Transportation Networks and the Geographic Concentration of Employment*

Dustin Frye†

May 17, 2022

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 Economics, Vassar College, 124 Raymond Ave, Box 76, Poughkeepsie, NY 12604 (e-mail: dufrye@vassar.edu; website: http://dustinfrye.github.io/)

Abstract

This paper examines the effect of expanding transportation networks on uneven spatial industrial growth across the United States from 1953 to 2016. The paper addresses the endogenous placement and timing of interstate construction by instrumenting for highway locations using a historic military map combined with a network theory algorithm to predict construction timing. Results indicate 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. The results also reveal positive spillovers occurred in later decades among adjacent counties along the metropolitan periphery.

1 Introduction

Industrial growth in the U.S. is characterized by its uneven spatial distribution. For example, over the last 60 years, nearly 80 percent of employment growth in the U.S. occurred in just 10 percent of counties.[1] Investments in transportation infrastructure are one form of place based policy used to promote regional growth and decrease spatial concentration. Beginning in 1956, construction of the Interstate Highway System in the United States 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, data and empirical limitations have prevented a comprehensive understanding of the spatial spillovers of interstates, and of how interstates impacted industrial growth over time.

To fill this gap, this paper estimates the impact of interstate highways on employment and establishment growth from 1956 to 2016, for the entire U.S., using a new empirical approach that addresses both the endogenous placement and the endogenous timing of interstate construction. I introduce a new methodology from an application of network theory, which is easy to implement and can be applied any setting where one needs to account for the endogenous allocation of resources can be patterned into a network. In addition to the new empirical design, this paper bridges the existing empirical literature that is either exclusively focused on metropolitan areas or exclusively on rural areas, by focusing on all counties in the lower-48 states.[2] Including this set of "missing middle"

[1] According to the County Business Patterns, between 1953 and 2016, the U.S. added 80.5 million jobs, with over 62 million added in the top 308 counties.

[2] For example seminal work by Michaels et al. (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.

counties, which are neither defined as metropolitan or rural in previous analysis, creates a comprehensive set of findings that are broadly applicable.

I estimate the effect of an interstate being present within a county on year over year changes in employment and establishments, using a reduced form analysis, where I instrument for the presence of an interstate to address 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). I address location endogeneity, by instrumenting for eventual interstate locations using a proposed 1920s military plan of high priority routes.[3]

Once routes were determined, interstate construction policy placed state and local officials in charge of determining when particular segments would be constructed, providing another means of encouraging local development. To address endogeneity in timing, I implement the Newman-Girvan algorithm on the 1920s military plan to predict the timing of interstate construction. The algorithm prioritizes constructing interstate segments 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 military plan.

Within the estimation strategy, I account for the concern that, independent of interstates, centrality in a network may be correlated with economic growth, by including a correction discussed in Borusyak and Hull (2021). I support the validity of the instrument by presenting an event study analysis that provides evidence for the parallel trends assumption holding in the reduced form and illustrates the endogeneity in the OLS.

This analysis, including the construction of the instrument, is made possible through the use of a county-year panel dataset spanning from 1953-2016. I constructed this dataset

[3]Note that some previous literature has used a historical map from 1947 (i.e. (Michaels et al., 2007; Baum-Snow, 2007; Duranton and Turner, 2012). In Section 4.2, I describe why this particular map is not well suited as an instrument in this setting.

by compiling County Business Pattern (CBP) data for every available year. The dataset contains county level employment and establishments counts for the lower-48 states. The CBP also includes the number of firms across eight different employment size categories. I complement this data with granular interstate construction timing information from the National Archives.

By using data that combines both urban and rural areas, I introduce new analysis to the literature on spatial spillovers. Results indicate that counties receiving an interstate highway experienced more year over year employment and establishment growth relative to counties that were not directly connected to the system. The establishment 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. These positive growth effects of an interstate are not confined to only interstate counties. Interstates generate spatial spillovers that spread into the non-interstate metropolitan periphery. I find growth in employment and establishments in both interstate counties as well as adjacent non-interstate counties that are within the same local labor market as a Metropolitan Statistical Area (MSA). This growth in the non-interstate metropolitan periphery is consistent with the decentralized suburban job growth in Baum-Snow (2020) and highlights the connection between interstates and agglomeration spillovers at the labor market level. I do not find evidence of positive or negative spillovers into other adjacent areas. This indicates that the negative spillover effects on interstate adjacent counties, found by Chandra and Thompson (2000), are confined to shifts between adjacent rural areas.

With a time varying instrument, and outcome data that varies annually, I am able to explore growth dynamics during three distinct eras. I consider the initial expansion era (1956-1975), the interstate completion era (1976-1995), and the post-construction era (1996-2016). Results from this era based approach show there is considerable variation over time that was previously inaccessible in other settings. Growth in employment and establishments among interstate counties is highest in the initial expansion era, but growth in the non-interstate metropolitan periphery is strongest in the later decades. These results indicate that despite being a largely disconnected system in the initial expansion era, there was considerable localized industrial growth due to the interstates.

With this analysis, I am able to comment on how an instrument that explicitly addresses the endogenous timing of construction compares to results if an instrument does not address timing. I find that using an instrument that only addresses location endogeneity underestimates the impact of transportation on growth, particularly in the period of interstate expansion. This indicates previous estimates may have understated the early growth due to interstates and introduces the possibility of similar endogeneity concerns in other settings. Furthermore, this analysis allows for the unpacking of endogeneity patterns over time. Results reveal that endogeneity in this context is dynamic over time. I find that endogeneity documented in previous literature, namely that lower performing urban areas received an interstate is most pronounced in early eras but fades over time.

In general, this paper contributes to our understanding of the spatial and temporal effects of transportation infrastructure on industrial growth.[4] With county-year level data that spans the entire U.S., and a new methodology that accounts for endogenous timing, I unite two distinct strands of literature and improve our understanding of the dynamic impacts of interstates.

2 A History of the Interstate Highway System

2.1 The Pershing Map and the National Interregional Highway Committee

Early proposals of interstate highway locations date back to the early 1920s. Following the First World War, the U.S. government began discussing the merits of a national highway

[4]For a comprehensive survey of this literature see Redding and Turner (2015)

system, similar to the system that existed in Europe. This led Congress and the Bureau of Public Roads to seek input from the War Department regarding a national system of interstate highways (Karnes, 2009). The War Department commissioned General John J. Pershing to provide a network map of high-priority military routes. The army did not value a "transcontinental road which merely crosses the continent", but rather wanted "roads connecting all our important depots, mobilization and industrial centers" (Swift, 2011, 76). The resulting map, depicted in Figure 1a, contained nearly 78,000 miles of highway that the War Department deemed as strategically important. The map emphasized "coastal and border defense and links to major munitions plants" (Swift, 2011, 76). The full Pershing Map contains three priority levels, the depicted map shows routes in the two highest priority levels spanning roughly 40,000 miles.[5] These routes were never built as superhighways but this map influenced future highway location decisions.

National interstate highway programs were reintroduced during the Great Depression as part of New Deal legislation. President Roosevelt formed the National Interregional Highway Committee "to investigate the need for a limited system of national highways to improve the facilities now available for interregional transportation" (U.S. Dept of Transportation, 1977, 273). Committee members included engineers, government officials, and highway planners. With the help of state highway departments, the committee produced a new 39,000 mile national highway plan. The committees objectives were to "provide highway transportation to serve the economic and social needs of the nation" (U.S. Dept of Transportation, 1977, 274). The highway network was intended to "serve the Nation's agricultural production, its mineral production, its forest production, its manufacturing centers and ... its population centers and defense establishments" (U.S. Dept of Transportation, 1977, 274). Interest groups on behalf of the farming and trucking industry "lobbied for their own plans to foster particular and local needs" (Rose, 1990, 16). The final plan, published in 1947 and depicted in Figure A.1.1, was the most comprehensive

[5] Section 4.3.4 uses the unbuilt priority three segments as a falsification exercise.

national network map that had been produced and served as the major guide of highway location decisions for the next decade.

[Figure 1 Here]

Highway construction plans were halted during the war and funding was restricted to high priority maintenance of current roads. Without adequate funding for repairs, the quality of highway infrastructure deteriorated rapidly. Prior to World War II, total road spending was about 1.4 percent of GNP and during the war this amount fell to about 0.2 percent (Karnes, 2009). As the quality of roads decreased, the demand for high quality roads increased rapidly. From 1945 to 1950 vehicle registrations increased nearly 60 percent (Swift, 2011). The Bureau of Public Roads determined that between the mid-1920s and early 1950s traffic had increased by 250 percent and highway demand had increased by a factor of eight (Rose, 1990). This put tremendous strain on the existing infrastructure, which was ill equipped to deal with new faster cars and heavier trucks. Travel times increased dramatically due to elevated levels of congestion and the increased probability of an accident (Kaszynski, 2000).

2.2 Federal Aid Highway Act of 1956

In the early 1950s, several Congressional Committees developed plans for funding and designing a new system of limited access interstate highways. President Eisenhower was influential in helping support these committees and invited Governors and heads of interest groups to participate in the planning process (Rose, 1990). Industry representatives from oil, trucking, and manufacturing were particularly influential in these discussions (Kaszynski, 2000). Congressional representatives were also influential in the system's design. Additional urban mileage was added to appease Congressional representatives with large urban constituencies (Boarnet, 2014).

In 1956, after several prior proposals, construction guidelines and financing terms were introduced. Congress ultimately agreed on legislation that included proposed locations and plans for funding. The system was approximately 90 percent federally funded and was financed with tax revenue from a variety of sources (Kaszynski, 2000). Eisenhower signed the Federal-Aid Highway Act of 1956 into law on June 29th. The final design, presented in Figure 1b, was "a culmination of decades of input and research from auto clubs, civil engineers, and state and federal highway officials" (Kaszynski, 2000, 167). The Highway Act of 1956 placed states in charge of construction and each state's funding was determined based on a formula of population, area, and highway mileage. This allowed state officials to build interstates at their own pace and prioritize different locations. Urban areas were often prioritized following strong federal encouragement (Lewis, 1997, 131). The solicitation of opinions from heads of industry and government officials for both the eventual location of interstate highways and the pace of construction have important consequences for empirically estimating the effects of interstate highways.

As interstate highways expanded through the 1960s and 1970s they became an increasingly important part of travel. In 1960, interstate highways constituted 0.3% of total paved roads in the U.S. and only 3.3% of total vehicle miles traveled (VMT) occurred on interstates. Completed interstate mileage nearly tripled during the 1960s and by 1970, interstates carried nearly 15% of total VMT (U.S. Department of Transportation Statistics, 1960, 1970). Today, interstates constitute only 1.1% of total road mileage but they support over 25% of total vehicle miles traveled (U.S. Department of Transportation Statistics, 2017). They have become an essential part of travel and commerce and have reshaped the spatial distribution of economic activity. According to the U.S. Census Bureau (2022), over 80 percent of employment is within 15 miles of an interstate highway.[6]

[6]Appendix Figure B.1.1 plots the distribution of employment by distance to the nearest interstate from the LEHD Origin-Destination Employment Statistics (LODES).

3 Data and Descriptive Evidence

To investigate the effects of interstate highways on industry dynamics, my empirical analysis relies on county level annual employment and establishment data collected by the Census Bureau and published in the County Business Patterns (CBP) from 1953 to 2016. These data are combined with contemporary and historical transportation network information, which allows me to estimate the relationship between the expansion of interstates and employment and establishment growth.

3.1 County Business Patterns

In 1946, the United States Census Bureau began publishing industry-level employment and establishment size counts by establishment size. I collected previously undigitized records from 1953 to 1964 using a combination of OCR scanning and hand collection.[7] Data from 1964 to 1970 are published for a limited number of industries on ICPSR (Ody and Hubbard, 2011).[8] The remaining years are available from the US Census Bureau and the National Archives.[9] 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).[10]

To generate the primary outcomes of interest, I construct employment and establishment counts for every county and for each available year from 1953 to 2016. 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.[11]

[7]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.

[8]This series had a few omissions, which I hand collected.

[9]Appendix C.1 provides more detail regarding the construction of the CBP dataset.

[10]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.

[11]For each industry, I regress the county sectoral employment on the full set of eight establishment count

The County Business Patterns consistently tabulate 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.[12]

3.2 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 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.[13] 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.[14] 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 1b shows the current interstate highway

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.

[12]Employment size groups vary by year, but typically include eight categories: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, and above 500 employees. The boundaries of the smaller categories change over time, so I combine them in the analysis to 1-19 employees. Note, one limitation of the County Business Patterns data is that it does not include establishments with zero employees.

[13]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.

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

locations.

3.3 Supplemental Data

To account for factors that are correlated with changing industrial dynamics and the expansion of interstate highways, I supplement the CBP and highway 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 5.2, 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 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.

3.4 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. As a preliminary comparison between interstate and non-interstate areas, Figure 2 plots the mean changes in total employment and establishments 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 across both measures. Appendix Table B.1.1 reports summary statistics for the full period and shows that on average, employment in interstate counties grew at about 2.2 percent per year compared to 1.7 percent growth in non-interstate counties. Establishments grew at slightly lower averages of 1.5 percent per year for interstate counties and 0.9 percent per year in non-interstate counties.

