Did Highways Cause Suburbanization?
Author(s): Nathaniel Baum-Snow
Source: *The Quarterly Journal of Economics*, Vol. 122, No. 2 (May, 2007), pp. 775-805
Published by: Oxford University Press
Stable URL: https://www.jstor.org/stable/25098858

DID HIGHWAYS CAUSE SUBURBANIZATION?*

NATHANIEL BAUM-SNOW

Between 1950 and 1990, the aggregate population of central cities in the United States declined by 17 percent despite population growth of 72 percent in metropolitan areas as a whole. This paper assesses the extent to which the construction of new limited access highways has contributed to central city population decline. Using planned portions of the interstate highway system as a source of exogenous variation, empirical estimates indicate that one new highway passing through a central city reduces its population by about 18 percent. Estimates imply that aggregate central city population would have grown by about 8 percent had the interstate highway system not been built.

I. INTRODUCTION

Between 1950 and 1990, the aggregate population of 1950 geography central cities in the United States declined by 17 percent despite population growth of 72 percent in metropolitan areas as a whole. This paper demonstrates that the construction of new limited access highways has contributed markedly to this central city population decline. This explanation is motivated by the land use theory developed by Alonso [1964], which predicts that faster commuting times push up the demand for space in suburbs relative to central cities. A complementary explanation for suburbanization focuses on changes in the amenity value of suburbs relative to central cities. Others provide evidence that the resulting Tiebout sorting [Tiebout 1956] has led people to the suburbs because of racial preferences interacted with changes in central city racial composition [Boustan 2006], increases in central city problems like crime and blight [Cullen & Levitt 1999], and the desegregation of central city school districts [Reber 2005]. Estimates presented below indicate that highways can explain about one-third of the change in aggregate central city population relative to metropolitan area population as a whole.

The empirical strategy used in this paper exploits the fact that an observable portion of the national road network was designed to link far away places and not to facilitate local commuting. As a result, some metropolitan areas (MSAs) received

* I thank the editor and three referees for their guidance in improving the content and exposition of this paper. In addition, I thank Gary Becker, Vernon Henderson, Matthew Kahn, Steven Levitt, Edwin Mills, Chad Syverson, and especially Derek Neal for helpful suggestions and encouragement. All errors are mine.

The Quarterly Journal of Economics, May 2007

more interstate highways than others simply because they are located nearer to other population centers. As such, the empirical work in this paper essentially compares rates of suburbanization in metropolitan areas that received many new highways between 1950 and 1990 to those receiving fewer during this period. In addition, I examine the extent to which highways can explain different rates of suburbanization across decades within MSAs. To account for the potential endogeneity of highway assignment, I instrument for the total number of highways built with the number of highways in a 1947 national interstate highway plan. The measure of highways used is the number of "rays" emanating from central cities. This measure is convenient in that it is somewhat comparable across urban areas with different spatial structures and is easily observed by inspecting a map.[1]

This paper demonstrates that innovations to the urban transportation infrastructure played a key role in influencing changes in the spatial distribution of the population in U. S. metropolitan areas between 1950 and 1990. Empirical estimates in this paper imply that each new highway causes constant geography central city population to decline by about 18 percent, all else equal. Using this estimate, I calculate that had the interstate highway system not been built, instead of declining by 17 percent, aggregate central city population would have grown by 8 percent.

Table I documents the evolution of the spatial distribution of the population in metropolitan areas between 1950 and 1990. Panel A presents trends in aggregate population growth within large MSAs and panel B presents analogous trends for large MSAs with central cities located at least 20 miles from a coast, major lake shore or international border. Table I shows that at 26 percent, the decline in 1950-definition aggregate central city population was considerably more pronounced in inland central cities than in the more complete sample. These declines came despite considerable net migration to metropolitan areas and sizable increases in total central city population. Aggregate population in central cities as defined by their political boundaries grew by 38 percent in inland MSAs and 14 percent in the full sample. Trends at the median in each of the two samples follow similar patterns as the aggregates.

The patterns in Table I suggest that a sizable fraction of urban population decentralization is independent of boundaries

1. I define rays as limited access highways connecting central cities' central business districts to the suburbs.

TABLE I
AGGREGATE TRENDS IN SUBURBANIZATION, 1950–1990

	1950	1960	1970	1980	1990	Percent change 1950–1990
Panel A: Large MSAs						
MSA population	92.9	115.8	134.0	144.8	159.8	72
Total CC population	44.7	48.5	51.3	49.2	51.0	14
Constant geography CC population	44.7	44.2	42.6	37.9	37.1	−17
N for constant geog. CC population	139	132	139	139	139	
Panel B: Large Inland MSAs						
MSA population	39.2	48.9	57.0	65.0	73.5	88
Total CC population	16.8	19.7	22.1	22.1	23.2	38
Constant geography CC population	16.8	16.5	15.4	13.3	12.5	−26
N for constant geog. CC population	100	94	100	100	100	
Total U. S. population	150.7	178.5	202.1	225.2	247.1	64

Notes: All populations are in millions. CC stands for central city. The sample includes all metropolitan areas (MSAs) of at least 100,000 people with central cities of at least 50,000 people in 1950. The sample in Panel B excludes MSAs with central cities located within 20 miles of a coast, major lake shore, or international border. MSA populations are for geography as of year 2000. Constant geography central city population uses 1950 central city geography. Census tract data are not available to build constant geography central city populations for some small cities in 1960. These cities are assigned a population of 0 for constructing the aggregates. Reported total U. S. population excludes Alaska and Hawaii.

determining the provision of local public goods. The fact that constant geography central city population fell significantly faster than central city population within political boundaries implies that space is an important factor needed to explain falling urban population density. The fact that large inland MSAs saw sharper declines in central city population than all large MSAs despite their faster total MSA population growth further supports this view. Tiebout sorting models imply that people moving further from the core would endeavor to form new communities that provide different levels of public services, rather than relocate within the same political jurisdiction. The aggregate data thus show that there is ample opportunity for a spatial mechanism to be an important driver of urban population decentralization, though it may partially interact with a Tiebout sorting mechanism.[2]

2. Mieszkowski and Mills [1993] make a similar point, noting that population dispersal is not just a post-WWII phenomenon in the United States and that suburbanization has been occurring around the world independent of the geography of local political jurisdictions. Nevertheless, there are clearly some public goods, like crime, that are more neighborhood-based than city-based. This argument is only relevant for services provided at the city level.

This paper proceeds as follows. Section II describes some relevant institutional details about the establishment of the interstate highway system, argues that a 1947 plan of the system provides useful exogenous variation in highway assignment across MSAs, and describes the data. Section III demonstrates that spatial patterns in residential location are consistent with predictions of standard land use theory. Section IV discusses the main estimation results. Finally, Section V concludes.

II. The Interstate Highway System

The first codified federal involvement in developing a national highway system came with the Federal Aid Highway Act of 1944. This legislation instructed the roads commissioner to formulate the first of what was to be a series of plans for a national system of interstate highways. The legislation stipulated that highways in the planned system should be ". . . so located as to connect by routes as direct as practicable, the principal metropolitan areas, cities, and industrial centers, to serve the national defense, and to connect at suitable border points with routes of continental importance in the Dominion of Canada and the Republic of Mexico. . ." In its report that shaped the 1944 legislation, the National Interregional Highway Committee considered the nationwide distribution of population, manufacturing activity, agricultural production, the location of post-World War II employment, a strategic highway network drawn up by the War Department in 1941, the location of military and naval establishments, and interregional traffic demand (in that order) to inform its recommended interregional highway system plan [U. S. House of Representatives 1944]. Furthermore the 1944 highway act does not mention anywhere that these highways should be designed to serve local commuting [U. S. Department of Transportation 1977, p. 277]. States were asked to submit proposals for their portion of the interstate highway system in response to the recommended national plan from the Interregional Highway Committee. Over the following three years, each state highway department negotiated with the federal roads administration over its portion of the comprehensive national plan. In 1947, roads commissioner MacDonald approved final proposals from the states for 37,324 miles of roadway to be built to federal interstate standards. This plan represents most of the system we see built today, some of it not to be finished until the 1980s. A map of this 1947 plan is presented in Figure I.[3]

Highways in the 1947 plan received minimal federal funding before the Federal Aid Highway and Highway Revenue Acts of 1956. Most highways that existed in 1956 and ended up as interstates were in the plan but funded with state-issued bonds to be paid off with toll revenues. The 1956 Interstate Highway Act expanded the mileage in the 1947 plan to a 41,000 mile interstate system and committed the federal government to pay 90 percent of the cost of construction. The 1956 plan incorporated some highways primarily meant for local commuting. With such a strong commitment by the federal government to build the interstate system, it becomes hard to argue that some of it was not built because of the lack of state funds. The state of the national road network in 1950 was such that there were only 341 miles of limited access highways that would be integrated into the interstate system. By 1990 the interstate system had grown to 43,420 miles and virtually the entire 1947 plan had been built. During this time, mileage in the primary sample of 139 metropolitan areas used in this paper grew from 161 to 16,716.

