Highways, Commuting and Trade: Unpacking Suburban Growth

Daniel Velásquez*

University of Michigan

Job Market Paper

Most updated version here

November 18, 2024

Abstract: I quantify the effect of U.S. Interstate Highway System on suburbanization. I leverage variation in highway construction dates and driving time reductions to show that new highways affect both commute costs and trade costs. I find that, on average, a rise of one standard deviation in market access raises population and employment by about 10% after 40 years at the county level. However, a rise of one standard deviation in commuting access raises population and employment by about 1% after controlling for market access changes. I develop a quantitative model of trade, commuting, migration, and scale externalities. I map the model to the average effect of market and commuting access through indirect inference. Using the calibrated model, simulations show highways account for 15% of suburban growth and 33% of the decline in urban cores. I find that suburbs developed and urban cores declined not only because of reduced commute times, but also, because of trade costs reductions.

*Department of Economics, University of Michigan (email: danielvc@umich.edu). I am very grateful to Luis Baldomero-Quintana, Dominick Bartelme, John Bound, Agostina Brinatti, Sun Kyoung Lee, Andrei Levchenko, Sebastian Sotelo, Stephen Terry, and Dean Yang for comments and suggestions. The standard disclaimer applies.

1 Introduction

Suburbanization has become a defining feature of the U.S. since 1950, with suburban counties emerging as major population and employment centers, while urban cores have lost residents and jobs. This spatial reorganization has far-reaching consequences—contributing to environmental degradation, reduced density, and rising fiscal and maintenance burdens for cities.[1] Nonetheless, the forces driving this transformation remain debated.[2] The Interstate Highway System, a key feature of the postwar era, is often seen as a catalyst for suburbanization.[3] Much of the debate centers on how Highways lowered commuting costs to urban centers, facilitating suburban living by enabling access to urban jobs. Yet, as I argue in this paper, by reducing trade costs, highways may have also moved jobs away from urban cores, encouraging the growth of decentralized employment hubs. Against this backdrop, this paper addresses two main questions: To what extent did the construction of the Interstate Highway System contribute to the growth of suburbs and the decline of urban cores in the U.S.? And, to what extent can these effects be attributed to reduced trade costs, beyond just reduced commuting costs?

I revisit the impact of the Interstate Highway System and offer a complementary view. I argue that reductions in trade costs—alongside reductions in commuting costs—were pivotal in shaping suburban growth and urban decline. I highlight how city networks shape the effects of highways on population and employment by improving market access. For instance, from Baltimore's viewpoint, shorter driving times on I-95 to New York benefit firms along this route more than those along I-70 to Pittsburgh, given New York's sheer size. In addition, enhanced market access along I-95 attracts new residents to Baltimore's suburbs. Hence, as trade costs fall, jobs and people increasingly gravitate toward Baltimore's suburbs with direct access to I-95. We can be sure that these effects are driven by improved market access and not commute access. While trade cost reductions can impact bilateral relationships over long distances, commute cost reductions do not. The long distance between Baltimore and New York makes commuting impractical; hence, New York's hypothetical influence on Baltimore's suburban growth occurs through market access.

I also emphasize that the observed patterns of suburban growth and urban core decline can only

[1] See Brueckner (2000); Glaeser and Kahn (2004); Ewing et al. (2003).

[2] See Cullen and Levitt (1999) on crime; Boustan (2010) on racial tensions and white flight; Reber (2005) on schooling; Baum-Snow and Hartley (2020) on amenity value, see Baum-Snow (2007) on highway construction.

[3] See Baum-Snow (2007).

be understood by taking into account both trade and commute cost reductions. Accounting for both types of cost reductions is particularly important in settings where the agglomeration of people and businesses creates productivity and attractiveness gains—what economists call scale externalities. In these settings, scale externalities can amplify the effect of reducing trade and commute costs reductions. For instance, lower commute costs between Baltimore's suburbs and its core stimulate suburban growth as people begin to move in. This influx enhances suburban amenities due to scale effects. Subsequently, further reductions in trade costs to New York create even stronger incentives for people to move to Baltimore's suburbs. With suburban amenities already boosted from migration (triggered by the initial drop in commute costs), the suburbs become increasingly attractive in response to the trade cost reductions.

To answer these research questions, I proceed in two steps. First, I show empirically that highways substantially reduced trade costs between counties, beyond simply reducing commuting costs. Second, I use the results of the empirical exercise to discipline a general equilibrium model. I then quantify the role of highways in the rise of the suburbs and the decline of urban cores due to trade cost reductions, alongside the fall in commuting costs.

In the first step, I estimate the impacts of reducing trade and commute costs on population and employment at the county level. I start by constructing a new data set with dated and georeferenced placements of highways by mile. These data allows me to compute bilateral driving times between counties for each decade. Using these data, I leverage variation in highway construction dates and driving time reductions to estimate how changes in access to employment centers (i.e., commuting access) and in access to markets for shipping goods (i.e, trade access) affect both population and employment growth in U.S. counties between 1950 and 2020. I find that a one standard deviation increase in trade access leads to a 9.6% rise in population and an 11.8% increase in employment, while a one standard deviation increase in commuting access results in only a 0.03% rise in population and a 1.5% rise in employment. In other words, on average across counties—and regardless whether they are suburbs or not—trade cost reductions have a greater role than commute costs reductions in driving both population and employment growth. This result holds across a battery of robustness tests, including concerns about the endogeneity of highway placement.

In the second step, I develop a quantitative spatial model incorporating trade, commuting, myopic migration, and externalities in production and amenities, with both tradable and non-tradable sectors and a construction sector to add quantitative realism. I use the results of the empirical exercise to discipline the general equilibrium model via indirect inference. I estimate amenity externalities by targeting the reduced-form evidence from the first step. These externalities amplify migration responses to trade and commute cost reductions by making growing locations more attractive. Following Bartelme, Li, and Velasquez (2024), I assume amenity externalities are active only within a middle range of population sizes.[4] I refer to this assumption as heterogeneous amenity externalities, in contrast to the standard, constant elasticity framework.

