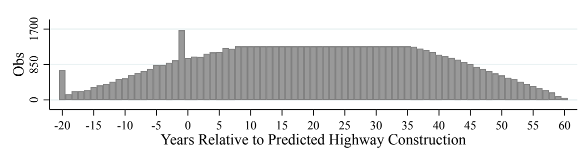
(c) Population Weighted Employment



(d) Population Weighted Employment



(e) Density of Actual Construction

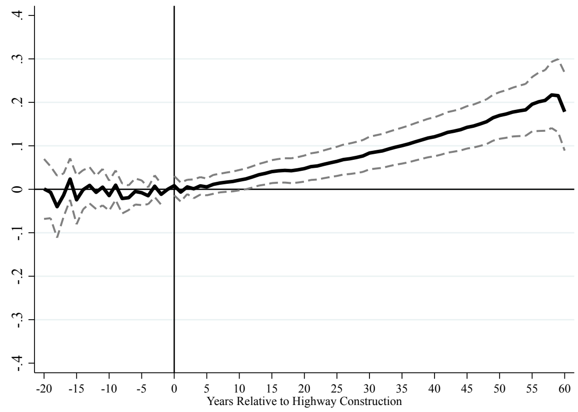


(f) Density of Predicted Construction

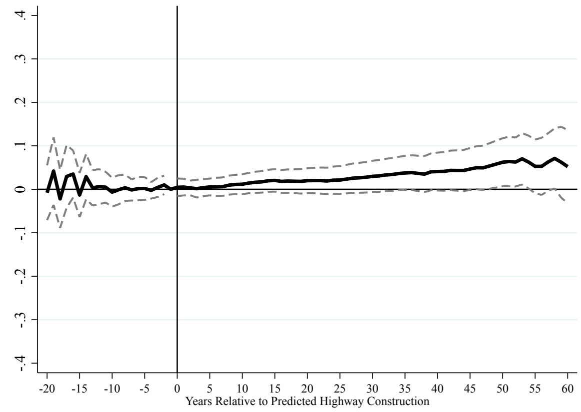
Notes: The figure plots event study (eqn 3) estimates of the effects of interstate highways (left column) and the Pershing reduced form locations (right column) on employment. Year $t-1$ is the excluded year. The top row presents unweighted estimates. The second row presents population weighted estimates using 1950 county population. The bottom row plots the density of points within each year bin.

Figure B.3.2: Establishment Changes as an Event Study with Interstate Highways and Proposed Pershing Highways

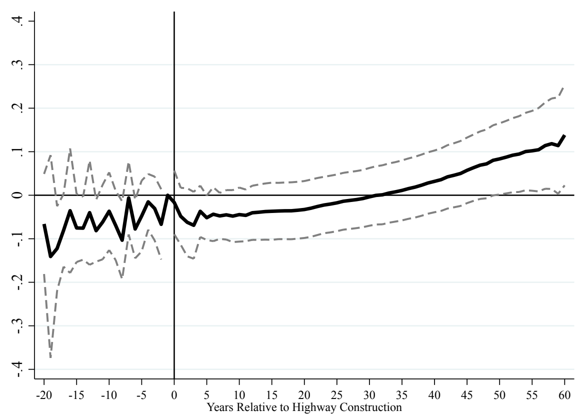
Interstate Highways | **Pershing Routes**