The 1956 legislation put in place a mechanism by which each state reported the completion month of each federally funded interstate highway segment within its borders. It is the resulting "PR-511" data that represents a key component of the longitudinal analysis done in this paper. The PR-511 data set allows for separate identification of segments funded as part of the 1956 Federal Highway Act, segments authorized and funded later by the federal government, locally funded segments, and toll segments. I combine the PR-511 data with a digital map of the interstate system to create a data set containing the miles of each interstate highway in each country in each year between 1950 and 1990.

The primary highway measure used in this paper I refer to as "rays." For this analysis, a "ray" is defined as a segment of road that connects the central business district (CBD) of the central city with the region outside the central city. If a highway passes

3. The 1944 report outlined a recommended plan of 33,920 miles of interregional highways that looks very similar to the 1947 plan. Michaels [2006] uses this 1944 plan as an instrument for rural highway construction. I choose the 1947 version because its map includes more detailed road alignments and more cities that are labeled. Results throughout this paper are not highly sensitive to the choice of plan used.

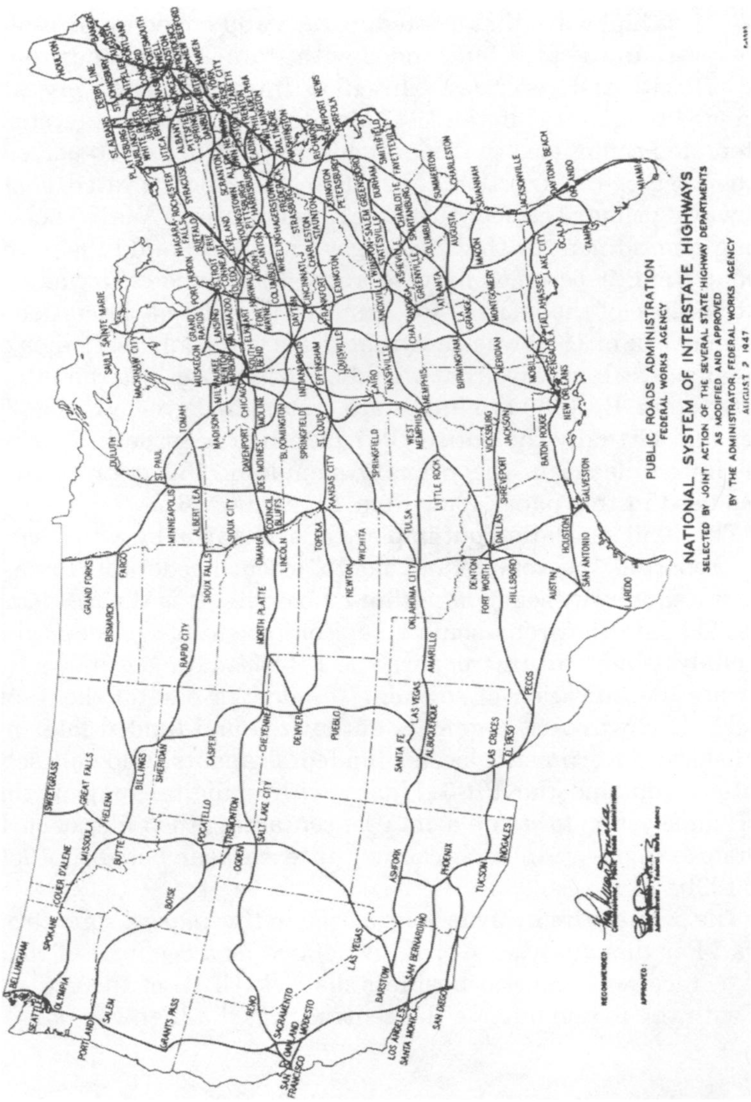


FIGURE I
The Projected System of Interstate Highways in 1947

through the central city, it counts as two rays whereas if a highway terminates in or near the central city it counts as only one. Rays must pass within one mile of the CBD and serve a significant area of the MSA outside the 1950-definition central city to be counted. Highways that split at or near the border of the 1950-definition central city count as multiple rays. Two highways that pass within one mile of the CBD and converge count as only one ray in the direction of convergence. Limited-access expressways that are not part of the interstate system but satisfy all of the listed criteria also count as rays, though only when considering changes in infrastructure between 1950 and 1990.[4] After identifying rays from printed maps, I use the PR-511 data to quantify construction progress. I count a ray as existing in an MSA in a given year if at least one mile was open to traffic. The median number of interstate rays among MSAs in the primary sample was 0 in 1950 and 1960 and 2 thereafter, with a mean of 2.34 in 1990. As a robustness check, I also consider highways that pass within four miles of the CBD as an alternate measure. The correlation between the growth in the two measures of rays between 1950 and 1990 is 0.76. The correlation between the growth in interstate rays and total rays is 0.91.

The fact that some urban highways were built without federal funding is evidence that state and local governments adjusted metropolitan area highway infrastructure at least partly in response to local commuting demand. That states constructed the urban sections of their planned federally funded highway system earlier than sections in rural areas is further evidence to this effect.[5] Highway construction is thus unlikely to have been randomly assigned across MSAs or within MSAs over time. Therefore, attaining consistent estimates of the effect of highways on suburbanization requires instrumenting for actual urban highway construction. As such, in long difference regressions the number of rays in the 1947 plan instruments for actual rays constructed between 1950 and 1990. In panel regressions

4. I only fully observe the stock of all non-interstate limited access highways in 1990. Only one of the thirty-nine non-interstate limited access highways in the primary sample existed in 1950. I could not find data on specific opening years for eighteen of these roads. Therefore, limited access non-interstate highways are excluded from all panel regressions.

5. In the primary sample, 22 percent of 1990 MSA interstate mileage had been completed by 1960, 71 percent by 1970, and 92 percent by 1980. However, only 16 percent of interstate mileage nationally had been completed by 1960, 68 percent by 1970, and 92 percent by 1980.

I instrument for the stock of highways with the (potentially fractional) number of rays that would have existed in each year had state governments allocated federally funded highway construction uniformly across federally planned interstate highways within their jurisdictions.

The validity of the 1947 plan as an instrument depends on the fact that the portion of the system in the plan was designed to facilitate trade and national defense, not to facilitate metropolitan area development. Further, it requires the availability of data on control variables that influence suburbanization and are correlated with the plan. Finally, it requires that planned rays be a good predictor of actual rays constructed.

Demographic data from 1940 and 1950 provide support for the claim that while actual urban highway construction responded to increased commuting demand, the 1947 plan did not respond to changes in commuting demand. A regression of the change in rays between 1950 and 1990 on the change in log MSA population between 1940 and 1950 yields a highly positive coefficient of 1.85 (se = 0.95). However, the same regression using planned rays as the dependent variable yields a coefficient that is not significantly different from 0 of −0.43 (se = 0.87). Because MSA population growth rates are highly serially correlated, a similar pattern emerges when examining the relationship between rays and population growth from 1950 to 1990. Moreover, non-white population and percent high school graduate in 1950 cannot predict either observed highway construction or planned rays.