Heterogeneous amenity externalities give the model greater flexibility to match the average effects of commute and trade access estimated in the first-step. Based on my findings, without amenity externalities, migration responses to trade cost and commute cost reductions are too weak to match the data. Under constant scale externalities, the model can match population responses to either trade or commute access but not both. Only with heterogeneous scale externalities, the model can match both responses. The intuition is as follows. According to the data, the drop in commuting costs was substantial in dense suburban areas, while trade costs reductions were so in less dense areas. To match the empirical fact that improvements in commute access generate minimal population responses, amenity externalities cannot be active in denser areas where commute cost reductions were substantial. In contrast, to match the fact that trade access improvements generate large responses on average, amenity externalities need to be active in less dense areas where trade cost reductions were significant. Empirically, I estimate that amenity externalities become inactive in areas with populations above roughly 200,000. Further, the data does not support a constant elasticity framework, as I can confidently reject a threshold 'close to infinity' for amenity externalities.

Once the model is calibrated to the U.S economy, I use it to quantify the extent to which highways contribute to suburban growth and urban core decline. I also use it to separately understand the contributions of trade and commute costs on suburbanization, while exploring their interaction. I simulate a counterfactual scenario without the Interstate Highway System, holding all other shocks between 1960 and 2020 constant. I evaluate three cases: highways reduce only commute costs, only trade costs, and both simultaneously. Comparing observed data to the scenario where highways reduce both trade and commute costs, I find that highways account for 15% of suburban growth and 33%

[4] One way to interpret this framework is by considering that congestion forces dominates in the largest locations, while the smallest locations lack enough population to generate endogenous amenities. TFP externalities, calibrated from the literature, remain constant, preserving large cities' productivity advantages.

of urban core decline. When highways reduce only commute costs, the model indicates that these reductions explain 88% of highways' impact on suburban growth. Yet commute cost reductions do not account for the decline in urban cores, as suburban expansion in this case draws primarily from rural areas. Conversely, adding trade cost reductions on top of commute cost reductions explains the remaining 12% of highways' impact on suburban growth and the entire effect of highways on urban core decline.

The intuition behind these results can be understood in three parts. First, when trade costs fall and commute costs remain unchanged, access to goods improves across long distances, making remote areas more attractive. Thus, rural areas grow in response to trade cost reductions (as shown in a counterfactual scenario where highways reduce only trade costs but not commute costs). Second, commute cost reductions, with trade costs held constant, decentralize activity over shorter distances: suburbs gain better access to jobs in core cities. Rural areas, too remote to benefit from reduced commute costs, lose out as suburbs grow at their expense due to stronger suburban-core connections. Finally, when both trade and commute costs decrease together, the combined effects steer growth toward suburbs. This shift toward suburbs occurs because migration in response to the initial drop in commute costs boosts local amenities in the suburbs, making them even more attractive when trade costs fall. Suburbs thus emerge as middle-ground hubs, benefiting from both trade and commute access. The interaction between costs reductions and amenities helps explain the observed data patterns—where suburbs grow while urban cores shrink.

In this research, I contribute to four strands of the literature. First, I advance the debate on the causes of suburbanization.[5] I argue that highways drive suburbanization not only by reducing commute costs (Baum-Snow, 2007; 2020; Baum-Snow et al., 2017) but also by reducing trade costs (Michaels, 2008; Faber, 2014; Duranton and Turner, 2014), which further accelerates suburbanization and contributes to urban core decline. Furthermore, I highlight the importance of city networks in shaping the trade access effects on suburbanization. Additionally, I offer a new quantification of highway effects that considers not only cores and suburbs but also smaller cities and rural areas. With a full geographical perspective, it becomes evident that commute cost reductions centralize activity around core-suburb links and away from rural areas. However, simultaneous trade and commute cost reductions are essential to explaining the observed suburban growth alongside core decline.

[5] See Cullen and Levitt (1999) on crime; Boustan (2010) on racial tensions and white flight; Reber (2005) on schooling; Baum-Snow and Hartley (2020) on amenity value, see Baum-Snow (2007) on highway construction.

Second, I contribute to the empirical and theoretical literature evaluating the impact of new transport infrastructure by examining the interaction between trade costs and commute costs.[6] While infrastructure is typically linked to reductions in either trade or commute costs, I analyze both dimensions together. Using my new data set of bilateral driving times, I show that, once trade access is accounted for, the average effect of commute access on population and employment growth becomes negligible. In contrast, trade access drives significant growth in these outcomes. I also introduce tools from macroeconomics, such as indirect inference, into spatial economics, allowing me to utilize both partial and general equilibrium responses to better inform parameter estimates. I argue that to capture a policy's general equilibrium effects, models can benefit from calibration based on empirically observed general equilibrium responses, which in some cases may offer advantages over partial equilibrium approaches. Lastly, the model is designed to capture the average reduced-form effects from the data. Because the introduction of highways induced heterogeneous changes in market and commuting access, the model then allows me to characterize the contributions of each factor separately and explore their interactions.

Third, I contribute to economic geography models of trade and migration by exploring how migration amplifies the effects of changes in trade and commute access.[7] My work builds on Monte et al. (2018), who emphasize commuting patterns, by providing direct evidence of how shifts in trade and commute access drive population and employment growth. I introduce amenity externalities, a key factor regulating migration responses, to match the model's predictions with observed patterns in population and employment. I further examine how the interaction of trade, commuting, and migration drives suburbanization in the U.S., demonstrating that migration intensifies the effects of both trade and commute access.

Fourth, there is a long tradition of research on scale externalities, commonly known as agglomeration economies.[8] Some of this literature focuses on amenity externalities.[9] I provide new estimates

[6] Trade: Michaels, 2008; Duranton et al., 2014; Faber, 2014; Allen and Arkolakis, 2014, 2023; Donaldson, 2018; Monte et al., 2018; Sotelo, 2020; Baldomero-Quintana, 2024; Frye, 2024. Commuting: Baum-Snow, 2007, 2020; Duranton and Turner, 2012; Monte et al., 2018; Heblich et al., 2020; Baum-Snow et al., 2020; Severen, 2021; Zarate, 2021; Brinkman and Lin, 2022; Tsivanidis, 2023; Weiwu, 2023, 2024; Chen et al., 2024; Velasquez, 2024.

