Transportation Networks, Institutions, and Regional Inequality

by

Dustin Frye

B.A., University of Montana, 2006

M.A., University of Colorado, 2011

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Economics

2015

This thesis entitled:
Transportation Networks, Institutions, and Regional Inequality
written by Dustin Frye
has been approved for the Department of Economics

Professor Murat Iyigun

Professor Jonathan E. Hughes

Date ____________________

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Frye, Dustin (Ph.D., Economics)

Transportation Networks, Institutions, and Regional Inequality

Thesis directed by Professor Murat Iyigun

Abstract

This dissertation focuses on long-run patterns of regional inequality by addressing two general themes: the importance of transportation networks for location choices of individuals and firms, and the role of institutions on economic development across Native American reservations. My findings highlight the significance of transportation networks, and local governance for regional economic development.

In the first chapter, I measure the effect of improvements in transportation infrastructure on industry growth and concentration. To address the endogenous placement of interstate highways, I instrument for eventual highway location using two proposed government plans. To address the endogeneity surrounding the timing of highway construction, I use a network theory algorithm to predict the timing of highway construction. The results indicate that the expansion of the Interstate Highway System (IHS) led to substantial employment growth in highway counties relative to non-highway counties. This employment growth was concentrated in a few industries. This paper demonstrates the importance of expanding transportation networks for the spatial arrangement of economic activity.

In my second chapter, I concentrate on the U.S. agricultural sector. The IHS altered the structure of transportation costs. This paper provides the first empirical analysis of the impact of new interstate highway infrastructure on farm property values and the portfolio of agricultural commodities produced. Estimates correcting for endogenous highway locations and construction timing indicate the value of land per acre fell in highway locations relative to non-highway locations. This loss appears driven by a declining value of agricultural products sold. Additional results find no evidence that highway counties are more specialized in their production than non-highway counties.

In the final chapter, I exploit the decentralization of governance across American Indian reservations and measure the long-run development differences for reservations that were granted less sovereignty through the Indian Reorganization Act (IRA). To mitigate selection concerns regarding IRA adoption, I exploit IRA voting results by restricting my analysis to narrowly determined elections. Results indicate that IRA adoption stifled economic development. Per capita income is over 40 percent lower on IRA reservations. Additional legislation in the late 1980s further decentralized IRA reservations; as a result income differences diminish by 2010.

Dedication

To my parents, whose love, support, and sacrifice allowed me to pursue my dream. Especially ...

to my dad, Dan Frye, for teaching me to work hard and laugh harder;

to my mom, Sami Mitchell, for filling my life with the perfect balance of art and little yellow math books;

to my step-dad, Mike Mitchell, for convincing me to take my first economics course and inspiring me to learn every day.

Acknowledgements

A special thank you to my advisor, Murat Iyigun, for guiding me as a researcher and helping me refine my ideas. I am grateful to Jonathan Hughes, Lee Alston, Carol Shiue, and Tony Cookson for all their helpful suggestions. A special thank you to Gisella Kagy, Edward Kosack, Steven Smith, and Zachary Ward for their feedback and support along the way. I would also like to thank Richard Weingroff for generously sharing the PR-511 reports and answering my many questions about it.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Transportation Networks and the Geographic Concentration of Industry

1.1 Introduction

Transportation costs are an integral component of the spatial arrangement of economic activity. Expanding transportation infrastructure impacts trade flows and alters the organization of cities by changing the cost of moving goods and commuting. New transportation infrastructure motivates firms and individuals to alter their location choices. The construction of the Interstate Highway System (IHS) in the United States, starting in 1956, introduced over 40,000 miles of new highways, which lowered the costs of moving goods and people. For example, from 1975 to 1985 shipping rates by truck fell by nearly 20 percent (Rose 1988). The IHS also led to changes in driving behavior. From 1966 to 1995 the percentage of total vehicles miles traveled along interstate highways increased from 10 percent to nearly 25 percent (FHWA 1997). These changes in costs and usage suggest that interstate highways could have altered the location choices of both firms and individuals. This paper uses the construction of the Interstate Highway System to understand the relationship between transportation infrastructure and industry concentration.

I measure the causal effect of having an interstate highway on industry growth and concentration using a reduced form analysis, where I instrument for the presence of a highway to address two types of endogeneity. The first endogeneity concern is the non-random placement of Interstate highways. Highways were often directed to struggling communities (Duranton and Turner 2012). The paper instruments for eventual highway location using two proposed government maps of high-priority routes. The second endogeneity concern I address is the endogenous allocation of funding by state politicians, which determined when particular segments of the IHS were constructed. To address the endogeneity surrounding the timing of highway construction, I use an algorithm from network theory to predict the timing of highway construction. The algorithm ranks predicted highway segments based on their importance for network connectivity and uses a simple social planner's problem to determine the order of predicted segment construction. With this method I construct an instrumental variable that predicts both where an interstate highway would locate and when it would be built by combining the location prediction and the predicted construction schedule.

I use a county-level panel dataset spanning from 1962 to 1996 to examine industry growth and concentration using several different measures. First, I compare differences in employment growth between highway and non-highway counties and find there were significant positive differences starting in the early-1980s. This growth was more pronounced in agriculture, retail sales, and the transportation and public utilities sector. I find very little evidence of growth in manufacturing employment. Next, I use two measures of industry concentration to determine whether the employment growth was concentrated in a fewer sectors. Results indicate there was substantial increases in employment concentration in highway counties relative to non-highway counties. To measure changes in the scale of firms by industry, I compare changes in the share of large firms in highway counties compared to non-highway counties as larger firm size is typically associated with increased concentration (Holmes and Stevens 2004). These results indicate that highways led to moderately larger manufacturing firms in highway counties relative to non-highway counties. Finally, I measure the full dynamic response of receiving an interstate highway. These results indicate that it takes between 15 and 20 years before highway counties significantly differ from non-highway counties. These results taken together suggest that the Interstate Highway System significantly contributed to industry concentration in highway counties.

My analysis is most directly related to the growing literature on relationship between transportation infrastructure and the organization of economic activity.[1] The majority of papers in this literature study the effect of highways in cities. Several papers document population and industry

[1] For a comprehensive survey of this literature see Redding and Turner (2014).

decentralization, and the growth of the suburbs (Baum Snow 2007, 2014; Baum-Snow *et al.* 2014; Rothenberg 2013). Duranton and Turner (2012) find employment increases in cities for several years after expansions in highway mileage. Duranton, Morrow, and Turner (2013) examine trade relationships between several major cities and find that cities with more highway mileage specialize in the production of heavier goods, but there was no difference in product value. Fernald (1999) finds that the construction of the Interstate Highway System led to a productivity boost with the majority of benefits being concentrated in vehicle intensive industries.

Michaels (2008) finds that interstate highways increase earnings in retail sales and trucking, both trade related activities, in rural counties within the United States. He also finds an increase in the demand for skilled labor; however he cannot identify an effect of highways on the industrial composition of employment. Chandra and Thompson (2000) examine the effect of interstate highways on earnings by industry using a distributed lag model for a subset of rural counties. They find that earnings increased for several industries and that counties adjacent to highways experienced a decline in earnings, a result they attribute to reorganization of economics activity and not growth.

My paper contributes to the relevant literature in several ways. This is the first paper in this literature to directly instrument for the timing endogeneity, which allows me to measure the effects of receiving a highway over a longer period of time and provides valuable insight into the political motives surrounding early highway construction. My outcomes of interest build on the employment findings from Baum-Snow (2014) and Duranton, Morrow, and Turner (2013). Both of these papers restrict their analysis to urban areas and are more interested in growth in the urban highway network than the broader national system. My paper is also the first to include both rural and urban counties in the analysis.

This paper relates to the literature identifying the consequences of shocks to the spatial equilibria of economic activity (Davis and Weinstein 2002; Redding and Sturm 2008; Redding, Sturm, and Wolf 2011). The construction of the IHS changed the distance between locations and altered the spatial equilibrium of employment and firm locations. It is important to understand the effect of changing the relative geography between locations because Allen and Arkolakis (2014) find that geographic location accounts for at least 20 percent of the spatial variation in income. The magnitude of their result indicates that changes in the relative distance between locations has important consequences for development.

The paper proceeds as follows. Section 1.2 gives a brief history of the Interstate Highway System, emphasizing the potentially confounding role that politicians and industrial leaders played in the design and construction of the IHS. Section 1.3 describes the data used in the empirical analysis and documents the pattern of industry growth and concentration that occurred between 1962 and 1996. Section 1.4 discusses the empirical strategy and the endogeneity issues associated with estimating the causal effects of transportation infrastructure on industry growth and concentration. Section 1.5 examines the role that highways played in employment and establishment growth in highway and non-highway counties. Section 1.6 discusses patterns of employment and establishment concentration induced by the IHS. Section 1.7 measures the dynamic effects of interstate highways. Section 1.8 provides two falsification exercises for robustness and Section 1.9 concludes.

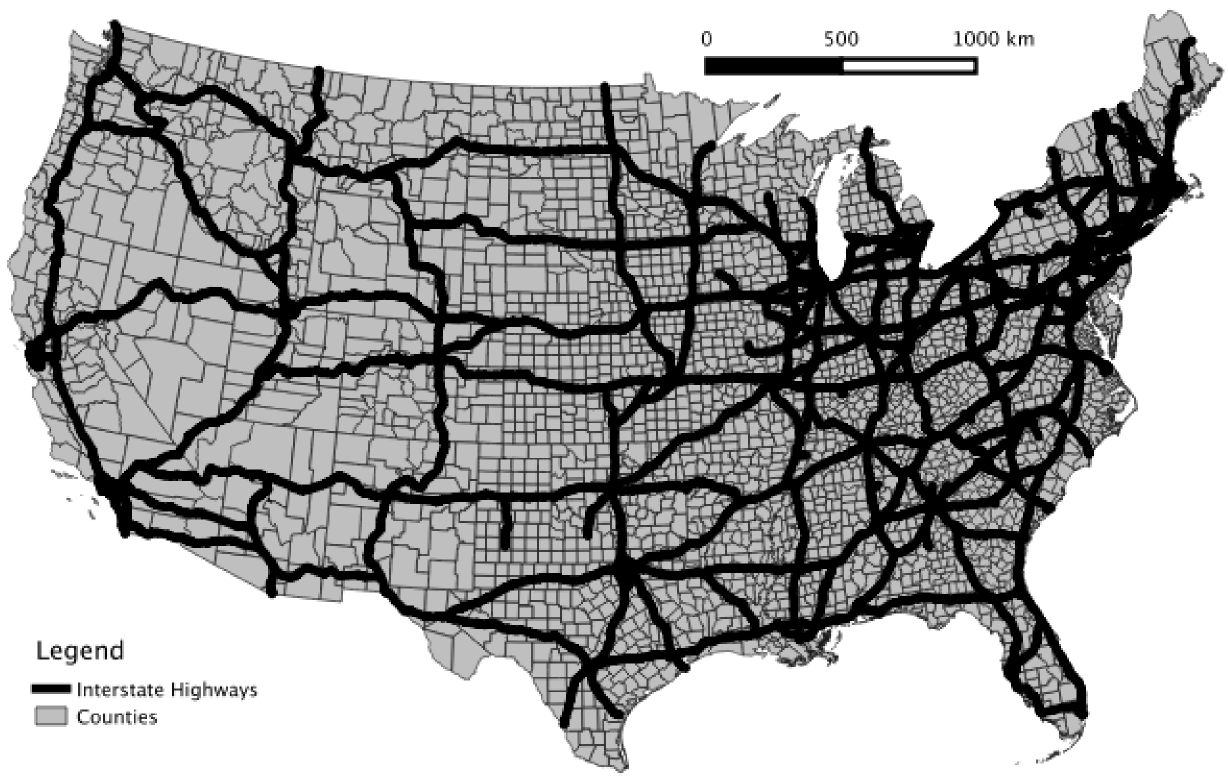
1.2 A History of the Interstate Highway System

1.2.1 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 some of 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).

In 1956, after several different plans, construction guidelines, and financing methods were introduced, the House and Senate ultimately agreed on an interstate highway plan. The plan was approximately 90 percent Federally funded and was paid for with taxes 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, as presented in Figure 1.1, was "a culmination of decades of input and research from auto clubs, civil engineers, and state and federal highway officials" (Kaszynski, 167, 2000). The Highway Act of 1956 placed states in charge of construction. Each state's funding was determined based on a formula of population, area, and highway mileage. This allowed states to build their segments of interstate highway when they wanted and at the pace they wanted. 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.

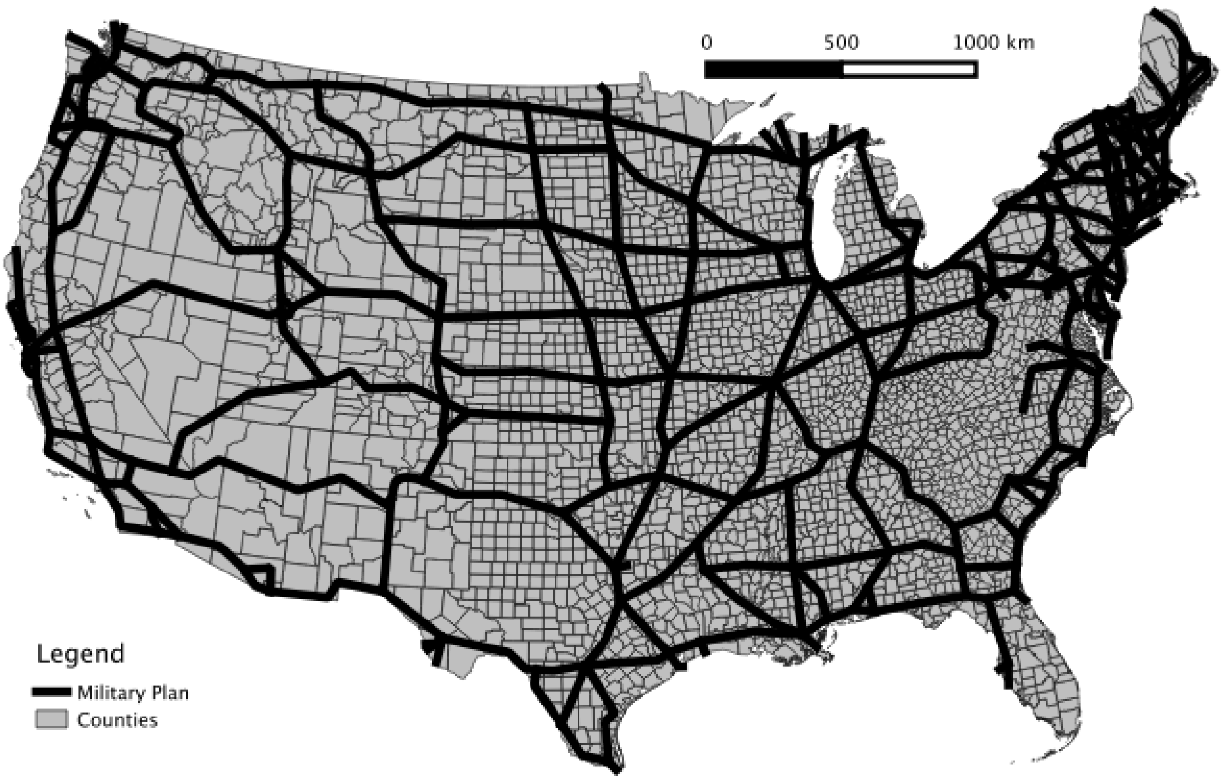
Figure 1.1: National System of Interstate and Defense Highways



1.2.2 The Pershing Map and the National Interregional Highway Committee

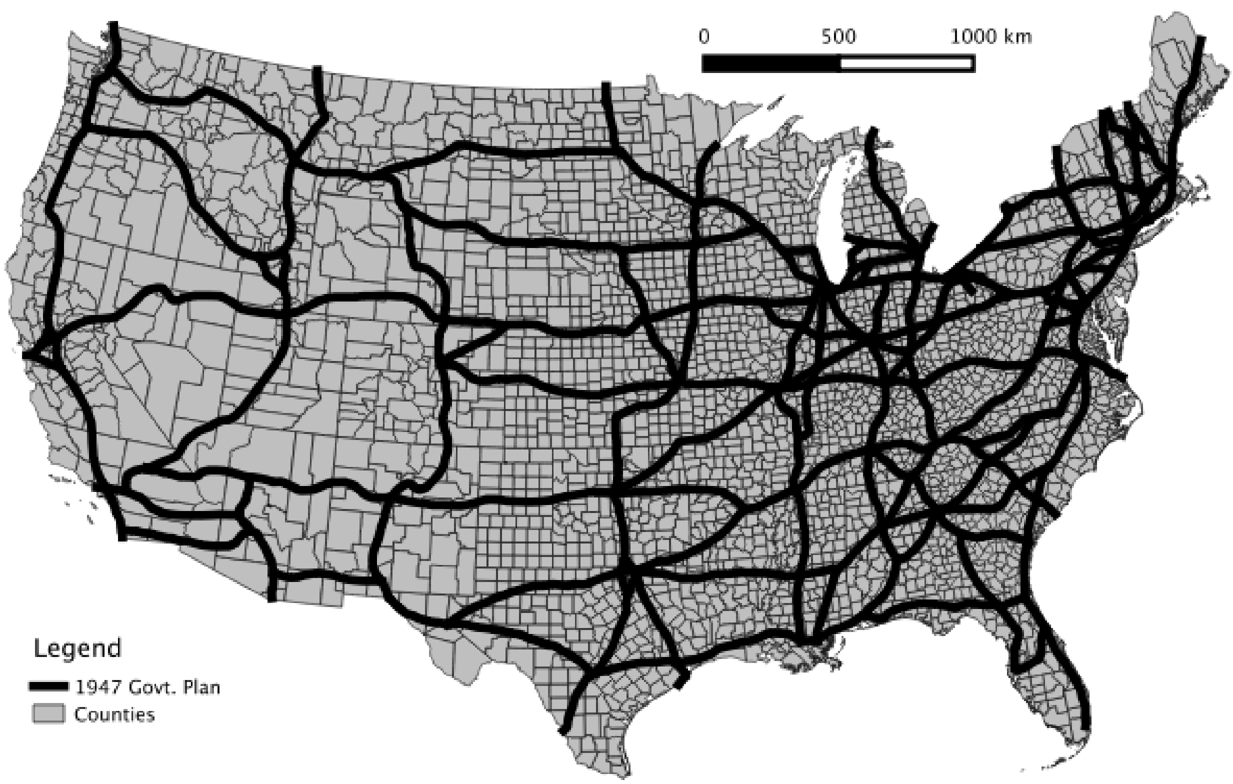
My empirical design requires that the interstate highway system was exogenously assigned to counties. Early proposals of interstate highway locations date back to the early 1920s, which may provide predictions of eventual IHS locations for my empirical strategy. Following the First World War the U.S. government began discussing the merits of a national highway 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, 76, 2011). The resulting map, depicted in Figure 1.2, 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, 76, 2011). These routes were never built as superhighways but this map influenced future highway location decisions.

Figure 1.2: Pershing Military Plan



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" (US DOT, 273, 1977). 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" (US DOT, 274, 1977). 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" (US DOT, 274, 1977). Interest groups on behalf of the farming and trucking industry "lobbied for their own plans to foster particular and local needs" (Rose, 16, 1990). The final plan, published in 1947 and depicted in Figure 1.3, was the most comprehensive national network map that had been produced and served as the major guide of highway location decisions for the next decade.

Figure 1.3: 1947 Plan from the Interregional Highway Committee



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 after 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).

1.3 Data and Preliminary Evidence

My empirical analysis uses a county-level panel dataset that spans from 1962 to 1996 for the contiguous United States. The primary outcomes of interest rely on annual employment and establishment data collected by the Census Bureau and published in the County Business Patterns. These data are combined with contemporary and historical transportation network information, which allows me to examine the relationship between transportation networks and the several measures of industry growth and concentration.

1.3.1 County Business Patterns

In 1962 the United States Census Bureau began publishing information regarding employment and the number of establishments for counties in the United States.[2] This paper uses the employment and establishment data for the primary Standard Industrial Classification (SIC) economic divisions: agriculture, construction, finance, manufacturing, mining, retail sales, services, transportation and public utilities, wholesale trade, and unclassified occupations.[3]

For each broad industry division, I observe the total number of establishments and the to-

[2] Prior to 1962, published establishment and employment information was combined for some counties in eight states. I exclude these counties, so the sample in 1962 consists of 2661 of the 3079 counties in the full sample.

[3] After 1996 the Census Bureau no longer used the SIC system, moving to NAICS. For classification consistency this paper concentrates on the period using the SIC system. The SIC experienced several modifications over this 30-year period, however the broad categories I am interested in were largely unaffected by these changes.

tal number of establishments in eight employment size groups.[4] One limitation of the County Business Patterns data is that it does not include establishments with zero employees.[5] For confidentiality purposes the Census Bureau censored the county-level employment data for some smaller industries. Similar to Duranton, Morrow, and Turner (2013), I impute employment values using the establishment count data.[6] The result is a county-level panel dataset spanning from 1962 to 1996 with employment quantities, establishment counts, and establishments counts by eight employment size groups for each of the ten SIC economic divisions. I also aggregate the ten SIC economic divisions to make a total category containing the employment, number of establishments, and establishment group counts for all sectors in the county.

1.3.2 Calculating Concentration Measures

To understand the relationship between highways and employment and establishment concentration I construct two measures of concentration. I use the following Herfindahl Index for employment concentration:

$$H_{ct} = \sum_i {s_{cit}}^2 \tag{1.1}$$

For each county c in year t, equation 1.1 sums the squared share of each division's employment in industry i. If employment is fully concentrated in a sector, then $H_{ct} = 10000$, and the index decreases as employment becomes more diverse. I construct the same measure using the number of establishments.

The Gini Specialization Index is an alternative concentration measure, used by Duranton and Puga (2004). This measure corrects for differences in local sectoral employment by comparing it to the national share of employment in the sector. Formally, the GSI is given by

[4] Employment size groups include: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, and above 500 employees.
[5] In the robustness section I will discuss what affect this omission has on the empirical results.
[6] For each industry I regress the county sectoral employment on the full set of eight establishment count groups and I use the resulting regression coefficients to impute the number of employees. The R^2 for each regression is between 0.945 and 0.999.

$$GSI_{ct} = \frac{1}{2}\sum_{i} |s_{cit} - s_{it}| \tag{1.2}$$

The value s_{cit} is the share of employment in county c in year t in industry i. The value s_{it} is the national employment share for industry i in year t. This index is closer to one if employment in a county is fully specialized in an economic division that has a very small employment share at the national level. The index is near zero if employment in the county and national employment are similarly distributed. I also construct the GSI using the number of establishments.

I supplement these two broad measures of concentration with two measures of industry-level concentration. These measures allows me to test for differences in the scale of firms across industries.[7] I use data on establishment counts within the eight employment size groups to construct two measures of firm size. I compare the fraction of firms with more than 20 employees and the fraction of firms with between one and four employees by industry. Combining these industry measures of concentration with the industry level patterns in employment growth provides insight into whether growth in concentrated in several large firms or is dispersed across several smaller firms. Distinguishing between these two results is important for understanding whether public infrastructure alters market power within an industry and whether it promotes entrepreneurship.

1.3.3 Interstate Highway System Maps

I use two data sources to construct an annual county-level panel dataset with Interstate Highway System information spanning from 1962 to 1996. The first is current highway location information from NationalAtlas.gov (2014). I combine this file with highway construction information from the PR-511 collection at the National Archives. 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.[8] I intersected this progress

[7] The results for firms in the smallest size bin may shed some insight onto the behavior of firms with zero employees, which were not included in the data.

[8] 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

with a map of county locations in 1980, which allows me to know the year a county was connected to the Interstate Highway System.[9] Figure 1.1 shows the current interstate highway locations overlaid on a map of county locations.

For each county, I determine whether an interstate highway intersects that county and the year that segment of highway was completed. I can use this data to determine two key measures for my empirical strategy, in each year I know whether a county had received an interstate highway and how many years ago that particular segment of highway was constructed.

1.3.4 Supplemental Data

In order to account for factors that are correlated with the economic growth, concentration, and location and funding of interstate highways, I supplement the economic and highway information with data covering population, historical economic data, and alternative methods of transportation.[10] I use county-level population data from the U.S. Census for every decade from 1910 to 1950. I combined this with information on the share of population living in cities larger than 25,000 people, the number of manufacturing establishments, and the number of farmers from the 1910 to 1940 censuses. I also collected information on the number of establishments and employees in manufacturing, wholesale trade, retail trade, and farming from the 1930 and 1940 censuses. This historical population and industry information is useful for supporting the exogeneity requirements of my instrumental variables. Lastly, I collected high school attainment information to help approximate the skill endowment of each county in 1950 (ICPSR 2005). This measure will allow me to look for evidence of heterogeneous effects of the interstate highway system based on the skill endowment of counties prior to highway construction.

I collected additional geographic information for alternative methods of transportation from

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.

[9] I adjust all of the county locations and data to be consistent with the 1980 county borders.

[10] Population and historical economic data are from the National Historical Geographic Information Systems (NHGIS) (2014).

NationalAtlas.gov (2014). I use GIS to construct an indicator that is equal to one if a county has a railroad.[11] For each county I calculate the Euclidian distance to the nearest coastal port and the nearest airport.