While planned rays are unrelated to MSA population growth, the level of MSA population is a good predictor of planned rays. Regressions of rays on log 1940 MSA population yield large and significant coefficients for actual rays constructed between 1950 and 1990 and planned rays alike. Together, these results give some indication that the federal government planned more highway construction in big cities as they existed in 1940, not to serve fast growing cities or cities that may have had pressures for racial or social segregation.[6] Because central city population is likely to be a function of MSA population, these results also indicate the

6. The level of MSA population in 1910 is also a good predictor of planned highways. However, MSA population growth between 1910 and 1950 cannot predict the number of rays in the plan.

TABLE II
FIRST STAGE RESULTS
LARGE MSAs IN 1950

Panel A: Long difference 1950–1990

	Change in Rays		
	1	2	3
Rays in the plan	.677 (.074)**	.526 (.076)**	.510 (.074)**
1950 Central city radius		.325 (.071)**	.306 (.072)**
Change in simulated log income			−.939 (1.819)
Change in log of MSA population			.856 (.279)**
Constant	.866 (.218)**	.218 (.247)	.463 (1.231)
Observations	139	139	139
R-squared	.38	.47	.50

Panel B: Full panel 1950–1990

	Rays		Smoothed Rays	
	1	2	4	5
Smoothed rays in plan	.826 (.029)**	.448 (.040)**	.833 (.023)**	.712 (.030)**
Log simulated income		1.563 (.234)**		.198 (.179)
Log MSA population		.170 (.200)		.591 (.154)**
MSA Fixed effects	Yes	Yes	Yes	Yes
Groups	132	132	132	132
R-squared	.55	.84	.67	.88

Notes: Panel A shows the first stage results for the regressions in Table IV. Panel B shows the first stage results for the regressions in Table VI. Listed specification numbers match those in the corresponding second stage tables. "Smoothed rays" is calculated by multiplying the stock of rays in 1990 in each MSA by the fraction of these rays' mileage that is completed at each point in time. "Smoothed rays in the plan" is calculated by multiplying the number of rays in the 1947 plan by the fraction of federally funded highway mileage in the 1956 Federal Aid Highway Act completed at each point in time. All coefficients remain significant when estimated using the more selected sample in Table I Panel B.

importance of holding population levels constant in an evaluation of the effect of rays on suburbanization.

First stage results reported in Table II show that the number of planned rays is a strong predictor of actual highway construction. Table II Panel A shows that conditional on control variables, each planned ray results in a precisely estimated average of 0.51 actual rays being built. Panel B shows that first stage results using the full panel variation in the data and including MSA fixed effects are similar, with coefficients of 0.45 for discrete rays and 0.71 for a measure of smoothed rays. That these coefficients are less than one is evidence that cities receiving many rays in the plan routed some away from their central business districts. Commensurate with the observed positive and significant first-stage relationship between MSA population growth and endogenous road construction, fast-growing metropolitan areas that received few rays in the plan substituted with highways that were locally funded or were added to the federally funded planned system after 1947.

Data for counties and central cities are taken from the *County and City Data Books* and aggregated to form MSAs. I hold MSA geography constant over time using definitions from 2000. I choose MSA definitions from the end of the study period so as to keep total MSA population as comparable as possible over time. I relegate all but the largest central city in each MSA, based on 1950 population, to be part of the suburbs. I use this central city definition in order to facilitate inclusion of central city data from 1950 in the analysis. Only the 139 MSAs of at least 100,000 inhabitants with central cities of at least 50,000 people in 1950 are included in the primary sample. As a robustness check, I replicate all results for all 240 MSAs for which there exist central city data in both 1950 and 1990.

One particularly vexing issue is that over half of the central cities in the sample annexed land over time as their populations decentralized. This phenomenon is especially widespread outside the Northeast and Midwest. It would be difficult to account for this problem with theory or econometrics given that city size expansion clearly was endogenous to suburbanization. As such, I used printed census maps from 1950 to match digital maps of 1970 and 1980 census tracts to 1950 geography. I then aggregated census tracts based on these visual comparisons such that the aggregate 1970 and 1980 areas matched as closely as possible to the land area of the 1950 city geography. In order to avoid understating post-1950 population, the post-1950 central city area was constructed to always exceed the reported 1950 area. The

Data Appendix contains a more complete description of the data construction.

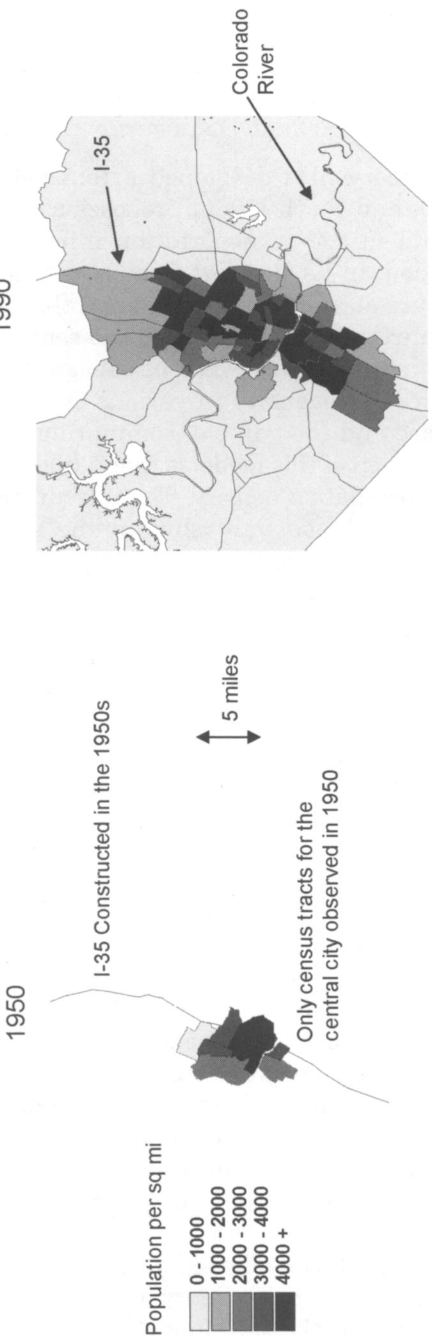
III. Highways and the Spatial Distribution of Residences

The land use theory first developed extensively by Alonso [1964], Mills [1967], and Muth [1969] proposes a mechanism by which improvements in the transportation infrastructure may cause suburbanization. In its simplest form, this theory assumes all employment occurs at a central location and the rental rate of land adjusts as a function of distance from the center to compensate for different commuting times of identical agents. A standard extension allows for heterogeneity in various factors that affect the demand for space and the value of commuting time.

One basic implication of this model is that a higher commuting speed implies lower population density. The introduction of a new highway ray results in an outward shift in the supply of land available for a given commuting time in a section of the city. This causes the equilibrium rental rate of land to fall throughout the metropolitan area, thereby reducing the population density through a price effect. Since the average commuting distance falls, individuals' income net of commuting cost rises. If space is a normal good, a new ray thus further reduces population density, and by extension central city population, through an income effect.

Baum-Snow [2007] demonstrates that this framework provides three implications that are evaluated in the empirical work in the remainder of this paper. First, population in metropolitan areas should spread out along new highways. Second, central city population should increase with metropolitan area population and the radius of the central city. Finally, central city population should decline with the number of highway rays. Commensurate with the second and third predictions and the aggregate data presented in Table I, the primary dependent variable used is the difference in log central city population, where the central city is defined by its 1950 geography. A monocentric land use model adapted to incorporate radial commuting highways implies that central city population is a function of the number of rays, the radius of the central city, MSA population, and the distribution of income.

Observed spatial patterns of residential development accurately follow predictions of this theory in many cities. As an example, Figure II shows the development of the Austin, TX,



Note: Each shaded region is a separate census tract.