[7] Caliendo et al., 2019; Tombe and Zhu, 2019; Fan, 2019; Caliendo et al., 2021; Heiland and Kohler, 2022; Cai et al., 2022; Pellegrina and Sotelo, 2024; Brinatti, 2024

[8] See Krugman (1991), Ciccone and Hall (1996), Glaeser and Maré (2001), Lucas and Rossi-Hansberg (2002), Duranton and Puga (2004), Combes, Duranton, and Gobillon (2008), Allen and Arkolakis (2014, 2023) , de la Roca and Puga (2017), Kline and Moretti (2014), Ahlfeldt, Redding, Sturm, and Wolf (2015).

[9] Bayer, Ferreira, and McMillan (2007), Diamond (2016), Leonardi and Moretti (2023), Almagro and Domínguez-Iino (2024), De la Roca, Parkhomenko, and Velasquez (2024).

of dynamic amenity externalities and evidence of heterogeneous scale elasticities using an indirect inference approach. Donaldson and Allen (2023) examine path dependence by showing how dynamic amenity and production externalities can create multiple steady states. Bartelme, Li, and Velasquez (2024) extend this framework by incorporating heterogeneous scale elasticities, showing how a temporary shock can have lasting impacts on mid-sized locations, but minimal effects on large ones, when agglomeration economies are strong but concentrated in mid-sized areas. My contribution lies in demonstrating that the heterogeneous scale externalities framework formulated by Bartelme, Li, and Velasquez (2024) not only provides realistic calibration in contexts where path dependence is expected following temporary shocks. My work also shows that this framework can produce realistic calibrations that match observed general equilibrium responses more effectively than the constant scale formulation, even in contexts where path dependence may or may not be expected.

2 Background, Data and Stylized Facts

2.1 Background and Data

Background. Though conceived in 1947 with a 37,324-mile plan—designed primarily for national defense, not urban development—the Interstate Highway system did not truly materialize until the Federal-Aid Highway Act of 1956 spurred construction of 42,800 miles of interconnected freeways in many construction waves. By 1990, virtually all of the original plan was executed, transforming the nation's road network from limited-access miles to a vast, integrated grid. Crucially, many highways were constructed following a "ray" pattern, connecting outlying areas directly to central city cores (Baum-Snow, 2007). This reshaped the structure of U.S. cities by reducing commute times between suburbs and urban cores. Similarly, highway segments connected many cities, effectively reducing trade costs between U.S. markets (Duranton et al., 2014).

Data. I compile several datasets. To study population characteristics at the county level, I use the Decennial Census and the American Community Survey Data (1950-2020).[10] To document the distribution of jobs, I rely on the County Business Patterns Database (1950-2020) by Eckert et al., (2021). To map out commuting flows over space and time, I rely on the Journey to Work Database

[10] To have a consistent definition of census counties, I interpolate them using Eckert et al. (2020) crosswalks. Elsewhere, I also rely on Census-tract level data. I interpolate census-tracts using crosswalks from Lee and Lin (2018) and the Longitudinal Tract Data Base.

(1970-2020).

To construct bilateral driving times between counties, I utilize the 2005 Highway Performance Monitoring System (HPMS), the PR-511 data set,[11] and various historical and planned transportation data. I merge PR-511 files with HPMS road data to georeference highway construction dates. Using the Dijkstra algorithm and a speed of 70 miles per hour, I calculate bilateral commuting times between county centroids for each decade from 1960 to 2000.[12] I also compute driving times assuming highways followed the exact paths of the 1947 Interstate plan, the 1920 Pershing Map, the 1528-1850 historical routes of exploration, and railroads around 1898 (Baumsnow, 2007; Duranton and Turner 2012; Michaels et al, 2019; Frye, 2023; Brinkman and Lin, 2024).[13]

After interpolating county-level data across decades and excluding Puerto Rico, Alaska, and other islands, the final dataset includes 3,092 counties for each decade from 1950 to 2020.

County classification. I use the 1990 NCHS Urban-Rural Classification Scheme for counties. This classification divides counties in large metropolitan areas (1 million+ population) into two groups: "core" metro counties (those containing all or part of the largest central city) and "fringe" metro counties (suburbs).[14] It also includes medium metros (population 250,000–999,999), small metros (50,000–249,999), and nonmetropolitan counties with cities (those with cities of 10,000 or more residents) and without cities (do not contain any part of a city of 10,000 or more residents).

This classification is ideal to study the spatial distribution of economic activity. First, the NCHS scheme captures not just population size but also functional relationships—such as the distinction between core and fringe counties in large metros—making it ideal for studying commuting patterns and suburbanization trends. Second, by categorizing both metropolitan and non-metropolitan areas, it allows for comparative analysis across a full continuum of urbanization, making it easier to assess how suburban areas differ from cores, medium and small metros, and rural areas. Moreover, every county in the U.S. is assigned a classification.

Figure 1 illustrates how the NCHS classification captures the spatial structure of metropolitan and non-metropolitan areas, distinguishing between core, suburban, and smaller metro regions. The figure

[11] The PR-511 data set identifies the year of construction of each highway segment (Baum-Snow, 2007; Baum-Snow, 2020; Weiwu, 2024).

[12] Centroids were determined using 2000s population density data from WorldPop.

[13] Procedures are detailed in Appendix A.3.