1.3.5 Summary Statistics

My completed county-level panel dataset contains employment and establishment information, highway location and construction information, historical population and economic data, and geographic measures of alternative methods of transportation infrastructure. Table 1.1 presents summary statistics for two groups: counties that eventually received an interstate highway and counties that never received an interstate highway. The table presents the number of observations, the mean, and the standard deviation for both groups for the full sample of years from 1962 to 1996. The last two columns calculate the difference between highway and non-highway counties. The most striking feature of the table is how different highway and non-highway counties are. Highway counties generally have more employment and establishments, and also are less concentrated. Highway counties are more likely to be near a MSA, have a railroad, and are generally closer to airports and ports. They were also have much larger populations in 1950 and their population grew much faster from 1940 to 1950.

[11] Due to data availability constraints I ignore railroad lines that were decommissioned following deregulation.

Table 1.1:

Table 1: Full Sample Summary Statistics By Highway Status

	Non-Highway Counties			Highway Counties				
	Mean	SD	N	Mean	SD	N	Difference in Means	SE
Employment (in 1,000s)	5.03	8.65	56,991	44.36	138.90	44,616	39.330***	0.583
Establishments (in 100s)	4.50	6.78	56,991	27.99	77.94	44,616	23.488***	0.328
Employment Herfindahl Index	2,729.48	1,007.64	56,991	2,521.14	776.25	44,616	-208.344***	5.773
Establishments Herfindahl Index	2,137.88	400.94	56,991	2,127.30	283.74	44,616	-10.581***	2.240
Gini Spec. Index for Employment	0.26	0.13	56,991	0.20	0.10	44,616	-0.059***	0.001
Gini Spec. Index for Establishments	0.15	0.07	56,991	0.11	0.05	44,616	-0.043***	0.000
Military Plan	0.18	0.38	56,991	0.32	0.47	44,616	0.147***	0.003
1947 Govt. Plan	0.06	0.23	56,991	0.81	0.39	44,616	0.749***	0.002
Distance to MSA (in km)	142.06	116.43	56,991	96.38	105.27	44,616	-45.673***	0.706
Railroad	0.43	0.50	56,991	0.77	0.42	44,616	0.339***	0.003
Airport Distance (in km)	60.21	29.92	56,991	42.29	27.59	44,616	-17.914***	0.183
Port Distance (in km)	363.98	291.64	56,991	291.93	275.50	44,616	-72.051***	1.799
1950 Population (in 1,000s)	18.38	16.87	56,991	82.89	237.52	44,616	64.513***	0.998
Percent Pop. Growth From 1940 to 1950	-0.01	0.21	56,991	0.12	0.27	44,616	0.131***	0.001

Notes: Data comes from the 1962 - 1996 County Business Patterns annual reports. Highway counties are those that ever received a highway. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, and Airport. *** p<0.01, ** p<0.05, * p<0.1

To preview the empirical strategy, Table 1.2 compares the differences between highway and non-highway counties in 1965 and 1996. The outcomes reported in the table are for the County Business Patterns employment and establishment count data, along with the concentration measures and the firm size measures. The differences between highway and non-highway counties are reported in the last two columns. Highway counties are significantly different from non-highway counties in both periods. Highway counties in 1965 have more employment and establishments, and are less concentrated. In 1996 the difference between highway and non-highway counties has grown for both employment and the number of establishments. The Herfindahl Index values for both employment and the number of establishments changes sign and now indicates that highway counties are more concentrated than non-highway counties. Comparing the difference in the means across the two time periods indicates that both employment and the number of establishments grew over the period. This growth was accompanied by increases in employment concentration, which suggests that a large portion of the job growth was concentrated in a few industries.

Table 1.2:

Table 2: Full Sample Summary Statistics of Growth and Concentration Measures
By Highway Status in 1965 and 1996

	Non-Highway Counties			Highway Counties				
	Mean	SD	N	Mean	SD	N	Difference in Means	SE
Panel A: 1965 Outcomes for Counties that Ever Receive a Highway								
Employment (in 1,000s)	3.10	4.65	1,727	28.46	100.91	1,352	30.627***	2.708
Establishments (in 100s)	3.41	3.96	1,727	19.46	58.41	1,352	18.758***	1.572
Employment Herfindahl Index	2,966.38	1,206.49	1,727	2,770.42	1,022.74	1,352	-180.261***	45.875
Establishments Herfindahl Index	2,520.48	527.53	1,727	2,360.18	321.83	1,352	-148.795***	18.302
Gini Spec. Index for Employment	0.29	0.14	1,727	0.23	0.12	1,352	-0.061***	0.005
Gini Spec. Index for Establishments	0.18	0.09	1,727	0.13	0.06	1,352	-0.051***	0.003
Panel B: 1996 Outcomes for Counties that Ever Receive a Highway								
Employment (in 1,000s)	7.31	12.72	1,727	62.63	169.96	1,352	55.319***	4.104
Establishments (in 100s)	6.12	9.72	1,727	39.66	99.75	1,352	33.546***	2.415
Employment Herfindahl Index	2,589.01	728.07	1,727	2,445.76	486.13	1,352	-143.253***	22.998
Establishments Herfindahl Index	2,042.79	258.31	1,727	2,149.44	228.71	1,352	106.654***	8.924
Gini Spec. Index for Employment	0.23	0.11	1,727	0.18	0.09	1,352	-0.050***	0.004
Gini Spec. Index for Establishments	0.14	0.06	1,727	0.10	0.05	1,352	-0.037***	0.002

Notes: Data comes from the 1962 - 1996 County Business Patterns annual reports. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, and Airport. *** p<0.01, ** p<0.05, * p<0.1

1.4 Empirical Strategy

1.4.1 Static Identification

To investigate the effect of the Interstate Highway System on employment growth and industry concentration, I exploit variation in the location of interstate highways at different points in time. I use a county-level panel dataset to estimate the following specification:

$$Y_{cit} = \sum_{d} \beta_d (hwy_{ct} \times YearBin_d) + \delta_{rt} + \gamma_c + X' \rho_{ct} + \epsilon_{cit} \tag{1.3}$$

where Y_{cit} is the outcome of interest in county c, in industry i at time t. The variable hwy_{ct} is an indicator variable that is equal to one if an interstate highway intersects county c at time t. The coefficients of interest are the set of β_d's, which measures the effect of the interstate highway system during the d different periods.[12] I include census region $\times$ year fixed-effects, δ_{rt}, county-fixed effects, γ_c, additional controls, $X'\rho_{ct}$, and ϵ_{cit} is the error term. The controls include alternative methods of transportation infrastructure, 1950 population, 1940 to 1950 population growth, and distance to closest Metropolitan Statistical Area (MSA). I also two-way cluster the standard error by county and state/year to account for serial correlation and spatial correlation in the error term. I estimate equation 1.3 using Two-Stage Least Squares (TSLS).[13]

This specification identifies the effect of being a highway county compared to non-highway at different points in time, β_d. This model does not allow me to separately identify the effect of highways on growth and relocation, but rather the effect of both. The county fixed-effects account for any county characteristics that are not varying between 1962 and 1996 that may be correlated with economic growth and the location and construction of interstate highways. By including the census region $\times$ year fixed-effects, the treatment effect of an interstate highway is only identified from variation within a census region in a year. Including this set of fixed-effects allows me to account for any region wide changes that affect employment, the opening or closing of businesses, or promote growth in specific industries that change over time and are correlated with the construction of interstate highways.[14]

I include transportation infrastructure, population, and geographic controls to account for differences between highway and non-highway counties prior to highway construction. In order to use these time-invariant controls I interact them with an indicator variable in each year to create a "trend" for each control. I include alternative transportation infrastructure controls to account for evolving trends in substitutability or complementarity between the alternative methods and interstate highways. For example, both airlines and railroads faced deregulation over this

[12] I estimate each β_d from a separate regression where I partial out all the fixed-effects and reduce the comparison to a bivariate regression within each period d. The periods of interest are from 1962-1966, 1967-1971, 1972-1976, 1977-1981, 1982-1986, 1987-1991, and 1992-1996.

[13] The results are nearly identical when I estimate the regression using Limited Information Maximum Likelihood (LIML).

[14] I have also considered replaced the region $\times$ year fixed-effects with state $\times$ year fixed-effects. One limitation of this specification is that it excludes non-highway counties in states without any highways until at least one county receives an interstate highway. In that sense I am losing potentially valuable counterfactual information, so I elected to use the census region $\times$ year fixed-effects. The results with state $\times$ year fixed-effects are slightly smaller in magnitude, but tell a similar story to the results presented below.

period, which likely altered the dynamics between highways and the deregulated industries. The population controls account for different trends both in the level of population as of 1950 and the growth in population between 1940 and 1950. The level of population or changes in population are likely correlated with whether a county receives an interstate highway and any future growth in employment. I control for the Euclidean distance from each county centroid to the nearest MSA to account for changes in the relative distance between locations due to technology improvements. These distance time trends account for automobile safety or speed improvements that may affect rural, suburban, and metropolitan areas differently.

1.4.2 Addressing Highway Endogeneity

Measuring the differences between highway and non-highway counties will likely result in biased estimates because counties selected to receive a highway and when they receive the highway are likely to differ along unobservable dimensions that are correlated with economic growth. The history of highway construction indicates that the placement 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 politicians 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 highway 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 address endogeneity concerns regarding highway location, I use two historical government proposals for a national highway system as separate instrumental variables to predict eventual highway location. The first is the military plan proposed by General Pershing in 1921 commonly referred to as the Pershing Map. Proposed highway location data are based on the digitized Pershing Map from the Bureau of Public Roads collection at the National Archives. I intersect the digitized highway locations with a county map from 1980 to determine the set of counties that received the proposed military routes. Figure 1.2 depicts the highly prioritized routes drawn in the Pershing Map.[15] The Pershing Map is relatively new in the literature and has only been used by Michaels et al. (2013).

The second is the proposed map from the National Interregional Highway Committee published in a 1947 report. I similarly digitize the 1947 Plan and identify the set of counties that received proposed highways. Figure 1.3 shows the 1947 Interregional Highway Committee plan. This map is visually very similar to the map of eventual highway locations. Table 1 confirms this result; 81 percent of highway counties were designated to receive a highway by the 1947 Plan compared to only 32 percent for the Pershing Map. The 1947 Plan is the most commonly used location instrument in the literature (Baum-Snow 2007, 2010, 2014; Michaels 2008; Duranton and Turner 2012; Duranton, Morrow, and Turner 2013). I include this instrument in order to position my results in the context of the prior literature.

I address the endogenous timing of highway construction using an application from network theory to predict the optimal timing of highway construction. I borrow from the Newman-Girvan Algorithm (Girvan and Newman 2002, 2004; Newman 2001, 2004) to prioritize each segment of the proposed highway networks. This algorithm was originally used to identify important connections in biological and social networks. To my knowledge this is the first application of this algorithm in the economics literature. In order to apply the algorithm to the each of the historical highway network plans, 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 it's length. The Newman-Girvan Algorithm calculates the edge-betweenness for each edge by determining the shortest path from each node to every other node in the system and then counting the number of shortest paths that move along that edge. Edges with the largest betweenness value are more important for connecting nodes in the network, therefore these edges of the networks should have been built earlier.

[15] The full Pershing Map contains three priority levels, the depicted map shows routes in the two highest priority levels. Priority three routes are shorter in length and appear to be designed to reach specialized locations, like military installations.

My algorithm sequentially builds the network edges with the highest betweenness value subject to an annual construction budget. I derive this constraint from the construction costs of the entire network. I calculate total construction costs by aggregating the construction cost of each edge. Construction costs are based on weighted average costs of the urban and rural mileage. I use construction cost estimates for urban and rural cost per mile from a 1955 Congressional highway proposal. Urban mileage had an estimated cost of $2,431,818 per mile, while rural costs are significantly lower at $378,787 per mile.[16] Contemporary cost estimates of adding new rural and urban highway mileage are consistent with this urban to rural cost ratio.[17] I use historical cost estimates instead of current cost estimates because it better approximates the decision a social planner would have made at the time of construction.

I calculate the total cost of construction for each entire network using the computed cost of each segment of the proposed network. I then calculate the annual construction constraint by dividing the total network construction cost over a twenty-five year construction period, which roughly approximates the timeframe of actual highway construction. 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. The algorithm allows me to assign a construction year for each edge, which results in a highway instrument that predicts both the location of an interstate highway and the year of construction.

1.4.3 Instrument Validity

1.4.3.aStatic Model Inclusion Restriction To test whether each proposed network with predicted construction timing sufficiently predicts whether a county will have an interstate highway at time t, I estimate the following first-stage regression using a Linear Probability Model.

[16] These construction cost estimates include the actual cost of construction as well as the cost of acquiring land.
[17] The ratio of construction costs is more important to the model than the actual costs.

$$hwy_{ct} = \theta Plan_{ct} + \psi_{rt} + \lambda_c + V'\pi_{ct} + \upsilon_{cit} \tag{1.4}$$

The variable $Plan_{ct}$ is an indicator for whether a county c is predicted to have a highway from the proposed network in year t. I also include the covariates from the second-stage, ψ_{rt} are the census region × year fixed-effects, λ are the county fixed-effects, $V'\pi_{ct}$ are the infrastructure, population, and geographic controls, and υ_{cit} is the error term.

Figures 1.4 and 1.5 present the first-stage regression results by year along with the corresponding F-statistics. The F-statistics in these figures only approximate the true F-Statistics used in the paper because the regressions estimate the treatment effects for the 5 year bins. Clustering the error terms by county and state/year alters the i.i.d. assumption associated with the standard first-stage F-statistic calculation. To test the inclusion restriction, I use Kleibergen-Paap F-statistics that adjusts for clustering the error term (Stock and Yogo 2005). The Kleibergen-Paap F-statstic in the static model ranges between 20 and 170 using the Pershing Map and 140 and 1700 using the 1947 Plan, which indicates that using either proposed system of roads with predicted construction timing is a sufficient instrument for both the location and timing of interstate highway construction.

Figure 1.4: First-Stage Coefficients and F-Statistics by Year for the 1947 Plan

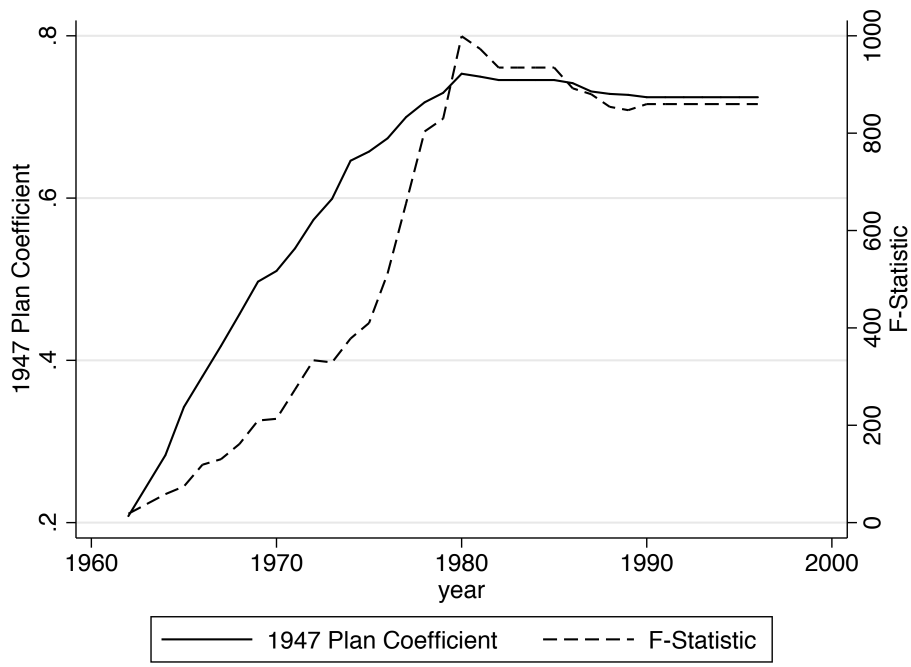


Figure 1.5: First-Stage Coefficients and F-Statistics by Year for the Military Plan

1.4.3.bExclusion Restriction

Using planned transportation networks to instrument for eventual highway location is consistent with several recent empirical papers examining the effects of transportation networks. The 1947 Plan is the most commonly used location instrument in the literature (Baum-Snow 2007, 2010, 2014; Michaels 2008; Duranton and Turner 2012; Duranton, Morrow, and Turner 2013). The primary objective of the 1947 was to *"connect by routes as direct as practicable the principal metropolitan areas, cities, and industrial centers, to serve the national defense and to connect suitable border points with routes of continental importance in the Dominion of Canada and the Republic of Mexico"* (United States Federal Works Agency 1947). The results in Section 8, confirm that the plan was not drawn as a result of growth in population or employment in agriculture and manufacturing. I control for both the level of 1950 population and population growth from 1940 to 1950 because planners were connecting population centers.

The validity of the Pershing system as a suitable instrument hinges on the degree to which military motives in 1921 are orthogonal to employment growth and industry concentration in the latter part of the 20th century. In other words, the Pershing predictions should only influence industry growth and concentration through their ability to predict actual highway construction. One concern is that routes proposed in 1921 may have directly influenced industry growth, employment growth, or population growth. One advantage of using the Pershing system is the strong military influence and the lack of input from outside political and economic agents. These military motivations are evident in the lack of proposed routes extending into southern Florida and the emphasis in roads along the coasts and the borders. Another advantage is that the Pershing system was connected with straight lines. These straight line connections remove the possibility of manipulating the route in order to pass through a specific county.

If the military designed the network around the potential growth of industrial centers, this might result in biased estimates. To test for this, I regress the Pershing system on changes in population and employment in both agriculture and manufacturing between 1910 and 1940, with the same set of fixed-effects and controls as equation 1.3 and I do not find any evidence that the military was choosing areas with high growth rates in either industry or in population. Section 1.8 elaborates further on these results.

1.5 Employment and Establishment Growth

1.5.1 Total Employment and Establishment Growth

To measure whether the Interstate Highway System changed industry concentration, I start by determining if there is a difference in the size of employment and the number of establishments for highway counties compared to non-highway counties and whether or not the difference is changing over time. Table 1.3 shows growth patterns for both employment and establishments using the OLS and TSLS specifications. The coefficient estimates, β_d, compare highway to non-highway counties measured in five year intervals. The coefficients can be interpreted as the difference in highway and non-highway counties in period d.

Table 1.3:

Table 3: The Effect of Highways on Total Employment and Total Establishments

	Ln(Employment)			Ln(Establishments)		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
Hwy X 1962-1966	-0.0506***	-0.0920	-0.00844	-0.0686***	-0.112**	-0.0545
	(0.0149)	(0.0594)	(0.0766)	(0.0116)	(0.0448)	(0.0540)
Hwy X 1967-1971	-0.0523***	-0.0289	-0.104**	-0.0515***	-0.0274	-0.0421
	(0.0115)	(0.0268)	(0.0455)	(0.00950)	(0.0214)	(0.0351)
Hwy X 1972-1976	-0.0390***	-0.0423***	-0.0735**	-0.0273***	-0.0257***	-0.0122
	(0.00756)	(0.0125)	(0.0307)	(0.00538)	(0.00905)	(0.0190)
Hwy X 1977-1981	-0.0133*	-0.0160*	-0.0220	-0.00591	-0.0118**	-0.00906
	(0.00694)	(0.00915)	(0.0281)	(0.00360)	(0.00533)	(0.0142)
Hwy X 1982-1986	0.0264***	0.0263**	0.0411	0.0223***	0.0228***	0.0276
	(0.00812)	(0.0104)	(0.0329)	(0.00553)	(0.00755)	(0.0208)
Hwy X 1987-1991	0.0586***	0.0795***	0.137***	0.0453***	0.0595***	0.0331
	(0.0106)	(0.0144)	(0.0423)	(0.00829)	(0.0116)	(0.0305)
Hwy X 1992-1996	0.0704***	0.101***	0.181***	0.0625***	0.0828***	0.0789*
	(0.0137)	(0.0193)	(0.0525)	(0.0109)	(0.0155)	(0.0412)

Notes: All estimates are from a panel of counties from 1962 - 1996 that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

By the early 1990s, employment was 7.04 percent higher in highway counties compared to non-highway counties. TSLS results for both instruments indicate positive employment growth occurred at a similar time but was substantially larger than the OLS results suggest. After the mid-1980s, the TSLS highway interaction terms are all substantially larger than the OLS. Considering the same period in the early 1990s, employment was 10-18 percent higher in highway counties relative to non-highway counties. Duranton and Turner (2012) find that, within US cities, a 10 percent increase in highway mileage leads to a 1.5 percent increase in total employment over 20 years. The results presented in Table 1.3 are consistent with their result: I find that highway counties gain between 10 and 25 percent more employment every 20 years compared to non-highway counties.

Columns 4 through 6 present the OLS and TSLS regression results for establishment growth.

The OLS and TSLS results indicate that growth in establishments was roughly monotonic. The patterns are similar to the employment results but are typically smaller in magnitude. Taking the employment and establishment results together, it suggests that highways led to employment growth on both the extensive and intensive margins.

The general pattern of growth is consistent with Michaels (2008), where the benefits of highway infrastructure occurs after the mid-1970s. The first three columns of Table 3 show the results for employment. The OLS results indicate that highways counties have lower employment in the early periods and larger employment differences in the later periods. The early employment differences suggest that highways led non-highway counties to have more employment. One explanation for this difference is that non-highway counties that are currently constructing their segments of interstate highway may have an influx of employment. I test this theory by comparing the employment of places that just received their highways to counties that are about to receive their highways.[18] The results suggest that between 20 and 40 percent of the difference between highway and non-highway counties can be explained by these soon to be highway counties.

The difference between the OLS and both TSLS estimates highlights two potential forms of bias consistent with politicians and lobbyists using interstate highways as place-based economic development policies for growth and directing interstate highways to negatively selected counties. Recall that the OLS estimates could be biased for two different reasons; the endogenous placement of highways and the endogenous funding of highways. The OLS estimates in the early periods suffer from both forms of endogeneity, whereas the estimates after 1990 primarily suffer from location endogeneity. The difference between the OLS and both TSLS estimates suggests the location endogeneity induces a negative bias on the estimates, which is consistent with planners and government officials assigning interstate highways to lower-quality locations. This result is consistent with the interstate highway literature and the literature on other place-based development interventions (Duranton and Turner 2012).

The difference between the OLS and both TSLS estimates in the early years indicates that

[18] These results are available by request.

the estimates are positively biased. The difference in the direction of the bias comes from differences in the predicted timing of highway construction. Figures 1.6 presents a map for actual interstate highway construction progress in 1965. Figures 1.7 and 1.8 present the maps for predicted construction progress using the Pershing Map and 1947 Plan respectively. The biggest differences between the maps is the disjoint nature of the IHS construction compared to the predicted construction plans. The predicted construction plans build the highway networks progressively. The number of small segments in the map of actual highway construction suggests that areas were targeted. This targeting was done specifically based on the quality of location. A comparison of the raw data supports this hypothesis: areas targeted earlier for highway construction had higher levels of employment and more establishments than areas targeted later. This bias is not present in the IV.