FIGURE II
Development Patterns in Austin, TX.

MSA between 1950 and 1990. Austin is a convenient case study because it is one of the largest cities to have only 2 rays, making it easier to evaluate whether spatial patterns in population decentralization follow highways, and it lies in an area largely unimpeded by natural and political boundaries. Interstate 35 opened through Austin in the 1950s. Even as the population of the metropolitan area grew rapidly, the downtown area north of the river lost population between 1950 and 1980, after which its population grew slightly. Furthermore, nearly all of the new residential development between 1950 and 1990 occurred along the north-south highway and not in the east-west direction.[7]

Table III presents a more systematic description of the residential structure of metropolitan areas in the data. Each column in Panel A lists census tract-level regression results from estimating equation (1) using cross-sectional data in 1970 and 1990 and two different samples.

$$\log P_{ij} = a_i + \beta \, \text{dis}_{ij}^{cbd} + c \, \text{dis}_{ij}^{hwy} + \varepsilon_{ij} \tag{1}$$

P_{ij} denotes the population per square mile in census tract j of MSA i. dis^{cbd} is the distance from the tract centroid to the central business district and dis^{hwy} is the distance to the nearest interstate highway. The functional form in equation (1) has been used extensively for estimation of urban density functions [Clark 1951; McDonald 1989] and in the absence of highways it is implied by a monocentric city model with quasilinear utility. The first-difference version reported in Panel B includes the distance to the CBD and to any highways that existed in 1970 as controls. Observations are weighted by the fraction of MSA population they represent in order to weight each MSA equally. Estimates of β from the two cross-sections are in line with estimates from earlier studies.[8]

Table III presents results showing that highways are associated with higher population density both in levels and first-differences. Panel A shows that conditional on CBD distance, population density is greater near highways in both the 1970 and 1990 cross-sections. In the more complete sample, population density is 1 to 2 percent smaller on average each additional mile further from the nearest highway. The magnitude of this

7. The city of Austin expanded from 32 to 218 square miles between 1950 and 1990. Only census tracts in the city proper are observed in 1950.
8. I use 1970 instead of an earlier year because previous years had less complete census tract data.

TABLE III
THE SPATIAL DISTRIBUTION OF METROPOLITAN AREA POPULATIONS

Panel A: 1970 and 1990 Cross-Sections

Sample		Log population density 1970	Log population density 1990
Large MSAs in 1950 (36,250 tracts, 139 MSAs)	Distance to CBD	−.132 (.001)**	−.114 (.001)**
	Distance to highway	−.014 (.002)**	−.019 (.002)**
Large MSAs in 1950 with central cities at least 20 miles from a coast or border (17,336 tracts, 100 MSAs)	Distance to CBD	−.134 (.002)**	−.117 (.001)**
	Distance to highway	−.055 (.003)**	−.054 (.003)**

Panel B: Evolution between 1970 and 1990

Sample		ΔLog population density
Large MSAs in 1950 (36,250 tracts, 139 MSAs)	Distance to CBD	.021 (.000)**
	ΔDistance to highway	−.015 (.002)**
Large MSAs in 1950 with central cities at least 20 miles from a coast or border (17,336 tracts, 100 MSAs)	Distance to CBD	.021 (.001)**
	ΔDistance to highway	−.008 (.003)**

Notes: Each pair of entries lists coefficients and standard errors from a regression of log population density on the listed variables at the census tract level. All regressions include MSA fixed effects. Regressions in Panel B also include the distance to the nearest highway in 1970. Estimated coefficients on distance to the nearest highway in 1970 are between −0.002 and 0.004. Regressions using the distance to planned highways as an instrument for the distance to observed highways yield similar results. When standard errors are clustered by MSA, results for the larger sample in Panel B and results for the smaller sample in Panel A remain significant at the 5 percent level. Other results are not statistically significant with clustering. Regressions are weighted by the fraction of MSA population that is represented in the tract. Analogous unweighted regressions produce highway distance coefficients that are larger in absolute value. All distances are in miles.

estimated gradient is much greater in the sample, including more centrally located MSAs, likely because the less restricted space in these MSAs generates equilibria with less population restricted to live near the highways. Panel B shows that conditional on CBD and 1970 highway distance, portions of MSAs near highways built between 1970 and 1990 had faster population growth than other areas. Population density decreased by about 1 percent for each additional mile from the nearest new highway built. These patterns in the data are consistent with highways influencing residential location patterns.[9]

Consistent with this observation about the equilibrium relationship between the location of highways and the spatial distribution of the population within cities, highway rays are also associated with declines in central city population shares. In the sample used in Panel B of Table I there exists a robust monotonic relationship between the decline in log central city population share and the construction of highways between one and seven rays. In a simple regression of Δ log (central city population/MSA population) on Δray, the average additional ray is associated with a decline in log central city population share of approximately 0.07 if MSAs with zero rays are included and 0.10 if MSAs with zero rays are excluded from the sample. Regressions using a less restrictive ray definition in which highways within four miles of the CBD may count as rays yield OLS coefficients of -0.09 excluding 0-ray MSAs and -0.10 including 0-ray MSAs.

IV. Results

In this section, I present estimates of the causal effect of highways on central city population growth. In addition to 1950 to 1990 difference regressions, this section also presents panel estimation results that include MSA fixed effects using data from the intervening years. In some specifications, the number of rays in the 1947 national plan enters as an instrument for the actual number of rays constructed. In the panel estimation, planned rays smoothed using national construction rates over time enters as the instrument. Results presented in this section show that a range of specifications and estimation strategies all indicate that each additional ray causes approximately a 9 percent decline in central city population. Moreover, I show that analogous regressions using more limited data from 1910 and 1950 indicate that highways planned in the 1940s cannot predict changes in central city population during the first part of the 20th century.

9. These results are not consistent with highways causing blight since migration likely reflects a revealed preference for land near highways. Analogous regressions describing the spatial distribution of housing values at the census tract level also yield negative coefficients on distance to highway in the two cross-sections.

IV.A. Long Difference Estimates

Equation (2) details the primary specification used in estimating the effects of highways on suburbanization. The unit of observation is the MSA and differences are between 1950 and 1990 levels:

(2)

$$\Delta \log N_i^c = \delta_0 + \delta_1 \Delta \text{ray}_i + \delta_2 r_{ci} + \delta_3 \Delta \tilde{w}_i + \delta_4 \Delta \log N_i^{\text{MSA}} + \delta_5 \Delta \tilde{G}_i + \varepsilon_i .$$

N^c represents 1950-definition central city population, r_c denotes central city radius, and N^{MSA} denotes MSA population. $\tilde{w}_i$ represents mean log annual personal income, and $\tilde{G}_i$ represents the Gini coefficient of the income distribution for MSA i. I control for these elements of the income distribution in order to account for potential differences across MSAs in changes in the distributions of demand for living in suburban communities relative to central cities, whether through a land use or a Tiebout sorting mechanism. In order to account for the potential endogeneity of the income distribution to suburbanization, I use a simulated income distribution based on employment shares in 1940 by one-digit industry for each MSA. I take the distribution of skill prices nationally from the census PUMS excluding the states that encompass each MSA in question for each industry of employment and combine them using the reported 1940 employment shares. The skill prices are built using wage and salary income of individuals working at least 48 weeks in the year prior to each census. Self-employed individuals are excluded. Thus the factors that drive the change in the measured income distribution for an MSA are the evolution of skill prices at the national level and the 1940 employment mix in the MSA. The inclusion of population and characteristics of the income distribution as controls in the regressions is intended primarily as a check on the exogeneity of the plan. If the plan is exogenous, the estimated effect of highways should not change with their inclusion.