Panels (c) and (d) of Figure 2 focus specifically on interstate counties and plot the change in employment and establishments relative to the initial construction year. Within these counties, I separately plot 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 built prior to the mid-1970s, but was much slower among those counties that built in the later decades.

[Figure 2 Here]

This faster growth among highway counties is coupled with significantly higher levels of initial employment. Appendix Table B.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.

Appendix Table B.1.3 summarizes the distribution of establishment sizes in 1953 and shows how each establishment size category changes over time. In 1953, an overwhelming majority establishments are in the smallest size category, with roughly 95 percent of establishments having fewer than 20 employees. Over time establishments get larger, so the share of establishments in the smallest category falls to 88 percent of firms. Panel (b) reports the growth differences by eventual interstate status. The panel shows interstate counties experienced a larger reduction in the share of firms in the smallest establishment size bin with increases in the fraction of larger firms across multiple categories. These differential increases in establishment size by interstate status suggest there are potential scale benefits of being in close proximity to interstate highways. These differences reinforce the selection concerns regarding interstate locations.

3.5 Defining Spatial Adjacency

In this paper, I assess the importance of adjacency and spillovers in space by leveraging the national coverage of the data and by considering spillovers within labor markets, which are defined by commuting zone boundaries according to Economic Research Service (2019). Figure 3 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. I define lettered counties that share a border with a light gray interstate counties as spatially adjacent. I define counties labeled with the letters A and B as interstate adjacent within the commuting zone. These counties are in the same local labor market as interstate treated counties. These are different than the counties labeled with the letter C, which is 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.

[Figure 3 Here]

4 Empirical Strategy

4.1 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 interstate highways using the following first-difference specification,

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

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.[15] The coefficient of interest is β, which estimates the average effect of the interstate highway

[15]In Section 6 I introduce distance based measures that capture the intensity of interstate treatment.

system. The specification includes state × year fixed-effects, δ_{st}, so the treatment effect of an interstate highway is identified using variation within a state in a year. Including this set of fixed-effects accounts for any state wide changes that affect employment, the opening or closing of businesses, or promote growth over time and are correlated with the construction of interstate highways.[16] I also include market size × year fixed-effects, θ_{mt}, to flexibly account for ways that metropolitan and rural areas grow over time. The specification includes CBP adjustment × year fixed-effects, γ_{dt}, to account for minor differences in CBP reporting each year. $X'\rho_{ct}$ is a full set of controls discussed in Section 3.3. Standard errors are two-way clustered by county and commuting zone × year to account for serial correlation and spatial correlation in the error term (Kelly, 2019, 2020).[17]

Even with the broad set of flexible time-varying fixed-effects, there still may be concern that county characteristics are correlated with employment and establishment growth. To account for this, I include a detailed list of controls that account for pre-existing differences. The controls, which were outlined in Section 3.3 and are presented in Appendix Table B.1.2, include baseline employment, establishments, and population measures, as well as additional measures to capture pre-highway levels of urbanization, transportation infrastructure, and geographic characteristics. Section Appendix C.2 describes each measure and its construction in more detail. These covariates are associated with constructing an interstate highway and industrial growth after 1953. 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 locations.[18]

[16] Appendix Table B.1.4 reports results under varying spatial fixed-effects including, commuting zone and census region and finds similar effects across specifications.

[17] Appendix Table B.1.5 presents OLS results using spatially weighted standard errors proposed by Conley (2010) and Hsiang (2010). These spatially adjusted standard errors are nearly identical across the alternative distances compared to the preferred two-way clustering.

[18] For recent examples across different forms of transit infrastructure and in multiple settings see Donaldson and Hornbeck (2016); Baum-Snow et al. (2017); Alder et al. (2017); Bartelme (2018); Alder (2019); Jaworski and Kitchens (2019); Jaworski et al. (2020); Jedwab and Storeygard (2020); Fajgelbaum and Schaal

4.2 Addressing Highway Endogeneity

Estimating differences between counties with and without an interstate will result in biased estimates because counties selected to construct an interstate, and when they construct the interstate, are likely to differ along unobservable dimensions that are correlated with changing employment or establishments. The history of highway construction indicates that the placement, timing of construction, and funding of highways was an intensely political process. 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. If these outside contributors viewed highway construction and development 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 historic military plan, described in Section 2, 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 are based on the digitized Pershing Map from the Bureau of Public Roads collection at the National Archives. Figure 1a depicts the highly prioritized routes drawn in the Pershing Map. An alternative proposed map from the National Interregional Highway Committee published in 1947 is the most commonly used location instrument in the U.S. interstate literature.[19] In Section 4.3.4 I will discuss the limitations of this map in a

(2020); Herzog (2020); Rothenberg (2013).

[19] Prior work has focused on measuring urban rays (Agrawal et al., 2017; Baum-Snow, 2007, 2020) or urban mileage (Duranton and Turner, 2011, 2012; Duranton et al., 2014), or has used focused on rural settings, mostly relying on the plan as an extension of the inconsequential units approach (Herzog, 2020; Michaels, 2008).

national setting with interstate treatment along the extensive margin.

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 of the proposed interstate networks.[20] In order to apply the algorithm to the historical interstate network plan in the Pershing Map, I decompose each planned road system 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. I then weight each edge by its length in kilometers to approximate a travel cost. The Newman-Girvan Algorithm calculates the edge-betweenness for each edge by determining the shortest path between every pair of nodes in the network. The algorithm then counts the number of times each edge is used to complete these shortest path trips. The total number of times an edge is used to complete these node-to-node trips is the edge-betweenness value. Edges with larger betweenness values are more commonly traveled and therefore are mathematically more important for connecting nodes in the network.

Figure 4 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

[20] This algorithm was originally used to identify important connections in biological and social networks. To my knowledge, this is the first application of the algorithm in the transportation economics literature. For specifics regarding the algorithm see Girvan and Newman (2002); Newman and Girvan (2004); Newman (2001, 2004).

thickness of the edge, where thicker edges have higher betweenness scores.

[Figure 4 Here]

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 Congressional highway proposal.[21] 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.[22] I use historical cost estimates instead of realized costs because it better approximates the decision a social planner would have made at the time of construction. Actual construction costs changed over time for several reasons as detailed by Brooks and Liscow (2019).

I calculate the annual construction constraint by dividing the total network construction cost over a twenty-five year construction period, which approximates the time-frame of actual interstate construction for the mileage of the Pershing Military Plan. Once I have an annual construction constraint, I rank the proposed networks edges with the highest betweenness scores first and build them in that order until the total amount spent on construction equals the annual construction constraint. Unbuilt edges are carried over to the next year and the process repeats. This procedure allows me to assign a construction year for each edge, which results in an interstate instrument that predicts both the location of an interstate and the year of construction.

To visually compare the construction year and mileage of the instrument to the actual interstates, Appendix Figure A.1.2 illustrates how the proposed construction horizon

[21] Estimates derived from House Document 120, submitted to the 84th Congress during the first session.

[22] The ratio of construction costs is more important to the model than the actual costs.

compares to the timing of construction for the Interstate Highway System. The figures show that the twenty five year construction horizon more closely matches the actual construction horizon.[23] Appendix Table A.1.1 shows the overlap between the instrument and the Interstate Highway System. A majority of observations fall in the matched categories of (No, No) or (Yes, Yes). Importantly the table illustrates the balance in the off-diagonals. The plan shows a balanced likelihood of predicting an interstate where an interstate was not built and building an interstate where the instrument does not predict it should be placed. This is especially important given that the Pershing map proposed less mileage than the actual interstate system built.

4.3 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 and how the different components of the exclusion restriction are plausibly satisfied. In Section 4.3.4, I perform three empirical falsification tests that support using the Pershing Map as a valid instrument.

4.3.1 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:

$$IHS_{ct} = \alpha \cdot Plan_{ct} + \lambda_{st} + \tau_{mt} + \sigma_{dt} + V'\mu_{ct} + \upsilon_{ct} \tag{2}$$

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

[23] Appendix Table B.1.6 presents results for the main specifications using a 35-year construction horizon. The findings across specifications are very similar between 25 and 35-year horizons.

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.

4.3.2 Exclusion Restriction: Pershing Route Locations

The validity of the Pershing system as a suitable instrument to address location endogeneity hinges on the degree to which the authors motives in 1921 are orthogonal to employment growth and industrial development in the latter part of the 20th century. The combination of the unique features of the Pershing system and a rich set of covariates, included in all specifications, leads to a plausible defense of the exclusion restriction.

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. I construct four controls to explicitly account for military interests in the early 1920s. The first focuses on the nodes identified in the Pershing map. These occur at endpoints and intersections within the map and are often near places of interest. To address potential nodal targeting of locations I include time varying nodal fixed effects specifically for counties where the Pershing Map contained a node. Second, I control for the distance from each county seat to the nearest node. Third, I digitized and geo-located nearly 700 World War One posts, camps, and stations in the US in 1918 (Center of Military History, 1931). Using these locations, I calculate the proximity from each county seat to the nearest World War One installation. Finally, I include distances to the nearest military conflict prior to World War One.

Another design advantage is that the Pershing system was connected with straight lines and was designed to avoid passing through the center of metropolitan areas. This creates a network style graph with straight line connections akin to those evaluated by Banerjee et al. (2020), Faber (2014), and Morten and Oliveira (2018). Straight line connections remove the possibility of local officials subtly manipulating the locations of the plan.

Even with the unique qualities of the Pershing Map one may still be concerned that pre-existing characteristics of a county may influence the outcomes independent of the location of the roads. To account for this, 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. 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 locations are orthogonal to employment growth and industrial development.

4.3.3 Exclusion Restriction: Centrality and Timing

The second layer of the exclusion restriction that the instrument must satisfy is that the timing component of the instrument must be orthogonal to employment growth and industrial development in the latter part of the 20th century. Predicted construction timing is determined by centrality within the proposed interstate network. One advantage of the approach is that it makes construction decisions over the full network, intentionally abstracting away from the state level decision making process that introduces timing endogeneity.

Borusyak and Hull (2021) raise the concern that centrality is an important missing variable when quantifying the relationship between transportation expansion and regional economic growth. To address this concern, I construct and include in the main specification, a centrality fixed effect that is in spirit of the correction proposed by Borusyak and Hull (2021). To construct the centrality fixed effect, I digitized the 1947 Rand McNally atlas of Federal and State Highways and apply the Newman-Girvan Algorithm to the full pre-interstate highway network and assign each county its maximum betweenness value.[24] I then take this continuous betweenness measure and partition it into 20 bins, which I interact with year dummies to flexibly account for the role of centrality over time. With these centrality fixed effects in all specifications, the comparisons are identified from variation in interstate treatment within similarly central counties.

To address this same conceptual issue, in another way, I include a continuous measure of market potential. I construct market potential following Harris (1954), using 1950 population and straight line distances between every pair of county seats. I include this control in all specifications.

4.3.4 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)[25] and estimate equation 1, using a time-invariant binary interstate highway indicator, with the full set of controls and fixed-effects outlined in Section 4.1 and the new covariates and fixed-effects introduced in Section 4.3. 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 wholesale trade. Table 1 reports coefficient estimates for employment and bank deposits and Appendix Table A.1.2 reports the establishment results.

[Table 1 Here]

Within each outcome, the first two columns report the OLS and IV specifications of

[24] Figure A.1.3 includes both the full route map and the resulting centrality mapped onto counties.
[25] Coverage includes each decade from the Decennial Census and the 1954 County Data Book.

interest. The coefficients reflect the estimated effect of interstates prior to their actual construction. The OLS estimates in columns (1) and (4) of Table 1 are positive and statistically significant, reflecting possible selection of faster growing counties or anticipatory growth based on expectations of highway development for both employment and bank deposits. Reassuringly, coefficient estimates using the Pershing locations as an instrument in columns (2) and (5) are not statistically significant. Appendix Table A.1.2 finds a similar positive relationship between interstates and wholesale trade establishments and retail sales establishments, but no significant relationship using the Pershing location instrument.

Appendix Table A.1.3 performs similar analysis using the 1947 Interregional Highway Committee Plan introduced in Section 2. The results 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.

As noted in Section 4.3.3, there is concern that centrality may have a direct effect on economic growth. To directly look for this, I replace the binary interstate indicator with a continuous measure of Pershing centrality to test whether centrality can predict economic growth prior to the construction of the interstate system. Results from this specification are reported for employment and bank deposits in columns (3) and (6) of Table 1 and for establishment growth in columns (3), (6), and (9) in Appendix Table A.1.2. 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 the construction of the Interstate Highway System.