Figure 1.6: Interstate Highways Constructed in 1965

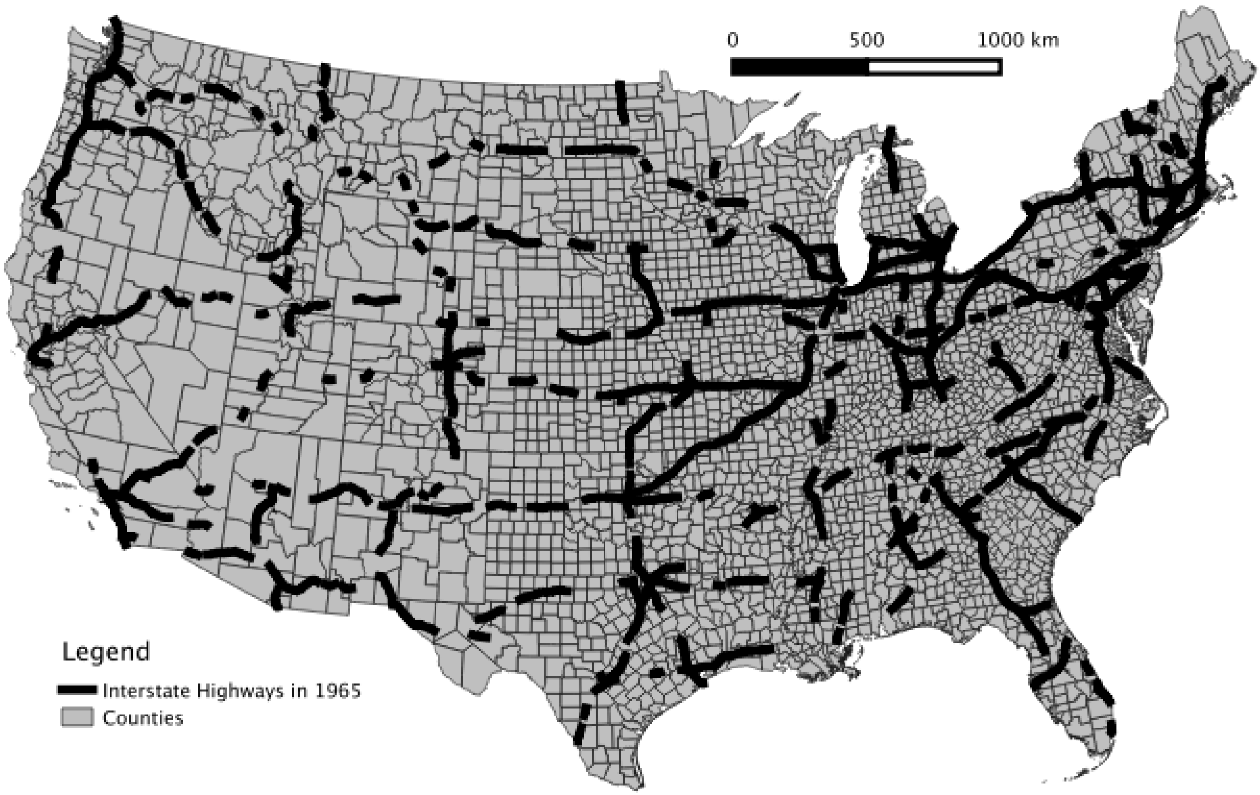


Figure 1.7: Proposed Military Plan Construction in 1965

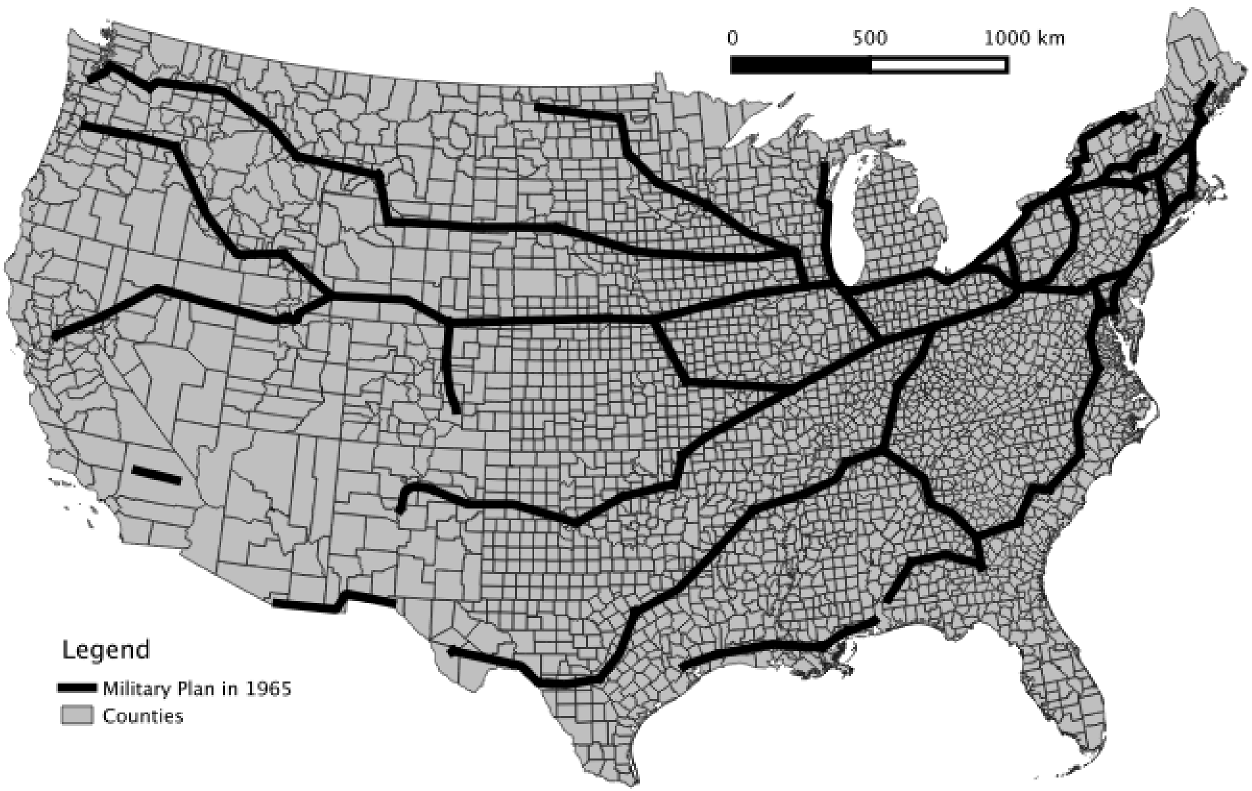
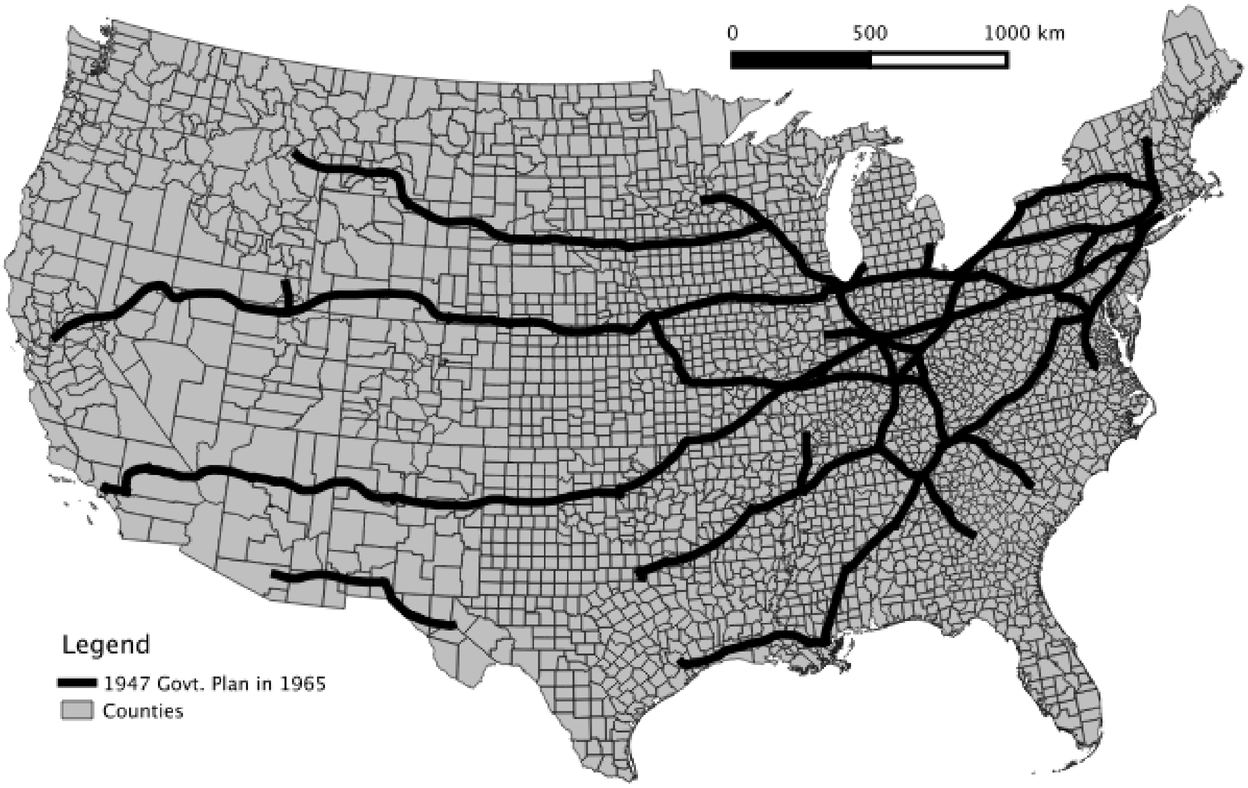


Figure 1.8: Proposed 1947 Plan Construction in 1965



Putting the two forms of endogeneity together, interstate highways were assigned to lower performing locations but within this group of locations they were constructed in the highest performing places first. The combination of these two forms of bias results in a positive bias in the early OLS estimates and an negative bias in the later estimates. The early results also indicate the importance of the positive timing bias, which is substantially larger than the negative location bias.

1.5.2 Employment Growth by Industry

Next, I determine if the employment growth observed in the previous section varies across sectors. Table 1.4 shows employment growth results for 4 of the 10 industry classifications, agriculture, manufacturing, retail sales, and transportation and public utilities.[19] These four industries generally follow the patterns found in total employment growth the the previous section. The results indicate employment grew the most in the agricultural sector and the transportation and public utilities sector. By the mid-1990s, employment in both sectors was between 17 and 27 percent higher in highway counties compared to non-highway counties. The large gains in employment in agriculture are consistent with the results found in Frye (2015a).

[19] Regression results for all 10 industry classifications are available by request. Employment growth across the ten industries is mostly consistent with earnings growth found by Chandra and Thompson (2000).

Table 1.4:

Table 4: The Effect of Highways on Employment Growth by Industry

	Agriculture			**Manufacturing**		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
Hwy X 1962-1966	-0.145***	-0.333**	-0.403**	-0.00842	0.0930	0.199
	(0.0328)	(0.131)	(0.165)	(0.0323)	(0.108)	(0.156)
Hwy X 1967-1971	-0.124***	-0.183**	-0.299**	0.00508	0.0815	-0.0186
	(0.0307)	(0.0721)	(0.127)	(0.0265)	(0.0582)	(0.109)
Hwy X 1972-1976	-0.0841***	-0.0883**	-0.0782	-0.00687	-0.0216	-0.0436
	(0.0239)	(0.0364)	(0.0939)	(0.0211)	(0.0314)	(0.0905)
Hwy X 1977-1981	-0.0356	-0.0309	-0.0145	-0.0153	0.00426	0.0685
	(0.0253)	(0.0336)	(0.0975)	(0.0206)	(0.0298)	(0.0989)
Hwy X 1982-1986	0.0396*	0.0531*	0.102	0.00872	0.0273	-0.0218
	(0.0230)	(0.0309)	(0.0964)	(0.0236)	(0.0307)	(0.102)
Hwy X 1987-1991	0.106***	0.148***	0.186*	0.00804	0.0265	0.0128
	(0.0247)	(0.0340)	(0.104)	(0.0257)	(0.0336)	(0.114)
Hwy X 1992-1996	0.151***	0.186***	0.247**	0.0254	0.0217	0.0440
	(0.0281)	(0.0375)	(0.120)	(0.0268)	(0.0362)	(0.118)

	Retail Sales			**Transp/Utilities**		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
Hwy X 1962-1966	-0.0879***	-0.171***	-0.0636	-0.121***	-0.306***	-0.328**
	(0.0150)	(0.0567)	(0.0699)	(0.0249)	(0.0983)	(0.134)
Hwy X 1967-1971	-0.0751***	-0.0728***	-0.0568	-0.0937***	-0.117**	-0.241**
	(0.0118)	(0.0263)	(0.0461)	(0.0221)	(0.0520)	(0.0985)
Hwy X 1972-1976	-0.0375***	-0.0536***	-0.0273	-0.0150	-0.0404	-0.0578
	(0.00707)	(0.0111)	(0.0265)	(0.0175)	(0.0312)	(0.0732)
Hwy X 1977-1981	0.00604	-0.00109	0.00911	-0.0262	-0.0271	-0.0222
	(0.00572)	(0.00803)	(0.0240)	(0.0193)	(0.0273)	(0.0824)
Hwy X 1982-1986	0.0321***	0.0256**	0.0395	0.0418**	0.0456*	0.0696
	(0.00763)	(0.0105)	(0.0326)	(0.0188)	(0.0257)	(0.0878)
Hwy X 1987-1991	0.0650***	0.0905***	0.0166	0.101***	0.132***	0.257**
	(0.0113)	(0.0155)	(0.0423)	(0.0238)	(0.0327)	(0.102)
Hwy X 1992-1996	0.0867***	0.116***	0.0513	0.0951***	0.175***	0.279**
	(0.0141)	(0.0195)	(0.0524)	(0.0266)	(0.0382)	(0.116)

Notes: Dependent variable is the log of employment for each industry. All estimates are from a1962 - 1996 panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Growth in manufacturing employment follows a similar monotonically increasing pattern but the results are not significantly distinguishable from zero. These muted gains in highway counties are consistent with the findings for manufacturing earnings in Chandra and Thompson (2000). Baum-Snow (2014) finds that increasing interstate highways in SMSAs led manufacturing jobs to move to rural areas or abroad. Combining my results with the results in Baum-Snow (2014) suggests that manufacturing jobs moved from urban areas to rural counties with interstate highways, because we do not see a net change in manufacturing employment in highway counties relative to non-highway counties.

Employment growth in retails sales follows a similar monotonically increasing pattern to overall employment growth for both the OLS and using the 1947 Plan as an IV. The TSLS results using the Pershing Map are slightly smaller in magnitude and are estimated with less precision. As a result we cannot rule out that there is no difference between retail sales employment in highway and non-highway counties. It is also worth noting that the estimated effects for employment are similar in magnitude to the growth in retail sales per capita found by Michaels (2008).

The degree of difference between the OLS and both TSLS estimates varies considerably across industries. The bias is most pronounced in agriculture and in transportation and public utilities, which is consistent with the historical accounts of industrial involvement in the planning of interstate highways. The bias is much smaller in retail sales and is similar to results found by Michaels (2008), which showed little difference between OLS and IV estimates when measuring the effect of highways on retail sales per capita.

1.6 Employment and Establishment Concentration

1.6.1 Concentration Across-Industries

Results from the previous section established that highways led to significant employment and establishment growth differences between highway and non-highway counties. The findings also indicate this growth was not equally distributed across industries. Unequal growth both across space and across industries suggests that highways may induce changes in regional specialization. In this section, I measure the degree to which interstate highways led to differential specialization in employment and the number of establishments. To empirically measure specialization I will use the Herfindahl Index and the Gini Specialization Index described in equation 1.1 and 1.2. Larger values for both of these measures indicate a higher degree of concentration where a larger share of employment is in fewer sectors. Table 1.5 presents OLS and TSLS results for the different concentration measures. The dependent variable in Panel A is the Herfindahl Index and the dependent variable in Panel B is the Gini Specialization Index.

Table 1.5:

Table 5: The Effect of Highways on Industry Concentration

Panel A: Herfindahl Index

	Employment			Establishments		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
Hwy X 1962-1966	60.52**	55.45	273.8*	-29.07***	-24.23	1.525
	(25.45)	(92.06)	(141.8)	(9.382)	(38.97)	(51.92)
Hwy X 1967-1971	32.67	47.13	-129.9	-17.82**	-3.830	-27.89
	(21.31)	(45.80)	(83.92)	(8.454)	(19.75)	(36.98)
Hwy X 1972-1976	-25.45*	21.01	-235.6***	-6.570	-6.361	-19.29
	(15.32)	(27.97)	(62.67)	(4.806)	(8.242)	(22.21)
Hwy X 1977-1981	-28.51*	-23.72	-159.8**	5.364	3.961	44.67**
	(15.17)	(20.42)	(64.32)	(5.155)	(6.905)	(22.13)
Hwy X 1982-1986	-22.31	-43.18**	62.33	2.458	-3.863	51.65**
	(15.05)	(20.04)	(66.09)	(5.791)	(7.794)	(24.95)
Hwy X 1987-1991	6.047	-11.71	312.6***	13.49**	4.056	17.90
	(19.68)	(25.35)	(93.33)	(6.166)	(8.348)	(28.20)
Hwy X 1992-1996	29.22	12.51	382.6***	30.52***	24.29**	-28.72
	(24.63)	(31.89)	(101.6)	(7.620)	(10.60)	(34.43)

Panel B: Gini Specialization Index

	Employment			Establishments		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
Hwy X 1962-1966	0.00194	-0.0196*	-0.00349	-0.00447***	-0.00435	-0.0125**
	(0.00291)	(0.0102)	(0.0153)	(0.00119)	(0.00435)	(0.00626)
Hwy X 1967-1971	-0.000893	-0.00919*	-0.00589	-0.00170	0.00264	-0.00251
	(0.00239)	(0.00534)	(0.00954)	(0.00111)	(0.00235)	(0.00459)
Hwy X 1972-1976	-0.00355**	-0.00180	-0.0201***	0.00110	0.00178	0.000502
	(0.00174)	(0.00296)	(0.00702)	(0.000772)	(0.00135)	(0.00314)
Hwy X 1977-1981	-0.00221	-0.00262	-0.00833	0.00271***	0.00222**	0.00199
	(0.00174)	(0.00227)	(0.00730)	(0.000843)	(0.00108)	(0.00361)
Hwy X 1982-1986	-0.000901	-0.00134	0.00234	0.000603	0.000659	0.00696*
	(0.00173)	(0.00228)	(0.00763)	(0.000906)	(0.00126)	(0.00400)
Hwy X 1987-1991	-0.000319	0.00384	0.0153	-0.000169	-1.27e-05	0.00529
	(0.00229)	(0.00315)	(0.0101)	(0.00102)	(0.00142)	(0.00418)
Hwy X 1992-1996	0.00106	0.00834**	0.00920	0.000248	-0.000306	0.00277
	(0.00265)	(0.00354)	(0.0112)	(0.00111)	(0.00156)	(0.00472)

Notes: All estimates are from a 1962 - 1996 panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

The concentration results using the Herfindahl Index indicate employment was more concentrated in highway counties in the early years of highway construction, then became more diverse, before finally becoming more concentrated again. The explanation for this pattern may be similar to the explanation for employment. If employment in non-highway counties grow, particularly in very few sectors, then the herfindahl index would likely rise initially. The results from the TSLS specification with the Pershing Map indicates this shift is only temporary and by the early-1970s employment in highway counties is less concentrated than in non-highway counties. By the 1990s, highway counties are substantially more concentrated. The TSLS estimates using the Pershing Map indicate employment in highway counties was 18 percent more concentrated at the mean than non-highway counties.[20] The concentration results using the Herfindahl Index for the number of establishments shows fewer statistically significant results. The results using the 1947 Plan as an IV suggest there may have been limited establishment concentration by the mid-1990s. The results using the Gini Specialization Index tell a similar story for employment. Highway counties appear to diversify their employment relative to non-highway counties in the mid-1970s. Employment then becomes more concentrated in the late 1980s, although the large standard errors makes inference difficult.

1.6.2 Skill-Endowments and Concentration

I exploit the introduction of the Interstate Highway System to quantify the role of skill endowments in changes in industry concentration. In a simple two-factor Heckscher-Ohlin trade model with two economies, lowering trade costs and removing trade barriers through the expansion of transportation networks should lead skill-abundant areas to shift production to skill-intensive industries and low-skill areas to shift production to low-skill industries. This suggests that areas in the tails of the skill distribution should be more likely to specialize in particular industries. I test this theory by determining whether or not employment in extreme skill places is more likely to be concentrated following the introduction of the Interstate Highway System. Empirically, I interact

[20] This value is the coefficient estimate divided by the average Herfindahl Index for employment across all years.

the highway indicator variables in equation 1.3 with a binary indicator for extreme skill. I define extreme skill as places in the top 25th and bottom 25th percentiles of 1950 skill distribution, where I approximate the skill distribution with the percent of people over the age of 25 with at least a high school diploma.

Table 1.6 presents the regression results measuring the effect of extreme skill on employment concentration. The coefficients of interest are the interaction term between highway counties in each year and the extreme skill dummy variable. Based on the Heckscher-Ohlin model, I expect these interaction terms to have a positive coefficient, indicating that places in the tails of the skill distribution are more likely to specialize following the introduction of interstate highways. The results seem to weakly support this theory although the coefficient estimates in many cases are not statistically different from zero.

Table 1.6:

Table 6: The Effect of Highways on Employment Concentration by Highway Age

	Herfindahl Index			**Gini Specialization Index**			F-Stat	
	OLS	TSLS	TSLS	OLS	TSLS	TSLS		
		1947 Plan	Military Map		1947 Plan	Military Map	1947 Plan	Military Map
Hwy X 1962 - 1973	34.665	40.285	83.649	-0.001	-0.005	-0.010*		
	(24.647)	(87.114)	(51.480)	(0.003)	(0.010)	(0.006)	230.4	78.67
Hwy X 1962 - 1973 X Extreme Skill Dum	-4.840	-73.911*	-32.190	0.000	0.002	0.002		
	(23.582)	(43.908)	(34.802)	(0.002)	(0.005)	(0.004)		
Hwy X 1974 - 1983	-40.653**	-205.708***	-30.113	-0.003	-0.015**	-0.003		
	(17.267)	(66.106)	(24.520)	(0.002)	(0.007)	(0.003)	790.6	65.36
Hwy X 1974 - 1983 X Extreme Skill Dum	15.003	40.053*	29.551*	0.001	0.001	0.002		
	(12.851)	(22.422)	(16.060)	(0.002)	(0.003)	(0.002)		
Hwy X 1984 - 1996	15.006	377.075***	-19.090	-0.000	0.011	0.007*		
	(27.595)	(106.915)	(37.485)	(0.003)	(0.012)	(0.004)	1108	62.28
Hwy X 1984 - 1996 X Extreme Skill Dum	-7.194	22.485	3.078	0.000	0.005	-0.001		
	(19.551)	(34.078)	(22.820)	(0.002)	(0.004)	(0.002)		

Notes: Kleibergen-Paap F-Statistics are reported. All estimates are from a 1962 - 1996 panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

1.6.3 Establishment Scale

This section considers the role of interstate highways in promoting changes in the size of firms. I measure scale changes using two similar metrics, the share of firms with more than 20 employees and the share of firms with between one and four employees. Understanding the effect of highways on the size of firms is informative for several reasons. It connects to a literature on the relationship between employment growth and firm size. It also has implications for market power within industries and the role of infrastructure in promoting entrepreneurship. My empirical analysis follows from equation 1.3 and I estimate the effects of interstate highways on the scale of firms for each industry. All of the results are available by request.

The results indicate the effect of interstate highways on the share of larger firms varies considerably by industry. When the dependent variable is the share of firms with more than 20 employees, the results indicate there was considerable variation across industries. The results indicate the percentage of firms in agriculture, construction, wholesale trade, and the unclassified industries show no difference in the percentage of firms with over 20 employees between highway and non-highway counties. Only firms in finance had a smaller proportion of medium and large sized firms in highway counties relative to non-highway counties. Combined with the employment growth results from Section 5.2, this suggests most of the employment growth in finance occurred among smaller firms. The proportion of medium and larger firms grew for several industries, including mining, retail sales, services, and transportation and public utilities, in highway counties relative to non-highway counties. When these results are considered with the employment growth results by industry, it appears that most of the employment growth occurred in medium and large firms for these industries.

Next I consider the effect on the smallest firm size category, firms with between one and four employees. Changes to these firms may give some insight into the impact of interstate highways on firms with no employees, which are not observed in the data. The results indicate the proportion of tiny firms changed for only a couple industries over this period as a result of interstate highways.

The percentage of tiny firms in retail sales, services, and transportation and public utilities all fell considerably after the expansion of interstate highways. Considering highways only affect three of the ten industries, these results indicate that small businesses with no employees are not likely to substantially change the results. These results also suggest interstate highways are not useful for decreasing market power or promoting entrepreneurship in small businesses, in fact taken with the prior results the opposite appears to be true.

1.7 Dynamic Effects of Interstate Highways

The prior two sections measure the differences between highway and non-highway counties at different points in time. Now I focus on a more dynamic model for measuring the effects of the Interstate Highway System, which map out the full dynamic response of the outcomes of interest to receiving an interstate highway.

1.7.1 Dynamic Identification

The static model allows me to estimate the causal difference between highway and non-highway counties at different points in time. However, it does not allow me to separate the effects of recently constructed highways from newly constructed highways. To understand the dynamic effects of having a highway a certain number of years after construction I need to use a more dynamic approach. I adjust the prior specification to identify the effects of the Interstate Highway System by measuring the evolution of the effects over time.

$$Y_{cit} = \phi_c + \rho_{rt} + \sum_a \beta_a HwyAgeDum_{ct}^a + X_{ct}'\mu + \epsilon_{cit} \tag{1.5}$$

Similar to equation 1.3, Y_{cit} is the outcome of interest in county c, in industry i at time t. The variable $HwyAgeDum_{ct}$ refers to a series of dummy variables set equal to one if a county received an interstate highway a years ago. The coefficients of interest are the set of β_a's, which measures the effect of the interstate highway system the stated number of years ago. These coefficients map out the full dynamic response of the outcomes of interest to receiving an interstate highway. I continue to include region $\times$ year fixed-effects, δ_{rt}, county-fixed effects, γ_c, additional controls, $X'\rho_{ct}$, and ϵ_{cit} is the error term. I include controls for alternative methods of transportation infrastructure, 1950 population, and 1940 to 1950 population growth, and distance to closest Metropolitan Statistical Area (MSA) because these are likely correlated with whether a county receives a highway and when they start building that highway. I two-way cluster the standard error by county and state/year to account for serial correlation and spatial correlation in the error term. I estimate equation 1.5 using Two-Stage Least Squares (TSLS).[21]

1.7.1.aDynamic Model Inclusion Restriction To test whether each proposed network with predicted construction timing sufficiently predicts the age of each segment of interstate highway, I estimate the following first-stage regression using a Linear Probability Model.

$$HwyAgeDum^{a}_{ct} = \alpha_c + \rho_{rt} + \gamma PlannedHwyAgeDum^{a}_{ct} + X'_{ct}\delta + \upsilon_{ct} \tag{1.6}$$

The variable $PlannedHwyAgeDum_{ct}$ is an indicator for whether county c is predicted to have a highway from the proposed network in year t that is age a years old. I also include the covariates from the second-stage, ρ_{rt} are the census region $\times$ year fixed-effects, α are the county fixed-effects, $V'\pi_{ct}$ are the infrastructure, population, and geographic controls, and υ_{cit} is the error term. The first-stage regression results are available in Table 1.7. The instrument predicts sufficiently well for both the Pershing Map and the 1947 Plan with the exception of 0-4 year highways using the Pershing Map.