Table IV presents regression estimates from various specifications of (2). The first column presents ordinary least squares estimates using the primary specification, and the remaining five columns present instrumental variables estimates from a range of specifications. Standard errors are clustered by the state in which the central city is located to account for potential state-level unobservables like environmental laws that could affect

TABLE IV
LONG-DIFFERENCE REGRESSIONS OF THE DETERMINANTS OF CONSTANT GEOGRAPHY CENTRAL CITY POPULATION GROWTH, 1950–1990

	Large MSAs in 1950					
		Change in log population in constant geography central cities				
	OLS3	IV1	IV2	IV3	IV4	IV5
Change in number of rays	−.059 (.014)**	−.030 (.022)	−.106 (.032)**	−.123 (.029)**	−.114 (.026)**	−.101 (.046)*
1950 central city radius	.080 (.014)**		.111 (.023)**	.113 (.023)**	.106 (.023)**	.125 (.021)**
Change in simulated log income	.084 (.378)			.048 (.417)	−6.247 (6.174)	−.137 (.480)
Change in log of MSA population	.363 (.082)**			.424 (.094)**	.374 (.079)**	.405 (.108)**
Change in Gini coeff of simulated income					−23.416 (23.266)	
Log 1950 MSA population						−.062 (.062)
Constant	−.640 (.260)*	−.203 (.078)*	−.359 (.076)**	−.588 (.281)*	4.580 (5.091)	−.611 (.265)*
Observations	139	139	139	139	139	139
R-squared	.39	.00	.01	.30	.33	.37

Notes: In columns IV1–IV5, the number of rays in the 1947 plan instruments for the change in the number of rays. Standard errors are clustered by state of the MSA central city. Standard errors are in parentheses. ** indicates significant at the 1 percent level, * indicates significant at 5 percent level. Summary statistics are in the Appendix Table. First stage results are in Table II.

suburbanization. Coefficients on rays in all specifications have the expected negative sign and are usually statistically significant and sizable. OLS results indicate that conditional on control variables, each additional ray is associated with about a 6 percent decline in central city population. In this primary specification, instrumental variables estimates imply about a 12 percent decline in central city population for each additional ray. The only control variable that significantly influences the coefficient on rays is the central city radius. A positive correlation between the central city radius and the number of rays, planned and actual, accounts for the absolute increase in the coefficient on rays when the radius enters as a control.[10] Using central city area as a control instead produces similar results.

10. A land use model would imply that the appropriate functional form may include an interaction between rays and central city radius. The resulting coeffi-

As is predicted by land use theory, the coefficients on central city radius and metropolitan area population growth are positive. The estimated coefficients on log simulated income and Gini coefficient are unstable and not statistically significant. The panel results discussed below give more precise estimates of these parameters. OLS regressions produce the same signs as the instrumental variables results reported in Table IV in all cases, with rays coefficients attenuated relative to the IV results. A quadratic term in the change in the number of rays gives a decreasing and concave function with the linear and quadratic terms jointly significant at the 1 percent level, though not significant individually. Unfortunately, the data are not sufficiently populated over the support of rays to precisely estimate a functional form. I also investigated a variety of alternative highway measures, including mileage and density. All estimated coefficients have the expected sign, and most are statistically significant. The first stage results presented in Table II show that conditional on the control variables, planned rays are a good predictor of actual rays. Including MSA radius as an additional control produces small insignificant coefficients with little change in the other coefficients. This check indicates that the MSA definitions from 2000 are spatially inclusive enough to include all relevant suburban areas.

Though not an implication of the data generating process proposed in Section III, it is possible that the initial level of metropolitan area population may influence central city population growth. For example, larger metropolitan areas may have more communities, providing more options for Tiebout sorting. Alternatively, MSA size may influence spatial patterns in local agglomeration economies, causing firms' relocation choices to respond differently as cities grow. The results in Specification 5 show that controlling for log 1950 population implies little change in the IV estimates.[11]

Table V presents results using the same specifications as Table IV for an additional more select sample and with the inclusion of census division fixed effects. The coefficient on rays

cient is not significantly different from 0. Therefore, only the simpler linear specification is reported. The 1950 central city radius is calculated as $\sqrt{A/\pi}$, where A is 1950 central city area.

11. For similar reasons, initial MSA population density may be an appropriate control. Results using this measure are similar to those using log 1950 MSA population.

TABLE V
COEFFICIENTS FROM LONG-DIFFERENCE REGRESSIONS, 1950–1990

	Census division fixed effects		Specification 1	2	3	4	5
All large MSAs in 1950	No	OLS	.005	−.032	−.059	−.055	−.040
			(.016)	(.015)*	(.014)**	(.014)**	(.016)*
	Yes	OLS	−.012	−.039	−.052	−.051	−.024
			(.014)	(.014)*	(.014)**	(.014)**	(.014)
	Yes	IV	−.026	−.080	−.104	−.104	−.065
			(.020)	(.023)**	(.021)**	(.022)**	(.031)*
All large MSAs in 1950 with central city at least 20 miles from an international border or coast	No	OLS	−.000	−.040	−.062	−.063	−.032
			(.018)	(.019)*	(.017)**	(.016)**	(.019)
	No	IV	−.014	−.101	−.142	−.132	−.098
			(.026)	(.030)**	(.038)**	(.030)**	(.110)
	Yes	OLS	−.019	−.044	−.053	−.053	−.018
			(.017)	(.018)*	(.016)**	(.016)**	(.017)
	Yes	IV	−.028	−.086	−.134	−.130	−.058
			(.023)	(.027)**	(.030)**	(.028)**	(.065)

Notes: Each entry in this table represents the coefficient on rays from a separate regression. The number of rays in the 1947 plan instruments for the change in the number of rays in the IV regressions. The square root of 1950 central city area enters as a control in Specifications 2–5; the change in simulated log income and the change in log MSA population enter in Specifications 3–5; the change in the Gini coefficient enters in Specification 4; log 1950 MSA population enters in Specification 5. Sample sizes are 139 for the top block and 100 for the bottom block. Standard errors are clustered by state of the MSA central city. Standard errors are in parentheses. ** indicates significant at the 1 percent level, * indicates significant at 5 percent level.

remains consistent across samples and is robust to the inclusion of fixed effects, though it is measured imprecisely when 1950 log population is included in the regression. The census division fixed effects estimates are a bit smaller but still significant when the full array of controls is included. OLS estimates of the coefficient on rays in Specification 3 are about −0.05, and IV estimates are about −0.10 in the primary sample. The results in Table V indicate that selecting the sample to exclude the thirty-nine central cities that are within twenty miles of a major lake shore, coast, or international border slightly strengthens the estimated effect of rays on central city population. This is consistent with predictions from a land use model in which the space into which the metropolitan area can expand is exogenously restricted. Results using the most complete sample, including all 240 metropolitan areas for which I could construct data, yield estimated coefficients on rays that are within 0.02 of those reported in Table IV.

One notable pattern in the results reported in Tables IV and V is that estimated instrumental variables coefficients exceed ordinary least squares coefficients in absolute value. If the only source of endogeneity of actual rays to changes in central city population is that state and local governments built highways to serve areas to which people were expected to suburbanize for reasons other than highways, a valid instrument would produce IV estimates that are smaller than OLS estimates in absolute value.

The reverse ordering is consistent with the existence of an additional bias due to a misspecification of the transportation infrastructure in the estimating equation. In particular, most of the IV-OLS gap can be explained by the fact that rays serving central business districts do not fully capture all relevant highway infrastructure for commuting between suburbs and central cities. One test for the existence of this misspecification is to evaluate the IV-OLS gap using the sum of the change in suburban rays and rays serving the central city as the key explanatory variable. This procedure is likely to reduce the bias due to highway misspecification because suburban rays no longer appear in the error term. The difficulty with this procedure is that the measure of suburban rays used is imperfect, as it represents many different road configurations. The OLS estimate of the effect of all rays analogous to that in Table IV Specification 3 is -0.044 while the IV estimate is -0.069, both remaining statistically significant. OLS and IV estimates in the more select sample that excludes coastal central cities are smaller at -0.045 and -0.062, respectively.[12] The similar results given by the two estimators is evidence that the bias due to endogenous highway construction is likely to be small.