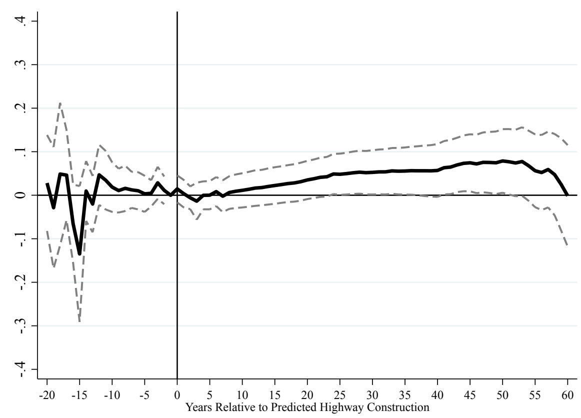
(a) Unweighted Establishments



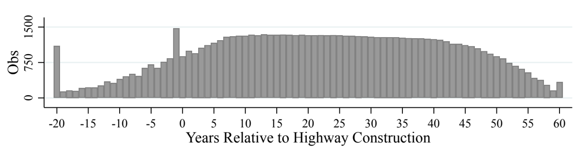
(b) Unweighted Establishments



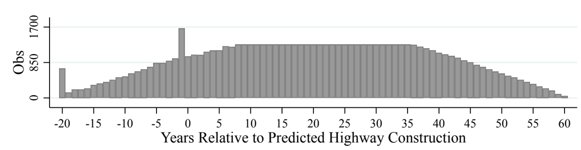
(c) Population Weighted Establishments



(d) Population Weighted Establishments



(e) Density of Actual Construction



(f) Density of Predicted Construction

Notes: The figure plots event study (eqn 3) estimates of the effects of interstate highways (left column) and the Pershing reduced form locations (right column) on the number of establishments. Year $t-1$ is the excluded year. The top row presents unweighted estimates. The second row presents population weighted estimates using 1950 county population. The bottom row plots the density of points within each year bin.

Appendix C Table Appendix

Table C.1.1: Summary Statistics by Eventual Highway Status

	(1) Non-IHS	(2) IHS	(3) Difference
Employment	5,848.777	54,950.207	49,101.430***
	[11,620.702]	[165,709.078]	(4,309.363)
Establishments	509.727	3,447.571	2,937.844***
	[872.994]	[9,514.575]	(246.298)
Δ Ln(Employment)	0.017	0.022	0.005***
	[0.120]	[0.084]	(0.001)
Δ Ln(Establishments)	0.009	0.015	0.006***
	[0.062]	[0.049]	(0.000)
Highway (0/1)	0.000	0.894	0.894***
	[0.000]	[0.307]	(0.004)
Pershing IV 25yr (0/1)	0.230	0.532	0.302***
	[0.421]	[0.499]	(0.015)
1947 Plan IV 25yr (0/1)	0.050	0.694	0.643***
	[0.219]	[0.461]	(0.011)
Interstate KMs per County Sq KM	0.000	0.026	0.026***
	[0.000]	[0.027]	(0.001)
Pershing IV KMs per County Sq KM	0.005	0.014	0.009***
	[0.011]	[0.019]	(0.001)
1947 Plan KMs per County Sq KM	0.001	0.017	0.017***
	[0.004]	[0.017]	(0.000)
Observations	95,872	76,272	172,144

Notes: Employment and Establishment data from 1953–2016 County Business Patterns annual reports. Columns (1) and (2) report means and standard deviations in brackets. Column (3) presents the difference in means, with standard errors in parentheses. The standard errors are clustered by county. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table C.1.2: Differences in Covariates by Interstate Status

	(1) Non-IHS	(2) IHS	(3) Difference
Ln(Employment in 1953)	7.053	8.478	1.425***
	[1.292]	[1.716]	(0.056)
Ln(Establishments in 1953)	5.190	6.288	1.098***
	[1.012]	[1.357]	(0.044)
Ln(1950 Population)	9.476	10.433	0.957***
	[0.878]	[1.224]	(0.039)
Ln(1940 Population)	9.502	10.340	0.837***
	[0.842]	[1.143]	(0.037)
Ln(1930 Population)	9.472	10.267	0.796***
	[0.823]	[1.130]	(0.037)
Ln(1920 Population)	9.427	10.161	0.734***
	[0.874]	[1.095]	(0.036)
Ln(1910 Population)	9.366	10.052	0.686***
	[0.876]	[1.063]	(0.036)
Pct of Pop in Urban Area in 1950	0.193	0.391	0.197***
	[0.218]	[0.280]	(0.009)
Market Potential with 1950 Pop	167.223	198.112	30.889***
	[56.718]	[107.254]	(3.213)
Area in sq mi	933.262	1,034.421	101.160**
	[1,114.182]	[1,549.751]	(49.884)
Latitude	38.349	38.210	-0.138
	[4.989]	[4.682]	(0.175)
Longitude	-92.811	-90.529	2.282***
	[10.730]	[12.225]	(0.421)
Latitude Sq	1,495.508	1,481.943	-13.565
	[385.351]	[355.130]	(13.392)
Longitude Sq	8,728.927	8,344.860	-384.067***
	[2,074.673]	[2,347.685]	(80.998)
Mean Elevation	470.110	412.169	-57.941***
	[525.661]	[486.795]	(18.313)
Ruggedness	73.730	78.449	4.720
	[115.504]	[117.105]	(4.226)
1911 Railroad KMs per sq mi	0.148	0.235	0.087***
	[0.107]	[0.178]	(0.005)
KM to 1947 Highway System	78.161	104.883	26.721***
	[71.494]	[96.721]	(3.139)
KM to 1911 RR	1,007.970	1,155.971	148.002***
	[514.978]	[550.670]	(19.430)
KM to Nearest 1955 Port	443.168	356.377	-86.791***
	[302.016]	[303.020]	(10.986)
KM to Nearest 1955 Airport	60.470	39.994	-20.476***
	[35.063]	[31.750]	(1.208)
KM to 1918 Military Highways	564.506	574.137	9.631
	[391.897]	[406.890]	(14.535)
KM to MSA	121.258	77.785	-43.473***
	[80.839]	[70.910]	(2.740)
KM to Mexican War Battle	1,298.721	1,495.521	196.799***
	[609.662]	[702.162]	(24.064)
KM to American Rev. Battle	703.216	620.200	-83.015***
	[643.406]	[708.294]	(24.701)