Event Study: It would strengthen the credibility of the instrument to confirm that the interstate locations, and timing, predicted by the instrument are uncorrelated with growth prior to predicted construction during the construction era. This is akin to showing that the reduced form satisfies a parallel-trends assumption. To test this, I leverage the recent advances in event-study methodologies by Goodman-Bacon (2021), Schmidheiny and Siegloch (2020), and Freyaldenhoven et al. (2021) in conjunction with the county-year panel data used in this paper. Specifically, I estimate the following linear panel model with dynamic highway treatment effects:

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

The estimated parameters of interest, β_m, separately identify 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 4.3. 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 A.1.4 and A.1.5 plot the sets of β_m and 95% confidence intervals that result from estimating equation 3 where the outcomes of interest are log of employment and log of establishments, respectively. I present sub-figures for both the actual interstate locations and the predicted ones using the instrument, and for both population weighted and unweighted outcomes. All figures show a positive effect of interstates, or predicted interstates, on the outcome after construction.[26] Importantly, when using the instrument, estimated effects on the outcomes prior to the year of construction are not significantly different from zero. These

[26]The confidence intervals grow as you move closer to the construction year endpoints. This is largely driven by the reduction in the number of within-construction year observations as illustrated by panels (e) and (f) of Figures A.1.4 and A.1.5. The reduction in pre-constructions years is due to data gaps caused by the fact that the CBP was published every three years between 1953 and 1962.

results suggest the parallel trends assumption holds, and provides some additional support for the validity of the instrument.

It is important to note that the instrument is not correlated with growth prior to construction, when the specification is run with and without population weights (Figures A.1.4 and A.1.5). 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. When looking at actual interstate location and timing, when the outcomes are population weighted, we see a negative 'effect' of actual interstate construction prior to their arrival, consistent with negative selection among higher population counties. Importantly this negative selection is not visible when the instrument is used. The event study figures show the instrument is not correlated with significant employment/establishment growth prior to predicted construction, even when the outcomes are population weighted.

Planned and Never Built Segments: An additional concern is that employment or economic activity was in some way influenced by the Pershing routes, through a mechanism unrelated to interstate construction. As introduced in Section 4.2, the full Pershing Map was designed with three priority levels and the constructed instrument in this paper uses the two highest priority levels. As an additional check on the instrument, I isolate the unbuilt segments that are priority level three and compare them to the rest of the non-interstate counties by estimating equation 1, using these unbuilt segments as the highway treatment. Using the full time span of data from 1956-2016, but only including counties that never received an interstate or were Pershing priority level 3, Appendix Table A.1.4 reports OLS results for year over year employment and establishment growth. The coefficient estimates are very small and statistically insignificant, supporting the assertion that the Pershing map locations did not directly influence changes in employment or establishments.

5 Results

5.1 Interstates and Firm Dynamics

Table 2 reports results from estimating equation 1 after adding the new covariates and fixed-effects described in Section 4.3 that address military motives and centrality concerns. Panel A presents first difference results for total county employment and panel B presents first difference results for total county establishments. These outcomes in tandem give us a more comprehensive view of the industrial growth dynamics that accompanied the expansion of the interstate system. The first two columns report results using a binary interstate highway indicator for whether the county received an interstate and the third and fourth columns add a measure of interstate density.[27] For ease of interpretation, the interstate density is re-scaled to be mean zero for the set of interstate counties. Within each specification pair, the table reports OLS and IV results. The lower section of each panel reports the Kleibergen-Paap F-Statistic, which indicates the instruments are sufficient predictors of highway status, and column (2) presents the single coefficient estimate of the binary highway treatment from the first-stage regression.

Interstates induced a significant increase in employment growth, Table 2 Panel A columns 1 and 2. Specifically, year over year employment increased 0.40–0.69 percent faster, on average, in interstate counties relative to counties without an interstate. Given the average employment growth rate of 2.2 percent for interstate counties, this suggests that between 18 and 31 percent of year over year 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

[27] Interstate density is calculated as the completed kilometers of interstate highway divided by county area in square kilometers.

that interstates were responsible for between 2,400 and 4,200 jobs. Relative increases in employment are driven by counties being connected to the interstate network at the extensive margin with no additional gains at the intensive margin (columns 3 and 4).[28]

On average, the number of establishments grew 0.34–0.61 percent faster among interstate counties relative to non-interstate counties, Table 2 Panel B columns 1 and 2.. 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. This suggests that interstates played a significant role in shaping firm location decisions and creating local environments that encourage urbanization economies.

[Table 2 Here]

In addition to encouraging growth in total establishments, interstates may lead to potential distributional differences in the size of establishments that grow if interstates promote economies of scale. I modify equation 1 to estimate the effect of interstates on year to year changes in the share of firms within each of the six employment size categories. Panel A of Table 3 presents the OLS results and Panel B presents the IV results, where each column is a separate employment size category. The results indicate interstates promoted scale increases in employment, which is in line with the trends in Appendix Table B.1.3. The share of establishments with fewer than 20 employees fell sharply as they were replaced with much larger firms. The results in Panel B indicate that interstates induced significant increases in the share of establishments with 50–99 employees and those with 250–499 employees.

[Table 3 Here]

Tying the results together from Tables 2 and 3, interstate highways led to faster growth in both employment and the number of establishments in interstate counties relative to

[28]In Section 6, I re-estimate the specification using only a density based interstate treatment. The results reveal a positive relationship between interstate highway density and employment and establishment growth. Appendix Table B.1.7 shows a slightly stronger relationship between interstate density and growth in establishments.

non-interstate counties. This employment growth altered the size distribution of establishments, leading to a reduction in the fraction of small firms and an increase in the fraction of medium and large firms.

Consistent with prior work on interstate highways and other large transit investments, the results in Table 2 are larger in magnitude in the IV specifications compared to OLS results. As outlined in Section 4.3.4, this bias is driven by negative selection among higher population counties. The difference between OLS and IV estimates in Table 2 reflect the combination of endogenous location selection and construction timing. We know from prior literature that some of these differences are driven by counties around lower performing metropolitan areas that received more mileage, like those described in Duranton and Turner (2012). Section 5.3 will explore the extent to which construction timing and location targeting among the set of non-MSA counties contribute to the observed differences.

5.2 Interstate Treatment Across Space

Decomposing the effects of interstates in space is important for understanding the spatial extent of transportation network benefits. I analyze the different relevant definitions of spatial adjacency, visualized in Figure 3, by adding binary adjacency treatments into equation 1. Table 4 builds on the interstate effects in Table 2 and adds three different adjacency treatments. The first is whether or not a county is spatially adjacent to a treated highway county, which is similar to the spatial comparison in Chandra and Thompson (2000) who focus on only rural counties. This specification includes the binary interstate treatment and another binary treatment for all directly adjacent counties, making the comparison group those counties farther away from the interstate system.

The second is whether a county is adjacent within a commuting zone. This specification introduces a single binary treatment for county types A and B from Figure 3. The third partitions 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 both of these specifications, 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. Table 4, columns (1)–(2), which include an interaction for general interstate adjacency, show there is no change in either employment or establishments for adjacent counties once endogeneity has been addressed. When the definition of adjacency is altered to be non-interstate counties within an interstate treated labor market, there is still no impact on establishments or employment, Table 4 columns (3)–(4).

Findings show significant positive spillovers occur in 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 include an MSA, Table 4 columns (5)–(6), we see significant positive impacts on industrial growth for both interstate counties and non-interstate counties in the metropolitan periphery.

[Table 4 Here]

5.3 Time Path of Treatment

Constructing a time-varying instrumental variable for interstate highways allows me to explore the time path of interstate treatment. To do this, I partition the 60 year sample into three 20 year eras: the expansion era (1956–1975), the completion era (1976–1995), and the post-construction era (1996–2016). During the expansion era, the interstate network remained largely disconnected because construction decisions were made at the state-level.[29] In 1975 roughly 75 percent of the system was complete and remaining segments of the original network plan were officially completed in 1992.[30] Post-construction interstate

[29] Appendix Figure A.1.6 illustrates the fractured nature of the system and contrasts the predicted Pershing construction in the same year.

[30] Appendix Figure A.1.2 plots the time path of highway construction.

expansion mostly consists of additional lane mileage (Turner et al., 2020).

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

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

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 5, columns (1) and (2), report the three era specific coefficients of interest for OLS and IV specifications. OLS estimates in column (1) suggests that interstates led to stable gains in employment and establishments, with very little change in the coefficients across the three eras. A different pattern emerges once I address location and timing endogeneity. IV estimates reveal the largest employment and establishment 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.

Next, I consider the time path of spatial spillovers in the metropolitan periphery. I do this by extending equation 4 to include era interactions of the non-interstate counties within interstate treated labor markets that contained an MSA, Table 5 columns (3) and (4).[31] 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.

[Table 5 Here]

[31] See type A in Figure 3.

Results show that failing to address the endogenous timing of interstate construction leads to understating the early era growth gains from interstates. To determine how important it is to address the endogenous timing of interstate construction, I replace the time varying interstate measure with a fixed measure of interstate treatment based on whether or not a county would eventually be connected to the interstate system. Results are reported in Table 5 columns (5) and (6). Estimates using the fixed highway status are smaller in magnitude, with the most pronounced differences during the two decades of significant expansion.

This analysis makes unpacking endogeneity patterns over time possible. The larger IV estimates in the early eras, Table 5 within columns (2) and (6), are consistent with interstate construction targeting lower performing areas for both interstate routes *and* early construction. This result was previewed in the event study figures in Section 4.3.4 and is supported by the historical narrative that federal officials prioritized urban interstate development.[32] 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. This early targeting of urban locations likely contributed to the highway revolts of the 1960s (Brinkman and Lin, 2019).

The endogeneity pattern flips for employment in the post-construction era, but not for establishments. This reveals that endogeneity in this context is multidimensional and dynamic. Part of the changing dynamics is the potential for changes in the set of compliers over time. Following the methodology proposed in Marbach and Hangartner (2020), Appendix Table A.1.5 presents the results from characterizing compliers for a subset of pre-interstate county characteristics.[33] Column (1) presents means for the full sample of

[32]The first two columns of Appendix Table B.1.8 confirm this result by comparing the pre-construction economic characteristics between counties with interstates constructed in the first decade compared to the second decade. The differences indicate that state and local officials prioritized constructing interstates in larger, more urban counties, with better access to transportation infrastructure. These differences are less pronounced when comparing across predicted construction periods.

[33]Their approach requires assuming that the characteristics of the subpopulations are independent of

counties, column (2) presents the mean characteristics for the set of compliers in the full sample. Compliers appear to be drawn from slightly more rural, lower population areas. The next three columns present only complier subpopulation means from the separate twenty year eras. The results show that over time the compliers shift toward more urban, higher population areas. The most striking difference is in the fraction of compliers in 1960 MSA boundaries. In the early era, only 7 percent of compliers were from an MSA, whereas by the post-construction era that has increased to nearly 18 percent. This is a significant shift given that across the entire sample only 11 percent of counties are in MSAs. This evidence is consistent with different endogeneity concerns over time.

Another potential explanation for the changing endogeneity pattern across eras relates to the main structural differences between the Pershing Map and eventual interstate system. The Pershing system was designed to connect the country and not to facilitate urban transit, therefore it lacks the additional radial segments built to facilitate metropolitan commuting. The endogeneity pattern seen in Table 5 is consistent with these radial or extra commuting based segments being built in response to demand from better performing areas. Given that the conventional endogeneity pattern holds for establishments, it is possible larger, high employment firms are locating near these radial segments built in response to growth. Overall, exploring the time path of treatment highlights the complexity of the endogeneity patterns and emphasizes the need to address both timing and location endogeneity.

6 Findings in the Context of Prior Literature

This paper is most closely related to the work by Duranton and Turner (2012), which estimates the impact of metropolitan interstate mileage on employment growth during the 1980s and 1990s. Their canonical takeaway is that 10 percent more mileage in an MSA

treatment status. This is not the case here, but I still believe this is a useful exercise that provides suggestive evidence to explain the changing coefficient dynamics by era.

leads to 1.5 percent more employment over a 20 year period. While my paper does not directly estimate elasticities, I can use estimates found in Table 5 to determine a back of the envelope comparison. If I focus on the closest era to Duranton and Turner (2012), 1976-1995, and compound the 0.008 percent growth difference over 20 years, the result is 14 percent more employment. This is similar to the expected employment gains from applying a 100 percent increase in metropolitan mileage to the results in Duranton and Turner (2012).[34] This finding highlights that my estimates for a specific era are similar to previous literature, however I add to our overall understanding of the dynamic nature of the impacts of interstates by showing that employment growth was substantially larger in the initial expansion era (1956–1975) and drops off considerably in the completion and post-construction eras, 1976–1995 and 1996–2016 respectively.

The rich county-year panel data set used in this paper, in conjunction with an instrument that accounts for location and timing endogeneity, gives more insight into the dynamic nature of endogeneity and how it compares to endogeneity identified in previous literature. The IV coefficients in Duranton and Turner (2012) are roughly 2 to 2.5 times larger than the OLS estimates. The difference in magnitudes in Table 5 are similar during the era that Duranton and Turner (2012) are discussing, the completion era of 1976–1995. The magnitude difference between IV and OLS estimates is 3.25 in the initial expansion era (1956-1975), likely reflecting the increased endogeneity in the early construction process. This highlights the importance of addressing timing endogeneity explicitly, and suggests prior literature may be underestimating impacts.