[21] The results are nearly identical when I estimate the regression using Limited Information Maximum Likelihood (LIML).

Table 1.7:

Table 7: The Effect of Highways on Total Employment and Total Establishments by Highway Age

	Ln(Employment)			Ln(Establishments)			F-Stat	
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map	1947 Plan	Military Map
0 - 4 Years	-0.0306***	-0.137	-1.309	-0.0315***	-0.128	-0.623	9.928	0.706
	(0.00833)	(0.201)	(1.825)	(0.00658)	(0.148)	(0.925)		
5 - 9 Years	-0.0293***	-0.147	-0.235	-0.0330***	-0.190*	-0.121	11.88	12.01
	(0.00824)	(0.127)	(0.176)	(0.00660)	(0.0994)	(0.112)		
10 - 14 Years	-0.0164**	0.0194	0.0529	-0.0187***	-0.0599	0.0440	19.67	21.42
	(0.00729)	(0.0866)	(0.105)	(0.00567)	(0.0577)	(0.0743)		
15 - 19 Years	0.00409	0.0941	0.222*	-0.00325	0.0309	0.114	22.79	20
	(0.00708)	(0.0722)	(0.117)	(0.00549)	(0.0531)	(0.0779)		
20 - 24 Years	0.0220***	0.153**	0.265**	0.0160***	0.107**	0.0537	40.55	18.95
	(0.00830)	(0.0660)	(0.129)	(0.00616)	(0.0491)	(0.0790)		
25 - 29 Years	0.0460***	0.143**	0.200*	0.0404***	0.131***	0.0512	53.12	25.96
	(0.00936)	(0.0676)	(0.102)	(0.00756)	(0.0492)	(0.0717)		
30 - 34 Years	0.0690***	0.181***	0.150*	0.0653***	0.165***	0.0412	60.87	48.10
	(0.0121)	(0.0632)	(0.0798)	(0.00940)	(0.0488)	(0.0563)		
35 - 39 Years	0.0691***	0.0284	0.00997	0.0855***	0.125***	0.0499	65.03	52.26
	(0.0182)	(0.0607)	(0.0810)	(0.0153)	(0.0474)	(0.0610)		

Notes: Kleibergen-Paap F-Statistics are reported. All estimates are from a 1962 - 1996 panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

1.7.2 Employment and Establishment Growth

Table 1.7 presents the regression results for the full dynamic response of employment and the number of establishments to receiving an interstate highway. These results are consistent with the prior findings of the effects of interstate highways on employment: both TSLS results indicate substantial employment and establishment growth takes between 15 and 20 years to be realized. This explains why many of the positive benefits of interstate highways are not evident in the static model until the late 1970s. In 1996, the average highway was about 30 years old, which indicates the average highway community experienced between 15 and 18 percent more employment than non-highway counties.

1.7.3 Employment and Establishment Concentration

The dynamic response of industry concentration as measured by the Herfindahl Index and the Gini Specialization Index are presented in Table 1.8. The two measures of concentration give slightly competing results. The Herfindahl Index results for employment suggest that the longer highways are in a county the more likely that county is to diversify. This is contrasted with the results from the Gini Specialization Index, which shows that the longer a highway is in an area the more like it is to specialize. Both concentration measures for the number of establishments appear to support that the longer an area has an highway the more likely it is to specialize although these estimates are imprecisely estimated, which makes inference difficult.

Table 1.8:

Table 8: The Effect of Highways on Industry Concentration by Highway Age

Panel A: Herfindahl Index

	Employment			Establishments		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
0 - 4 Years	16.37	162.8	-4278	-14.94***	-27.7	-941.4
	(16.1)	(311.8)	(5459)	(5.787)	(130.3)	(1334)
5 - 9 Years	24.35*	198.2	-92.11	-10.91**	-167.1*	28.07
	(14.63)	(239.7)	(329)	(5.407)	(91.57)	(136.3)
10 - 14 Years	2.006	192.7	125.7	0.791	-141.3**	76.62
	(13.67)	(173.4)	(212.3)	(4.824)	(62.21)	(90.01)
15 - 19 Years	6.573	-32.29	314.9	8.153*	-75.1	88.17
	(12.53)	(135.8)	(228.4)	(4.566)	(53.33)	(86.95)
20 - 24 Years	7.58	-18.22	928.2***	10.29**	35.14	126.4
	(13.81)	(110.3)	(317)	(5.222)	(43.44)	(93.93)
25 - 29 Years	-7.265	-124.5	293.5	13.27**	53.24	20.75
	(16.21)	(112.2)	(178.6)	(5.746)	(37.46)	(65.79)
30 - 34 Years	11.68	-217.7**	22.44	16.36**	63.07*	-36.01
	(20.64)	(104.3)	(143.9)	(6.469)	(37.63)	(52.04)
35 - 39 Years	-24.19	-336.8***	-128.8	10.24	39.29	-23.98
	(27.21)	(108.1)	(108.6)	(9.349)	(37.46)	(37.08)

Panel B: Gini Specialization Index

	Employment			Establishments		
	OLS	TSLS 1947 Plan	TSLS Military Map	OLS	TSLS 1947 Plan	TSLS Military Map
0 - 4 Years	0.000189	0.00623	0.0587	-0.00119	-0.00522	-0.152
	(0.00183)	(0.0357)	(0.233)	(0.000819)	(0.0170)	(0.206)
5 - 9 Years	-0.00198	-0.0426	-0.0818*	-0.000206	-0.00260	-0.0257
	(0.00166)	(0.0300)	(0.0425)	(0.000731)	(0.0108)	(0.0193)
10 - 14 Years	-0.00311**	-0.0353**	-0.0377	0.000850	0.000264	-0.00772
	(0.00152)	(0.0179)	(0.0254)	(0.000680)	(0.00747)	(0.0131)
15 - 19 Years	-0.00152	-0.00847	0.00669	0.000214	0.000730	0.0198
	(0.00147)	(0.0146)	(0.0260)	(0.000664)	(0.00774)	(0.0137)
20 - 24 Years	0.000165	0.00436	0.0328	-2.75e-05	0.00694	0.0400***
	(0.00159)	(0.0133)	(0.0283)	(0.000778)	(0.00656)	(0.0151)
25 - 29 Years	0.000994	-0.00194	0.00560	0.000147	-0.00543	0.0108
	(0.00177)	(0.0128)	(0.0207)	(0.000829)	(0.00531)	(0.00993)
30 - 34 Years	0.000705	0.0248*	0.00441	0.000981	0.00255	0.00116
	(0.00222)	(0.0137)	(0.0171)	(0.000942)	(0.00485)	(0.00677)
35 - 39 Years	-0.00136	0.0225*	0.00147	0.00122	0.00163	0.000727
	(0.00311)	(0.0119)	(0.0127)	(0.00136)	(0.00490)	(0.00530)

Notes: All estimates are from a 1962 - 1996 panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

1.8 Robustness

1.8.1 Effects Prior to Construction

One threat to the empirical strategy is that the military or the Interregional Highway Committee may have targeted areas to receive highways that were growing already, were expected to grow, or had time varying characteristics that made them more likely to grow. Table 1.9 empirically tests for this possibility by measuring the effect of the Pershing Map and the 1947 Plan on several economic outcomes prior to the construction of the Interstate Highway System. I construct a panel dataset from the U.S. Census that includes information on population, urbanization, and the two dominant industries, agriculture and manufacturing, that covers from 1900 to 1940. Using this dataset I estimate the following regression, which is similar to equation 1.3, to determine the likelihood that the government targeted specific areas for growth potential:

$$Y_{ct} = \sum_{d} \beta_d (Plan_c \times YearBin_d) + \delta_{rt} + \gamma_c + X'\rho_{ct} + \epsilon_{ct} \tag{1.7}$$

where Y_{ct} is the percent growth in the outcome of interest in county c between time t and time $t-1$. The variable $Plan_c$ is an indicator variable that is equal to one if a county c was supposed to receive a planned highway. The coefficients of interest are the set of β_d's, which measures the effect of the planned highway during the d different periods. The interaction term in 1910 is the excluded year. I include the same set of controls and fixed-effects as the prior regressions. I also two-way cluster the standard errors by county and state/year.

Table 1.9:

Table 9: The Effect of a Planned Highway on Historical Census Outcomes Prior to Construction

	(1) Population	(2) Urbanization	(3) Manuf. Estab	(4) Manuf. Employ	(5) Number of Farms	(6) Farmers
Panel A: Planned 1947 Map						
1947 Plan X 1920	0.00631	-0.0523	-0.0162	0.145	-0.00363	-0.00363
	(0.0218)	(0.0971)	(0.0801)	(0.100)	(0.0228)	(0.0227)
1947 Plan X 1930	0.0384*	0.0221	-0.00541	0.141**	0.00206	0.00308
	(0.0220)	(0.0822)	(0.0456)	(0.0690)	(0.0183)	(0.0184)
1947 Plan X 1940	0.0290	-0.108	0.0357	0.102	0.00118	0.00311
	(0.0251)	(0.0732)	(0.0439)	(0.0727)	(0.0240)	(0.0241)
Panel B: Planned Military Map						
Military Plan X 1920	-0.0127	-0.00202	0.0660	0.140	0.00931	0.00931
	(0.0139)	(0.0866)	(0.0594)	(0.0880)	(0.0187)	(0.0187)
Military Plan X 1930	-0.0173	0.0344	0.000243	0.0362	-0.0157	-0.0151
	(0.0119)	(0.105)	(0.0313)	(0.0628)	(0.0140)	(0.0139)
Military Plan X 1940	-0.0266**	-0.143*	-0.000865	-0.0193	-0.0141	-0.0128
	(0.0125)	(0.0775)	(0.0281)	(0.0642)	(0.0155)	(0.0156)
Observations	11,244	11,244	10,460	10,160	11,244	11,225
Number of fips	2,811	2,811	2,641	2,623	2,811	2,811

Notes: All estimates are from a panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

The outcomes of interest from the regression are two general measures, population and urbanization, and two measures of industry similar to the metrics used in the paper, establishments and employment. Panel A of Table 1.9 presents the results for the 1947 Plan. There is some evidence that the 1947 plan may have been influenced by the growth potential in 1930, however these effects appear to diminish by 1940. Panel B presents the results for the Military Plan. These results look better in the years immediately around the proposed plan. The only statistically significant difference is that Pershing Map is negatively associated with population and urbanization in 1940. This may be by design, the original Pershing Map was intentionally designed to run near but not through urban areas. Mechanically this could create a negative relationship between Pershing Map counties and growth if the routes were drawn to intentionally avoid growing areas. This does not appear to be affecting the measures of industry growth. Overall, these results seem to indicate neither plan targeted locations that were poised to grow.

1.8.2 Planned but Unbuilt Highway Segments

To verify that counties assigned to receive highways in the Pershing Map did not grow because they were assigned routes, which would violate the exclusion restriction, I examine whether planned but unbuilt routes in the Pershing Map affected growth. The full Pershing Map, contained three priority levels of routes, many of which were never constructed. I exploit these unbuilt routes to verify that it is actually receiving a route that benefits a location, not have a planned route that was never built. I estimate the following regression to determine whether these unbuilt routes predict growth in employment and the number of establishments:

$$Y_{ct} = \sum_{d} \beta_d (Unbuilt_c \times YearBin_d) + \delta_{rt} + \gamma_c + X'\rho_{ct} + \epsilon_{ct} \tag{1.8}$$

where Y_{ct} is either the log of employment or the number of establishments in county c at time t. The variable$Unbuilt_c$ is an indicator variable that is equal to one if a county c was supposed to receive a segment of the Pershing Map and never received any highway. The coefficients of interest are the set of β_d's, which measures the effect of the unbuilt segment during the d different periods. The interaction term in 1962-1966 is the excluded year. I include the same set of controls and fixed-effects as in the prior models. I also two-way cluster my standard errors by county and state/year. I restrict the sample to counties that never received an interstate highway, so the comparison is between non-highway counties and non-highway counties that contain any unbuilt portions of the Pershing Plan. Figure 1.9 shows the sample of counties with proposed routes that did not receive them. The results are presented in Table 1.10 and suggest that the unbuilt segments of the Pershing Map have no impact on employment and the number of establishments. This result supports the exogeneity requirements for the Pershing Map.

Figure 1.9: Counties with Unbuilt Segments of the Proposed Military Plan

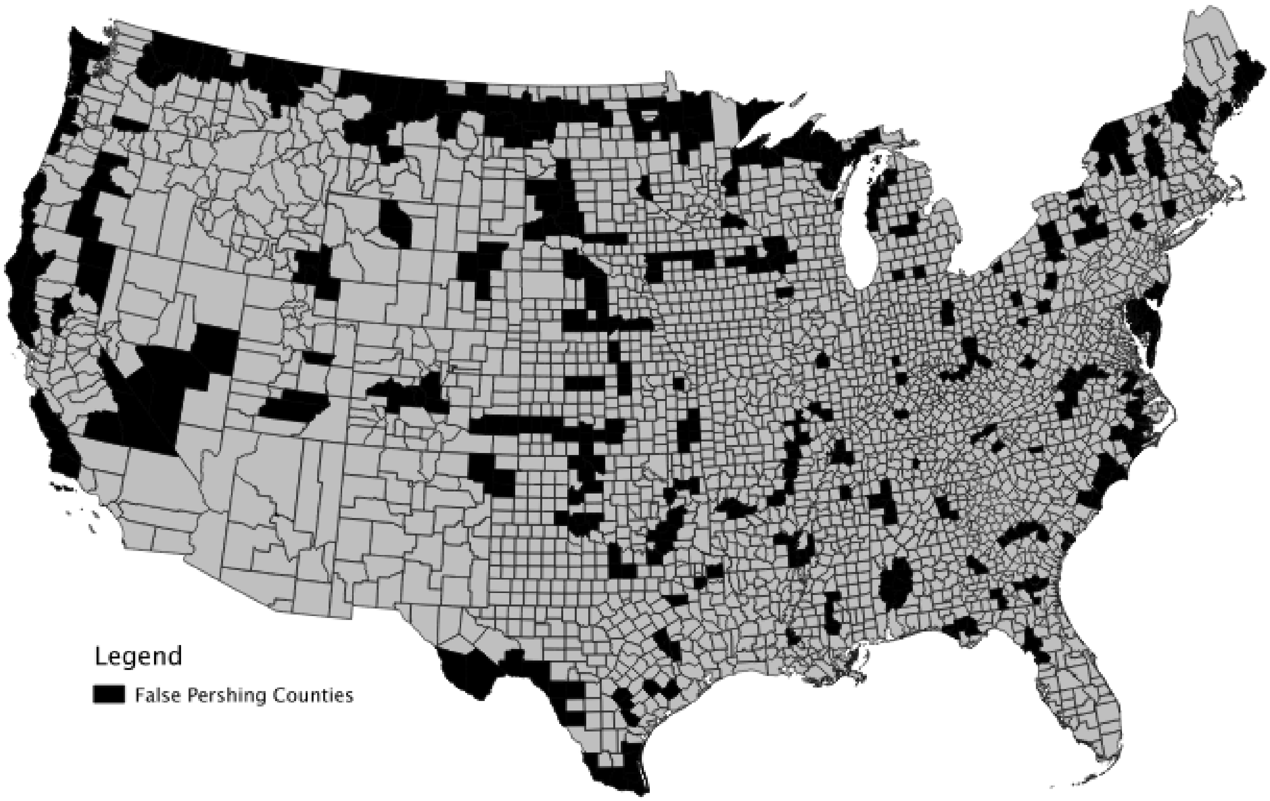


Table 1.10:

Table 10: The Effect of Unbuilt Military Routes on Total Employment and Total Establishments

	Ln(Employment) OLS	Ln(Establishments) OLS
Hwy X 1967-1971	-0.00373	0.00785
	(0.0124)	(0.00967)
Hwy X 1972-1976	-0.0103	0.00439
	(0.00849)	(0.00519)
Hwy X 1977-1981	-0.00905	-0.00421
	(0.00782)	(0.00386)
Hwy X 1982-1986	0.00290	-0.00195
	(0.00859)	(0.00554)
Hwy X 1987-1991	0.00871	-0.00785
	(0.0113)	(0.00774)
Hwy X 1992-1996	3.57e-05	-0.0148
	(0.0133)	(0.0107)

Notes: All estimates are from a 1962-1996 panel of counties that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Port, Airport. *** p<0.01, ** p<0.05, * p<0.1

1.9 Conclusions

This paper examines the causal effect of interstate highways on the geographic concentration of industry. The paper addresses two major forms of endogeneity regarding the placement and timing of highway construction by using historic government proposed national highway network plans and network theory. The bias induced by timing endogeneity is salient to the literature on other government infrastructure projects that are rolled out over time and show the need to account for the temporal variation in the allocation of fundings.

Results indicate the expansion of transportation infrastructure led to substantial employment growth in highway counties relative to non-highway counties. This employment growth was concentrated in a few industries, which led highway counties to specialize more after the expansion of interstate highways. I also find evidence that highways caused a difference in the scale of firms away from very small firms towards large firms. This paper demonstrates that expanding transportation networks are important for reshaping the spatial arrangement of economic activity.

Chapter 2

The Impact of Transportation Networks on Farm Structure and Agricultural Production

2.1 Introduction

Trucking is essential for American agriculture. According to the 2007 Commodity Flow Survey, nearly 75% of the total market value of agriculture was transported exclusively by truck (CFS 2007). The importance of trucking in longer distance shipping has grown substantially over the last 60 years. In 1955, 58% of interstate fruit and vegetable shipments moved by truck, by 2013 that number had increased to nearly 95% (USDA 1955; USDA 2013). This structural change in the transportation of agricultural products coincides with the construction of over 40,000 miles of new interstate highway infrastructure, which started construction in 1956. The Interstate Highway System (IHS) was a significant upgrade on the existing highway infrastructure in the United States, which altered the structure of transportation costs, lowering both the freight costs and time costs associated with transporting agricultural goods, and increased the market access for farmers.

This paper explores the causal link between the upgrade in transportation infrastructure in the US, brought about by the expansion of the Interstate Highway System, and two aspects of American agriculture: agricultural property values and specialization in agricultural production. I elaborate on the mechanisms behind changes in farm property values by evaluating the impact of the IHS on the market value of agricultural production and changes to the amount of land devoted to farming. To assess changes to agricultural production portfolios, I measure whether farming near interstate highways is more specialized than farming further away from interstate highways. I measure specialization using a Herfindahl Index constructed from share of farmland devoted to producing 65 crops. This paper focuses on long-run changes to agriculture by examining the evolution of farm behavior from 1950 to 2002.

I estimate a fixed-effects specification to determine the effect of the Interstate Highway System on farm property values and specialization in agricultural production. Because areas that receive interstate highways are likely to differ along unobservable dimensions from areas that do not receive interstates, measuring differences between areas that receive interstate highways and those that do not will likely result in biased estimates. To address the endogeneity concern regarding the non-random assignment of interstate highways, I use a proposed military route plan from the early 1920s to instrument for eventual highway location. This plan was designed to connect high priority depots, industrial centers, and allow for improved military resource mobility (Swift 2011), making it less susceptible to the economic and political influence of actual highway construction. A second potential source of endogeneity is the timing of highway construction. State governments determined the timing of highway construction and likely prioritized segments for unobservable reasons. To address the endogeneity surrounding the timing of highway construction, I use the Girvan-Newman (Girvan and Newman 2002, 2004; Newman 2001, 2004) algorithm from network theory to predict the timing of highway construction. The algorithm ranks predicted highway segments based on their importance for network connectivity and uses a social planner's problem to determine the predicted order of construction.[1] Conditional on a set of geographic and population controls, the instrumental variable allows me to determine the causal impact of the Interstate Highway System on agricultural property values and specialization in agricultural production.

Results indicate that during the period of early highway construction, interstate highways improved agricultural property values. However, these benefits erode during the 1980s and by the early 1990s, the value of agricultural land and buildings per acre is significantly lower in highway counties compared to non-highway counties. To better understand the mechanisms behind these results, I examine the effect of the IHS on the market value of agricultural production per acre and

[1] This methods was previously applied by Frye (2015b).

the amount of land devoted to farming. It appears most of the difference in property value is being driven by a decline in the market value of agricultural production. The difference between the OLS and IV estimates indicates that highway planners chose locations with the best opportunities for farming and the worst opportunities for industrial employment growth. The results also suggest that state officials prioritized high potential industrial locations and not high potential farming locations. Empirical estimates of the effect of the IHS on specialization in agricultural production indicate that interstate highways promoted very little specialization prior to 1980 and may have promoted more diversity in production by the mid-1990s. Taken together these results do not suggest that the IHS induced farmers to specialize more in their production.

To better understand spatial differences in how interstate highways impact the agricultural sector, I extend my empirical analysis to exploit variation in distance from the interstate highway. Results indicate that farmers within 60 km of an interstate highway behave differently from those farther than 60 km from an interstate highway. The results indicate that farmers within 60 km of a highway have lower property values, which is largely driven by producing lower value crops, and having larger farms.

This paper contributes to the broader literature on the effects of transportation networks on the spatial distribution of economic development.[2] The majority of this literature focuses on industrial development in urban areas and suburban growth (Baum Snow 2007, 2014; Baum-Snow *et al.* 2014; Rothenberg 2013; Duranton and Turner 2012). Another strand of literature focuses on the trade implications of the Interstate Highways System. Duranton, Morrow, and Turner (2013) and Michaels (2008) look for evidence of specialization in manufacturing following the introduction of the IHS. Duranton, Morrow, and Turner examine shipments of manufactured goods between major cities and find that cities with more highway mileage specialize in producing heavier goods, but there was no difference in the value of the manufactured goods. Michaels finds evidence that earnings in rural American counties grow more in trade related activities following the introduction of the IHS. These papers highlight the importance of the IHS for the spatial distribution of economic

[2] Redding and Turner (2014) provide a comprehensive survey of this literature.

activity and the patterns in regional specialization. My paper contributes to our understanding of how the agricultural sector fits into these patterns of regional specialization.

Within the American agricultural sector, Chandra and Thompson (2000) find that earnings in rural highway counties fell dramatically after the introduction of the IHS. They are not able to distinguish whether their result is due to to the relocation of agriculture to other areas or the decline of the agricultural sector. Frye (2015b) looks at employment growth and specialization across the United States. He finds employment in agriculture increased in highway counties following the introduction of the IHS. One goal of this paper is to elaborate on these findings and provide a more comprehensive assessment of the benefits of interstate highways to the agricultural sector and provide some insight on the different production decisions interstate highways induce farmers to make.

There is a well documented positive relationship between infrastructure and agricultural productivity in developing countries (Fan and Zhang 2004; Gollin et al. 2010; Mamatzakis 2003; Teruel and Kuroda 2005). The literature is thinner regarding impacts of infrastructure on individual farm decisions. Rao et al. (2006) show that road density is correlated with the production location of perishable, high value crops in India. Changes to transportation costs also directly impact land use. Weinhold and Reis (2008) finds that lower transportation costs influenced land clearing behavior in the Brazilian Amazon. While the structure of farming and the preexisting transportation infrastructure in the United States is different than in many developing countries, this paper contributes to a more general understanding of how falling trade costs and market integration can induce farmers into more specialized agricultural production.

More specifically, I contribute to the literature in three ways. This is the first paper to estimate the causal effect of the Interstate Highway System on farm property values and agricultural production portfolios. By examining the specialization behavior induced by the IHS, I am able to speak to a larger literature on falling transportation costs and regional specialization. Second, this paper makes an empirical contribution by using distance from an interstate highway as a measure of highway treatment. This allows me to understand the relative importance that distance from a highway plays in value of farmland and the production decisions of farmers. Finally, this paper extends our understanding of the political economy decisions underlying the construction of the Interstate Highway System. By analyzing the two sources of endogeneity and placing them in the context of the existing literature, I am able to elaborate more on the activity level and importance of different industries in the planning and construction of the IHS.

The paper proceeds as follows. Section 2.2 gives a brief history of the relationship between agriculture and highways in America, highlighting the potential confounding role of politicians and agricultural lobbyists in the design and construction of the IHS. Section 2.3 describes the data used in the empirical analysis and describes the general differences between highway and non-highway locations. Section 2.4 introduces the empirical strategy, formalizes the endogeneity concerns, and describes the instrumental variable used to overcome these issues. Section 2.5 examines the effect of the Interstate Highway System on farm property values and agricultural production portfolios. Section 2.6 introduces a second distance based empirical strategy for analyzing the effect of highways and discusses the results of this distance based measure. Section 2.7 shows that the major findings of the paper are robust to exclusion of major urban centers and Section 2.8 concludes.