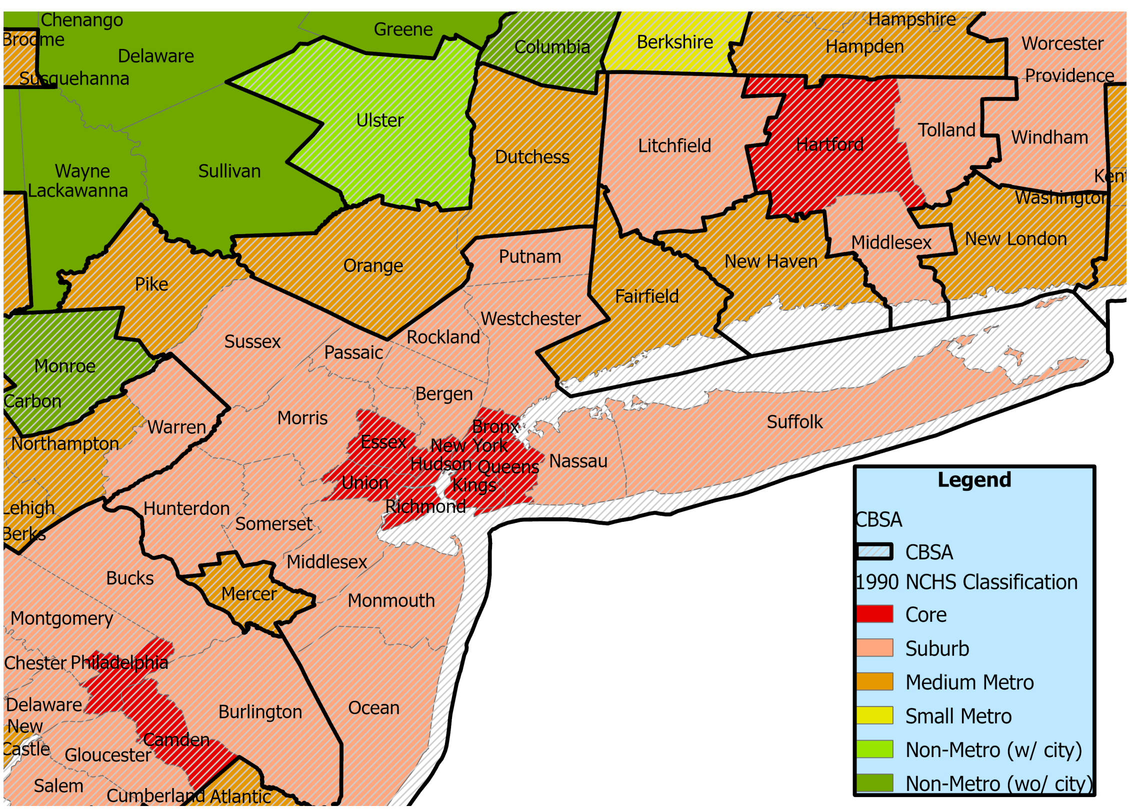
[14] This refers to counties that either: (i) are in metropolitan areas of 1 million or more but are not classified as large central metro, or (ii) are in areas with less than 1 million population but are adjacent to a large central metro county in a neighboring metropolitan area.

shows the New York and Hartford metropolitan areas, highlighting how the scheme differentiates New York's core counties (e.g., New York, Kings, and Bronx) from suburban counties like Bergen and Rockland. Similarly, we can differentiate Hartford's core and fringe areas.

2.2 Stylized Facts

In this section, I highlight two key facts. First, suburbs have emerged as both employment and population centers. This paper shows that this was, in part, due to reductions in commute costs, and crucially, the fall in trade costs. Second, commuting to suburbs has grown. A key insight of this paper is that lower trade costs decentralize jobs to suburbs. This effect intensifies when workers can commute easily, enabling suburbs to scale up their workforce as demand for their goods rises when trade costs are reduced. Therefore, it is essential to study trade, commuting, and migration together to fully understand suburbanization in the U.S.

Figure 1: New York and Hartford



Notes: 2010 County shapefiles merged with 1990 NCHS Rural-Urban Classification.

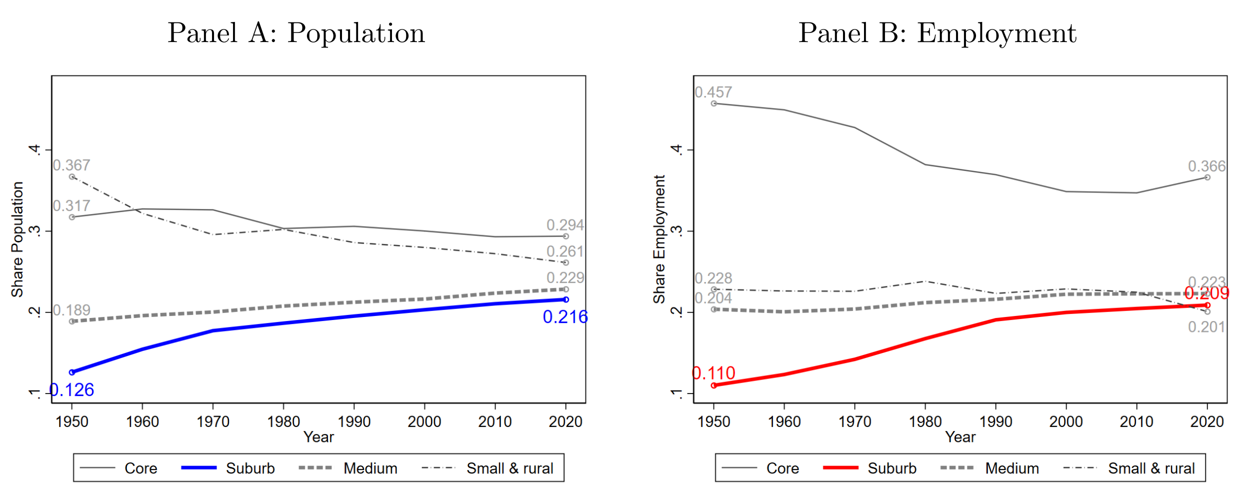
Fact 1: Suburbs emerged as population and employment centers at the expense of rural and core counties, respectively.

Suburbs have emerged as major population and employment centers. Initially accounting for about one-tenth, suburbs now represent roughly one-fifth of the total population and employment. However, while population losses were primarily concentrated in rural areas, employment declines were more pronounced in core counties.

Figure 2 shows the nationwide share of population (Panel A) and employment (Panel B) from 1950 to 2020 across different urban classifications as defined by the NCHS classification. I focus on four groups: core, suburbs, medium metros and rural (which also include small metros). In Panel A, population in suburbs rose from 12.6% to 21.6%. This shift was accompanied by a modest decrease in core counties (31.7% to 29.4%), a significant decline in rural counties (from 36.7% to 26.1%), and a modest rise in medium metros (18.9% to 22.9%). In contrast, in Panel B, the rise of employment in suburbs (11.0% to 20.9%) coincided with a sizable decline in core counties (45.7% to 36.6%), a modest decline in rural areas (22.8% to 20.1%) and rise in medium metros (20.4% to 22.3%).