Given that the average effect of all rays should be smaller than just that of those connecting the central city with the suburbs, it is prudent to view the estimates of the effect of all rays as a lower bound on the true value of δ_1 from equation (2). It should be noted that this omitted variable problem afflicts the IV estimate in addition to the OLS estimate if the plan is correlated with suburban ray infrastructure, conditional on the control variables in (2). Assuming that suburban rays represent a linear part of the

12. A Hausman test reveals that we cannot reject that the difference in IV and OLS coefficients in the more select sample is only due to sampling error, while the gap in the more complete sample is significant at the 3 percent level. A somewhat less satisfying way to evaluate this misspecification problem is by adding a measure of suburban rays as a control to the regressions. This procedure similarly reduces the IV-OLS gap.

error term ε and an upper bound for the coefficient on suburban rays of -0.05, I calculate a bound on the misspecification bias to δ_1. The omitted suburban rays variable biases the OLS estimate in Table IV Specification 3 toward 0 by 0.01 while it biases the analogous IV estimate away from 0 by 0.03. The biases go in opposite directions because conditional on central city radius, the partial correlation between central city and suburban ray construction is negative whereas the plan predicts positive suburban ray construction.

The remaining IV-OLS discrepancy likely reflects the fact that the planned rays that ended up being built to serve CBDs had a greater effect on central city population decline than endogenously constructed rays. Many endogenous rays may have been built in response to decentralization that had partly occurred by 1950 or to better serve newly suburban areas established in response to constructed planned highways. In addition, many rays built with local funding do not protrude out to the edge of metropolitan areas, and thus do not have the opportunity to serve as large a portion of the suburbs as planned rays connecting to other metropolitan areas. Finally, the discrepancy may reflect the fact that exogenous rays were generally built before endogenous rays, allowing them to have a larger effect on central city population decline.

IV.B. Panel Estimates

There are several distinct advantages to evaluating the effect of highways using long differences. The highway data in 1950 and 1990 are more complete than in the intervening years because of the inclusion of both interstate and non-interstate rays. Moreover, the 1950–1990 differences provide sufficient time for people to move in response to changes in the transportation infrastructure and thus are more likely to accurately reflect long-run equilibrium responses. Unfortunately, there are some drawbacks to using the long differences. While I try to use selection rules to make the sample more homogenous, metropolitan areas are unique in many ways. Evaluating the effects of highways by comparing central city depopulation as a function of highways over different decades within MSAs is thus potentially valuable in allaying concerns about omitted variable bias.

Using the stock of rays as an explanatory variable presents some difficulties in a panel setting. Given the possibility of adjustment costs causing slow transitions to new equilibria due to durable housing, for example, it is not clear what the appropriate timing of the highway measure is for the intervening years of the panel. If, for example, it takes twenty years for residential location patterns to respond fully to changes in highway infrastructure, panel results without somehow smoothing rays or instrumenting will be significantly attenuated towards 0 because they allow for adjustment over at most a ten-year period.[13] If it takes time for individuals to adjust to new transportation infrastructure, and this new infrastructure is arriving continuously over time as it was in 1960 and 1970, we can also think of highways as being measured with error. Two observations for 1970 may represent one city that got its rays in 1969 and another that got its rays in 1961. Such inconsistent measurement of timing is likely to bias the estimated coefficient toward 0. Furthermore, highways often opened in stages such that in a given year only part of a given ray may have been available for use. I evaluate the speed of convergence to new equilibria by examining results analogous to those reported in Table IV using 1950–1970 differences instead. The resulting coefficient on highway rays is somewhat smaller but still significant. OLS estimates of the coefficient on rays using Specification 3 is −0.03 while the IV coefficient is −0.06, about one-half of the estimates from the full 1950–1990 differences. This evidence is consistent with Rappaport [2004], who shows through simulation that even small migration frictions can mean it takes many years for an equilibrium with new residential locations across metropolitan areas to take hold.

Table VI repeats the regressions from Tables IV and V in levels including MSA fixed effects to account for unobserved metropolitan area characteristics. Each observation represents an MSA in a census year between 1950 and 1990. Standard errors are clustered by the state of the central city, allowing for MSA-level serial correlation and cross-sectional intrastate correlation in the error term. Due to limitations in available data, the variable measuring the number of rays in the panel regressions only reflects interstate highways, not other limited access highways. I count roads with at least one mile open to traffic in a given metropolitan area towards the MSA's stock of rays at any given point in time. The number of MSAs in the sample is smaller than

13. Although slow convergence rates toward new spatial equilibria are likely, it is not clear what highway construction year should be used. If agents have rational expectations, they should react when announcements of new highways are made. Specification checks show, however, that using the timing of planning, instead of completion, does not change results much.

TABLE VI
PANEL IV REGRESSIONS OF THE DETERMINANTS OF CONSTANT GEOGRAPHY CENTRAL CITY POPULATION, 1950–1990

	Large MSAs in 1950					
	Log central city population					
	1	2	3	4	5	6
Number of rays	−0.111	−0.142	−0.140			
	(0.016)**	(0.026)**	(0.028)**			
(1990 Rays) × (Fraction of Ray miles completed at t)				−0.097	−0.089	−0.086
				(0.016)**	(0.012)**	(0.013)**
Log simulated income		−0.083	−0.061		−0.288	−0.229
		(0.117)	(0.109)		(0.075)**	(0.077)**
Log MSA population		0.266	0.263		0.294	0.286
		(0.104)*	(0.105)*		(0.100)**	(0.098)**
Simulated Gini coefficient			−0.623			−1.415
			(1.106)			(0.847)
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
R-Squared	0.20	0.22	0.22	0.14	0.56	0.57

Notes: The instrument used is (rays in the plan) × (MSA mileage of highways running through the central city at time t)/(MSA mileage of highways running through the central city in 1990). Standard errors are clustered by the state of the central city. Standard errors are in parentheses. ** indicates significant at the 1 percent level, * indicates significant at the 5 percent level. First stage results are in Table II. Each regression includes 132 MSAs with five observations each, one for each year 1950–1990. There are fewer MSAs in this sample than that in Table IV because of lack of census tract data for seven MSAs in 1960.

in the long-difference regressions because 1950-definition central city population is not available for all cities in all years.

Panel OLS regressions using discrete rays as an explanatory variable produce coefficients of less than −0.02. These small coefficients reflect adjustment costs and measurement problems. To get at estimating a long-run effect, I also estimate the panel regressions using a measure of smoothed rays. I define smoothed rays as the stock of rays in 1990 multiplied by the fraction of the rays' mileage completed in each MSA at each point in time. Smoothing the rays increases the coefficients in Specification 3 to −0.04, about 0.02 smaller in magnitude than the OLS coefficients in Tables IV and V.

Table VI presents instrumental variables results with MSA fixed effects. Specifications 1 to 3 use discrete rays as the explanatory variable while Specifications 4–6 use smoothed rays. The panel IV results are consistent with results from the long differences whether endogenous rays are smoothed or not. The coefficient on discrete rays is -0.14, and the coefficient on smoothed rays is -0.09. By using only the part of the rays correlated with the aggregate flow in federally funded nationwide construction, IV smooths out the discrete rays measure used in Specifications 1–3. Allowing rays to take on fractional values allows even small suburbanization effects to show up in the coefficient. The fact that the IV results are similar for the two different highway measures supports the claim that measurement error is the likely culprit to explain the small OLS coefficients for discrete rays. IV results using other samples of MSAs are similar.

The coefficient on log MSA population is between 0 and 1 in both the long difference and panel regressions, though it is smaller in the panel regressions. This is consistent with predictions of theory.[14] The coefficient on simulated log income is negative, statistically significant and more stable in the panel regressions. The observation that central city population decline was more rapid in decades that had faster income growth is consistent with current land use patterns in which the poor live in cities. This coefficient may reflect employment opportunities in the suburbs that differ by income. A mean-preserving spread in the simulated income distribution is also estimated to decrease the population of the central city. This result may reflect Tiebout sorting or a land use mechanism.