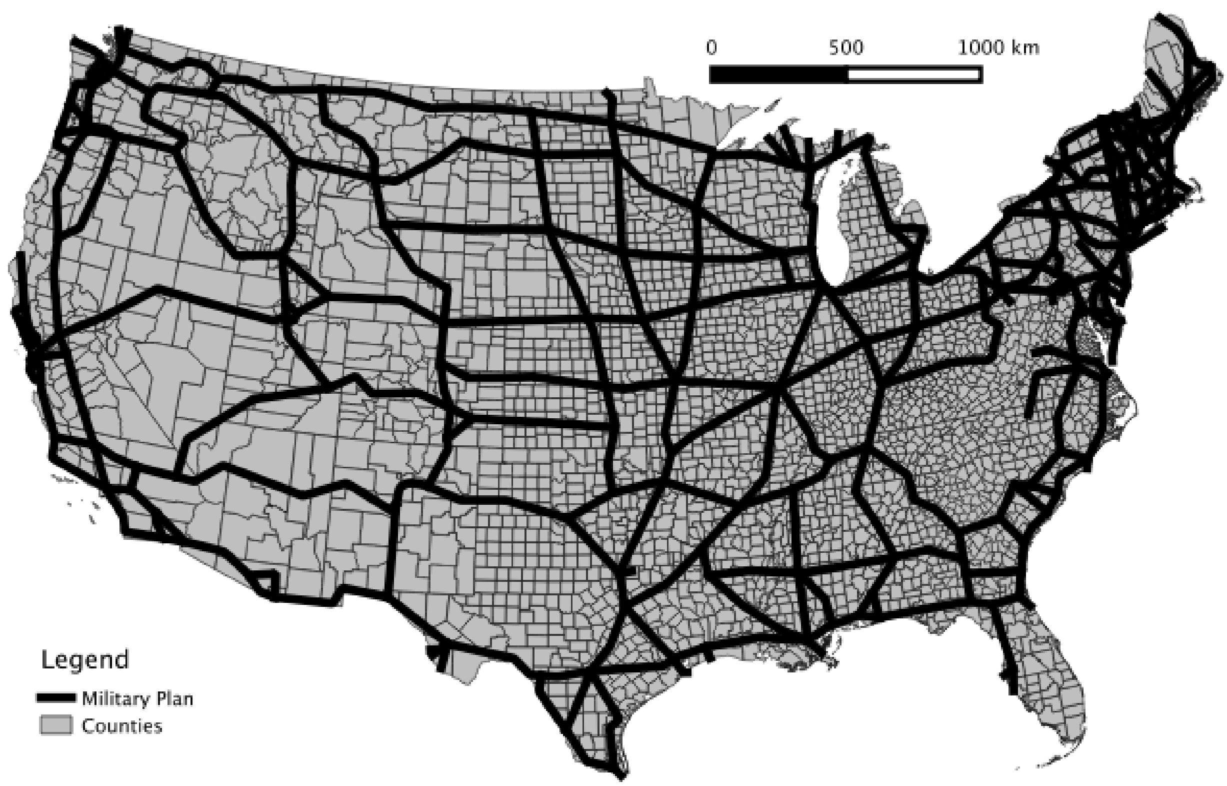
2.2 Highways and American Agriculture

Highway infrastructure became fundamentally important to American agriculture following improvements in the quality of trucks and refrigeration technology. Shipping agriculture by truck over short distances grew in popularity in the 1920s, particularly for perishable crops and dairy products. Shipping over longer distances was done primarily by rail and by water. Prior to World War I, a national system of limited access interstate highways was largely ignored by the Bureau of Public Roads because coast-to-coast travel was exceedingly rare (Karnes 2009). Considering most travel was local and regional, early transportation networks were largely developed by state and local governments, and private venture (Kaszynski 2000).

Following the First World War the U.S. government discussed the merits of a national system, similar to what it saw 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 map contained nearly 78,000 miles of highway deemed as strategically important. The army did not value a "transcontinental road which merely crosses the continent", but rather they wanted "roads connecting all our important depots, mobilization and industrial centers" (Swift, 76, 2011). These intentions are most noticeable in southern states, which were largely ignored because the military determined that a southern invasion was highly unlikely (Swift 2011).[3] Figure 2.1 depicts the set high priority routes proposed in the Pershing Map. These routes were never built as superhighways but this map influenced future highway location decisions.

[3] In fact southern Florida was not given any highways because it was to swampy for an invasion.

Figure 2.1: Pershing Military Plan

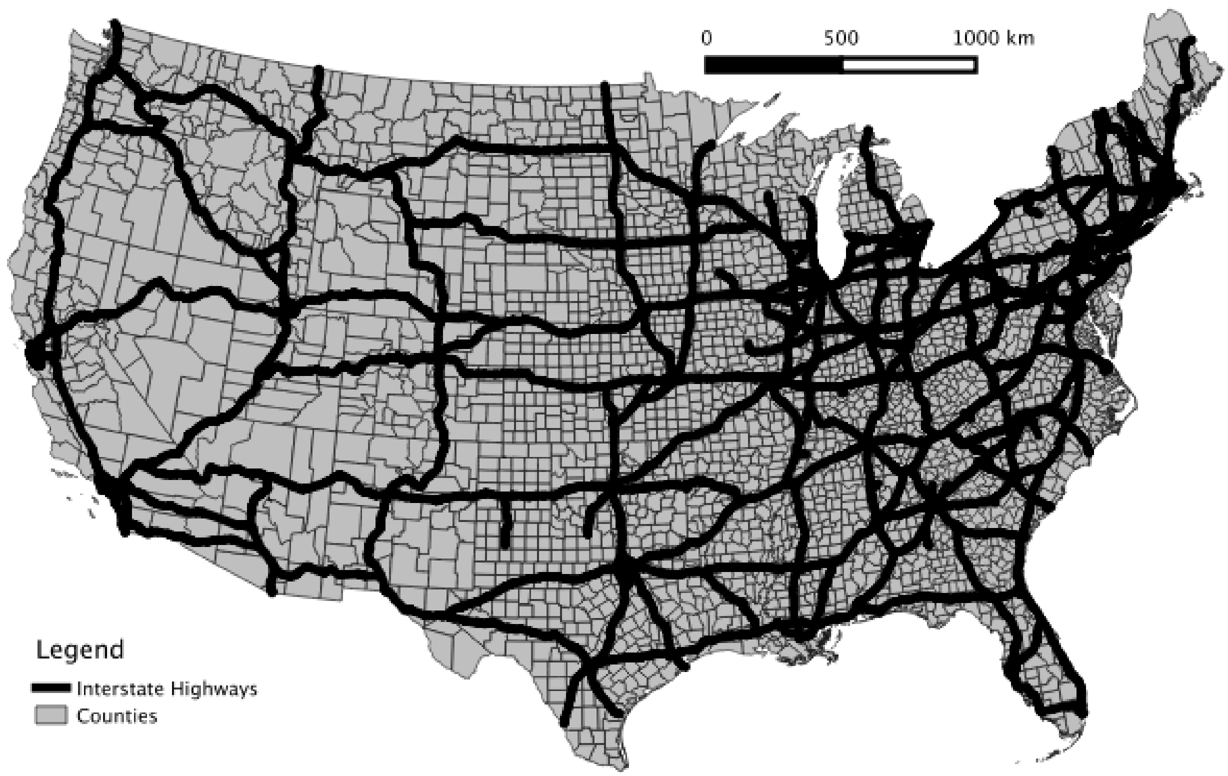


Construction plans for any new highway routes were halted during the WWII and highway funding was restricted to high priority routes. Without adequate funding for repairs the quality of highway infrastructure deteriorated rapidly. Prior to World War II total road spending was about 1.4 percent of GNP and after the war this amount fell to about 0.2 percent (Karnes 2009). As the quality of roads decreased the demand 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 that 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). Congestion and waiting in traffic also contributed to higher incidences of food spoilage (Rose 1990). The combination of congested roads and deteriorating road conditions inhibited growth in interstate shipping of agricultural products by truck.

Several Federal Highway Acts were introduced to combat the rising tide of congestion and accidents but the funding allocations were barely able to cover highway repairs. 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 some of these committees and invited Governors and heads of interest groups to participate in the planning process (Rose 1990). Industry representatives from oil, trucking, farming, and manufacturing were particularly influential in these discussions (Kaszynski 2000).

In 1956, after several different plans, construction guidelines, and financing methods were introduced, the House and Senate ultimately agreed on a interstate highway plan. The plan was approximately 90 percent Federally funded and was paid for with a mixture taxes (Kaszynski 2000). Eisenhower signed the Federal-Aid Highway Act of 1956 into law on June 29th. The final design, as presented in Figure 2.2, was "a culmination of decades of input and research from auto clubs, civil engineers, and state and federal highway officials" (Kaszynski, 167, 2000). The implementation of the Federal-Aid Highway Act of 1956 was left to states. Each state was in charge of construction, which allowed each state control over when they would build the predetermined routes within their borders. This introduces two potentially important confounding factors for estimating the effects of interstate highways, how locations were determined and how states elected to allocate the funding for highway construction.

Figure 2.2: National System of Interstate and Defense Highways



2.3 Data and Summary Statistics

My empirical analysis uses a county-level panel dataset of the contiguous United States that spans from 1950 to 2002.[4] The outcomes of interest rely on aggregate county-level farm information from the Agricultural Census. I combine these data with interstate highway location information, which allows me to examine the relationship between the Interstate Highways System and farm property values and agricultural production specialization.

[4] My final sample contains over 3000 counties with adjusted county borders to match their 1980 locations.

2.3.1 Agricultural Census Data

This paper uses county-level aggregate data on farming and agricultural production from 1950 to 2002. These data were published approximately every five years in the County Data Books and the Census of Agriculture compiled by the U.S. Census Bureau (Haines 2010). The data contain information on farm organization and production, including the average value per acre of land and buildings, the market value per acre of farm production, the amount of acreage devoted to farming, and the land devoted to production for several crops.[5]

2.3.2 Agricultural Specialization

I use data describing the amount of land devoted to production of several crops to construct a measure of concentration to better understand the relationship between highways and agricultural specialization. I measure specialization using a standard Herfindahl Index:

$$H_{ct} = \sum_i {s_{cit}}^2 \tag{2.1}$$

For each county c in year t, equation 2.1 sums the squared share of production devoted to crop i. If production is fully concentrated in a single crop, then $H_{ct} = 1$, and the index decreases as production becomes more diverse. From 1950 to 2002, the Agricultural Censuses detail the amount of farmland devoted to 65 crops.[6] The share of production, s_{cit}, is the share of land devoted to each crop i, in county c, in year t.

2.3.3 Interstate Highway System

2.3.3.aData Sources I combine information on the location of the Interstate Highway System and the timing of highway construction to construct a county-level panel dataset that

[5] Both the average value per acre of land and buildings and the market value per acre of farm production are inflation adjusted by the Consumer Price Index into 2000 dollars.

[6] A detailed list of the crops used, see Table A1 in the Appendix.

indicates the presence of an interstate highway within a county and the distance to the closest interstate highway. Interstate highway location information is from NationalAtlas.gov (2014). I combine the highway location information with highway construction information from PR-511 collection at the National Archives, which allow me to know the year each segment of interstate highway was completed. Figure 2.2 shows the current IHS locations.

2.3.3.bMeasuring Highway Treatment I measure the effect of a county having an interstate highway in two ways. First, I use an indicator variable for whether or not an interstate highway passes through the county in the given year. Second, I create a distance based measure that uses the distance from the county seat of a given county to the nearest segment of constructed interstate highway. Using these two measures of treatment separately, allows me to understand the dynamics of how regions differently.

2.3.4 Summary Statistics

To foreshadow the empirical design, Panel A and Panel B of Table 2.1 present summary statistics by highway status. Panel A presents the set outcomes in 1950. Prior to highway construction, counties that will eventually receive highways have more valuable farmland, produce products with a higher market value per acre, and have more land devoted to farming. Highway counties are also slightly more diversified in their mixture of crop production, as the Herfindahl Index is slightly smaller in highway counties. Panel B presents the set of outcomes in 2002, the difference in the value of land per acre and the market value of crops produced per acre had grow between highway and non-highway counties. This growth occurred despite land leaving agricultural production at a faster rate in highway counties compared to non-highway counties. Both highway and non-highway counties appear to be slightly more specialized by 2002.

Table 2.1:

Table 1: Summary Statistics of Agricultural Outcomes, Geographic and Population Controls, and Measures of Highway Treatment

	Non-Highway Counties			Highway Counties				
	Mean	SD	N	Mean	SD	N	Difference in Means	T-Stat
Panel A: Outcomes in 1950								
Value of Land and Property Per Acre	419.54	44.66	1714	787.49	50.87	1321	367.95	5.44
Market Value of Goods Sold Per Acre	1214.61	431.64	1714	3030.01	491.67	1321	1815.40	2.77
Total Acreage Devoted to Farming	362874.90	10214.68	1715	401064.50	11634.32	1322	38189.59	2.47
Herfindahl Index for Land Use	0.39	0.00	1715	0.36	0.00	1322	-0.03	-5.31
Panel B: Outcomes in 2002								
Value of Land and Property Per Acre	1626.01	84.15	1711	2939.23	95.80	1320	1313.22	10.30
Market Value of Goods Sold Per Acre	2780.81	487.17	1691	5041.64	555.41	1301	2260.84	3.06
Total Acreage Devoted to Farming	304841.20	8973.33	1698	299067.50	10235.70	1305	-5773.69	-0.42
Herfindahl Index for Land Use	0.41	0.00	1715	0.38	0.01	1322	-0.03	-3.40
Panel C: Geographic and Population Controls								
Distance to Closest MSA (km)	142.21	2.71	1715	97.20	3.08	1322	-45.01	-10.97
Railroad (0/1)	0.43	0.01	1715	0.77	0.01	1322	0.34	19.96
Distance to Closest Airport (km)	60.23	0.70	1715	42.70	0.80	1322	-17.53	-16.54
Distance to Closest Port (km)	364.58	6.89	1715	296.52	7.85	1322	-68.06	-6.52
Distance to Closest Military Base (km)	106.74	1.59	1715	81.01	1.81	1322	-25.74	-10.71
Population in 1950	18375.63	3713.10	1715	81849.14	4229.15	1322	63473.52	11.28
Population in 1920	16698.38	2355.60	1715	53562.84	2682.98	1322	36864.46	10.33
Population Growth Between 1940 and 1950	-1.06	0.57	1715	11.86	0.65	1322	12.92	15.02
Panel D: Measures of Highway Treatment								
Interstate Highway System	0.00	0.00	1715	1.00	0.00	1322	1.00	.
Pershing Map	0.18	0.01	1715	0.32	0.01	1322	0.15	9.53
Distance to Closest Interstate Highway	61.60	0.75	1715	8.84	0.85	1322	-52.76	-46.63
Distance to Closest Pershing Map Route	57.13	1.09	1715	31.97	1.24	1322	-25.16	-15.28

Notes: Data comes from the 1950-2002 Agricultural Censuses, Decennial Censuses, and County Data Books. All distances are calculated from the county centroid to the centroid of the nearest Metropolitan Statistical Area, Military Base, Port, and Airport. *** p<0.01, ** p<0.05, * p<0.1

2.4 Empirical Strategy: Binary Highway Treatment

2.4.1 Ordinary Least Squares

To measure the effect of the Interstate Highway System on farm property values and production specialization, I exploit variation in whether or not a county has an interstate highway across time. Similar to Frye (2015b), I estimate the following regression using a county level panel dataset:

$$Y_{ct} = \sum_t \beta_t (hwy_{ct} \times YearDum_t) + \delta_{rt} + \gamma_c + X' \rho_{ct} + \epsilon_{ct} \tag{2.2}$$

where Y_{ct} is the outcome of interest in county c in year t. The variable hwy_{ct} is an indicator variable that is equal to one if an interstate highway exists in county c at time t. The coefficients of interest are the set of β_t's, which measure the effect of the Interstate Highway System at the different years in the data.[7] I include census region $\times$ year fixed-effects, δ_{rt}, county-fixed effects, γ_c, additional controls, $X'\rho_{ct}$, and ϵ_{cit} is the error term. The controls include alternative methods of transportation infrastructure, 1950 population, 1940 to 1950 population growth, distance to the closest military base, and distance to closest Metropolitan Statistical Area (MSA). Standard errors are two-way clustered by county and state/year to account for serial correlation and spatial correlation in the error term.

This specification identifies the treatment effect of a county being connected to the Interstate Highway System. I weight each regression by the number of farms in 1950, so the coefficients can be interpreted as the average treatment effect for a farm. The model includes county fixed-effects to account for any time-invariant county characteristics, as well as census region $\times$ year fixed-effects to address any regional changes that are correlated with agriculture and the construction of interstate highways.

The regression includes several covariates to account for potential time varying confounding factors that may be correlated with the agricultural sector and the dispersion of the Interstate Highway System. The controls include alternative transportation infrastructure, historical population, and geographic characteristics.[8] Because these controls are time-invariant, I interact them with a year dummy variable to create a "trend" for each covariate. The alternative transportation infrastructure controls include an indicator for whether a county has a railroad, the distance from the closest major port, and the distance from the closest airport. Adding these covariates accounts for changes in the substitutability or complementarity of alternative methods of transportation infrastructure. The regression controls for the population in 1950 and population growth between 1940 and 1950. It's probable that both the level and growth of population influenced eventual high-

[7] I estimate each β_t from a separate regression where I partial out all the fixed-effects and reduce the comparison to a bivariate regression within each year t.

[8] See the Appendix for a full description of the data sources for the covariates.

way location and are likely related to the agricultural production. I also control for the Euclidean distance from the center of the county to the nearest MSA. This allows me to account for the fact that areas closer to major cities are more likely to receive interstate highways and those areas are less likely to be engaged in agriculture. Finally, I control for the distance to the nearest military base. One of the central motivations for the IHS was national defense. It is possible that highway placement were influenced by military base locations and military base locations may have some influence over local economic activity, including farming.

2.4.2 Addressing Highway Endogeneity

Despite the fixed-effects and the covariates, equation 2.2 is likely to produce biased estimates because the construction of interstate highways is likely to be correlated with an unobservable variable related to the structure of farms or production. Farming interest groups were some of the most politically active advocates involved in the highway planning process (Rose 1990). If these interest groups were able to manipulate the location or timing of highway construction it would bias the OLS estimates. To address this potential source of endogeneity, I construct an instrumental variable that address both the location and timing bias.

I address the location endogeneity concerns by using a historical military proposal for a national highway system to predict eventual highway locations. This historical plan, referred to as the Pershing Map, was created in 1921 as the first proposed national system of interstate highways.[9] Figure 2.1 shows the major routes of the Pershing Map.[10] I account for timing endogeneity by applying the same network theory approach as Frye (2015b). I use the Girvan-Newman algorithm (Girvan and Newman 2002, 2004; Newman 2001, 2004) to calculate the relative importance of each segment of the Pershing Map and determines the order that each route should be constructed. This allows me to assign a predicted construction year for each route of the Pershing Map, which results

[9] For more information on the Pershing Map see Michaels et al. (2013) and Frye (2015b).

[10] The Pershing Map divided it's routes into three priority levels. I exclude priority level three routes from the analysis because they seem to target more specialized locations. This exclusion also allows the mileage from the IHS to more closely match the mileage from the Pershing Map.

in an instrumental variable that predicts both the location of a predicted interstate highway and the year that segment should be built.[11]

2.4.3 Instrument Validity

2.4.3.aInclusion Restriction To test whether the Pershing Map with predicted construction timing sufficiently predicts whether a county will have an interstate at time t, I estimate the following first-stage regression using a Linear Probability Model:

$$hwy_{ct} = \zeta Pershing_{ct} + \delta_{rt} + \gamma_c + V'\pi_{ct} + \upsilon_{cit} \tag{2.3}$$

The variable $Pershing_{ct}$ is an indicator for whether a county c is predicted to have a highway from the proposed network in year t. I also include the covariates from the second-stage, ψ_{rt} are the census region $\times$ year fixed-effects, λ are the county fixed-effects, $V'\pi_{ct}$ are the infrastructure, population, military base, and geographic controls, and υ_{cit} is the error term.

Figure 2.3 presents the F-Statistics of the first-stage regression in equation 2.3. I test the inclusion destruction using Kleibergen-Paap F-Statistics, which adjust for clustering the standard errors (Stock and Yogo 2005). The Kleibergen-Paap F-Statistics vary between 16.2 and 47.8, which indicates that using the Pershing Map with predicted construction timing is a sufficient instrument for the locations of interstate highways and the timing of construction.

[11] For more details regarding the implementation of this network theory algorithm see Frye (2015b).

Figure 2.3: F-Statistics by Year for the Pershing Plan for Binary Highway Treatment

I estimate equation 2.2 using Two-Stage Least Squares (TSLS) using equation 2.3 as the first-stage equation.

2.4.3.bExclusion Restriction Satisfying the exclusion restriction requires that the military motives behind the location choices of the Pershing routes are orthogonal to changes in the Agricultural sector after the 1950s. The Pershing Map was designed specifically for national defense, which is evident in the emphasis in roads along coasts and borders and connections between major cities. The county fixed-effects help account for many of the geographic motives, like whether a location is along a coast or border. Including controls for distance to the nearest port also helps address any concerns about changes in the importance of ports or coasts over time.

Most of the nodes in the Pershing Map occur near major population centers. The industrial composition and demand for agricultural products in high population areas is another potentially confounding factor that is positively correlated with the likelihood of receiving an interstate highway. In order to address this issue, I include controls for 1920 population, 1950 population, and growth in population between 1940 and 1950. By controlling for 1920 population, I am able to account for the fact that high population areas were more likely to receive Pershing routes. The regressions also include controls for distance from each county centroid to the closest MSA, which also help account for the influence that major cities have on the economic environment of surrounding communities. One major advantage of the Pershing Map is that the system was designed with straight line connections between the nodes. These connections are beneficial because they remove any political manipulation to reach communities that are not on these lines.

Another potential threat to the validity of the Pershing Map is that the routes proposed in 1921 directly influence population growth or the industrial mix. Frye (2015b) shows that the Pershing Map locations did not have any influence on population growth or the industrial mix between 1910 and 1940. Conditional on the set of controls, the Pershing Map appears to satisfy the exclusion restriction. This is consistent with several recent empirical papers examining the effects of highways, which use planned transportation systems as an instrument for eventual highway locations (Baum-Snow 2007, 2010, 2014; Michaels 2008; Duranton and Turner 2012; Duranton, Morrow, and Turner 2013; Frye 2015b).

2.5 Results

2.5.1 Property Values

Assessing changes in property values induced by interstate highways provides an initial indication of whether or not highways are beneficial for the agricultural sector. Table 2.2 presents both the OLS and Two-Stage Least Squares regression results describing the effect of the Interstate Highway System on the per acre value of farmland and buildings, the market value per acre of farm production, and the amount of acreage devoted to farming. Columns (1) and (2) show the effect of the IHS on the per acre value of farmland. Starting with the period of major highway construction into the early 1980s, it appears that highways were beneficial for agricultural property values. Taken in 1978, the coefficient estimates indicate that agricultural property values were 13.3% higher on farms in highway counties compared to farms in non-highway counties. These elevated property values coincide with the speculative rise in farmland prices that characterized the 1970s. Taken in the context of these speculative land prices, it is unclear whether interstate highways actually raise the value of land or helped fuel the speculative bubble. By the mid-1990s, the value of land and buildings per acre were significantly less in highway counties relative to non-highway counties.

Table 2.2:

Table 2: The Effect of Highways on Agricultural Land Values, Market Value of Agricultural Production, and Acres of Farmland

	Ln(Land Value Per Acre)		**Ln(Market Value Sold Per Acre)**		**Ln(Farmland)**	
	OLS	TSLS	OLS	TSLS	OLS	TSLS
	(1)	(2)	(3)	(4)	(5)	(6)
Hwy X 1959	0.024	0.227**	0.051**	0.385**	-0.024	-0.224**
	(0.035)	(0.098)	(0.020)	(0.150)	(0.019)	(0.088)
Hwy X 1964	0.006	0.048	0.026**	0.115	0.003	-0.023
	(0.015)	(0.075)	(0.012)	(0.105)	(0.011)	(0.047)
Hwy X 1969	0.009	0.096*	0.025**	0.024	0.001	0.072**
	(0.011)	(0.053)	(0.010)	(0.051)	(0.006)	(0.030)
Hwy X 1974	-0.010	0.063	-0.009	0.060	-0.003	0.061***
	(0.007)	(0.041)	(0.012)	(0.048)	(0.005)	(0.021)
Hwy X 1978	-0.005	0.133*	-0.023**	-0.039	0.022***	0.083**
	(0.015)	(0.073)	(0.012)	(0.044)	(0.008)	(0.041)
Hwy X 1982	-0.011	0.100	-0.019	-0.026	0.010	0.054
	(0.016)	(0.063)	(0.014)	(0.055)	(0.009)	(0.033)
Hwy X 1987	0.013	-0.032	-0.001	-0.047	0.009	-0.018
	(0.018)	(0.069)	(0.021)	(0.107)	(0.012)	(0.051)
Hwy X 1992	0.015	-0.091	-0.041**	-0.225	0.011	-0.029
	(0.019)	(0.080)	(0.018)	(0.138)	(0.008)	(0.035)
Hwy X 1997	0.028	-0.071	-0.051***	-0.267**	0.000	-0.022
	(0.018)	(0.088)	(0.019)	(0.133)	(0.015)	(0.063)
Hwy X 2002	-0.014	-0.311**	-0.057***	-0.332**	-0.018	-0.108*
	(0.029)	(0.142)	(0.017)	(0.141)	(0.014)	(0.055)
Observations	3,035	3,035	3,035	3,035	3,035	3,035

Notes: All estimates are from a panel of counties from 1950 - 2002 that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest MSA, Port, Airport, or Military Base. *** p<0.01, ** p<0.05, * p<0.01.