These facts represent the phenomenon this paper explains through the reduction in commute and trade costs driven by highway expansions.

Figure 2: Population and employment share out of the nationwide total



The y-axis shows the nationwide share of population or employment by classification according to the 1990 NCHS Urban-Rural Classification Scheme for counties. Small metros and nonmetropolitan counties were pooled together. For example, $Share_{k,t} = \frac{\sum_{i \in k} Pop_{i,t}}{\sum_{i}^{S} Pop_{i,t}}$, where S denotes total number of counties, i indexes counties, and k indexes categories (e.g. "Fringe"). Population data comes from the Decennial Census and the American Community Survey Data. Employment data comes from the County Business Patterns Database. To make sure counties definition over time are consistent, I interpolate county level data using Eckert et al. (2020) crosswalks.

Fact 2: Suburban counties have become increasingly important as commuting destinations.

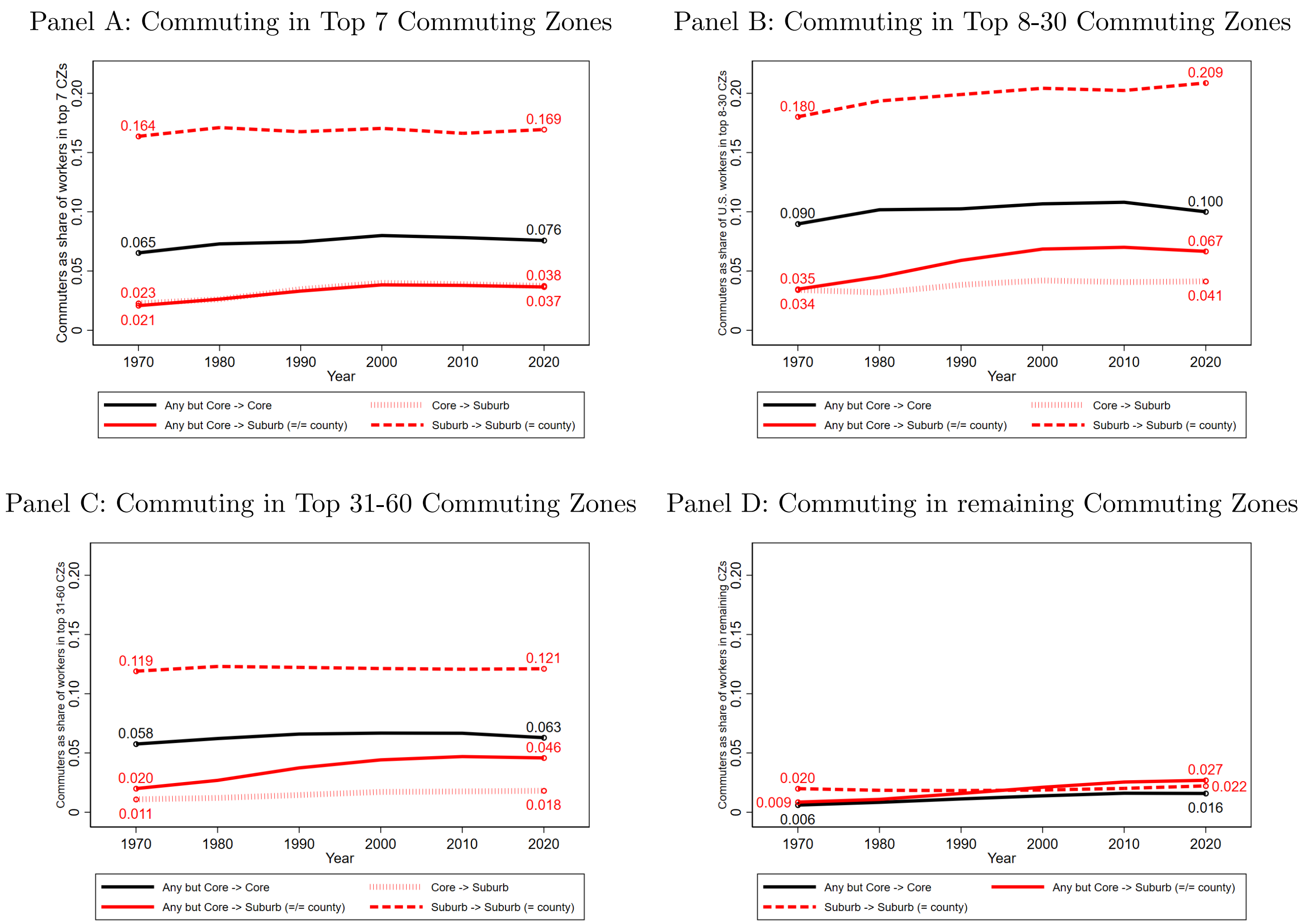
Figure 3 contains four panels, each representing commuting patterns within a specific group of commuting zones. The sample is divided into four groups: the top 7 zones, zones ranked 8-30, 31-60, and the remaining zones, with each group accounting for roughly 25% of the U.S. population. Each panel shows the share of workers commuting between different origins and destinations within these zones (e.g., core to suburb).

Suburban counties have grown increasingly important as commuting destinations. Two key patterns stand out. First, commuting from non-core areas to suburban locations has risen significantly, particularly in mid-sized commuting zones (Top 8–30 and Top 31–60), as shown by the solid red line in the plots. This increase surpasses the rise in non-core to urban core commuting (black lines). Second, within suburban locations, the share of individuals living and working within the same county (dashed red lines) has remained steady or increased, further underscoring their role as central hubs for both residence and employment.

This paper argues that the growing prominence of suburbs as commuting destinations stems in part from highway construction, which reduced trade costs and shifted employment away from urban cores. While commuting access also plays a role, its effect is distinct: mid-sized suburban zones primarily draw workers from nearby non-core areas rather than redirecting them toward urban centers. This dynamic enables suburbs to scale their workforce as demand for locally produced goods rises in response to reduced trade costs.