IV.C. Additional Robustness Checks

While I have demonstrated that the estimated effect of highways on central city population growth is robust to an array of samples, specifications, and several estimators, it is still possible that the number of rays assigned in the 1947 highway plan may have been developed in response to central city decline in the early part of the 20th century. If accurate, this story implies that serial correlation in central city population growth rates could generate a spurious relationship between the plan and central city population growth rates after 1950. Alternatively, the 1947 plan could have been developed in response to very good forecasts about future commuting demand. To evaluate these possibilities, I perform a falsification exercise by examining the extent to which we can predict changes in central city population between 1910 and 1950 using

14. In a monocentric model with quasilinear utility, it is straightforward to show that at any number of rays, the elasticity of central city population with respect to MSA population is bounded between 0 and 1.

the 1947 plan. This exercise is limited in several dimensions by data availability. In particular, there is no income data in 1910 and only limited data on central city area. Among the sixty-three central cities whose areas I observe in 1910, average area more than doubled by 1950. I do not have a good way of accounting for the associated change in central city geography. Regressions analogous to those reported in Table IV absent the income variables and using rays in the 1947 plan as an instrument for future construction yields a statistically insignificant coefficient of -0.003. Results using the reduced form are of similar magnitude. This 0 result is robust to a range of specifications, including controlling for central city radius in 1910 and 1950. A similar exercise using data from 1940 and 1950 yields similarly small estimates.

To this point, I have maintained the assumption that MSA population is exogenous. Considering a system of cities, however, requires endogenizing MSA population. A model incorporating a system of cities suggests that holding the infrastructure in all other cities constant, more rays would attract new migrants to a city because they would cause the equilibrium utility level to increase. Thus, the effect of a ray on central city population, holding MSA population constant, should be larger in magnitude than the effect if it is allowed to vary due to migration. That is, accounting for endogenous MSA population should increase the magnitude of the estimated effect of highway rays on central city population. Following Glaeser, Kolko, and Saiz [2001], I instrument for population growth between 1950 and 1990 with mean January temperature and annual precipitation in the central city. While not ideal, in the panel regressions I instrument for the level of MSA population with its ten-year lag. IV coefficients on rays when instrumenting both for rays and MSA population using weather variables are within 0.01 of those reported in Tables IV and VI and remain precisely measured. Conditional on the other exogenous variables, these two weather variables are good predictors of MSA population. Results are similarly robust to instrumenting for MSA population using an industry shift share.

Another strategy for handling the potential endogeneity of MSA population is to measure suburbanization as log central city population share. I choose not to measure suburbanization this way in order to straightforwardly connect the empirical analysis to the aggregate trends shown in Table I and because it would complicate interpretation of the coefficient on rays. This measure would make it impossible to separate out the effect of highways on migration within versus between metropolitan areas without making an assumption about the elasticity of central city population with respect to MSA population. However, for some purposes log central city population share is a more intuitive measure of suburbanization. Regressions using this alternate measure of suburbanization as the dependent variable yield estimates that are slightly larger in absolute value to those reported in Table IV. For example, the analogous OLS estimate to that in Table IV Specification 3 is -0.10 while the analogous IV estimate is -0.14. These larger estimates reflect the fact that MSAs with more new highways had greater net in-migration.

V. Conclusions

Urban transportation improvements are an important element in explaining post-World War II suburbanization in the United States. This paper uses unique data to estimate the effect of new highway rays on suburbanization using a 1947 plan of the interstate highway system as a source of exogenous variation. Taken together, long-difference and panel results imply an effect of one new highway ray on the change in log constant geography central-city population of about -0.09. Population in large 1950 definition central cities fell on average by 28 percent between 1950 and 1990. If the average city had not received one exogenously assigned highway (2 rays), estimates thus indicate that its population decline would have been cut by more than half. Empirically, cities received 2.6 rays on average between 1950 and 1990, though not all of these were randomly assigned.

I evaluate the importance of highways for explaining central city population decline by examining the counterfactual evolution of aggregate central city population where no highways were constructed. A coefficient of -0.09 implies that had the interstate highway system not been built, aggregate central city population in MSAs in the primary sample would have increased by 8 percent between 1950 and 1990. This aggregate effect is larger than the average effect because cities with higher populations received more rays on average. An additional three percentage point increase in aggregate central city population would have occurred had limited access highways that are not part of the interstate system not been built. These counterfactual increases come in contrast to the actual decline of 17 percent in aggregate central city population for the primary sample. In the more select sample of MSAs considered, counterfactual population would have changed by between −4 percent without interstate highways and −1 percent without any highways. Based on these calculations, highways account for about one-third of the decline in aggregate central city population relative to that in entire metropolitan areas between 1950 and 1990. The Appendix details the procedure for making these calculations.

Although this paper provides estimates of the treatment effect of highways on central city population, it has less to say about the mechanism by which highways cause suburbanization. Beyond the pure residential land use framework proposed in Section III, one important potential mechanism is that new highways allow firms to move to the suburbs. Indeed, urban employment decentralization was more rapid than residential decentralization between 1950 and 1990. In 1950, 20 of 40 million MSA jobs were in central cities while by 1990 only 27 of 87 million jobs remained in central cities. However, as seen in Table I, 45 of 93 million MSA residents lived in central cities in 1950 when compared with 51 of 160 million in 1990. In addition to chasing suburbanizing workers, firms may have relocated because highways free manufacturing firms from shipping through a port or downtown rail hub or because highways allow local agglomeration economies to operate over longer distances. Estimates presented in this paper should thus be interpreted as reflecting changes in equilibrium land use patterns that have occurred through a potentially complicated set of interactions between firms and workers. The potential existence of road congestion may also influence interpretation of the results. However, data collected by the Texas Transportation Institute indicates that congestion is probably not a major concern. The average ratio of free-flow to congested highway speeds was 1.16 in the primary sample and 1.12 in the more select sample of MSAs in 1992.

Two broad topics related to the questions addressed in this paper are left for future research. The first is to gain a better understanding of the mechanisms by which innovations to the urban transportation infrastructure have caused the changes in land use patterns documented in this paper. Second, it is evident that there remains much more to be learned about the causes of suburbanization. The importance of the changing income distribution, changes in the distribution of household size and structure, and changes in local public goods provision for explaining changes in the spatial distribution of population and employment in U. S. metropolitan areas all invite future investigation.

Appendix

A. Data

The *County and City Data Books* (CCDB) report decennial census data aggregated to counties and cities of at least 25,000 inhabitants. I aggregate counties to form metropolitan areas based on their definition in 2000, with a few minor alterations. I combine PMSAs with no central city to adjacent PMSAs with a central city and I combine PMSAs with adjacent central cities. In New England, I use census-defined NECMAs. I use the same metropolitan area geography for all years, aggregating data reported at the county level. I use the CCDB city population numbers for all central cities in 1950. In later years, I use only the CCDB for the population of central cities whose areas reported in the CCDB increased by less than 10 percent when compared with 1950. I assign each metropolitan area only one central city, using the census-defined central city with the largest population in 1950. The CCDB is also the source of weather variables for central cities.

Printed maps of central cities that expanded geographically over time are taken from the *1950 Census of Population and Housing*. I identify census tracts that fall within 1950 central city geography using these maps and digital maps of census tracts in 1970 and 1980. Reported conversion factors map 1980 tracts to 1990 tracts. Tracts from 1960 whose centroids fall inside the constructed 1950 geography as approximated with 1980 tracts were aggregated to form the 1960 cross-section. I use the latitude and longitude of 1960 tract centroids reported in the *Elizabeth Mullen Bogue File* [Bogue 1975]. The *Neighborhood Change Database* provides consistent census tract geography in 1970 and 1990 for Table III.