To better understand the mechanisms that are driving these results, I examine the market value of agricultural products sold per acre and the amount of land devoted to farming. These results are presented in columns (3) - (6) of Table 2.2. Interstate highways impact the market value of production per acre in nearly the same way they affected property values. This is consistent with the fact that changes to agricultural production are a key component behind the fluctuations in agricultural land values. Analyzing the effect of the IHS on amount of land devoted to agricultural production reveals similar cyclical patterns. The summary statistics indicate that the amount of acreage devoted to farming has been falling over time. However, according to the regression results, during the late-1960s and 1970s relatively more land was retained in agricultural production in highway counties compared to non-highway counties. The estimates indicate that in 1978, farms in highway counties were 8% larger than farms in non-highway counties. This occurred at a time where there were higher agricultural land prices in highway areas as well. Taking these two results together it may indicate that highway counties had more agricultural opportunities, which led these counties to devote more land to agriculture than non-highway counties. However, these opportunities do not appear to be driven by increases in the value of production.

The OLS results and the instrumental variable results differ dramatically for all three outcomes. Understanding the direction and source of the bias is important for understanding the political economy behind the construction of the Interstate Highway System. Disentangling the location bias from the timing bias is important for understanding the decision making process at the federal and state level. The location bias is most apparent after all of the highways have been completed in the later periods, when the variation between the OLS and IV estimates is mostly based on variation in location. Looking at land value per acre and the market value of agricultural production after the mid-1980s, the difference between the OLS and IV estimates suggests there was a positive location bias. This implies that the IHS was designed to pass through the best farmland, which may indicate a strong influence from agricultural supporters in the design of the IHS. The location bias in the interstate highway literature consistently shows a negative location bias for employment and industry growth (Duranton and Turner 2012; Baum-Snow 2014; Frye 2015b). These two results appear to be consistent, areas with the best farmland are likely the same areas that have poor employment and industrial growth opportunities.

The OLS bias during the early periods is negative, which is composed of both the timing and location bias. Given that the location bias is positive, it implies that the timing bias is negative. This indicates that state highway planners built the worst locations for farming first, which is also consistent with the previous literature. Frye (2015b) finds that state planners targeted areas with high employment and growth potential to start highway construction. These high growth potential areas were likely more industrial and presumably worse for farming. Understanding these biases and the motivations behind them is important for understanding the political economy of transportation infrastructure projects.

2.5.2 Production Specialization

As transportation costs fall following the expansion of the Interstate Highway System, farmers may be induced to specialize according to their locational comparative advantage. In this section, I test whether this assertion was true following the introduction of interstate highways. I measure specialization in agriculture using a Herfindahl Index of the shares of farmland devoted to producing 65 crops, as described in Section 3.2. Both the OLS and TSLS regression results are presented in Table 2.3. The results show limited evidence of specialization. The TSLS coefficients are consistently positive from the late-1950s through the early 1980s, however they are not statistically distinguishable from zero. By the late 1990s it appears that highway counties are slightly more diversified than non-highway counties. Overall, these results do not seem to indicate that the Interstate Highway System induced farmers to specialize more in their production.

Table 2.3:

Table 3: The Effect of Highways on Crop Diversification

	Herfindahl Index	
	OLS	TSLS
	(1)	(2)
Hwy X 1959	0.001	0.025
	(0.010)	(0.029)
Hwy X 1964	-0.006	0.005
	(0.005)	(0.021)
Hwy X 1969	0.001	0.045
	(0.013)	(0.036)
Hwy X 1974	-0.004	0.035
	(0.014)	(0.026)
Hwy X 1978	-0.011	0.029
	(0.009)	(0.026)
Hwy X 1982	-0.002	0.013
	(0.006)	(0.022)
Hwy X 1987	-0.001	-0.024
	(0.005)	(0.024)
Hwy X 1992	0.003	-0.001
	(0.006)	(0.022)
Hwy X 1997	0.010	-0.042
	(0.006)	(0.032)
Hwy X 2002	0.011	-0.076
	(0.009)	(0.046)
Observations	3035	3035

Notes: All estimates are from a panel of counties from 1950 - 2002 that include county fixed-effects, region X year fixed-effects, and the full set of covariates. Each entry in the table comes from a separate regression. Robust standard errors are two-way clustered by both county and state/year. All distances are calculated from the county centroid to the centroid of the nearest MSA, Port, Airport, or Military Base. *** p<0.01, ** p<0.05, * p<0.01.

2.6 Distance Based Highway Treatment

2.6.1 Empirical Specification

To better understand the spatial heterogeneity in how the Interstate Highway System impacts the agricultural sector, I alter the binary measure of highway treatment to exploit how far a county is from an interstate highway. To measure this distance, I calculate the Euclidean distance from each county's county seat to the nearest interstate highway in each year. For simplicity, I split these distances into three classifications: close, medium, and far. Close counties are defined as those where the county seat is within 30 km of an interstate highway. I define medium counties as having a county seat within 30 to 60 km of an interstate highway and far counties as being farther than 60 km from the interstate highway.[12]

To measure the effect of the IHS on farm property values and production specialization, I exploit variation in the distance a county is from an interstate highway across time. I estimate a similar regression to equation 2.2:

$$Y_{ct} = \sum_t \beta_t(hwyclose_{ct} \times YearDum_t) + \sum_t \theta_t(hwymed_{ct} \times YearDum_t) + \delta_{rt} + \gamma_c + X'\rho_{ct} + \epsilon_{ct} \quad (2.4)$$

where Y_{ct} is the outcome of interest in county c in year t. The variable $hwyclose_{ct}$ is an indicator that is equal to one if an interstate highway is within 30 km of the county seat of county c at time t. The variable $hwyfar_{ct}$ is an indicator that is equal to one if an interstate highway is between 30 km and 60 km from the county seat of county t in year t. The excluded set of distances are those counties more than 60 km from an interstate highway. The coefficients of interest are the set of β_t's, which measure the effect of being within 30 km of an interstate highway at the different years I observe in the data, and the set of θ_t's, which measure the effect of being between 30 and 60 km of an interstate highway.[13] I include the same set of fixed-effects and controls as in equation 2.2.

[12] I selected these distance cutoffs based on the distribution of distances in 2002. Roughly 1/3 of counties in 2002 fall into each of these distance bins.

[13] For each year, I estimate the β_t and θ_t pair from a separate regression where I partial out all the fixed-effects and reduce the comparison to a regression within each year t.

I also two-way cluster the standard error by county and state/year to account for serial correlation and spatial correlation in the error term.

In the first-stage, I instrument for both $hwyclose_{ct}$ and $hwymed_{ct}$ using the predicted distances from the Pershing Map. Figure 2.4 shows the first-stage Kleibergen-Paap F-Statistics by year. The values range between 5.8 and 19.4, indicating that in the early years using the Pershing Map distance from a county is a slightly weaker predictor of actual highway distance from a county. According to the critical values produced by Yogo and Stock (2005), this is on the lower bound of being an acceptable instrument. To help address this weakness in the first-stage, I estimate equation 2.4 using Limited Information Maximum Likelihood (LIML).

Figure 2.4: F-Statistics by Year for the Pershing Plan for Close and Medium Distance Locations

2.6.2 Results

Figures 2.5 through 2.7 presents graphs of the β and θ coefficients of interest and their corresponding confidence intervals for the OLS and LIML instrumental variable regression results for each outcome. Figure 2.5 presents the effects of the proximity to the IHS on property values per acre. The top two graphs of Figure 2.5 show the OLS results for counties defined as close, where the county seat is less than 30 kilometers from the nearest interstate highway, and counties defined as medium distance, where the county seat is between 30 and 60 kilometers from the nearest interstate highway. Both sets of OLS regression coefficients suggest that there were not property value differences between counties close to interstate highways or a medium distance from interstate highways and counties farther than 60 km from an interstate highway. The instrumental variable regression results depict a stronger difference. Both close counties and medium counties experience similar declining property values over time, compared to farther away counties. Given how similar the coefficient estimates are for close and medium counties, it suggests that farmers notice very little difference once they are within 60 km of an interstate highway.

Figure 2.5: Effect of Highways on Property Values per Acre for Close and Medium Distance Locations

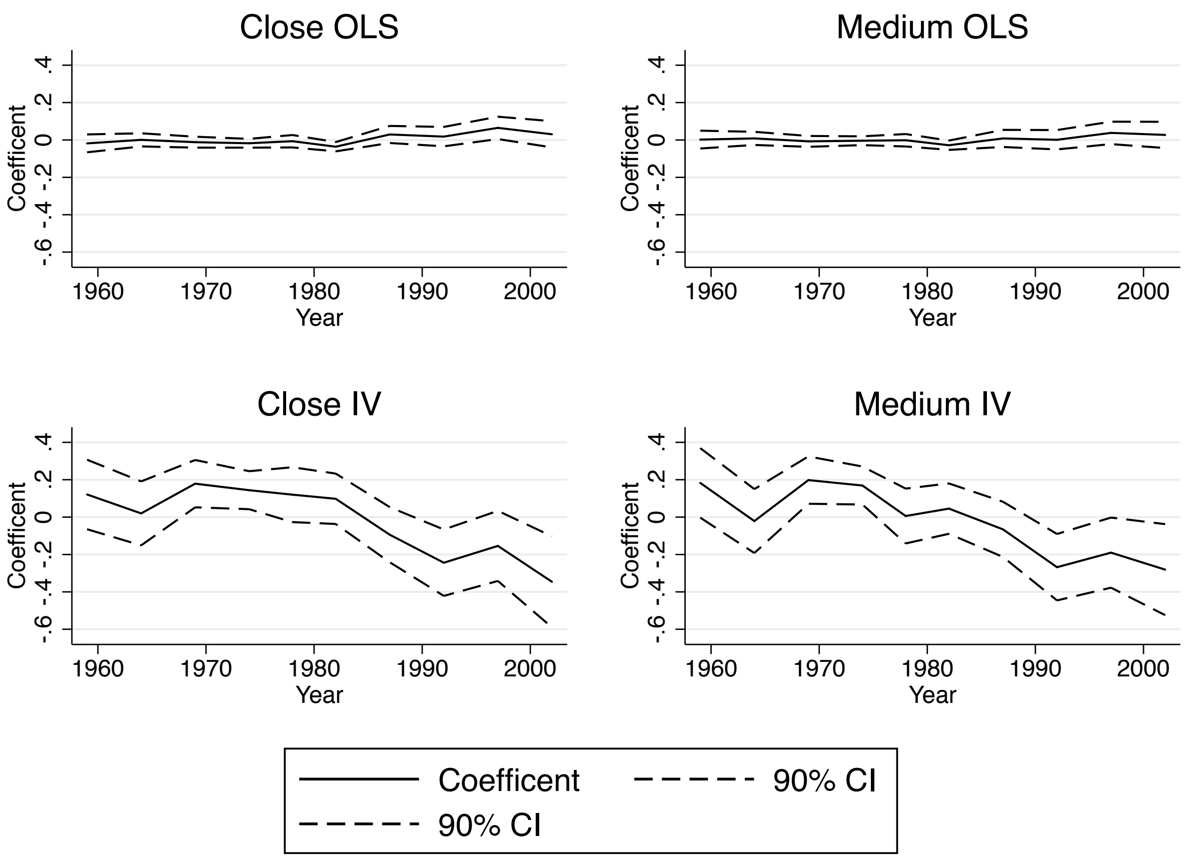


Figure 2.6 shows a similar set of graphs where the outcome of interest is the market value of production per acre. These graphs show a very similar relationship to those seen in Figure 2.5. There is very little difference visible in the OLS regressions, but a much stronger trend emerges in the instrumental variable specification. These results indicate that both close counties and medium counties saw declining market values of agricultural products sold per acre relative to farther away counties. It appears that close counties have a slightly lower market value than medium counties, although this difference is not statistically significant. Figure 2.7, which examines the effect of the IHS on the amount of farmland, has a different pattern over time. According to these results, both close and medium distance counties have slightly more land devoted to farming. These

graphs indicate that medium counties have a slightly higher percentage of land being retained in farming. Considering these results together, it appears that distance from the interstate is important. Counties farther than 60 km from the interstate have higher value agricultural land, produce more high value crops, and have less land devoted to farming than areas closer to the interstate highway.

Figure 2.6: Effect of Highways on Market Value of Production per Acre for Close and Medium Distance Locations

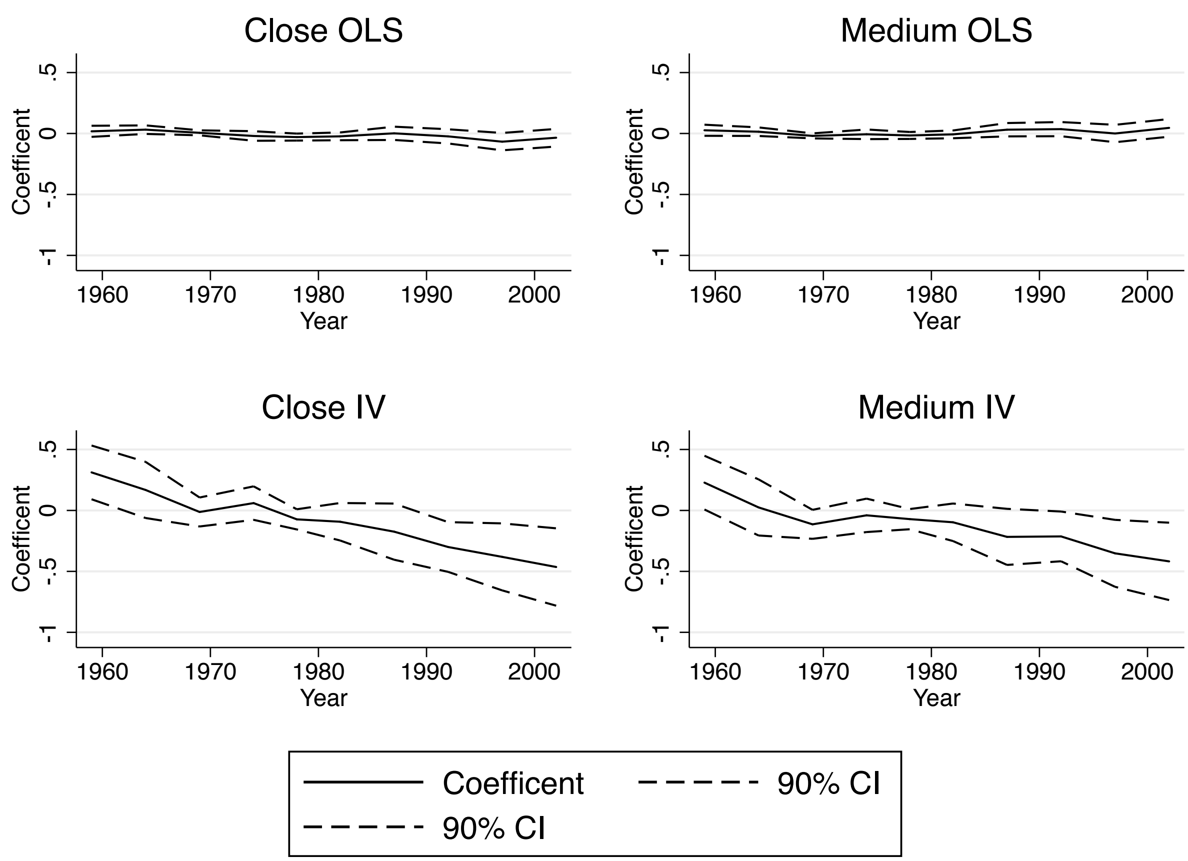
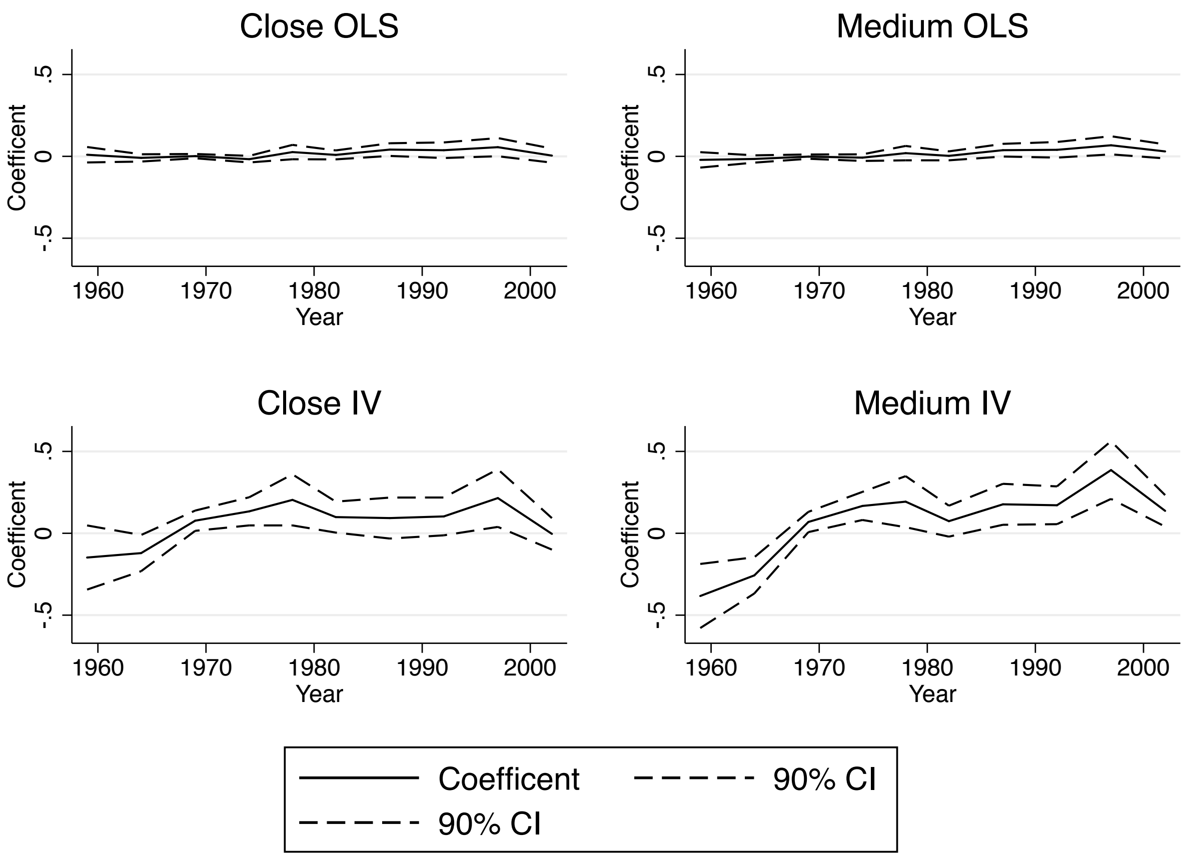


Figure 2.7: Effect of Highways on Acres of Farmland for Close and Medium Distance Locations



2.7 Robustness

2.7.1 Excluding Major Urban Centers

As a robustness check, I verify that the major results are not being driven by agricultural activity in major cities. For this model, I run the regression from equation 2.2 after dropping the major urban areas. The regression results are consistent with my prior intuition that major urban counties are not driving the results.[14]

[14] Regression results are available upon request.

2.8 Conclusion

This paper examines the causal effects of the Interstate Highway System on farm property values and agricultural production portfolios. The paper addresses endogeneity concerns for both the location of interstate highways and the timing of highway construction using an instrumental variable based on a proposed military plan and network theory. The paper also incorporates distance based measures of highway treatment to better understand spatial differences in how interstate highways impact the agricultural sector. The results indicate that highway counties have slightly lower property values, mostly due to a declining value of agricultural products sold.

Additionally, I find no evidence that crop production in highway counties is more specialized than in non-highway counties. These findings contradict the literature examining the impacts of road infrastructure in developing countries. This suggests the need for a more comprehensive theory connecting transportation networks and agricultural specialization that can incorporate the differences in the agricultural sectors of economies of different sizes and at different stages of development.

Chapter 3

The Indian Reorganization Act, Tribal Sovereignty, and Economic Development

3.1 Introduction

Over the last 30 years, decentralized governance has increased in popularity across developing countries. More recently, World Bank support prompted formally centralized economies to experiment with decentralizing fiscal responsibilities, administration, and the delivery of services (World Bank 2000, 2001). These changes have produced mixed results (Bardhan 2002; Thornton 2007; Zhang and Zou 1998; Lin and Liu 2000; Akai and Sakata 2002). Econometric estimation issues, particularly concerning endogeneity, are one major challenge that has limited our understanding of the economic effects of decentralization. To address these limitations, this paper exploits the decentralization of governance across American Indian reservations and measures the long-run development differences for reservations that were granted more sovereignty.

In 1934 the United States government passed one of the most important pieces of legislation governing American Indian reservations, the Indian Reorganization Act (IRA). Adoption of the IRA was voluntary and each reservation had 18 months to vote on whether or not to adopt the IRA. If adopted, IRA reservations were subject to more administrative oversight from the Secretary of Interior and the Bureau of Indian Affairs (BIA) (Clow 1987; Philp 1999). These additional constraints limited the sovereignty of tribes (Legters and Lyden 1994). Tribes that did not adopt the IRA maintained their own tribal governments and constitutions, free from BIA oversight. This created two types of decentralized tribal governments across American Indian reservations. This paper empirically measures the impact of these two different types of decentralization on current reservation economic conditions by comparing IRA and non-IRA reservations.

Comparing contemporary economic outcomes of adopters and non-adopters of the IRA is problematic because tribes may have adopted the IRA for several reasons that may be correlated with contemporary economic development, resulting in biased empirical estimates. In order to mitigate these selection concerns, I exploit IRA voting results from the mid-1930s by restricting the sample to tribes that held narrowly determined IRA elections. Presumably, the decision to vote for or against the IRA by a small fraction of voters should influence current economic conditions only through the tribal adoption of the IRA, thus providing plausibly exogenous variation in the initial adoption of the IRA.

My empirical specification exploits the narrow IRA voting results in a regression discontinuity (RD) framework to estimate the effect of the IRA on the outcomes of interest. Regression results using reservation-level data from the 1990 U.S. Census indicate early adoption of the IRA stifled economic development among reservations that held close IRA elections. Per capita income reservation income is over 40 percent lower among IRA reservations on average. Similarly, the fraction of the population receiving income from public assistance was over 55 percent higher among reservations that adopted the IRA. Lower education levels among IRA reservations are one source of the income disparity. The fraction of college-educated individuals on IRA reservations is nearly 35 percent lower, suggesting either lower educational attainment or high skill migration. Another difference between IRA and non-IRA adopters is the level of racial integration. IRA reservations have a significantly higher proportion of the population identifying as Native American. The combination of educational differences and the disparity in racial integration explain a large fraction of the income differential between IRA and non-IRA reservations.

The primary mechanism driving these results is the administrative oversight imposed by the Bureau of Indian Affairs. A series of federal laws reduced the severity of BIA oversight in the late 1980s and early 1990s (Legters and Lyden 1994). If BIA oversight imposed by the IRA significantly slowed development than this reduction in administrative oversight should have been more beneficial for IRA reservations. Results examining differences in reservation income in 2000 and 2010 support this assertion, implying that BIA oversight is partially responsible for the differences in economic development across reservations. These findings indicate that the weak decentralization received by IRA reservations limited economic growth. Decentralization has been shown to be particularly beneficial in cases where there are informational or political constraints (Bardhan 2002; Oates 1999). This may be particularly true on American Indian reservations, where cultural variation introduces informational barriers.

This paper makes two major contributions to the literature on decentralization and economic growth. First, the decentralization programs of the last few decades have only produced short-term results. This environment provides the opportunity to understand the long-run impacts of the decentralization. Second, many of the empirical decentralization studies, particularly cross-country studies, struggle to overcome endogeneity biases. The unique process that allowed tribes to vote on the adoption of the Indian Reorganization Act allows me to identify a causal effect of decentralizing BIA authority to tribal governments.

My work contributes to a growing literature highlighting the importance of institutional quality on American Indian reservations for development. The empirical findings in this literature have primarily focused on the economic impacts of property rights and jurisdictional quality.[1] The role of sovereignty in development on reservations is less well developed. The structure of tribal constitutions has been show to be important for development (Akee et al. 2012). However, this literature has not yet addressed the role that formative decentralization played in reservation development.

This paper also contributes to the literature analyzing the long-term impacts of colonial institutions on long-run economic development. The formal and informal institutions instituted by colonizing countries have been shown to have persistent economic effects.[2] My paper is most closely related to Dippel's (2014) work, which finds that indigenous bands that were forced to

[1] (Akee 2009; Anderson and Lueck 1992; Anderson and Parker 2008; Parker ; Cookson 2010; Cookson 2014)
[2] (Acemoglu et al. 2001; Nunn 2008, 2009, 2010, 2011; Dell 2010, 2012)

share a reservation in the late 19th century have substantially lower contemporary incomes. My paper examines a similarly important period in the formation of contemporary American Indian reservations and finds that differences in decentralization across reservations created large income differences that persisted over time.