Overall this facts serves as a motivation to set up a model with a realistic geography that captures commute and trade (and migration) linkages.

Figure 3: Share of workers by type of commute out of total number of workers in size bin



This figure is divided in four panels or size bins. Each account for about 25% of the population. Panel A depicts the top 7 commuting zones. Panel B, the top 8-30 commuting zones. Panel C, the top 31-60 commuting zones. Panel D, the remaining commuting zones. The y-axis shows the share of commuters by 'type' of commute according to origin and destination out of the total number of workers in a given size bin, i.e., $Share_{k,bin,t} = \frac{\sum_{i\in(k,bin)} Workers_{i,t}}{\sum_{i\in(bin)} Workers_{i,t}}$, where i indexes counties, k types of commute, and bin size bin. The graph depicts these 4 types of commutes: noncore-to-core, noncore-to-suburb, core-to-suburb, suburb-to-suburb (within the same county). These shares do not sum one because there are omitted categories. Commuting flow data comes from the Journey To Work Database (1970-2020). To make sure counties definition over time are consistent, I interpolate commuting flows using Eckert et al. (2020) crosswalks.

3 Reduced form: The Effect of Commuting and Trade Access

In this section, I provide direct evidence of the commuting and trade access effects of highways by leveraging the fact that highways reduced both commute times and trade costs, but did so in an heterogeneous way across locations and dates. The impact of a new highway segment depends on the connections it creates: it can link a county to major employment hubs, increasing commuting access, or to large consumer markets where firms can ship their goods, enhancing trade access (or both).

3.1 Measurement of commuting and trade access

To understand the impacts of highways through commute and trade, I introduce a measure of commuting access and of trade access.

Commuting access. I measure commuting access by weighting changes in commute costs by commute flows, and taking an average:

$$\Delta \log \Phi^W_{i,t} \quad = -\textstyle\sum_j \pi^0_{ij|i} \cdot \Delta\tau_{i,j,t} \tag{1}$$

where subscripts are: origin i, destination j, and decade t. $\pi_{ij|i,1970}$ is the share of workers living in i that commute to j in 1970. $\tau_{i,j,t}$ are driving times between counties i and j at decade t. Δ indicates a decade-by-decade change. The commuting access measure, $\Delta \log \Phi^W_{i,t}$, indicates that a location i's commuting accessibility to jobs will increase to the extent that commuting times are reduced along the pairs (i, j) that are used by a greater share of workers, as measured by the commuting flows, $\pi_{ij|i,1970}$. A standard model of commuting as in Ahlfeldt et al. (2015) and Tsivanidis (2023) features the a relationship between changes in a location's commuting accessibility, and changes in commute times as the one I propose in this section.

Trade Access. To compute trade access, I weigh changes in trade costs by trade flows:

$$\Delta \log \Psi^T_d \propto -\sum_o \lambda_{od|o} \Delta \log \tau_{o,d,t}$$

where $\lambda_{od|o}$ represents the share of sales from o to county d. Unfortunately, data on trade flows $\lambda_{od|o}$ for the earliest periods of my sample are unavailable. Therefore, I follow Donaldson and Hornbeck (2016) and approximate trade shares as $\lambda_{od|o} \approx \frac{L_d(\varsigma_{od})^{-\varepsilon}}{\Psi^T_o}$ where market access is approximated as $\Psi^T_o \approx \sum_{d=1}^N (\varsigma_{od})^{-\varepsilon} L_o$ and $\varsigma_{od} = \tau^{\varkappa}_{od}$ denotes the trade costs between o and d. This expression indicates that the sales share $\lambda_{od|o}$ will be larger when the destination county is large adjusted for trade costs, relative to the average discounted size of other counties. Consequently, we have:

$$\Delta \log \Psi_d^T \propto -\sum_o \underbrace{\frac{L_o \left(\varsigma_{od}\right)^{-\varepsilon}}{\Psi_d^T}}_{\tilde{\lambda}_{od}} \Delta \log \tau_{o,d,t}.$$

Here, $-\varkappa\varepsilon$ represents the trade elasticity, which can be calibrated to -2 (Boehm et al., 2023), -1.5 (Monte et al., 2018), or -1.4 (Duranton et al., 2014). For this analysis, I choose -2 and provide robustness tests. The trade access measure $\Delta \log \Psi_d^T$ indicates that as driving times between counties decrease, trade access will increase, making larger markets more accessible.

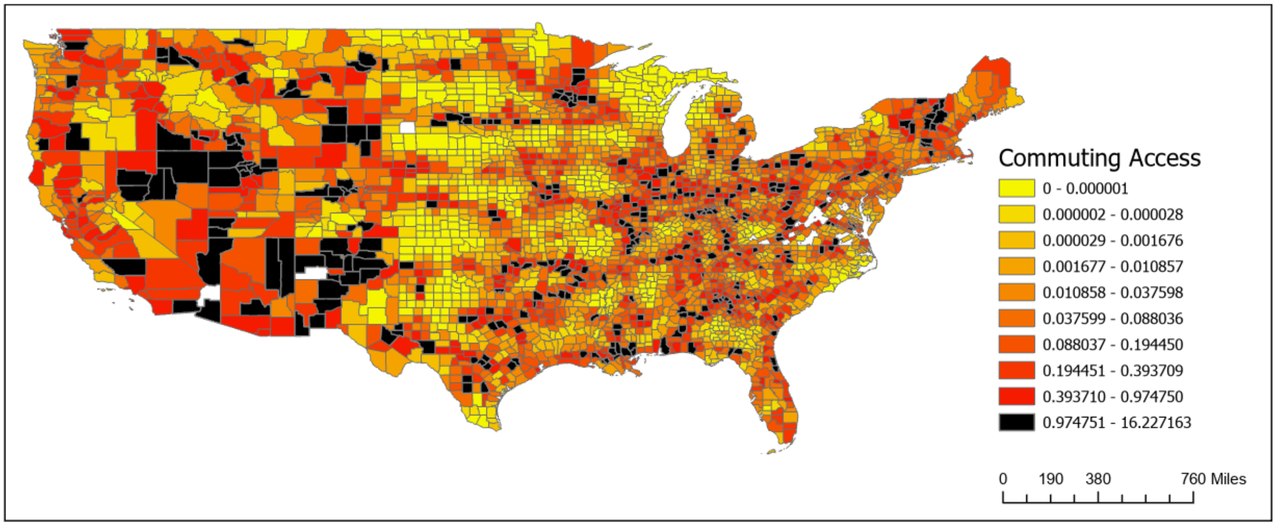
Comparing trade and commuting access. While one might expect a strong correlation between trade access and commute access, the reality is different. The correlation between these measures was only 23.3% from 1960 to 1970. The decade-by-decade correlation between these two was 32.8% from 2000 to 1960. This low correlation arises because trade access assigns positive weights to distant locations, unlike commute access. For example, while few people commute from Baltimore to New York, goods produced in Baltimore are still consumed in New York. Figure 4, highlights the distinct spatial patterns of changes in commuting (Panel A) and trade access (Panel B) from 1960 to 1970. Commuting access increases are more localized, with many regions showing access growth in at least one county. In contrast, trade access spans a wider area.