KM to Civil War Battle	366.089	332.371	-33.719**
	[400.943]	[385.674]	(14.252)
KM to French/Indian War Battle	1,305.409	1,168.498	-136.911***
	[796.151]	[894.995]	(30.957)
KM to Indian War Battle	216.142	208.248	-7.894
	[137.431]	[133.288]	(4.907)
KM to Insurrections	484.820	445.393	-39.427***
	[279.682]	[274.789]	(10.056)
KM to War 1812 Battles	655.234	607.431	-47.803**
	[583.842]	[652.614]	(22.623)
KM to WW1 Sites	123.956	79.992	-43.964***
	[77.833]	[69.987]	(2.671)
KM to Naval Bases	575.484	451.964	-123.519***
	[332.333]	[312.001]	(11.661)
KM to Airfields	99.187	78.854	-20.334***
	[71.297]	[60.170]	(2.372)
KM to Military Forts	122.980	93.962	-29.018***
	[78.792]	[67.286]	(2.636)
KM to Pershing Map Nodes	127.516	93.235	-34.281***
	[67.972]	[72.577]	(2.562)
1947 Hwy Betweenness	50,752.914	91,373.211	40,620.293***
	[76,019.500]	[103,086.656]	(3,343.281)
Rural County (0/1)	0.682	0.298	-0.384***
	[0.466]	[0.458]	(0.017)
Metro County (0/1)	0.062	0.383	0.320***
	[0.242]	[0.486]	(0.014)
MSA County in 1960 (0/1)	0.016	0.233	0.218***
	[0.125]	[0.423]	(0.012)
Boundary Adjustment (0/1)	1.001	1.007	0.005**
	[0.034]	[0.081]	(0.002)
Grouped County (0/1)	0.169	0.106	-0.063***
	[0.375]	[0.309]	(0.012)
Indep. City County (0/1)	0.002	0.015	0.013***
	[0.042]	[0.120]	(0.003)
Suppressed Employment (0/1)	0.002	0.000	-0.002*
	[0.042]	[0.000]	(0.001)
Observations	1,712	1,362	3,074

Notes: Detailed source and measurement information in Appendix A. Columns (1) and (2) report means and standard deviations in brackets. Column (3) presents the difference in means, with standard errors in parentheses. The standard errors are clustered by state × year. *** p<0.01, ** p<0.05, * p<0.1.

Table C.1.3: Interstate Highways and Growth in the Share of Establishments of Different Sizes

Panel A: OLS Estimates

	Employ. 1-19 (1)	Employ. 20-49 (2)	Employ. 50-99 (3)	Employ. 100-249 (4)	Employ. 250-499 (5)	Employ. 500+ (6)
Interstate (0/1)	-0.0136***	0.0078***	0.0035***	0.0014*	0.0002	0.0005
	(0.0028)	(0.0021)	(0.0011)	(0.0008)	(0.0004)	(0.0003)