The paper is organized as follows. Section 3.2 describes the history of the Indian Reorganization Act and elaborates decentralization of power from the BIA to individual tribes. Section 3.3 describes the data and controls used to quantify the effect of the IRA and discusses the preliminary differences that exist across reservations. Section 3.4 elaborates on the selection concerns associated with comparing IRA to non-IRA reservations and introduces the IRA voting records as a solution for overcoming the selection bias. Section 3.5 introduces the formal empirical strategy and discusses the results. Section 3.6 shows how limited decentralization, in the form of federal oversight, is the major source behind the economic differences between IRA and non-IRA reservations. Section 3.7 verifies the robustness of the estimation strategy and Section 3.8 concludes.

3.2 The Indian Reorganization Act and Tribal Government

The Indian Reorganization Act represented a dramatic change in federal Indian policy. In the early 1930s, at the urge of the new Commissioner of Indian Affairs, the IRA proposed restoring tribal self-governance marking a severe departure from the assimilationist policies that had dominated for nearly a century. The IRA ended the allotment of tribal lands, placing allotted and tribal lands in federal trust.[3] It established the authority of the Secretary of the Interior over matters of tribal lands and natural resources and established a fund that allowed tribes to restore their reservation land base. The IRA also established a revolving credit account for tribal governments and corporations in an effort to increase the availability of credit (Carlson 1981). Congress passed the IRA, also known as the Howard-Wheeler Act, on June 18, 1934.

Within 18 months of the IRA passing Congress, each tribe voted on whether or not to adopt

[3] For more information regarding land tenure on Indian reservation see Anderson and Luck (1992), Anderson and Parker (2009), and Frye (2012).

the IRA. Each reservation that adopted the IRA was required to form a new tribal constitution or charter, although in practice some did not. These constitutions were reviewed and amended by the BIA. In many instances the resulting IRA constitution imposed a model of tribal governance based on a corporate structure that differed from many of the traditional tribal democratic systems (Rusco 2000). IRA reservations were subject to more administrative oversight from the Secretary of Interior and the BIA (Clow 1987; Philp 1999). This administrative oversight occurred in several ways. First, any transactions involving land and natural resources or state and local governments required the approval of the Secretary of Interior. Also any tribal or corporate projects using the revolving credit funds were subject to close supervision from local bureau officials assigned to monitor the funds and minimize loses (Mekeel 1944). Given these administrative barriers several historians have described the IRA as granting tribes limited sovereignty (Legters and Lyden 1994) and claiming that IRA reservations were still under the federal government despite the promise of self-rule (Philp 1999). Lemont (2006) claimed that it was not until the early self-determination acts of the mid-1970s that IRA tribes had authority over their own reservations.

Tribes not electing to adopt the IRA maintained their own tribal governments and constitutions and were not subject to the same set of federal restrictions. Therefore, the decentralization of governance from the BIA to tribal governments occurred more dramatically for non-IRA reservations. This historical setting provides a unique opportunity to compare the long-run effects of two different types of decentralization on current economic conditions.

3.3 Data and Preliminary Evidence

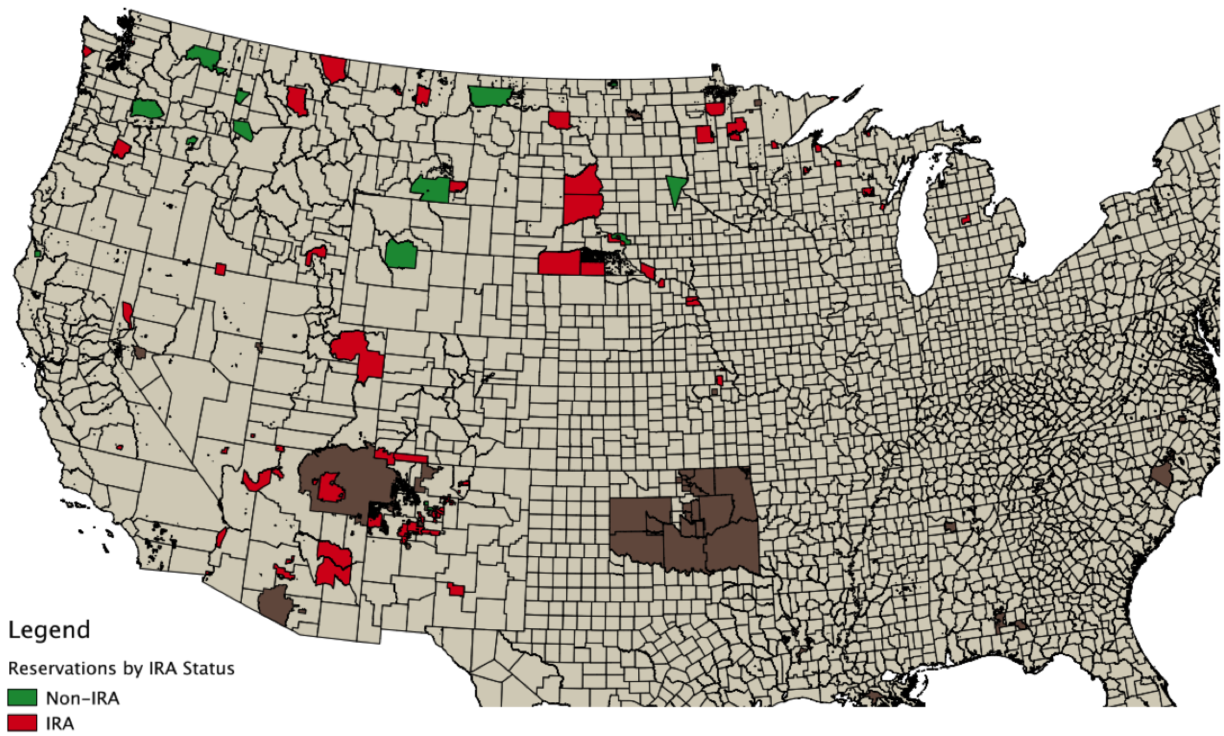
3.3.1 Measures of Reservation Development and Controls for Differences Across Reservations

To measure the long-run effects of the IRA, I use several data sources to create a reservation-level dataset that includes contemporary outcomes, historic census data, and spatial controls. Contemporary reservation level census data is from the National Historic Geographic Information System for 1990 (NHGIS 2011). I measure economic well-being using reservation-level per capita income. As secondary measures of economic well-being I also analyze the share of households receiving public assistance, and the share of individuals older than 25 that completed college or entered high school.

To address potential confounding factors that may be correlated with contemporary income and IRA status, my empirical analysis will include several historical and spatial controls. To help account for pre-IRA differences between IRA and non-IRA reservations I include historical census records from the 1% sample of the 1910 U.S. Census, which includes an oversample of Native Americans. This oversample includes 20 percent of the Native American population, which I aggregate to the reservation level based on household location (Haines 2010). These historical records include basic demographic information, literacy rates, labor force status, and occupational scores. Given that the IRA was introduced following the Allotment Era, I include allotment and land tenure characteristics from Indian Land Tenure, Economic Status, and Population Trends (OIA 1934). I also include differences in land quality, availability of natural resources, and proximity to urban centers to address any resource or market based differences between IRA and non-IRA reservations. Land quality data are from the FAO GAEZ, natural resources and urban location data is from National Atlas (FAO 2015; National Atlas 2014).

To create the final sample I chose to drop current reservations with less than 150 people. These reservations are so small that it is unclear whether or not tribal governments operate like larger reservations. I also drop reservations established prior to 1800. This restricts the analysis to reservations established in a relatively similar era. My final sample includes 119 current reservations, each with information regarding current economic conditions, geographic characteristics, 1910 reservation characteristics, 1934 allotment characteristics, and IRA voting records. The following map indicates the spatial distribution of IRA and non-IRA reservations. In general, IRA and non-IRA reservations are evenly distributed across the western states, with the exception of the southwest. As robustness, I run the analysis both with and without these southwestern reservations and do not see any substantial changes.

Figure 3.1: Map of Current Reservations by IRA Status



3.3.2 Summary Statistics

Table 3.1 presents summary statistics for the full sample of reservations. The table indicates the mean, standard error, and number of observations by IRA status. The final three columns present the difference between IRA and non-IRA reservations and tests whether or not there is a statistical difference between the two groups. Several of the outcomes are different between IRA and non-IRA reservations. IRA reservations have lower incomes, higher proportion of Indians, and lower housing values. Several of the geographic controls exhibit statistical differences as well. IRA reservations are much closer to coal deposits and have poorer surrounding economic environments.

Among the historical controls there appear to be differences in education, marriage and average family size, all of which are related to assimilation. IRA reservations were also less likely to have been allotted. These preexisting differences suggest that the adoption of the IRA is unlikely to be exogenous.

Table 3.1:

Summary Statistics for Full Sample

	IRA Reservations			Non-IRA Reservations			Difference		
	Mean	SE	N	Mean	SE	N	Mean Diff	SE	T-stat
Outcomes									
Per Capita Income	6757.94	375.90	88	7555.65	375.90	31	-797.70	475.28	-1.68
Median Household Income	17422.85	812.13	88	19913.26	812.13	31	-2490.41	1039.81	-2.40
Indian per Capita Income	4982.63	289.29	88	5859.61	289.29	31	-876.99	327.15	-2.68
Population	3722.67	1482.26	88	5330.77	1482.26	31	-1608.10	1564.94	-1.03
Share of Indians	68.50	5.15	88	53.49	5.15	31	15.02	5.92	2.54
Median House Value	46219.27	3716.16	88	54287.10	3716.16	31	-8067.82	4335.15	-1.86
Share Completed High School	63.39	2.62	88	62.40	2.62	31	0.99	2.91	0.34
Share Completed College	29.81	2.24	88	32.67	2.24	31	-2.86	2.65	-1.08
Geographic Characteristics									
Distance from MSA (in km)	202.56	36.21	85	177.71	36.21	31	24.85	39.33	0.63
Avg Suitability for Wheat	34.25	2.97	85	30.23	2.97	31	4.02	3.84	1.05
Distance to Coal Deposits (in km)	133.87	10.39	88	73.27	10.39	31	60.61	17.61	3.44
pc Income of Neighboring Counties	11328.02	334.30	83	12107.28	334.30	28	-779.26	417.24	-1.87
1910 Reservation Characteristics									
Share of Women	48.81	0.68	84	49.42	0.68	31	-0.61	0.86	-0.71
Fraction under 18 in school	44.22	2.93	84	52.48	2.93	31	-8.26	3.36	-2.46
Percent Literate	49.57	3.16	84	54.93	3.16	31	-5.37	3.95	-1.36
Percent in Labor Force	46.50	2.72	84	42.85	2.72	31	3.65	3.31	1.10
Percent Married	40.22	1.24	84	37.05	1.24	31	3.17	1.39	2.28
Percent Living on Farms	37.34	4.24	84	36.43	4.24	31	0.92	5.28	0.17
Avg Age	26.73	0.50	84	27.37	0.50	31	-0.64	0.66	-0.97
Avg Family Size	5.31	0.20	84	4.79	0.20	31	0.52	0.25	2.05
Percent White Blood	14.03	1.25	84	12.57	1.25	31	1.46	2.12	0.69
1934 Allotment Characteristics									
Percent Allotted	0.55	0.08	88	0.71	0.08	31	-0.16	0.10	-1.67
Avg Allotment Acreage (000s)	155.67	78.75	88	186.30	78.75	31	-30.62	90.29	-0.34
Avg Number of Allotments	951.67	277.07	88	993.42	277.07	31	-41.75	343.48	-0.12
Avg Amount of Surplus Land (000s)	89.83	75.17	88	158.34	75.17	31	-68.51	89.96	-0.76
Avg Acreage Alienated (000s)	57.80	13.60	88	41.78	13.60	31	16.02	22.16	0.72
IRA Voting Population	624.5909	83.00893	88	426.5484	83.00893	31	198.0425	120.4248	1.64

3.4 Selection Concerns and IRA Voting Records

3.4.1 Selective Adoption of the IRA

Comparing tribes that adopted the IRA to those that did not is likely to result in a biased estimate. Tribes voted to adopt the IRA for several reasons, which may be correlated with contemporary reservation development resulting in biased OLS estimates. For example, poorly organized tribes in 1934 may have adopted the IRA because of the high organization costs associated with forming their own constitution and government structure. This organizational dysfunction is likely to persist through time and decrease contemporary economic development. Therefore, poorly organized tribes would likely result in negatively biased OLS estimates.

Tribes that were more assimilated in 1934 may have found the structure of the IRA to be a more familiar form of government and therefore may have been more likely to adopt it, however historical assimilation is likely positively correlated with better economic performance today (Mekeel 1944). Due to the fast implementation of the IRA, the BIA sent several advocates to reservations to promote and educate tribes about the IRA (Mekeel 1944). Given the limited time and resources at the BIAs disposal they likely recruited in more receptive or developed areas and therefore have a higher probability of IRA adoption in these areas (Lemont 2006). If assimilation, receptiveness to federal programs, or development in 1934 is positively correlated with economic development then the OLS estimates will be positively biased.

3.4.2 Voting Records and Narrowly Determined Elections

In order to mitigate these selection concerns, I exploit IRA voting results from the mid-1930s by restricting the sample to tribes that held narrowly determined IRA elections.[4] Presumably, the decision to vote for or against the IRA by a small fraction of voters should influence current economic conditions only through the tribal adoption of the IRA, thus providing plausibly exogenous variation in the initial adoption of the IRA.

[4] I collected IRA voting results from *Ten Years of Tribal Government Under I.R.A.* (Haas 1947).

3.4.3 Summary Statistics Within Optimal Bandwidth

Table 3.2 compares IRA and non-IRA reservations within the restricted sample of narrowly determined elections to check whether or not there were any preexisting differences prior to the IRA election.[5] As described in Section 3.3.2, differences exist in the geographic data, census data, and the allotment data for the full sample. However, within the restricted sample these differences are much smaller. The final column of the table tests for a statistical difference between IRA and non-IRA reservations. The results indicate that only distance from urban areas is statistically different between IRA and non-IRA reservations within the restricted sample.

Table 3.2:

Summary Statistics of Controls within Optimal Bandwidth

	IRA Reservations			Non-IRA Reservations			Difference	
	Mean	SE	N	Mean	SE	N	Mean Diff	RD Diff
Geographic Characteristics								
Distance from MSA (in km)	188.50	45.44	49	183.96	45.44	22	4.54	***
Avg Suitability for Wheat	39.61	3.19	49	31.18	3.19	22	8.43	
Distance to Coal Deposits (in km)	159.36	13.81	51	70.26	13.81	22	89.10	
pc Income of Neighboring Counties	10974.69	404.48	48	11973.71	404.48	20	-999.01	
1910 Reservation Characteristics								
Share of Women	48.85	0.79	48	49.59	0.79	22	-0.74	
Fraction under 18 in school	45.78	3.12	48	53.47	3.12	22	-7.69	
Percent Literate	52.68	2.93	48	57.51	2.93	22	-4.83	
Percent in Labor Force	45.40	3.07	48	39.46	3.07	22	5.94	
Percent Married	39.15	1.40	48	36.66	1.40	22	2.49	
Percent Living on Farms	38.58	5.39	48	37.09	5.39	22	1.48	
Avg Age	26.43	0.53	48	27.01	0.53	22	-0.58	
Avg Family Size	5.45	0.25	48	4.89	0.25	22	0.55	
Percent White Blood	16.62	1.51	48	14.08	1.51	22	2.54	
1934 Allotment Characteristics								
Percent Allotted	0.71	0.07	51	0.86	0.07	22	-0.16	
Avg Allotment Acreage (000s)	236.04	66.63	51	168.68	66.63	22	67.36	
Avg Number of Allotments	1382.57	307.63	51	1131.68	307.63	22	250.89	
Avg Amount of Surplus Land (000s)	79.32	74.26	51	143.50	74.26	22	-64.18	
Avg Acreage Alienated (000s)	97.17	16.53	51	48.96	16.53	22	48.22	
IRA Voting Population	804.902	105.2806	51	514.1364	105.2806	22	290.7656	

[5] I define the criterion for the restricted sample in Section 3.5.

3.5 Empirical Strategy and Results

3.5.1 Regression Discontinuity

My empirical specification exploits the narrow IRA voting results in a regression discontinuity (RD) framework to estimate the effect of the IRA on the outcomes of interest. My preferred RD specification is of the form:

$$Y_i = \beta_0 + \beta_1 IRA_i + f(x_i) + RezChar_i'\gamma + Allot_i'\delta + Geo_i'\theta + \epsilon_i \tag{3.1}$$

$$\forall x_i \epsilon (c-h, c+h)$$

where Y_i is the outcome of interest, IRAi is the treatment, and h is the bandwidth. The running variable, x_i measures the difference in the IRA voting divided by the eligible voting population. In most cases not everyone cast a ballot for or against the IRA. A practical interpretation of the bandwidth is the fraction of individuals that need to change their votes to alter the IRA election outcome. I chose to include the eligible voters that abstain from voting because it seems more plausible to induce a smaller fraction of those voters to vote than change a larger proportion of individuals that cast votes.[6]

The coefficient of interest is β_1, which measures the effect of adopting the IRA conditional on the controls. I estimate this equation using a Local Linear Regression, which combines a suitable bandwidth and a linear control function, $f(x_i)$. I use the algorithm outlined by Imbens and Kalyanaraman (IK) (2011) to choose my optimal bandwidth. The results are robust to a variety of bandwidths including a newer optimal bandwidth algorithm from Calonico, Cattaneo, and Titiunik (CCT) (2014).

[6] I have run the analysis with $x_i = \frac{(yes-no)}{(yes+no)}$, where I only consider the individuals that voted and the results do not change significantly although the optimal bandwidths are much larger.

3.5.2 Results

My primary outcome of interest is per capita income. Figure 3.2 plots the log of per capita income and fits a 4th order polynomial to the data before and after the cutoff. Apparent from the figure is the large discontinuity around the IRA win margin. Reservations to the left of the margin did not adopt the IRA and Figure 3.2 indicates these reservations have significantly higher incomes. Table 3.3 presents the regression discontinuity results for per capita income under several different specifications. The first column presents the results using the IK optimal bandwidth. The results indicate that reservations who narrowly adopted the IRA have over 48 percent lower incomes measured in 1990. Column 2 reports results using the CCT optimal bandwidth, which is slightly more restrictive and finds a larger effect.

Figure 3.2: Regression Discontinuity Plot of Per Capita Income

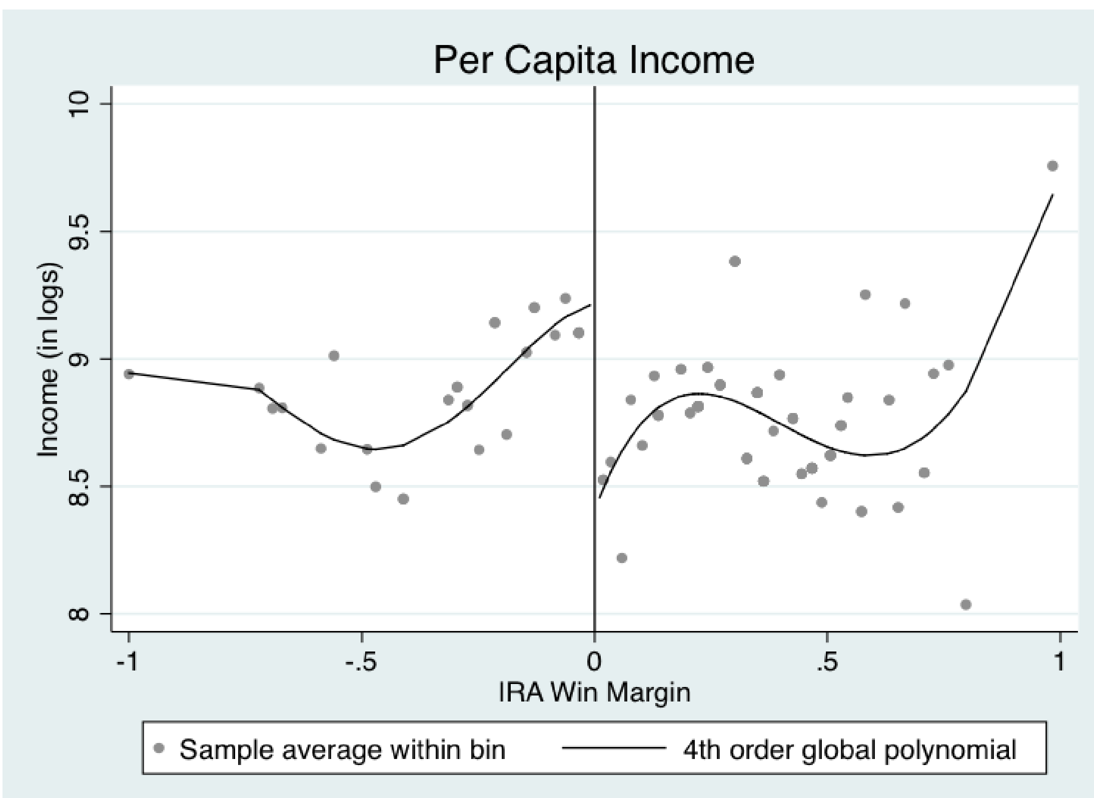


Table 3.3:

Indian Reorganization Act and Per Capita Reservation Income

	(1)	(2)	(3)	(4)	(5)
IRA Voting	-0.658***	-0.757***	-0.525***	-0.616***	-0.535**
	(0.193)	(0.258)	(0.187)	(0.177)	(0.208)
Geography Controls			x	x	x
Allotment Controls				x	x
1910 Controls					x
Bandwidth Type	IK	CCT	IK	IK	IK
Bandwidth	0.426	0.267	0.301	0.301	0.301
LATE in %s	-48.21	-53.09	-40.84	-45.99	-41.43
Observations	73	38	119	119	119

Per capita income is in logs. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Columns 3 through 5 step in the various controls for historical and contemporary reservation demographics, allotment and land tenure characteristics from 1934, agricultural land quality measures, and several spatial characteristics including the distance to natural resources and the distance to urban areas and major metropolitan areas. These controls are particularly helpful with small sample bias (Imbens and Lemieux 2008). Including these controls only reduces the point estimates slightly. The results indicate that narrow IRA adoption led to substantially lower contemporary incomes on Indian reservations.

Table 3.4 shows RD results for both the full reservation population and only those self-identifying as Native American. The results from column 2 indicate that the IRA less negatively impacts individuals identifying as only Native American on the census. One possible reason is that IRA reservations may have larger tribal governments and have preferential hiring toward Native Americans, which improves Indian incomes relative to other reservation residents.[7]

[7] I am currently looking for reservation level data on federal payments to tribal governments or tribal government employment data to test this assertion.

Table 3.4:

Regression Discontinuity Estimates of the Impact of the Indian Reorganization Act

	(1) Indian Per Capita Income	(2) Per Capita Income	(3) Percent Complete College	(4) Percent with Less than HS	(5) Share of Indians	(6) Share Receiving Pub. Assistance
IRA Voting	-0.295**	-0.658***	-10.62*	2.605	63.43***	12.36**
	(0.135)	(0.193)	(5.739)	(5.467)	(13.24)	(6.010)
Bandwidth Type	IK	IK	IK	IK	IK	IK
Bandwidth	0.901	0.426	0.410	0.424	0.364	0.524
Outcome Mean	5211.08	6965.75	30.56	14.14	64.59	22.14
LATE in %s	-25.55	-48.21	-34.75	18.42	98.2	55.83
Observations	117	73	69	71	63	93

Per capita income is in logs. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 3.3 and Figure 3.4 shows regression discontinuity graphs for two different measures of education. The first is the proportion of the population with a college degree. The figure seems to indicate that non-IRA reservations have a slightly higher percentage of college educated residents. The second figure repeats the plot for the proportion of reservation residents with less than a 9th grade education and appears to find no result. I would not expect pre-high school dropout rates to be strongly influenced by the IRA given the national trends in high school attendance and the fact that education policy is often set outside of local tribal governments. Columns 3 and 4 of Table 3.4 support the evidence presented in the figures. The fraction of college educated individuals is over 10 percentage points lower on IRA reservations, based on the sample mean that is a difference of nearly 35 percent. As expected there is no statistical difference in the proportion of the population with less than a 9th grade education.