Identifying spatial patterns in the spillovers induced by interstate highways is another area that the national dataset used in this paper has advantages over the prior literature. In a comparison among rural areas, Chandra and Thompson (2000) find that interstates

[34] Another way to think about this is to use the distance density results in Appendix Table B.1.7. From the coefficient estimates in column (2), a one standard deviation increase in interstate density leads to 0.002 percent faster growth in employment. Compounding that growth over 20 years results in roughly 4 percent more employment growth following a standard deviation increase in interstate density. This is a comparable estimate to a 25–30 percent increase in metropolitan mileage in Duranton and Turner (2012).

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 no evidence of negative spillovers.[35] The results in Table 4 instead point to positive spillovers in adjacent areas within labor markets that included a major metropolitan area.

7 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). The expansion of the Interstate Highway System over the second half of the twentieth century provides insights into how large scale transit infrastructure alters the location of economic activity in the context of a developed country. Evaluating these effects over time requires a new empirical approach that addresses the endogenous placement and endogenous timing of highway construction. Results from this paper indicate that, counties traversed by interstates experienced early and persistent employment and establishment growth relative to unconnected counties. The employment growth was concentrated among larger firms. The interstate system also induced positive spatial spillovers in employment and establishment growth among non-interstate counties adjacent to metropolitan areas in recent decades.

This paper reveals the importance of addressing the endogenous timing of interstate construction, which is complex and dynamic. This type of network based instrument has many applications including settings where the allocation of resources can be patterned according to a network. The most direct applications include other forms of infrastructure such as rail, airports, or electricity. It could potentially be applied in settings where firms are sourcing inputs allocating outputs to other firms or consumers. Additionally, the

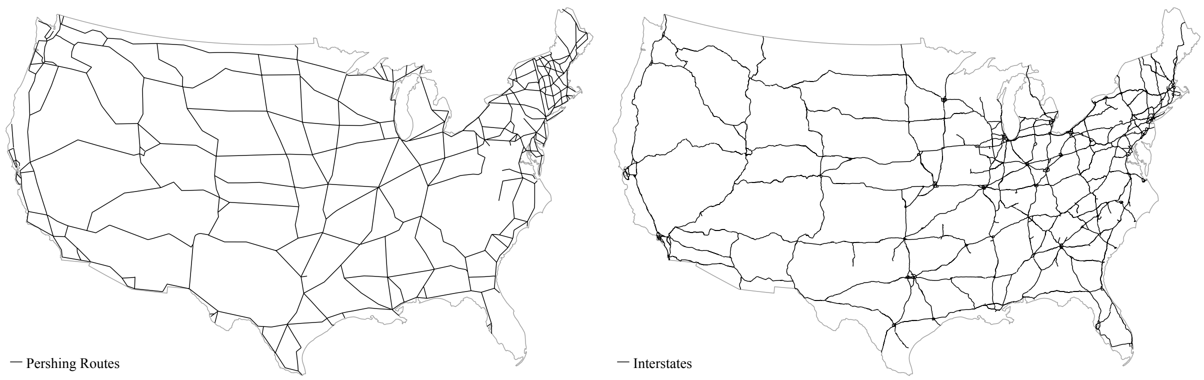
[35]This finding is consistent with Herzog (2020) who estimates a quantitative spatial equilibrium model.

broader classification of graph theory based algorithms are applicable in settings where governments or agencies are determining the distribution of project based spending over time.

The data and empirical methods developed in this paper open new opportunities to evaluate temporal and spatial patterns in the ways interstates shaped regional economies. Changes in trade and commuting costs are distributed differently across industries, which leads to a natural extension to evaluate how interstates altered industrial growth by sector and whether it contributed to spatial concentration in sectors.

8 Figures

Figure 1: Proposed and Constructed National Highway Network Plans

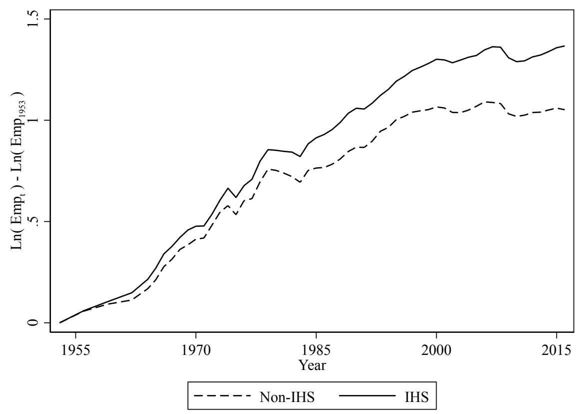


(a) Pershing Military Plan

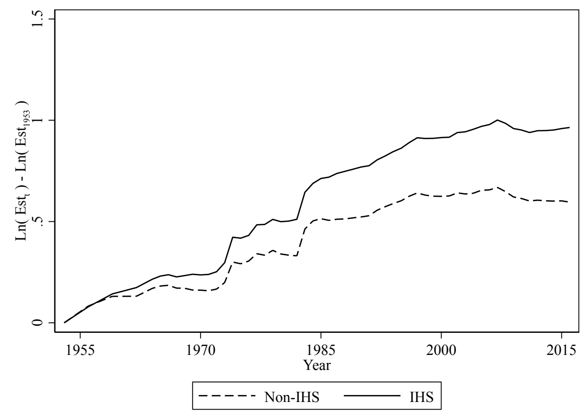
(b) Constructed Interstate Highway Locations

Notes: Proposed Pershing routes digitized from original map housed at the U.S. National Archives. Map of current Interstate Highway System locations from Federal Highway Administration.

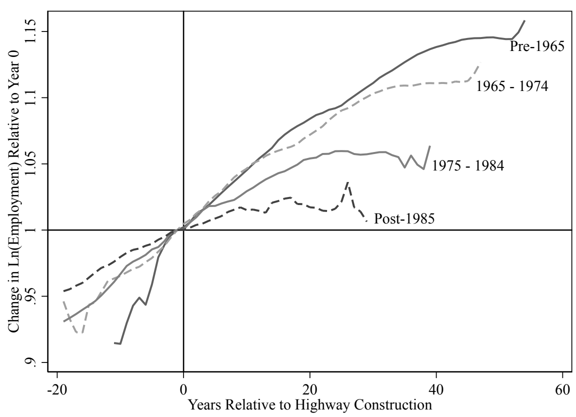
Figure 2: Illustrating Employment and Establishment Growth by Interstate Status and Construction Era



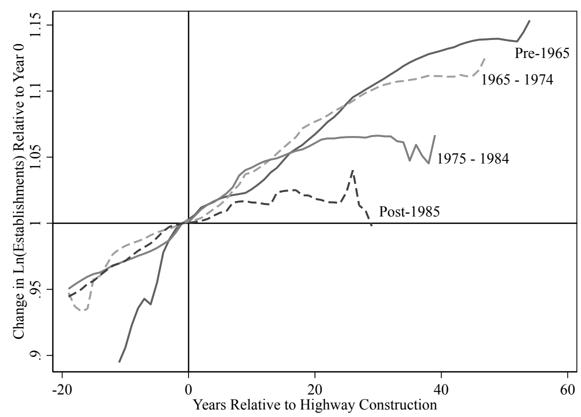
(a) Employment Growth by Interstate Status



(b) Establishment Growth by Interstate Status



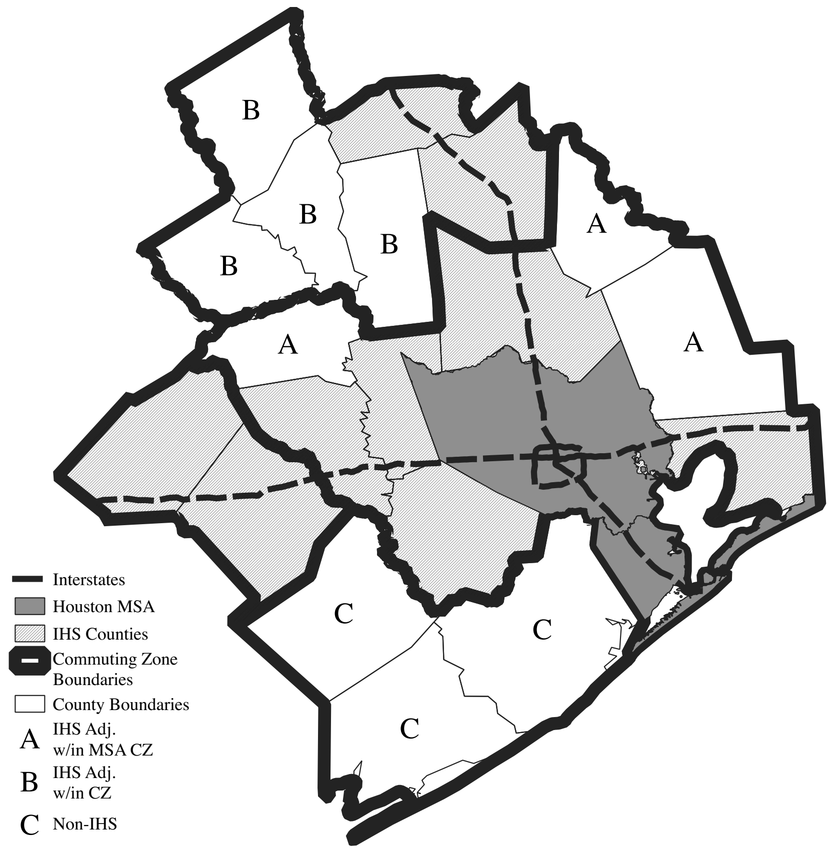
(c) Employment Growth by Interstate Construction Era



(d) Establishment Growth by Interstate Construction Era

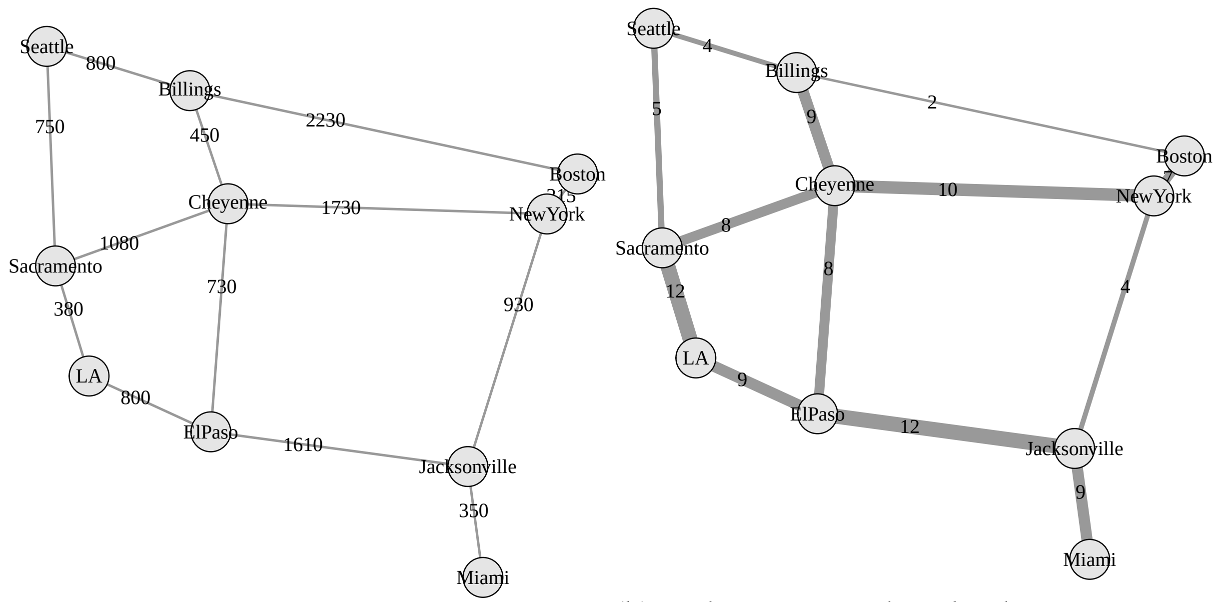
Notes: Employment and establishment data from 1953–2016 County Business Patterns annual reports. Highway designation based on highway status in 2016. Construction date information from PR-511 collection. Top row compares interstate and non interstate counties over time relative to their 1953 level of 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 3: 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 5.2

Figure 4: 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 4a presents the ten cities with approximate distances between cities listed along each edge. Figure 4b presents the edge betweenness calculation along each edge and adjusts the line width to reflect higher betweenness values.