Given the low correlation, we can leverage a variety of comparisons to identify the effects of commute and trade access on population and employment. For example, counties may experience increases in commuting access without corresponding increases in trade access, or vice versa. Some counties may benefit from both, while others see little to no change. Additionally, even if a new highway segment is built in a different county, it can enhance access from a commuting or trade perspective as the road network develops.

Figure 5 illustrates two examples. Gray lines show highway segments built by 1960, and blue lines show those built by 1970. Red counties represent large population centers (top 5% in 1950). Panel A focuses on the Atlanta metro area, highlighting Coweta County (southwest) and Jackson County (northeast). Both counties improved their commuting access to central Atlanta, but Jackson also gained access to a key trade route. As a result, both counties saw commuting access increases (94th percentile), but Jackson's trade access rose more (83rd percentile vs. 54th for Coweta). Panel B zooms out to the Greenville-Anderson metro area, where counties saw trade access improvements (80th percentile or higher) even without new local highway segments. This is thanks to network completion elsewhere, indicated by the black arrows.

Figure 4: Changes in commute and trade access between 1970 and 1960

Panel A: Commuting Access



Panel B: Trade Access

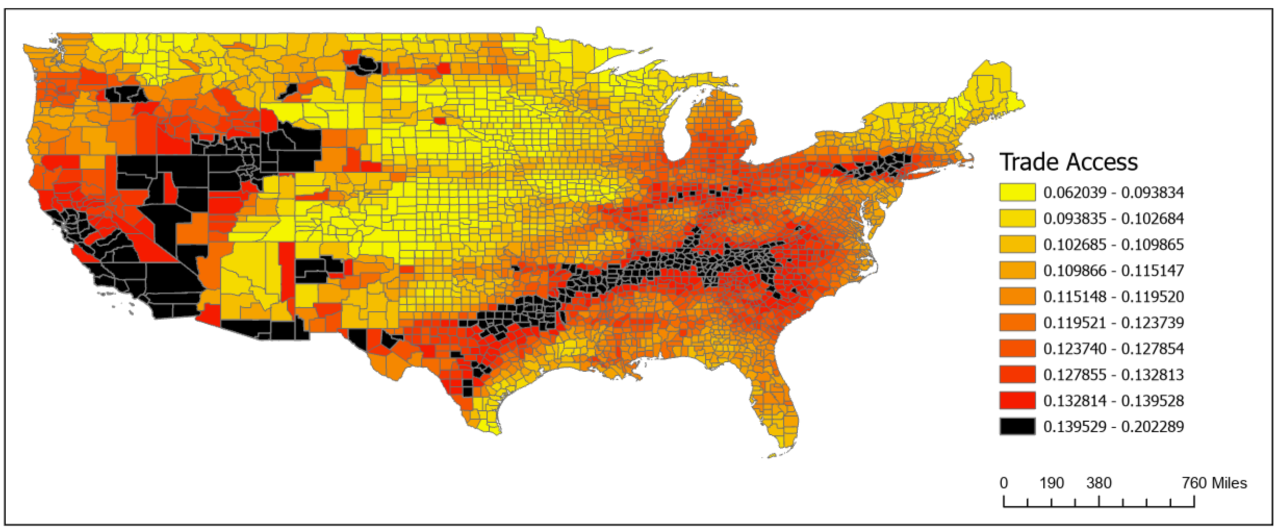
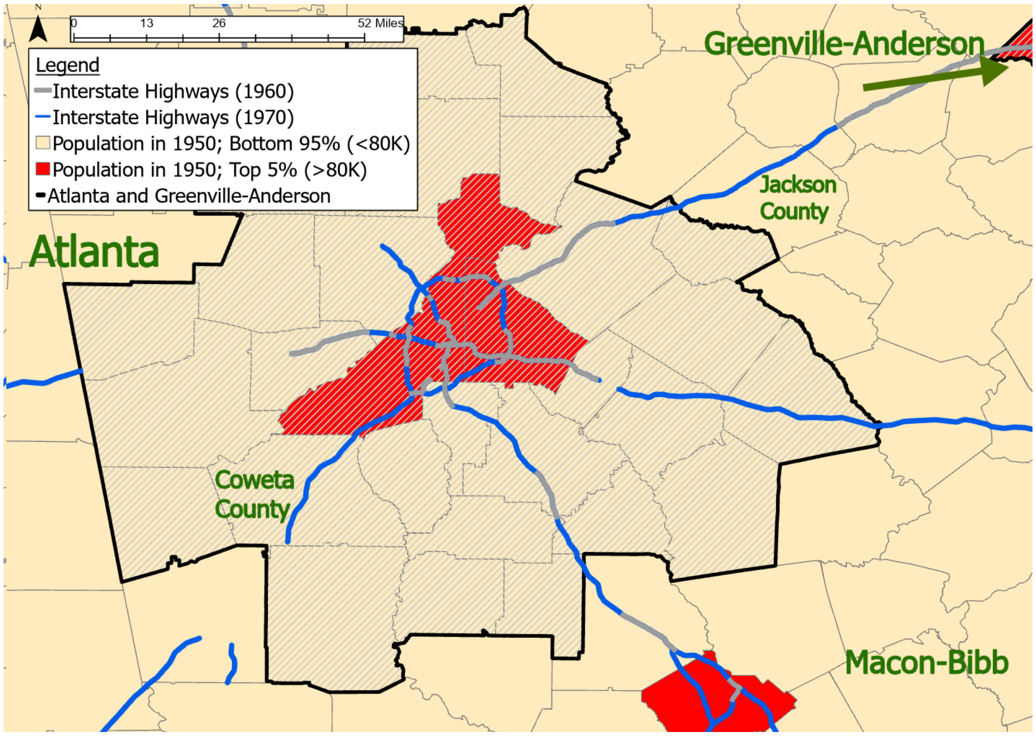
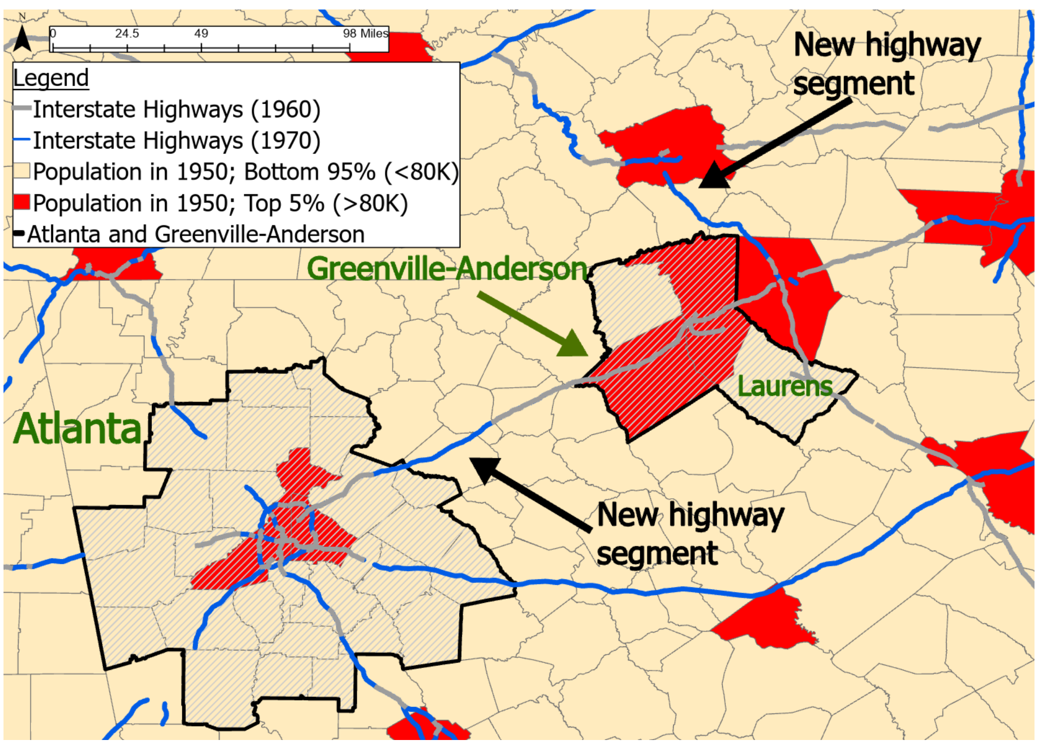