Figure 3.3: Regression Discontinuity Plot of College Educated Citizens

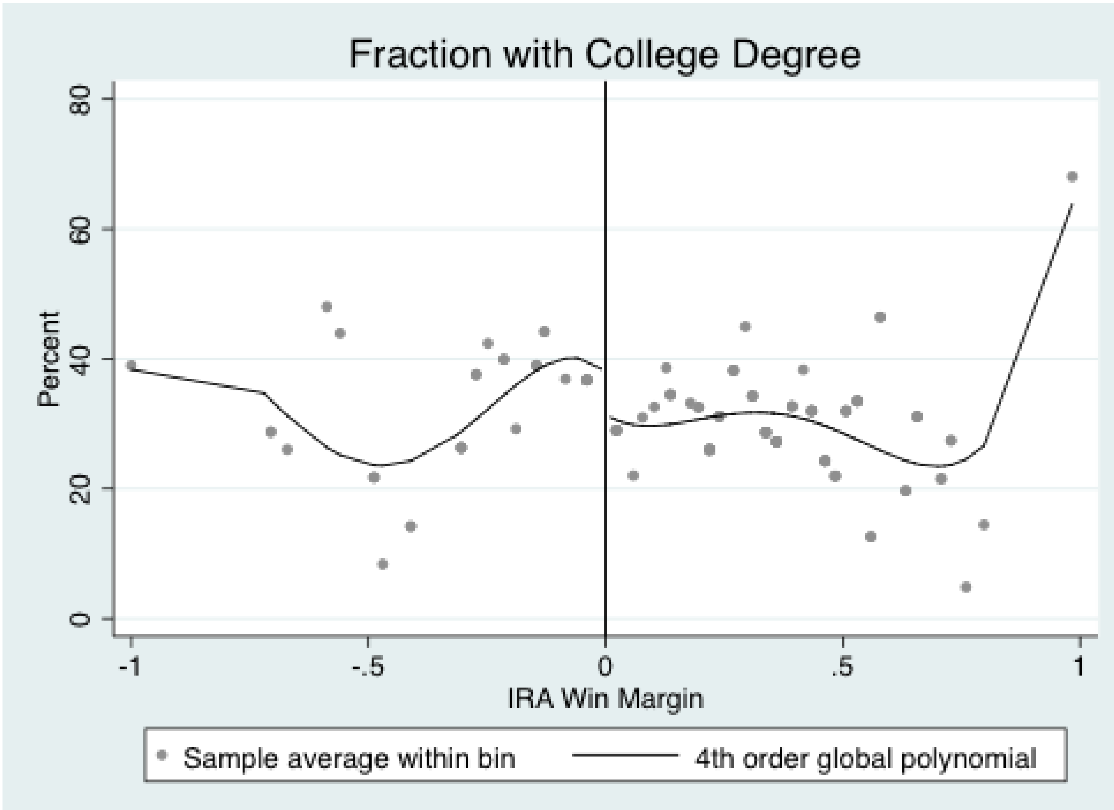
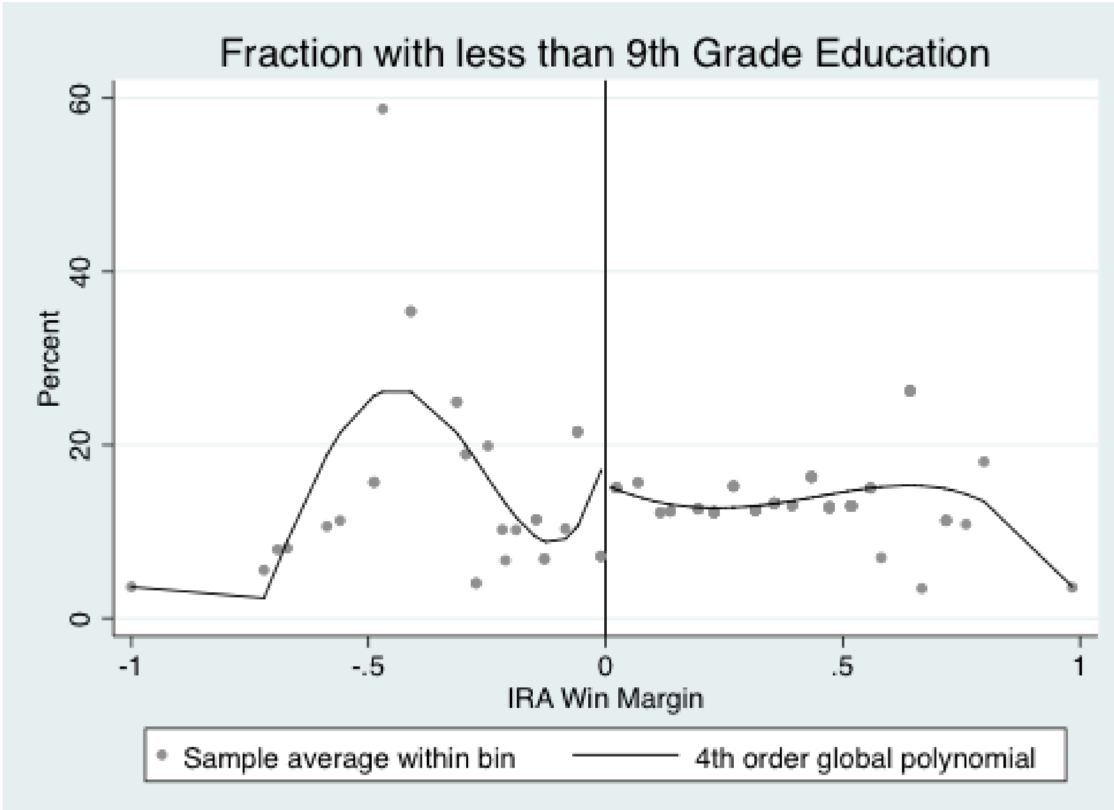


Figure 3.4: Regression Discontinuity Plot of Citizens with a High School Degree



Column 5 of Table 3.4 also reports differences in the level of racial integration between IRA and non-IRA reservations. IRA reservations have a significantly higher proportion of individuals self-identifying as single race Native American.[8] Column 6 examines the fraction of individuals using public assistance. The results are consistent with the earlier income results. Individuals living in IRA reservations are over 12 percentage points more likely to be receiving some type of public assistance. Given an average public assistance rate of 22 percent, individuals on IRA reservations are over 56 percent more likely to be receiving some type of public assistance. These findings indicate that the limited decentralization granted to IRA reservations inhibited economic growth.

[8] Migration selection differences between IRA and non-IRA reservations may be partially responsible for these differences. I am hoping to get data on migration by race to determine whether or not this is a result of emigration of Native Americans or immigration of non-Native Americans.

3.6 The Mechanism of Restrictive Federal Oversight

To verify that the growth differences between IRA and non-IRA reservations are driven by differences in the degree of decentralization and tribal sovereignty, I exploit legal changes in overall tribal sovereignty that occurred in the late 1980s and early 1990s that should have led to a convergence in the sovereignty between IRA and non-IRA reservations. In the late 1980s two important pieces of legislation increased tribal sovereignty, the Indian Gaming Regulatory Act and the Indian Self-Determination and Education Assistance Act (Dippel 2014). These laws reduced BIA oversight, which may have limited the benefits of being a non-IRA reservation. In order to test this assertion, I run a Difference-In-Difference (DID) specification that examines whether or not there where differential effects on per capita income of being an IRA reservation in 2000 and in 2010 compared to 1990. I estimate the following regression:

$$\begin{aligned} Y_{it} = \beta_0 + \beta_1 IRA_i + \beta_2 I_{2000} + \beta_3 I_{2010} + \beta_4 (I_{2000} \times IRA_i) + \beta_5 (I_{2010} \times IRA_i) \\ + f(x_i) + RezChar_i' \gamma + Allot_i' \delta + Geo_i' \theta + \epsilon_{it} \end{aligned} \tag{3.2}$$

The coefficients of interest are β_4 and β_5, which measure the effect of being an IRA reservation in 2000 and 2010. I expect IRA reservations to benefit more from relaxing administrative oversight; therefore I expect β_4 and β_5 to be positive. The other variables are the same as in the previous regression discontinuity specification. In an effort to address the selection issues from before, I restrict the sample to the set of reservations within the optimal bandwidth. The DID specification assumes that in the absence of federal changes, which allowed for increased tribal self-governance, the IRA and non-IRA reservations would have had equal trends. However, it is plausible that the non-IRA reservations would have slightly higher growth in the absence of the federal changes. This suggests that the coefficients of interest, β_4 and β_5 will be negatively biased. Table 3.5 presents the DID regression results for a variety of specifications.

Table 3.5:

Indian Reorganization Act and Per Capita Reservation Income Between 1990 - 2010

	(1)	(2)	(3)	(4)	(5)
IRA	-0.166**	-0.00356	-0.863***	-0.554*	
	(0.0752)	(0.0775)	(0.229)	(0.295)	
IRA x 2000	0.156***	0.123**	0.155***	0.123**	0.125**
	(0.0527)	(0.0552)	(0.0535)	(0.0561)	(0.0628)
IRA x 2010	0.0843	0.0301	0.0842	0.0301	0.0541
	(0.0687)	(0.0729)	(0.0697)	(0.0741)	(0.0628)
2000	0.162***	0.158***	0.162***	0.158***	0.162***
	(0.0398)	(0.0450)	(0.0404)	(0.0457)	(0.0541)
2010	0.296***	0.307***	0.296***	0.307***	0.296***
Geography Controls		x		x	
Allotment Controls		x		x	
1910 Controls		x		x	
Quartic Polynomial			x	x	
Res. Fixed-Effect					x
Bandwidth Type	IK	IK	IK	IK	IK
Bandwidth	0.545	0.545	0.545	0.545	0.545
Observations	282	282	282	282	282
R-squared	0.176	0.472	0.351	0.581	0.848

Per capita income is in logs. Standard errors in parentheses are clustered by reservation. *** p<0.01, ** p<0.05, * p<0.1

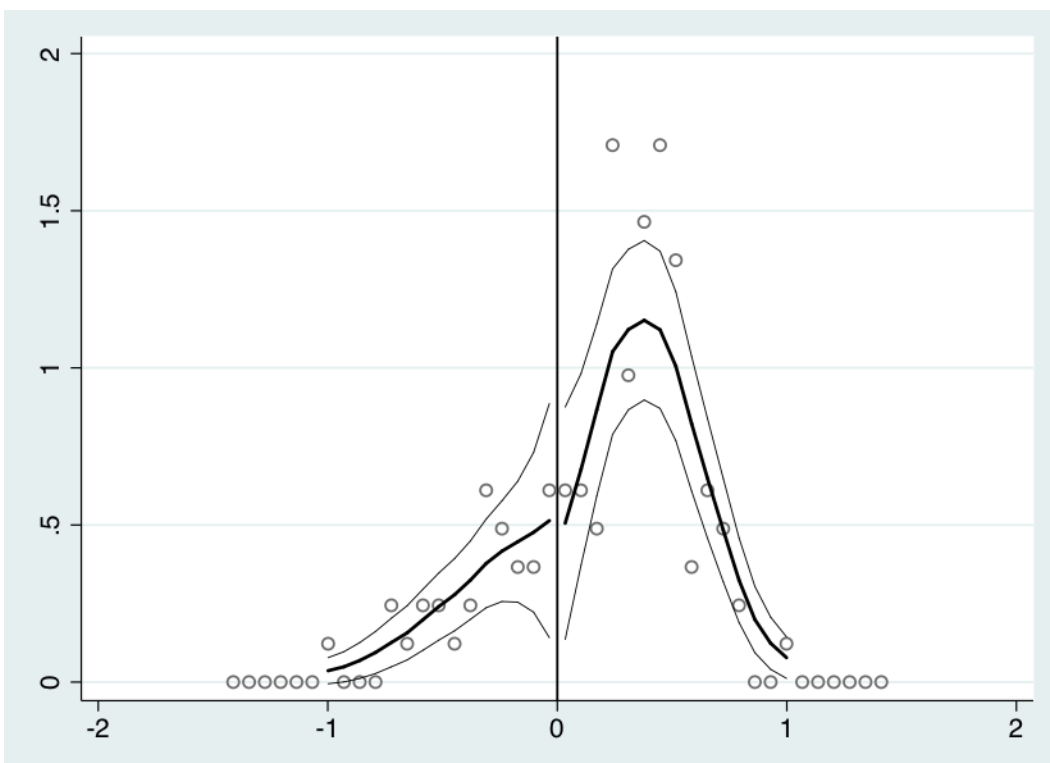
The results indicate the increased tribal sovereignty led to marginal improvements in incomes among IRA reservations relative to non-IRA reservations after 1990. Column 1 presents the standard DID regression results without any controls. The interaction term coefficients indicate that IRA reservations grew approximately 15 percent faster than non-IRA reservations between 1990 and 2000. These results taper off slightly by 2010. Columns 2 through 4 introduce time invariant controls for geography, allotment, and 1910 reservation characteristics. The final specification replaces the controls with reservation fixed-effects. The results are consistent across the different specifications and suggest that the federal oversight faced by IRA reservations was partially responsible for suppressing economic development over the 20th century. These results support the mechanism suggested previously, that federal oversight limited the sovereignty granted to tribes and thereby slowed the rate of economic development.

3.7 Robustness

3.7.1 Manipulation Around the Threshold

If individuals can manipulate whether or not the tribe passed the IRA, and therefore create a discontinuity around the voting threshold, then RD does not properly correct for the selection problem. One reason this might be problematic in the case of IRA voting are the anecdotal accusations that the BIA altered elections in favor of the IRA (Johansen and Pritzker 2007). McCrary (2008) developed a non-parametric test that measures whether or not a discontinuity exists around a threshold. Figure 3.5 presents the results from the McCrary Density Test. The coefficient estimate from the McCrary Density Test is -0.244 with a standard error of 0.674. Therefore I find no evidence of manipulation of the voting to the other side of the threshold.

Figure 3.5: McCrary Density Test with IRA Voting



3.7.2 Alternative Cutoffs

As robustness, I selected four different voting cutoffs and tested whether or not similar discontinuities were present and did not find any evidence of income differences at these different cutoffs.

Table 3.6:

Indian Reorganization Act and Per Capita Reservation Income with Varying Cutoffs

	(1)	(2)	(3)	(4)
IRA Voting	0.0634	-0.243	0.137	-0.102
	(0.167)	(0.158)	(0.142)	(0.196)
Voting Cutoff	-0.25	-0.1	0.2	0.4
Bandwidth	0.320	0.423	0.370	0.238
Observations	119	119	119	119

Per Capita Income is in logs. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.8 Conclusion

This paper measures long-run differences in economic development induced by the decentralization of governance between the Bureau of Indian Affairs and Native American tribal governments. I find that limited decentralization, in the form of the Indian Reorganization Act, was detrimental for economic development on American Indian reservations. The findings suggest that among reservations who held narrowly determined IRA elections the IRA led to lower incomes, a smaller fraction of the population with a college degree, less racial integration, and a larger reliance on public assistance. This paper contributes to a growing literature on the intersection of legal and social institutions for economic development, particularly focusing on the long-run benefits of political decentralization in a setting with strong cultural heterogeneity.

Bibliography

Acemoglu, D., S. Johnson, and J. A. Robinson 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. The American Economic Review 91(5), 1369–1401.

Akai, N., and M. Sakata 2002. Fiscal decentralization contributes to economic growth: evidence from state-level cross-section data for the United States. Journal of Urban Economics 52(1), 93–108.

Akee, R. 2009. Checkerboards and Coase: The Effect of Property Institutions on Efficiency in Housing Markets. Journal of Law and Economics 52(2), pp. 395–410.

Akee, R., M. Jorgensen, and U. Sunde 2012. Constitutions and Economic Development: Evidence from the American Indian Nations. IZA Discussion Paper.

Allen, T., and C. Arkolakis 2014. Trade and the Topography of the Spatial Economy.

Anderson, T., and D. Parker 2008. Sovereignty, Credible Commitments, and Economic Prosperity on American Indian Reservations. Journal of Law and Economics 51(4), pp. 641–666.

Anderson, T. L., and D. Lueck 1992. Land Tenure and Agricultural Productivity on Indian Reservations. Journal of Law and Economics.

Anderson, T. L., and D. P. Parker 2009. Economic development lessons from and for North American Indian economies *. Australian Journal of Agricultural and Resource Economics 53(1), 105–127.

Bardhan, P. 2002. Decentralization of Governance and Development. The Journal of Economic Perspectives 16(4), 185–205.

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

Baum-Snow, N. 2010. Changes in transportation infrastructure and commuting patterns in US metropolitan areas, 1960–2000. The American Economic Review 100(2), 378–382.

Baum-Snow, N. 2014. Urban transport expansions, employment decentralization, and the spatial scope of agglomeration economies.

Baum-Snow, N., L. Brandt, J. V. Henderson, M. A. Turner, and Q. Zhang 2014. Roads, railways and decentralization of Chinese cities.

Calonico, S., M. D. Cattaneo, and R. Titiunik 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. Econometrica 82(6), 2295–2326.

Carlson, L. 1981. Indians, Bureaucrats, and Land: the Dawes Act and the Decline of Indian Farming. Contributions in economics and economic history. Greenwood Press.

Chandra, A., and E. Thompson 2000. Does Public Instrastructure Affect Economic Activity? Evidence from the Rural Interstate Highway System. pp. 1–34.

Clow, R. L. 1987. The Indian Reorganization Act and the Loss of Tribal Sovereignty: Constitutions on the Rosebud and Pine Ridge Reservations. Great Plains Quarterly.

Commodity Flow Survey 2007. United States Census Bureau. Washington D.C.

Cookson, J. A. 2010. Institutions and Casinos on American Indian Reservations: An Empirical Analysis of the Location of Indian Casinos. Journal of Law and Economics 53(4), 651–687.

Cookson, J. A. 2014. Economic Consequences of Judicial Institutions: Evidence from a Natural Experiment. Available at SSRN 1893693.

Davis, D. R., and D. E. Weinstein 2002. Bones, Bombs, and Break Points: The Geography of Economic Activity. The American Economic Review 92(5), 1269–1289.

Dell, M. 2010. The persistent effects of Peru's mining mita. Econometrica 78(6), 1863–1903.

Dell, M. 2012. Path dependence in development: Evidence from the Mexican Revolution.

Dippel, C. 2014. Forced coexistence and economic development: evidence from Native American Reservations. Econometrica 82(6), 2131–2165.

Duranton, G., P. Morrow, and M. A. Turner 2013. Roads and Trade: Evidence from the US.

Duranton, G., and D. Puga 2004. From sectoral to functional urban specialisation. Journal of Urban Economics.

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

Fan, S., and X. Zhang 2004. Infrastructure and regional economic development in rural China. China Economic Review 15(2), 203–214.

Fernald, J. G. 1999. Roads to prosperity? Assessing the link between public capital and productivity. American Economic Review, 619–638.

Food and Agricultural Organization of the United Nations 2015. Global Agro-Ecological Zones.

Frye, D. 2012. Land Allotment Timing During the Dawes Act. Working paper.

Frye, D. 2015a. The Impact of Transportation Networks on Farm Structure and Agricultural Production. Working paper.

Frye, D. 2015b. Transportation Networks and the Geographic Concentration of Industry. Working paper.

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.

Gollin, D., R. D. Rogerson, and National Bureau of Economic Research 2010. Agriculture, roads, and economic development in Uganda [electronic resource].

Haas, T. H. 1947. Ten years of tribal government under I. R. A.

Haines, M., I. university Consortium for Political, and S. Research 2010. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002.

Holmes, T. J., and J. J. Stevens 2004. Chapter 63 Spatial distribution of economic activities in North America. Handbook of Regional and Urban Economics.

ICPSR 2005. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 Codebook. pp. 1–1546.

Imbens, G., and K. Kalyanaraman 2011. Optimal bandwidth choice for the regression discontinuity estimator. The Review of Economic Studies, rdr043.

Imbens, G. W., and T. Lemieux 2008. Regression discontinuity designs: A guide to practice. Journal of Econometrics 142(2), 615–635.

Johansen, B. E., and B. M. Pritzker 2007. Encyclopedia of American Indian History [4 volumes]. Abc-clio.

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.

Legters, L. H., and F. J. Lyden 1994. American Indian Policy: Self-Governance and Economic Development.

Lemont, E. D. 2006. American Indian Constitutional Reform and the Rebuilding of Native Nations.

Lin, J. Y., and Z. Liu 2000. Fiscal Decentralization and Economic Growth in China*. Economic Development and Cultural Change 49(1), 1–21.

Mamatzakis, E. C. 2003. Public infrastructure and productivity growth in Greek agriculture. Agricultural Economics 29(2), 169–180.

McCrary, J. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics 142(2), 698–714.

Mekeel, S. 1944. An Appraisal of the Indian Reorganization Act. American Anthropologist 46(2), 209–217.

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

Michaels, G., F. Rauch, and S. J. Redding 2013. Task Specialization in U.S. Cities from 1880-2000. NBER Working Paper Series.

Minnesota Population Center 2011. National Historical Geographic Information System: Version 2.0. http://www.nhgis.org.

National Atlas 2014. http://www.nationalatlas.gov Accessed: 2014-09-30.

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.

Nunn, N. 2008. The long-term effects of Africa's slave trades. The Quarterly Journal of Economics 123(1), 139–176.

Nunn, N. 2009. The importance of history for economic development.

Nunn, N. 2010. Religious conversion in colonial Africa. The American Economic Review 100(2), 147–152.

Nunn, N., and L. Wantchekon 2011. The Slave Trade and the Origins of Mistrust in Africa. American Economic Review.

Oates, W. E. 1999. An Essay on Fiscal Federalism. Journal of Economic Literature 37(3), 1120–1149.

Office of Indian Affairs 1934. Indian Land Tenure, Economic Status, and Population Trends. Report on Land Planning.

Parker, D. P. The Effects of Legal Institutions on Access to Credit: Evidence from American Indian Reservations. Technical report.

Philp, K. R. 1999. Termination Revisited: American Indians on the Trail to Self-Determination, 1933-1953. Lincoln, NE: University of Nebraska Press.

Rao, P. P., P. S. Birthal, and P. K. Joshi 2006. JSTOR: Economic and Political Weekly, Vol. 41, No. 26 (Jun. 30 - Jul. 7, 2006), pp. 2747-2753. Economic and Political Weekly.

Redding, S. J., and D. M. Sturm 2008. The Costs of Remoteness: Evidence from German Division and Reunification. The American Economic Review 98(5), 1766–1797.

Redding, S. J., D. M. Sturm, and N. Wolf 2011. History and industry location: Evidence from german airports. The Review of Economics and Statistics 93(3), 814–831.

Redding, S. J., and M. A. Turner 2014. Transportation costs and the spatial organization of economic activity. NBER Working Paper Series.

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.

Rusco, E. R. 2000. A Fateful Time: The Background and Legislative History of the Indian Reorganization Act.

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.

Teruel, R. G., and Y. Kuroda 2005. Public infrastructure and productivity growth in Philippine agriculture, 1974–2000. Journal of Asian Economics 16(3), 555–576.

Thornton, J. 2007. Fiscal decentralization and economic growth reconsidered. Journal of Urban Economics.

United States Department of Agriculture 1955. Fresh Fruit and Vegetable Shipments: By Commodities, States, and Months. United States Department of Agriculture: Agricultural Marketing Service. Washington D.C.

United States Department of Agriculture 2013. Fresh Fruit and Vegetable Shipments: By Commodities, States, and Months. United States Department of Agriculture: Agricultural Marketing Service. Washington D.C.

United States. Federal Highway Administration 1977. America's highways, 1776-1976: a history of the Federal-aid program. U.S. Dept. of Transportation, Federal Highway Administration : for sale by the Supt. of Docs., U.S. Govt. Print. Off.

Weinhold, D., and E. Reis 2008. Transportation costs and the spatial distribution of land use in the Brazilian Amazon. Global Environmental Change 18(1), 54–68.

World Bank 2000. Entering the Twenty-First Century.

World Bank 2001. Decentralization & Subnational Regional Economics.

Zhang, T., and H. Zou 1998. Fiscal decentralization, public spending, and economic growth in China. Journal of Public Economics.

Appendix A

A.1 Crop List

- Field Corn
- Sorghum
- Wheat
- Oats
- Barley
- Rye
- Rice
- Flaxseed
- Mixed Grains
- Proso Millet
- Safflower Seed
- Other Misc. Grains
- Alfalfa
- Small Grain Hay
- Grass Silage
- Wild Hay
- Other Misc. Hay
- Tomatoes
- Sweet Corn
- Cucumbers and Pickles
- Watermelons
- Snapbeans
- Green Peas
- Lettuce
- Asparagus
- Cantaloups
- Other Vegetables
- Oranges
- Grapefruit
- Other Citrus Fruit
- Walnuts
- Almonds
- Pecans
- Other Nuts
- Apples
- Peaches
- Pears
- Cherries
- Plums

- Grapes
- Other Non-Citrus Fruit
- Strawberries
- Other Misc. Berries
- Soybeans
- Peanuts
- Cotton
- Tobacco
- Irish Potatoes
- Sweet Potatoes
- Dry Field and Seed Beans
- Dry Lima Beans
- Dry Field and Seed Peas
- Sugar Beets
- Other Field Crops
- Sugar Cane
- Sunflower Seeds
- Other Misc Crops
- Alfalfa Seed
- Red Clover Seed
- Lespedeza Seed
- Timothy Seed
- Fescue Seed
- Ryegrass Seed
- Other Misc. Field Seed
- Nursery and Greenhouse Products

A.2 Data Sources

I supplement the economic and highway information with data covering population, historical economic data, and alternative methods of transportation.[1] I use county-level population data from the U.S. Census for every decade from 1910 to 1950. I collected additional geographic information for alternative methods of transportation from NationalAtlas.gov (2014). I use GIS to construct an indicator that is equal to one if a county has a railroad.[2] For each county I calculate the Euclidian distance to the nearest coastal port and the nearest airport.

[1] Population and historical economic data are from the National Historical Geographic Information Systems (NHGIS, 2011).

[2] Due to data availability constraints I ignore railroad lines that were decommissioned following deregulation.