9 Tables

Table 1: Effects of Interstate on Employment and Banking Prior to Construction

	Δ Employment			Δ Bank Deposits		
	(1)	(2)	(3)	(4)	(5)	(6)
Interstate (0/1)	0.0022***	-0.0049		0.0011*	-0.0034	
	(0.0007)	(0.0045)		(0.0007)	(0.0036)	
Pershing Centrality			0.0002			0.0002
			(0.0002)			(0.0002)
Observations	12,342	12,342	5,188	11,666	11,666	4,967
Counties	3,100	3,100	1,300	2,932	2,932	1,247
State X Years	192	192	192	188	188	188
KP F-Statistic		85.50			74.90	

Notes: The specifications in columns (1), (2), (4), and (5) report results from estimating a modified equation 1, which regresses a time invariant binary interstate highway indicator on average period over period changes in the log outcome. The specifications in columns (3) and (6) replace the binary interstate treatment with a continuous measure of the maximum centrality score associated with the Pershing plan in the county. Every specification includes state × year fixed effects, along with the full set of controls outlined in Appendix C.2. 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 is the total inflation adjusted value of deposits. 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 × year. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Interstate Highways and Growth in Employment and Establishments

Panel A: Employment

	(1) OLS	(2) IV	(3) OLS	(4) IV
Interstate (0/1)	0.0040***	0.0069***	0.0037***	0.0072***
	(0.0005)	(0.0026)	(0.0006)	(0.0027)
Interstate KMs per County Sq KM			0.0140	-0.0112
			(0.0140)	(0.0410)
Observations	171,940	171,940	171,940	171,940
Counties	3,071	3,071	3,071	3,071
State X Years	2,688	2,688	2,688	2,688
Pershing First-Stage		0.182		
		0.017		
KP F-Statistic		117.74		58.73

Panel B: Establishments

	(1) OLS	(2) IV	(3) OLS	(4) IV
Interstate (0/1)	0.0034***	0.0061***	0.0031***	0.0061***
	(0.0004)	(0.0022)	(0.0005)	(0.0023)
Interstate KMs per County Sq KM			0.0126	0.0037
			(0.0111)	(0.0338)
Observations	171,938	171,938	171,938	171,938
Counties	3,071	3,071	3,071	3,071
State X Years	2,688	2,688	2,688	2,688
Pershing First-Stage		0.182		
		0.017		
KP F-Statistic		117.74		58.73

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 both the binary indicator and a measure of interstate highway density, which is rescaled to be mean 0 when the binary measure is equal to 1. Every specification includes state $\times$ year fixed effects, along with the full set of controls outlined in Appendix C.2. 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 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)
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

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)
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
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

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 outlined in Appendix C.2. 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 4: Interstates and Growth by Spatial Adjacency

Panel A: Employment

	General Adjacency		Commuting Zone Adjacency		Commuting Zone Adj. by MSA Status	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
Interstate (0/1)	0.003***	0.007**	0.004***	0.006*	0.005***	0.007**
	(0.001)	(0.003)	(0.001)	(0.003)	(0.001)	(0.003)
IHS Adj.	-0.001*	-0.000				
	(0.001)	(0.003)				
IHS Adj. w/in CZ			-0.001*	0.001	-0.002***	-0.001
			(0.001)	(0.002)	(0.001)	(0.002)
IHS Adj. w/in MSA CZ					0.002***	0.003***
					(0.001)	(0.001)
KP F-Stat		44.50		46.35		30.52

Panel B: Establishments

	General Adjacency		Commuting Zone Adjacency		Commuting Zone Adj. by MSA Status	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
Interstate (0/1)	0.003***	0.006**	0.004***	0.005**	0.004***	0.006**
	(0.001)	(0.002)	(0.000)	(0.003)	(0.000)	(0.003)
IHS Adj.	-0.001*	-0.001				
	(0.000)	(0.002)				
IHS Adj. w/in CZ			-0.001	0.001	-0.001***	-0.002
			(0.000)	(0.002)	(0.000)	(0.002)
IHS Adj. w/in MSA CZ					0.003***	0.004***
					(0.000)	(0.001)
KP F-Stat		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, plus an indicator for directly adjacent counties. Columns (3)–(4) report results with both the binary indicator and an indicator for all adjacent counties within an interstate treated commuting zone. Columns (5)–(6) partitions the prior adjacency to distinguish adjacent commuting zones that include a MSA. Every specification includes state $\times$ year fixed effects, along with the full set of controls outlined in Appendix C.2. 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 $\times$ year fixed effects. Standard errors are two-way clustered by county and commuting zone $\times$ year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 5: 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 4.

References

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

Alder, S. (2019). Chinese roads in india: The effect of transport infrastructure on economic development.

Alder, S., M. Roberts, and M. Tewari (2017). The effect of transport infrastructure on india?s urban and rural development. *Unpublished manuscript, University of North Carolina, Chapel Hill*.

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.

Bartelme, D. (2018). Trade costs and economic geography: evidence from the us. *Work. Pap., Univ. Calif., Berkeley*.

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.

Boarnet, M. G. (2014). National transportation planning: Lessons from the u.s. interstate highways. *Transport Policy 31*, 73–82.

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 (2019). Freeway revolts!

Brooks, L. and Z. D. Liscow (2019). Infrastructure costs. *Available at SSRN 3428675*.

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. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics 92*(1), 1–45.

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. (2020). National transportation networks, market access, and regional economic growth.

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

Jaworski, T., C. T. Kitchens, and S. Nigai (2020). Highways and globalization. *NBER Working Paper 27938*.

Jedwab, R. and A. Storeygard (2020). The average and heterogeneous effects of transportation investments: Evidence from sub-saharan africa 1960-2010. Technical report, National Bureau of Economic Research.

Karnes, T. (2009). *Asphalt and Politics: A History of the American Highway System*. McFarland.

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.

Marbach, M. and D. Hangartner (2020). Profiling compliers and noncompliers for instrumental-variable analysis. *Political Analysis 28*(3), 435–444.

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., L. S. o. E. Performance, and P. S. C. for Economic (2007). The long-term consequences of regional specialization [electronic resource].

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.

Morten, M. and J. Oliveira (2018). The effects of roads on trade and migration: Evidence from a planned capital city. *NBER Working Paper 22158*, 43.

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.

Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E 70*(5), 56131.

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.

Rothenberg, A. D. (2013). Transport infrastructure and firm location choice in equilibrium: evidence from indonesia?s highways. *Unpublished Manuscript, Department of Economics, University of California, Berkeley*.

Schmidheiny, K. and S. Siegloch (2020). On event study designs and distributed-lag models: Equivalence, generalization and practical implications.

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.

U.S. Census Bureau (2022). LEHD Origin-Destination Employment Statistics (LODES).

U.S. Department of Transportation Statistics (1960). *Highway Statistics*. Federal Highway Administration.

U.S. Department of Transportation Statistics (1970). *Highway Statistics*. Federal Highway Administration.

U.S. Department of Transportation Statistics (2017). *Highway Statistics*. Federal Highway Administration.

U.S. Dept of Transportation (1977). *America's highways, 1776-1976: a history of the Federal-aid program*. U.S. Governmant Printing Office.

Appendix

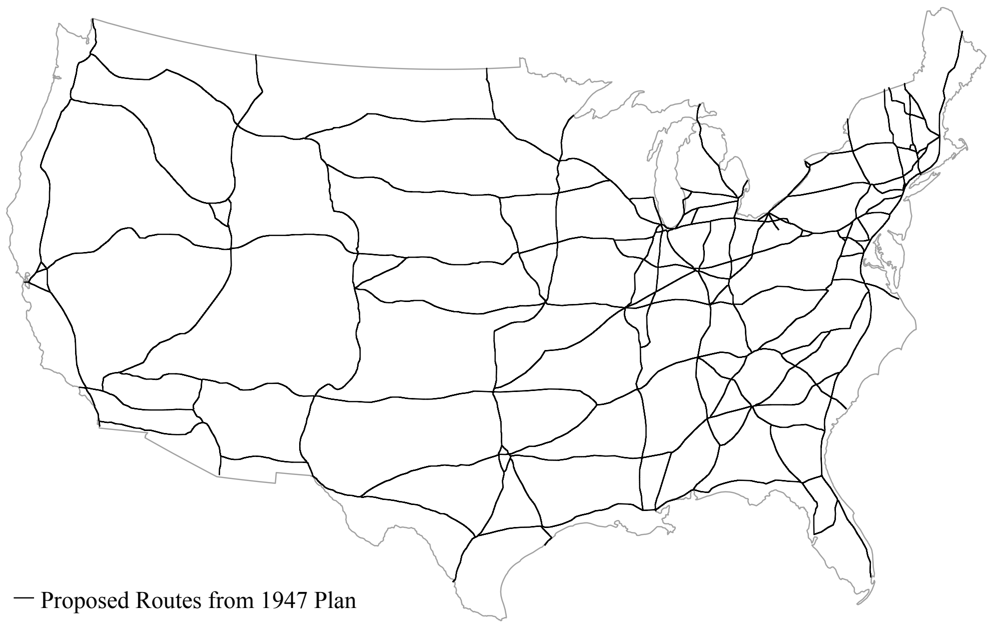
to

"Transportation Networks and the Geographic Concentration of Employment"

Appendix A Exhibits Relating to IV

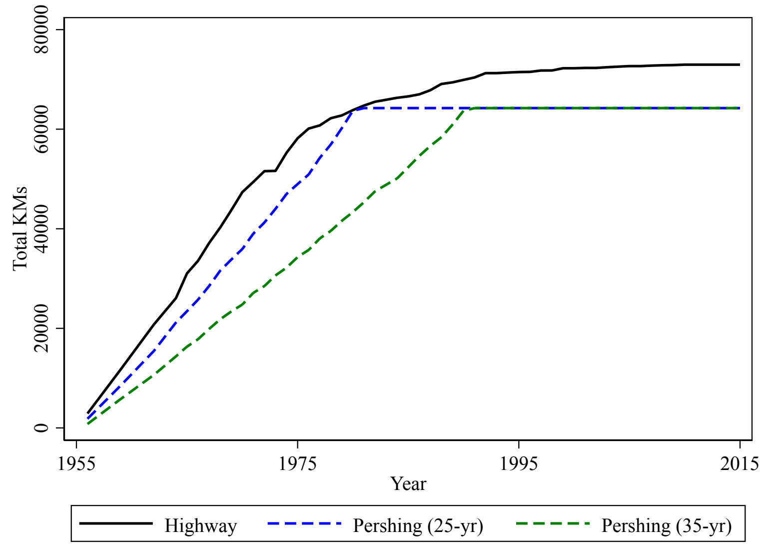
Appendix A.1 Figures: Highway Construction and Illustrating the Instrument

Figure A.1.1: 1947 Plan from Interregional Highway Committee

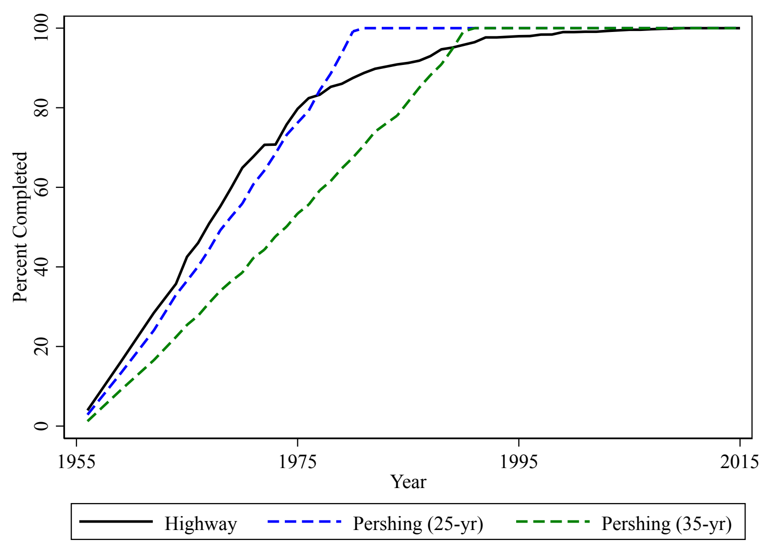


Notes: Interregional Highway Committee routes digitized from Public Roads Administration map.