I inspected each metropolitan area in a *Road Atlas* from 2004 to identify potential highway rays. Rays consist of interstate or non-interstate highways that pass within one mile of the central business district (as defined below) and connect to an area outside the 1950-definition central city. The *Form PR-511 Database* contains opening dates of each segment of the interstate highway system kept by the Federal Highway Administration. Each record lists the state, route number, milepost, segment length, funding source, and date(s) on which various stages of construction were completed.

Because the digital version of the database was lost, this data set was scanned and hand-edited. This data set includes all federally funded portions of the system and interstates that were in place in 1956. The data set includes only about half of the locally-funded mileage constructed after 1956. The remaining locally funded interstate mileage was added by cross-referencing Internet reports of "roadfans" with the "Route Log and Finder List" published by the Federal Highway Administration. Only one non-interstate limited access highway ray was found to exist in 1950. I match the PR-511 data to a digital map of the interstate highway system to create a data set of the miles of each route number in each county in each year. I aggregate counties appropriately to form highway mileage in each metropolitan area over time. For panel regressions, I count a ray as open in a given year if at least one mile of the road exists in the metropolitan area. Because of limited time series data on non-interstates, only interstate highways are used in the panel analysis.

Central business district definitions used for Table III are taken from the *1982 Economic Censuses' Geographic Reference Manual.* This book reports all census tracts that make up the CBD in the opinion of local businesspeople. All CBD distances are to the centroid of the agglomeration of these census tracts.

The 1950–1990 census Public Use Microdata one percent Samples (PUMS) are used to build the evolution of skill prices by one-digit industry at the state level. The distribution of skill prices is built using the annual wage and salary income of each individual working at least forty-eight weeks in the calendar year prior to each census. The self-employed are excluded. The simulated income distribution calculated for each MSA uses the full distribution of wages in each year weighted by the industrial mix in 1940 for all people living in states that do not include any part of the MSA in question.

The full data set used and the code used to build it and create the tables are available upon request from the author.

B. Construction of Counterfactual Population

I take the predicted number of rays from a separate first stage regression based on Specification 5 in Table II Panel B for each sample to determine the number of "exogenous rays" in each metropolitan area in each year. To build the time series for total rays I assume that non-interstate rays were built at the same rate as interstate rays in each MSA. I then first-difference the predicted rays from the first stage and multiply the resulting

APPENDIX TABLE:
SUMMARY STATISTICS

Variable	1950	1990	1990–1950	Min	Max
Fraction of population living inside 1950 geography central city	.47 (.16)	.22 (.11)	−.25 (.10)	−.51	−.03
Fraction of population living inside of central city	.47 (.16)	.35 (.19)	−.12 (.14)	−.42	.33
Log population of 1950 geography central city (millions)	−1.82 (.94)	−2.10 (.99)	−.28 (.28)	−.81	.72
Log population of politically defined central city (millions)	−1.82 (.94)	−1.64 (1.00)	.18 (.57)	−.81	2.22
Log MSA population (millions)	−.99 (.91)	−.43 (.98)	.56 (.40)	−.21	1.95
Total number of rays through central city	.07 (.33)	2.66 (1.70)	2.59 (1.66)	0	7
Number of interstate rays through central city	.07 (.33)	2.34 (1.58)	2.27 (1.54)	0	7
Number of planned rays	2.55 (1.51)	2.55 (1.51)	–	0	7
Total rays in the MSA	.14 (.50)	4.38 (2.40)	4.24 (2.36)	0	14
Central city radius (miles)	3.18 (1.63)	5.04 (2.77)	1.86 (2.27)	1.16[a]	11.98[a]
MSA radius (miles)	25.99 (10.42)	25.99 (10.42)	–	8.48	93.17
Simulated mean log income using 1940 employment distribution	9.61 (.12)	10.27 (.05)	.66 (.06)	.56	.89
Gini coefficient of simulated income distribution	.33 (.03)	.37 (.01)	.04 (.02)	−.02	.07
Median log household income	9.99 (.16)	10.60 (.15)	.60 (.16)	.27	.99

Notes: The sample size is 139 and consists of the same MSAs used in the regressions in Table IV. Standard deviations are in parentheses. Min and max refer to the 1950–1990 difference unless the variable does not change over time or has a superscript.
a. This number refers to the min or max of the 1950 level, not the change from 1950 to 1990.

number by 0.09 to derive the log difference in central city population that can be attributed to new exogenous highways during each decade between 1950 and 1990 for each MSA. I subtract the estimated log difference in central city population due to exogenous highway expansion (which is negative) from the observed log difference in central city population during each decade, calculate the counterfactual log levels in 1960–1990 based on the 1950 cross-section and exponentiate to put the estimates in terms of population, performing the appropriate Jacobian transformation. Finally, I sum the resulting populations over the MSAs in the relevant sample.

BROWN UNIVERSITY, DEPARTMENT OF ECONOMICS

REFERENCES

Alonso, William, *Location and Land Use*, MA (Cambridge, MA: Harvard University Press, 1964).
Baum-Snow, Nathaniel, "Suburbanization and Transportation in the Monocentric Model," *Journal of Urban Economics*. In press (2007).
Bogue, Donald, *Elizabeth Mullen Bogue Census Tract File, 1960* (Chicago: University of Chicago, 1975).
Boustan, Leah Platt, "Was Postwar Suburbanization 'White Flight?' Evidence from the Black Migration," Manuscript, 2006.
Censuses of Population and Housing, 1910 and 1940–1990, (Washington: U. S. Bureau of the Census).
Clark, Colin, "Urban Population Densities," *Journal of the Royal Statistical Society,* CXIV (1951), 490–496.
County and City Data Books 1947–1994 (Washington: U. S. Bureau of the Census).
Cullen, Julie Berry and Steven D. Levitt, "Crime, Urban Flight, and the Consequences for Cities," *Review of Economics and Statistics,* LXXXI (1999), 159–169.
Geographic Reference Manual: 1982 Economic Censuses, (Washington: U. S. Bureau of the Census, 1983).
Glaeser, Edward L., Jed Kolko, and Albert Saiz, "Consumer City," *Journal of Economic Geography,* I (2001), 27–50.
McDonald, John F., "Econometric Studies of Urban Population Density: A Survey," *Journal of Urban Economics,* LXX (1989), 361–385.
Michaels, Guy, "The Effect of Trade on the Demand for Skill—Evidence From the Interstate Highway System," Manuscript, 2006.
Mieszkowski, Peter, and Edwin S. Mills, "The Causes of Metropolitan Suburbanization," *Journal of Economic Perspectives,* VII (1993), 135–147.
Mills, Edwin S., "An Aggregative Model of Resource Allocation in a Metropolitan Area," *American Economic Review,* LVII (1967), 197–210.
Muth, Richard F., *Cities and Housing* (Chicago: University of Chicago Press, 1969).
Neighborhood Change Database, 1970–2000 (GeoLytics in association with the Urban Institute, 2004).
U. S. Department of Transportation, *America's Highways: 1776–1976: A History of the Federal-Aid Program* (Washington: U. S. Government Printing Office, 1977).
——, *Route Log and Finder List* (Washington: U. S. Government Printing Office, 2001).
U. S. House of Representatives, *Interregional Highways: Report of The National Interregional Highway Committee Outlining and Recommending a National System of Interregional Highways* (Washington: Federal Works Agency, 1944).
Rappaport, Jordan, "Why Are Population Flows So Persistent?," *Journal of Urban Economics,* LVI (2004), 554–580.
Reber, Sarah, "Court-Ordered Desegregation: Successes and Failures in Integration since Brown vs. Board of Education," *Journal of Human Resources,* XL (2005), 559–590.
Road Atlas: United States, Canada, Mexico, (Maspeth, NY: American Map, 2004).
Rose, Mark H., *Interstate Express Highway Politics 1939–1989* (Knoxville: University of Tennessee Press, 1990).
Status of Improvement of the National System of Interstate and Defense Highways File PR-511 (Washington: U. S. Federal Highway Administration, 1993).
Tiebout, Charles M., "A Pure Theory of Local Expenditure," *Journal of Political Economy,* LXIV (1956), 416–424.