Panel B: Pershing IV Estimates

	Employ. 1-19 (1)	Employ. 20-49 (2)	Employ. 50-99 (3)	Employ. 100-249 (4)	Employ. 250-499 (5)	Employ. 500+ (6)
Interstate (0/1)	-0.0285**	0.0087	0.0151***	0.0015	0.0038**	0.0001
	(0.0120)	(0.0093)	(0.0055)	(0.0037)	(0.0016)	(0.0010)
Pershing First-Stage	0.1979	0.1969	0.1979	0.1979	0.1977	0.1984
	(0.0163)	(0.0164)	(0.0163)	(0.0163)	(0.0164)	(0.0164)
KP F-Statistic	146.7371	145.0306	146.6861	146.8707	146.2116	147.1031
Observations	171,951	171,951	171,951	171,951	171,951	171,951
Counties	3,071	3,071	3,071	3,071	3,071	3,071
State X Years	2,688	2,688	2,688	2,688	2,688	2,688

Notes: Every specification reports results from estimating equation 1, where the outcome of interest is the year over year change in the share of firms within the size category specified in the column. Panel A presents OLS results and Panel B presents IV results using a binary interstate highway indicator. Every specification includes state $\times$ year fixed effects, along with the full set of controls. Employment and establishment data are from 1956–2016 County Business Patterns annual reports. Standard errors are two-way clustered by county and commuting zone $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Table C.1.4: Summary Statistics of Establishment Size by Eventual Highway Status

Panel A: Differences in 1953

	(1) Non-IHS	(2) IHS	(3) Difference
Share of Estab. w/ 0-19	95.142	93.167	-1.975***
	[3.308]	[3.068]	(0.115)
Share of Estab. w/ 20-49	3.270	4.385	1.114***
	[2.185]	[1.952]	(0.075)
Share of Estab. w/ 50-99	0.881	1.309	0.428***
	[1.292]	[0.859]	(0.039)
Share of Estab. w/ 100-249	0.469	0.730	0.261***
	[0.663]	[0.608]	(0.023)
Share of Estab. w/ 250-499	0.160	0.242	0.082***
	[0.330]	[0.406]	(0.014)
Share of Estab. w/ 500+	0.078	0.167	0.089***
	[0.491]	[0.359]	(0.015)
Observations	1,712	1,362	3,074

Panel B: Average Annual Changes Between 1956 and 2016

	(4) Non-IHS	(5) IHS	(6) Difference
Δ Share of Estab. w/ 0-19	-0.091	-0.110	-0.019***
	[1.428]	[0.963]	(0.002)
Δ Share of Estab. w/ 20-49	0.061	0.072	0.011***
	[1.373]	[0.906]	(0.001)
Δ Share of Estab. w/ 50-99	0.019	0.024	0.004***
	[0.740]	[0.477]	(0.001)
Δ Share of Estab. w/ 100-249	0.009	0.012	0.003***
	[0.465]	[0.307]	(0.000)
Δ Share of Estab. w/ 250-499	0.002	0.002	0.000
	[0.223]	[0.172]	(0.000)
Δ Share of Estab. w/ 500+	0.001	0.001	0.000
	[0.124]	[0.099]	(0.000)
Observations	95,872	76,272	172,144

Notes: Establishment size data from 1953–2016 County Business Patterns annual reports. Columns (1), (2), (4), and (5) report means and standard deviations in brackets. Columns (3) and (6) present the difference in means, with standard errors in parentheses. The standard errors are clustered by county. *** p<0.01, ** p<0.05, * p<0.1.

Table C.1.5: Interstate Highways and Growth: Constructed Interstate Density

Panel A: Employment				
	(1) OLS	(2) IV	(3) OLS	(4) IV
Interstate KMs per County Sq KM	0.0632***	0.0765	0.0140	-0.0112
	(0.0146)	(0.0479)	(0.0140)	(0.0410)
Interstate (0/1)			0.0037***	0.0072***
			(0.0006)	(0.0027)
KP F-Statistic		42.383		58.73
Panel B: Establishments				
	(1) OLS	(2) IV	(3) OLS	(4) IV
Interstate KMs per County Sq KM	0.0539***	0.0777**	0.0126	0.0037
	(0.0119)	(0.0391)	(0.0111)	(0.0338)
Interstate (0/1)			0.0031***	0.0061***
			(0.0005)	(0.0023)
KP F-Statistic		42.383		58.73

Notes: Every specification reports results from estimating a modified equation 1, where the outcome of interest is the year over year change in either employment or establishments, and interstate treatment is based on the completed kilometers of interstate per square kilometer of county area. Every specification includes state × year fixed effects, along with the full set of controls outlined in Appendix A. Employment and establishment data are from 1956–2016 County Business Patterns annual reports. Standard errors are two-way clustered by county and commuting zone × year. *** p<0.01, ** p<0.05, * p<0.1.