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



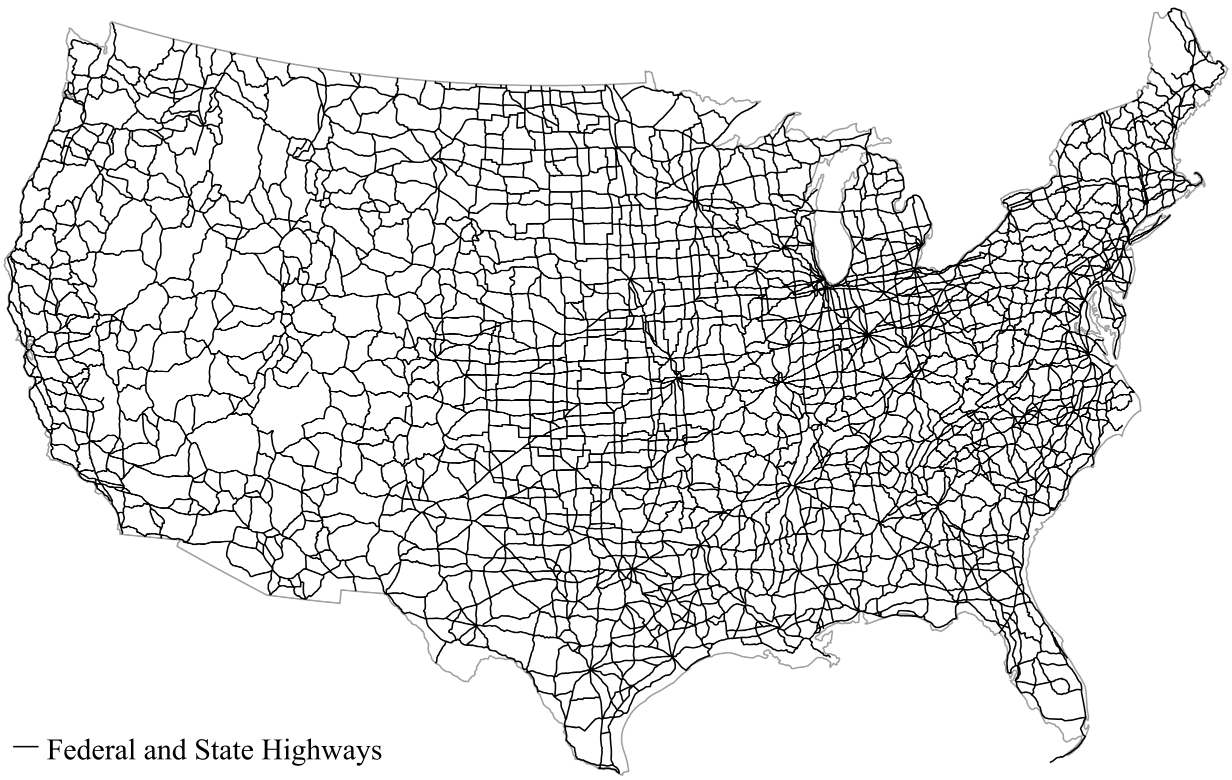
(a) Construction in KMs over 25 years



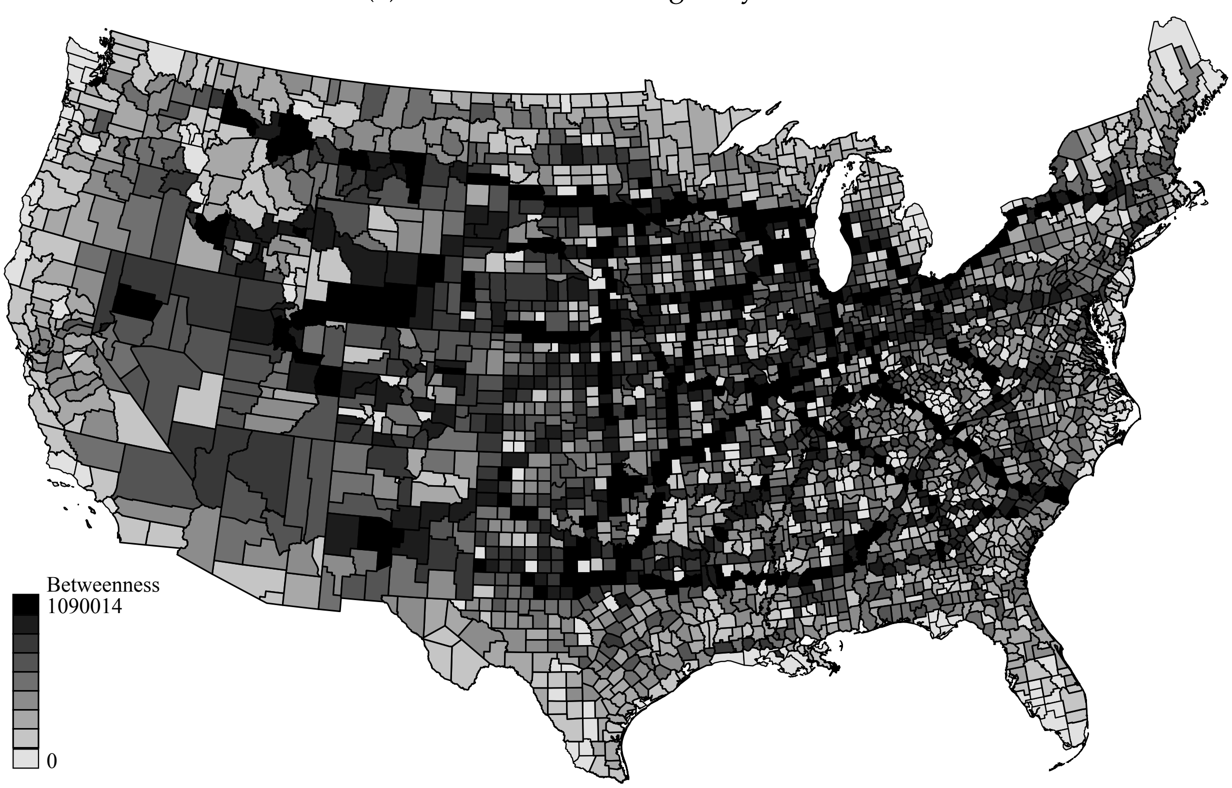
(b) Construction as a Share Completed over 25 years

Notes: Figure plots the expansion of interstate highways over time (solid black line) with the top figure presenting mileage and the bottom figure 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.

Figure A.1.3: Mapping Pre-Interstate Highway Centrality



(a) Federal and State Highways in 1947

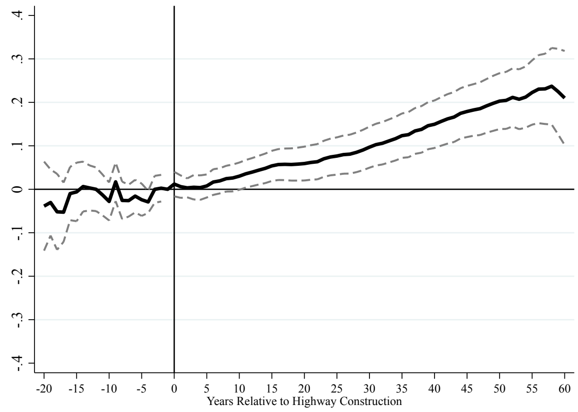


(b) Mapping Centrality from 1947 State and Federal Highways

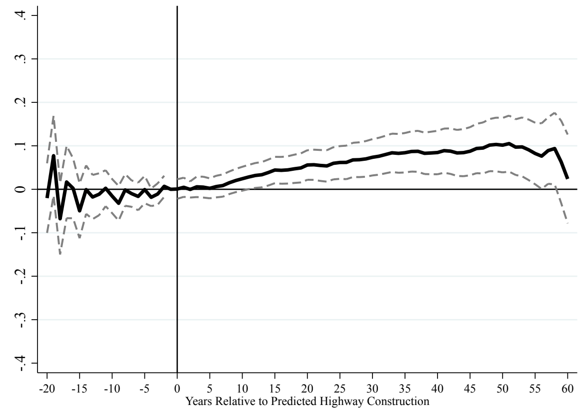
Notes: Panel (a) plots the full set of federal and state highways in 1947. Panel (b) maps centrality values for this same network based on Newman (2001) and Newman and Girvan (2004) into nine categories.

Figure A.1.4: Employment Changes as an Event Study with Interstate Highways and Proposed Pershing Highways

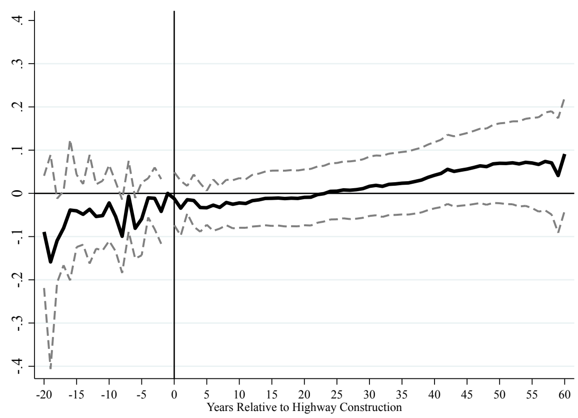
Interstate Highways | **Pershing Routes**



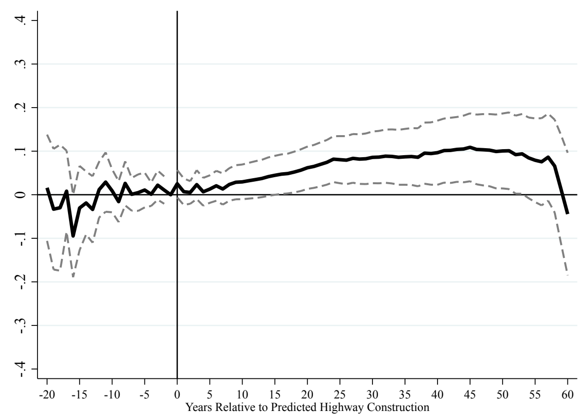
(a) Unweighted Employment



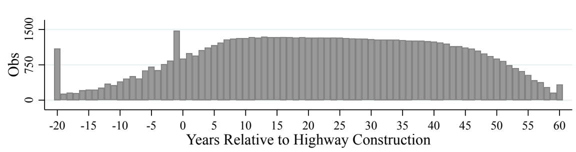
(b) Unweighted Employment



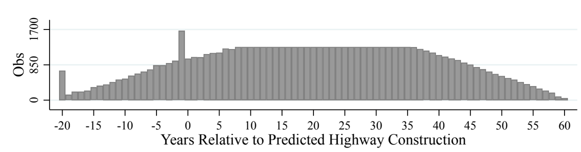
(c) Population Weighted Employment



(d) Population Weighted Employment



(e) Density of Actual Construction

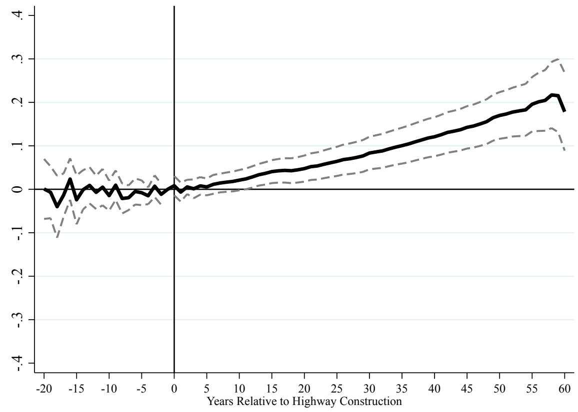


(f) Density of Predicted Construction

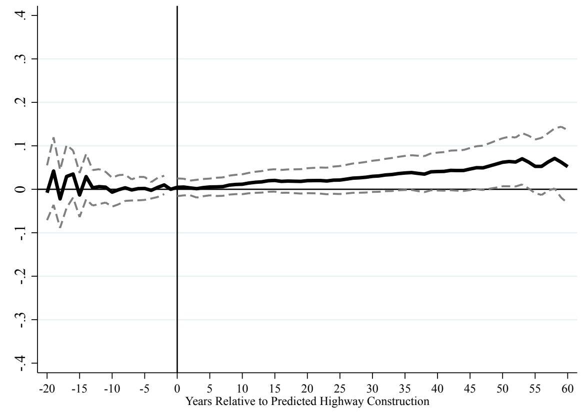
Notes: The figure plots event study (eqn 3) estimates of the effects of interstate highways (left column) and the Pershing reduced form locations (right column) on employment. Year $t-1$ is the excluded year. The top row presents unweighted estimates. The second row presents population weighted estimates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

Figure A.1.5: Establishment Changes as an Event Study with Interstate Highways and Proposed Pershing Highways

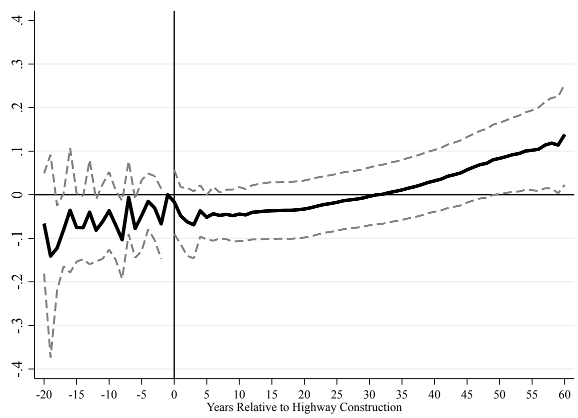
Interstate Highways | **Pershing Routes**



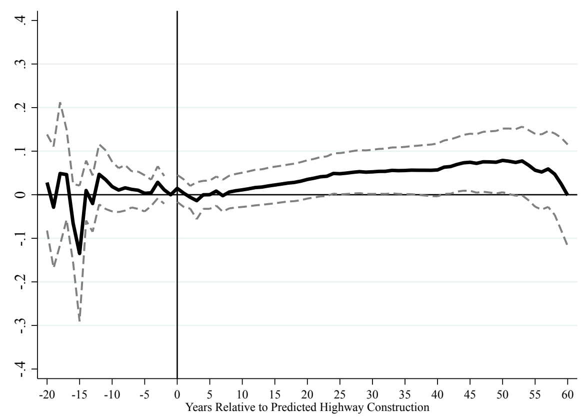
(a) Unweighted Establishments



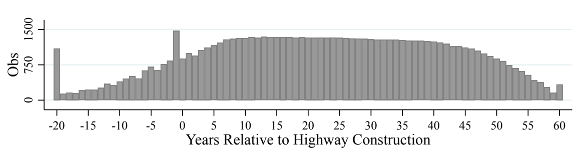
(b) Unweighted Establishments



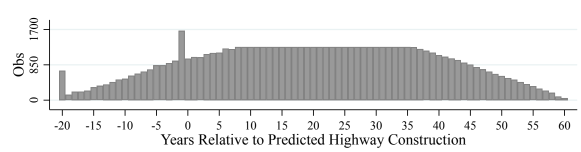
(c) Population Weighted Establishments



(d) Population Weighted Establishments



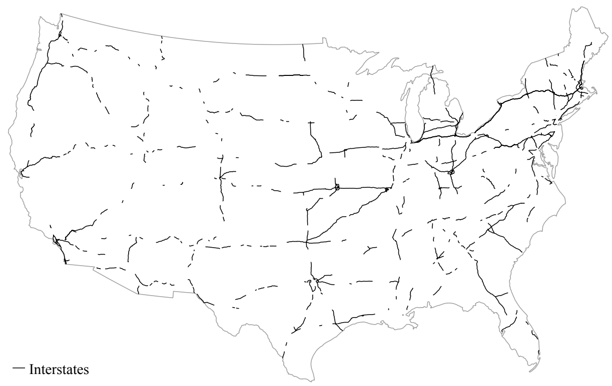
(e) Density of Actual Construction



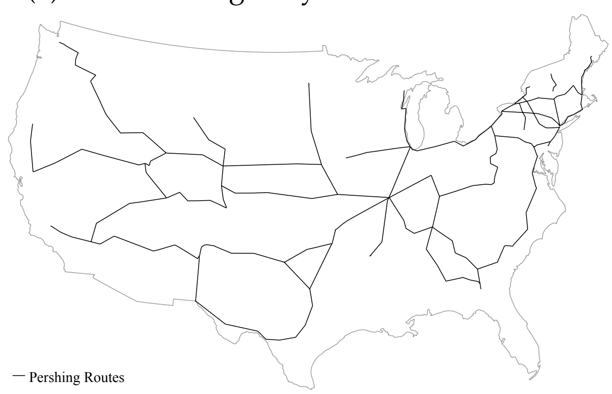
(f) Density of Predicted Construction

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.

Figure A.1.6: Highway Construction in 1965



(a) Interstate Highway Construction in 1965



(b) Predicted Pershing Construction in 1965

Notes: Figure maps completed and predicted segments of the Interstate Highway system in 1965. Sub-figures (a) plots the completed segments, sub-figure (b) plots the predicted Pershing segments.

Appendix A.2 Tables: Highway Construction and Illustrating the Instrument

Table A.1.1: Overlap Between Interstate Highway Assignment and Proposed Highway Maps

	Pershing Overlap		
	Interstate		
Pershing	**No**	**Yes**	**Total**
No	69,832	29,792	99,624
Yes	26,040	46,480	72,520
Total	95,872	76,272	172,144
No	72.84%	39.06%	57.87%
Yes	27.16%	60.94%	42.13%
Total	100.00%	100.00%	100.00%

Notes: Pershing routes digitized from original map housed at the U.S. National Archives.

Table A.1.2: Effects of Interstates on Industry Prior to Construction

	Manufacturing Estab			Wholesale Estab			Retail Estab.		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	IV	OLS	IV	OLS	OLS	OLS	IV	OLS
Interstate (0/1)	0.00006	0.00503		0.00151*	-0.00497		0.00133***	0.00003	
	(0.00100)	(0.00568)		(0.00077)	(0.00494)		(0.00048)	(0.00259)	
Pershing Centrality			-0.00009			0.00003			-0.00002
			(0.00019)			(0.00021)			(0.00009)
Observations	10,176	10,176	4,584	11,887	11,887	5,101	12,389	12,389	5,198
Counties	2,556	2,556	1,148	3,013	3,013	1,284	3,098	3,098	1,300
State X Years	192	192	192	192	192	192	192	192	192
KP F-Statistic		64.61			80.28			84.63	

Notes: The specifications in columns (1), (2), (4), (5), (7), and (8) report results from estimating a modified equation 1, which regresses a time invariant binary interstate highway indicator on average changes in the log outcome. The specifications in columns (3), (6), and (9) replace the binary interstate treatment with a continuous measure of the maximum centrality score associated with the Pershing plan in the county. Every specification includes state $\times$ year fixed effects, along with the full set of controls outlined in Appendix C.2. Outcome data from 1930–1950 decadal censuses and 1954 County Business Patterns reported at the county level. Manufacturing, Retail Trade, and Wholesale Trade all the 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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.1.3: Comparing the Effects of Highways Prior to Construction Using the 1947 Plan

Panel A: IV Specifications Using 1947 Plan

	Employment	Bank Deposits	Manufacturing	Retail	Wholesale
	(1)	(2)	(3)	(4)	(5)
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 B: Centrality Specifications Using 1947 Plan

	Employment	Bank Deposits	Manufacturing	Retail	Wholesale
	(1)	(2)	(3)	(4)	(5)
Centrality Score	-0.00008	0.00036	-0.00006	0.00024	-0.00082
	(0.00074)	(0.00054)	(0.00069)	(0.00031)	(0.00073)
Observations	4,784	4,607	4,290	4,804	4,752
Counties	1,201	1,155	1,075	1,201	1,192
State X Years	188	184	188	188	188

Notes: The specifications in Panel A report results from estimating a modified equation 1, which regresses a time invariant binary interstate highway indicator on average changes in the log outcome. The specifications in Panel B replace the binary interstate treatment with a continuous measure of the maximum centrality score associated with constructing 1947 Interregional Highway Plan in the county. Every specification includes state $\times$ year fixed effects, along with the full set of controls outlined in Appendix C.2. 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 is the total inflation adjusted value of deposits; Farms, Manufacturing, Retail Trade, and Wholesale Trade all the 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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.1.4: Effect of Proposed but Never Constructed Priority Three Military Plan

	Employment	Establishments
	(1)	(2)
Unbuilt Pershing (0/1)	0.0008	0.0012
	(0.0016)	(0.0013)
Observations	69,681	69,679
Counties	1,245	1,245
State X Years	2,352	2,352

Notes: The specification estimates the effect of unbuilt Pershing priority three segments on year over year employment and establishment growth. The sample is restricted to non-interstate counties, where the Unbuilt Pershing counties are the subset with proposed, but never built segments from original full Pershing Map. Regression includes state × year fixed effects, along with the full set of controls outlined in Appendix C.2. Employment data 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.

Table A.1.5: Comparing Compliers by Era

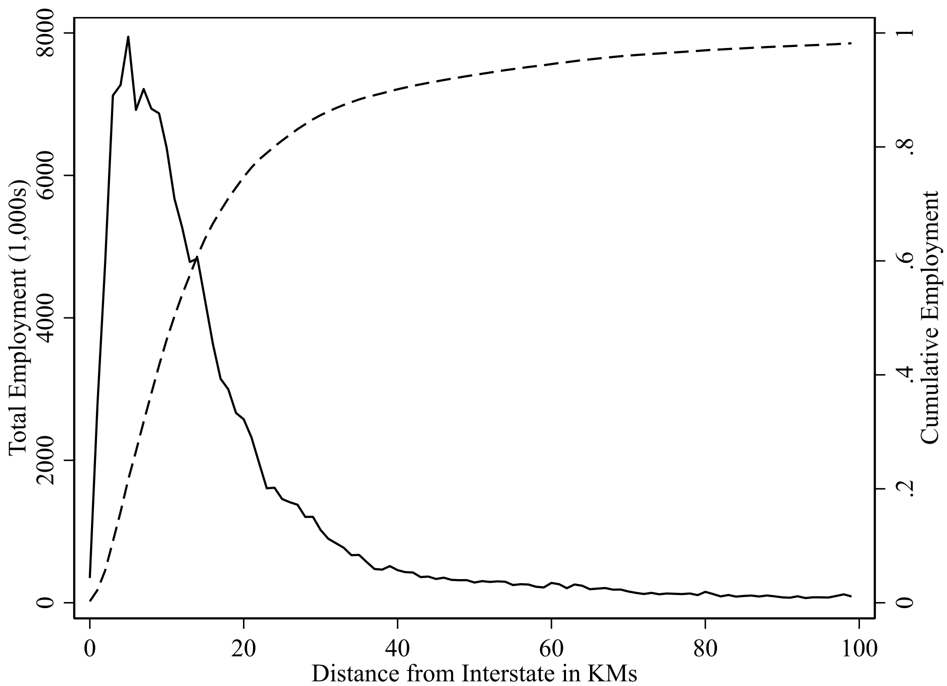
	Full Sample		Compliers By Era		
	(1)	(2)	(3)	(4)	(5)
	All	Complier	Pre-1975	1976–1995	Post-1995
Share of Counties	1	.357	.346	.345	.343
Counties	3,071	1,095	1,064	1,060	1,054
Ln(Population in 1950)	9.9	9.81	9.61	9.87	9.89
Pct of Pop in Urban Area	.281	.268	.217	.28	.293
1960 MSA County (0/1)	.112	.141	.0718	.165	.179
Metro County (0/1)	.204	.201	.107	.22	.245
Rural County (0/1)	.511	.575	.65	.564	.557
Ln(Employ in 1953)	7.68	7.53	7.28	7.6	7.63
Ln(Estab in 1953)	5.68	5.54	5.33	5.6	5.63
KM to MSA	102	101	109	100	101
1911 RR KMs per sq mi	.186	.185	.156	.192	.195
KM to 1947 Highways	90.1	90.7	86.9	94	95
1947 Hwy Betweenness	68,782	72,243	63,166	72,855	71,660

Notes: The table presents the means, by era, of several county level characteristics taken prior to interstate construction for compliers from applying the methodology in Marbach and Hangartner (2020).

Appendix B Table and Figure Appendix

Appendix B.1 Appendix Figures

Figure B.1.1: Employment and Distance from Interstate Highways



Notes: The figure illustrates total employment and the share of national employment within fixed distances from the Interstate Highway System. Employment is tabulated in 2016 from LODES at the Census Block Group and distance is measured from the Block Group centroid to the nearest interstate.

Appendix B.2 Appendix Tables

Table B.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 B.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 C.2. 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 $\times$ year. *** p<0.01, ** p<0.05, * p<0.1.

Table B.1.3: 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 B.1.4: Estimates With Varying Spatial Fixed-Effects

Panel A: Employment

	Flexible Covariates		CZ × Year FEs		Census × Year FEs	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
Interstate (0/1)	0.0040***	0.0069***	0.0037***	0.0071*	0.0039***	0.0060**
	(0.0006)	(0.0026)	(0.0006)	(0.0040)	(0.0005)	(0.0026)
Observations	171,940	171,940	167,124	167,124	171,996	171,996
Counties	3,071	3,071	2,985	2,985	3,072	3,072
State X Years	42,672	42,672	37,800	37,800	42,672	42,672
Pershing First-Stage		0.184		0.150		0.180
		0.017		0.019		0.017
KP F-Statistic		114.88		65.53		119.27

Panel B: Establishments

	Flexible Covariates		CZ × Year FEs		Census × Year FEs	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
Interstate (0/1)	0.0034***	0.0062***	0.0031***	0.0057**	0.0033***	0.0054**
	(0.0004)	(0.0022)	(0.0004)	(0.0029)	(0.0004)	(0.0022)
Observations	171,938	171,938	167,120	167,120	171,994	171,994
Counties	3,071	3,071	2,985	2,985	3,072	3,072
State X Years	2,688	2,688	2,744	2,744	42,672	42,672
Pershing First-Stage		0.184		0.150		0.180
		0.017		0.019		0.017
KP F-Statistic		114.88		65.53		119.27

Notes: Every specification reports results from estimating equation 1, which regresses a binary interstate indicator on average changes in either log employment or log establishments. Columns (1)–(2) replace time invariant covariates with time dummy interacted covariates. Columns (3)–(4) replace state × year fixed effects with commuting zone × year fixed effects. Columns (5)–(6) replace state × year fixed effects with a set of state fixed effects and census region × year fixed-effects. Employment and establishment data are from 1956–2016 County Business Patterns annual reports at the county level. Standard errors are two-way clustered by county and commuting zone × year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.1.5: Effects of Interstates with Spatially Adjusted Standard Errors

	Employment			Establishments		
	(1)	(2)	(3)	(4)	(5)	(6)
Interstate (0/1)	0.0040***	0.0040***	0.0040***	0.0033***	0.0033***	0.0033***
	(0.0005)	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0004)
Observations	171,940	171,940	171,940	171,938	171,938	171,938
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
Std. Error Type	Cluster	Conley	Conley	Cluster	Conley	Conley
Std. Error Radius	CZ × Yr	100 km	250 km	CZ × Yr	100 km	250 km

Notes: Every specification reports results from estimating equation 1, which regresses a binary interstate indicator on average changes in either log employment or log establishments, and includes the baseline set of controls and fixed-effects. Employment and establishment data are from 1956–2016 County Business Patterns annual reports at the county level. Standard errors in Columns (1) and (4) are two-way clustered by county and commuting zone × year. Columns (2)–(3) and (5)–(6) report standard errors following Conley (1999, 2010) and Hsiang (2010) under varying distances. *** p<0.01, ** p<0.05, * p<0.1.

Table B.1.6: Effects of Highways with 35 Year IV Construction Timeline

Panel A: Employment				
	(1)	(2)	(3)	(4)
Interstate (0/1)	0.0040***	0.0098***	0.0037***	0.0102***
	(0.0005)	(0.0028)	(0.0006)	(0.0029)
Interstate KMs per County Sq KM			0.0140	-0.0171
			(0.0140)	(0.0439)
Observations	171,940	171,940	171,940	171,940
Counties	3,071	3,071	3,071	3,071
State X Years	2,688	2,688	2,688	2,688
Pershing First-Stage		0.1784		
		0.0165		
KP F-Statistic		116.6989		58.2339
Panel B: Establishments				
	(1)	(2)	(3)	(4)
Interstate (0/1)	0.0034***	0.0076***	0.0031***	0.0075***
	(0.0004)	(0.0023)	(0.0005)	(0.0024)
Interstate KMs per County Sq KM			0.0126	0.0015
			(0.0111)	(0.0345)
Observations	171,938	171,938	171,938	171,938
Counties	3,071	3,071	3,071	3,071
State X Years	2,688	2,688	2,688	2,688
Pershing First-Stage		0.1784		
		0.0165		
KP F-Statistic		116.6998		58.2344

Notes: Table replicates results from Panel A of Table 2 using a 35 year construction timeline. Every specification reports results from estimating equation 1, where the outcome of interest is average year over year changes in the log of employment and the highway treatment is the binary interstate highway indicator. Columns (3) and (4) include the measure of interstate highway density. Every specification includes state and census region × year fixed effects, along with the full set of controls outlined in Appendix C.2. Employment data from 1956–2016 County Business Patterns annual reports. Standard errors are two-way clustered by county and state × year. *** p<0.01, ** p<0.05, * p<0.1.

Table B.1.7: Interstate Highways and Growth: Constructed Distance Treatment

	Employment		Establishments	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
Interstate KMs Density	0.0632***	0.0765	0.0539***	0.0777**
	(0.0146)	(0.0479)	(0.0119)	(0.0391)
Observations	171,940	171,940	171,938	171,938
Counties	3,071	3,071	3,071	3,071
State X Years	2,688	2,688	2,688	2,688
Pershing FS		0.339		0.339
		0.052		0.052
KP F-Statistic		42.383		42.383

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 C.2. 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.

Table B.1.8: Differences in Covariates by Interstate Status and Construction Era

	Interstate Counties			Predicted Pershing Counties		
	(1)	(2)	(3)	(4)	(5)	(6)
	Constructed: 1956 - 1965	Constructed: 1966 - 1975	Difference	Constructed: 1956 - 1965	Constructed: 1966 - 1975	Difference
Ln(Employment in 1953)	8.777	7.875	0.903***	8.663	8.289	0.375*
	[1.786]	[1.532]	(0.108)	[1.936]	[1.790]	(0.204)
Ln(Establishments in 1953)	6.542	5.798	0.744***	6.493	6.162	0.332**
	[1.411]	[1.197]	(0.083)	[1.546]	[1.404]	(0.159)
Ln(1950 Population)	10.639	10.030	0.609***	10.611	10.319	0.292*
	[1.297]	[1.062]	(0.075)	[1.401]	[1.271]	(0.160)
Ln(1940 Population)	10.521	9.986	0.535***	10.490	10.255	0.235
	[1.220]	[0.984]	(0.070)	[1.299]	[1.195]	(0.149)
Ln(1930 Population)	10.444	9.920	0.524***	10.415	10.202	0.213
	[1.210]	[0.970]	(0.071)	[1.270]	[1.181]	(0.145)
Ln(1920 Population)	10.310	9.859	0.452***	10.307	10.120	0.188
	[1.184]	[0.926]	(0.068)	[1.174]	[1.144]	(0.138)
Ln(1910 Population)	10.191	9.769	0.422***	10.165	10.033	0.132
	[1.136]	[0.922]	(0.073)	[1.152]	[1.113]	(0.144)
Pct of Pop in Urban Area in 1950	0.439	0.300	0.139***	0.429	0.377	0.053**
	[0.278]	[0.268]	(0.016)	[0.298]	[0.284]	(0.023)
Market Potential with 1950 Pop	197.698	188.823	8.875	183.071	197.974	-14.903
	[104.634]	[83.658]	(8.205)	[110.141]	[103.987]	(11.226)
Area in sq mi	1,223.577	775.212	448.365***	1,167.707	1,269.011	-101.304
	[1,879.231]	[776.075]	(159.995)	[1,344.354]	[2,111.006]	(193.131)
Latitude	38.627	37.465	1.162**	38.933	38.752	0.181
	[4.611]	[4.788]	(0.437)	[5.732]	[3.722]	(0.969)
Longitude	-91.288	-90.591	-0.697	-92.849	-91.945	-0.904
	[13.223]	[10.802]	(1.305)	[14.629]	[12.739]	(1.963)
Latitude Sq	1,513.304	1,426.522	86.782**	1,548.564	1,515.528	33.036
	[351.544]	[361.342]	(33.888)	[438.786]	[285.290]	(74.564)

Longitude Sq	8,508.067	8,323.076	184.992	8,834.503	8,615.832	218.671
	[2,544.001]	[2,080.345]	(253.347)	[2,840.793]	[2,439.329]	(383.840)
Mean Elevation	455.098	376.326	78.772*	395.108	519.941	-124.833
	[521.144]	[467.557]	(43.357)	[503.064]	[580.868]	(98.379)
Ruggedness	89.796	65.316	24.480**	98.845	94.394	4.451
	[126.581]	[109.334]	(11.248)	[145.103]	[126.998]	(18.786)
1911 Railroad KMs per sq mi	0.250	0.203	0.047***	0.232	0.243	-0.011
	[0.197]	[0.143]	(0.012)	[0.182]	[0.206]	(0.020)
KM to 1947 Highway System	119.620	84.476	35.144***	116.904	112.379	4.526
	[112.491]	[61.739]	(8.560)	[92.640]	[120.193]	(9.008)
KM to 1911 RR	1,194.344	1,082.038	112.306*	1,320.861	1,098.652	222.209**
	[587.855]	[474.873]	(56.010)	[609.362]	[587.305]	(97.326)
KM to Nearest 1955 Port	368.636	362.510	6.126	300.883	437.852	-136.969**
	[327.709]	[277.791]	(30.201)	[309.770]	[329.682]	(56.071)
KM to Nearest 1955 Airport	37.498	46.204	-8.706***	33.346	43.874	-10.528***
	[32.013]	[32.000]	(2.039)	[28.384]	[37.046]	(2.630)
KM to 1918 Military Highways	571.242	581.321	-10.079	548.443	572.818	-24.374
	[409.818]	[393.391]	(43.293)	[367.899]	[411.975]	(49.310)
KM to MSA	73.322	87.502	-14.179***	77.529	88.920	-11.391
	[75.169]	[67.871]	(5.234)	[78.751]	[80.841]	(10.450)
KM to Mexican War Battle	1,500.748	1,413.276	87.472	1,476.291	1,442.731	33.560
	[752.991]	[614.274]	(70.744)	[768.570]	[752.855]	(98.147)
KM to American Rev. Battle	672.465	604.732	67.733	845.504	649.097	196.408
	[763.917]	[647.308]	(78.567)	[853.337]	[732.407]	(118.406)
KM to Civil War Battle	366.502	299.768	66.733	471.697	336.472	135.225*
	[412.376]	[350.802]	(43.570)	[503.168]	[358.483]	(80.183)
KM to French/Indian War Battle	1,223.271	1,171.072	52.200	1,397.280	1,230.254	167.027
	[963.453]	[811.628]	(100.462)	[1,038.018]	[936.595]	(141.380)
KM to Indian War Battle	201.591	210.722	-9.131	195.892	208.268	-12.376
	[132.030]	[131.332]	(13.034)	[133.424]	[125.428]	(23.408)
KM to Insurrections	443.010	466.554	-23.544	510.258	427.292	82.966

	[292.026]	[246.647]	(27.364)	[338.060]	[238.620]	(51.627)
KM to War 1812 Battles	662.339	580.234	82.105	741.042	669.388	71.654
	[711.143]	[581.820]	(71.087)	[818.904]	[675.796]	(119.210)
KM to WW1 Sites	73.861	90.217	-16.356***	66.009	90.063	-24.054***
	[71.292]	[66.260]	(5.630)	[63.776]	[81.771]	(7.717)
KM to Naval Bases	461.901	468.758	-6.858	393.298	517.917	-124.619**
	[327.364]	[297.715]	(28.630)	[286.653]	[342.184]	(46.908)
KM to Airfields	73.471	85.361	-11.890**	73.089	77.632	-4.543
	[59.267]	[61.376]	(5.113)	[63.903]	[57.306]	(9.935)
KM to Military Forts	90.164	99.465	-9.302*	93.217	97.203	-3.987
	[67.555]	[66.274]	(5.012)	[69.281]	[73.835]	(8.954)
KM to Pershing Map Nodes	85.181	105.457	-20.276***	61.178	89.453	-28.276***
	[73.400]	[71.016]	(6.526)	[57.947]	[73.045]	(7.740)
1947 Hwy Betweenness	96,331.867	86,119.445	10,212.425	75,714.891	99,365.828	-23,650.938**
	[105,962.758]	[99,861.828]	(6,812.460)	[88,090.883]	[103,320.945]	(9,699.082)
Rural County (0/1)	0.234	0.426	-0.192***	0.276	0.361	-0.085
	[0.424]	[0.495]	(0.040)	[0.447]	[0.481]	(0.052)
Metro County (0/1)	0.456	0.265	0.190***	0.451	0.353	0.098**
	[0.498]	[0.442]	(0.036)	[0.498]	[0.478]	(0.044)
MSA County in 1960 (0/1)	0.305	0.117	0.187***	0.312	0.234	0.078**
	[0.460]	[0.322]	(0.024)	[0.464]	[0.424]	(0.037)
Boundary Adjustment (0/1)	1.004	1.005	-0.002	1.005	1.002	0.003
	[0.060]	[0.071]	(0.003)	[0.069]	[0.043]	(0.004)
Grouped County (0/1)	0.090	0.135	-0.045	0.062	0.110	-0.048
	[0.286]	[0.342]	(0.028)	[0.242]	[0.314]	(0.037)
Indep. City County (0/1)	0.007	0.005	0.002	0.010	0.004	0.006
	[0.085]	[0.071]	(0.002)	[0.098]	[0.061]	(0.006)
Suppressed Employment (0/1)	0.000	0.000	0.000	0.000	0.000	0.000
	[0.000]	[0.000]	(0.000)	[0.000]	[0.000]	(0.000)
Observations	834	392	1,226	417	535	952

Notes: Detailed source and measurement information in Appendix C.2. 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 state × year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix C Data Appendix

Appendix C.1 County Business Pattern

County Business Patterns (CBP) data was collected from three sources: Archival CBP reports for 1953 - 1962, ICPSR 25894 for 1964 - 1970, and US Census Bureau for 1970 - 2015. 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.

Due to confidentiality and reporting restrictions, some employment totals were redacted from the final CBP reports. I impute the missing redacted employment values similar to Duranton and Turner (2012). In cases where there were missing establishment counts. I impute those based on the information in the prior and following periods. 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.

I adjust for county boundary changes using 1950 boundary definitions following Hornbeck (2010). 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. I also aggregate independent cities, typically in Virginia, into their surrounding or neighboring counties similar to Jaworski and Kitchens (2019).

Appendix C.2 Covariate Descriptions

I compile county level covariate information from several sources to account for pre-interstate differences in market size, geography, and access to transportation infrastructure.

1. I compile county level population and urbanization data from 1910 to 1950 are available from National Historical Geographic Information System (2011). Using the 1950 population data, I construct a measure of market potential for each county based on the euclidean distance between every pair of county seats. For several constructed measures I rely on 1950 county seats instead of county centroids because in large rural counties they provide more information regarding the most economically relevant location in the county.

2. I include three controls for county type, by designating each county to either overlap with a 1950 MSA boundary (National Historical Geographic Information System, 2011) or be classified as either rural or metropolitan using the definitions from Hines et al. (1975). In this case the excluded category is the set of non-MSA, non-rural, and non-metropolitan counties. As an additional spatial control, I control for the distance from the nearest MSA centroid to each county seat.

3. 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.

4. I control for several measures of existing transportation infrastructure and in some cases the proximity from each county seat to the infrastructure.

   (a) I construct two measures of access to railroads from the records provided by Atack (2016). The first measures railroad density as the total kilometers of track within the county. The second measures distance from the nearest railroad to the county seat.

   (b) I measure access to major historic highway networks using a 1918 map, which pre-dates the Pershing Map, using distance from nearest route to county seat.

   (c) To account for access to airports and ports, I determined the coordinates of each type of location from (U.S. Census Bureau, 1958) and measure the geographic distance to county seats.

5. To account for changes in other forms of road development, I construct to annual measures of spending on state highways from the Annual Survey of State and Local Government Finances from 1943 to 2016. The first measure is the most recent years capital outlay and the second accumulates the prior five years.