Figure 5: Illustrative examples

Panel A: Atlanta, Georgia



Panel B: Atlanta and Greenville-Anderson Metro areas



3.2 Empirical strategy

Preamble. Consider a naive specification where I run a simple Two Way Fixed Effect (TWFE) regression relating outcomes such as population and employment to trade and commute access as defined earlier, conditioning on fixed effects and controls:

$$y_{it} = \varrho\Phi_{it} + \varphi\Psi_{it}^{T} + FE_i + FE_t + \beta' X_{it} + \varepsilon_{it} \tag{2}$$

This approach has a few limitations. First, GIS-based driving time measurements might be inaccurate due to varying traffic speeds, unlike the assumed constant speed of 70 mph. There could also be measurement error when merging PR-511 data with the HPMS roads shapefile due to the lack of segment identifiers. If measurement error is classic, OLS would be biased towards zero. Moreover, highway placement is influenced by political factors, which might lead to biased results if highways are strategically placed in areas that are growing less (OLS would provide a biased toward zero estimate) or in areas that are growing more (OLS would over estimate the true impact). Second, the impact of highways could be persistent over time. As many authors have pointed out, a simple TWFE regression is prone to bias from heterogeneous impacts over time (Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; Dube et al., 2023; Callaway et al., 2024). This is because, if a unit was treated in the past and the effect of the treatment is persistent, then this unit is 'contaminated' and should not be used as a comparison unit for units that are treated later in the sample.

To address the first issue, I use an instrumental variables estimation with four instruments previously used in the literature: the 1947 Interstate plan, the 1920 Pershing Map, the 1528-1850 historical routes of exploration, and railroads around 1898 (Baumsnow, 2007; Duranton and Turner 2012; Michaels et al, 2019; Frye, 2023; Brinkman and Lin, 2024). I further discuss these instruments later in the text. To address the second issue, I apply the local projection-based differences-in-differences (LP-DiD) method from Dube et al. (2023), which uses local projections to estimate dynamic effects and avoid bias by dropping 'unclean' observations. Details are expanded upon next.

Local projection differences in differences (LP-DiD). I apply the LP-DiD approach (Dube et al., 2023), which combines local projections with a difference-in-differences framework to estimate dynamic impulse responses from panel data. The key innovation of LP-DiD is the "clean-control" condition, which ensures unbiased estimates by carefully selecting control groups. Specifically, counties that received treatment in earlier periods are excluded from serving as controls in later periods. This avoids contamination that could arise if previously treated areas were used as comparisons, ensuring that the control group remains unaffected by earlier shocks and the estimated responses reflect the true causal effect.

A challenge in this setting is that all locations are affected to some degree due to the continuous nature of access measures. Each wave of highway construction—occurring in 1970, 1980, 1990, and 2000—alters driving times, influencing every county to some degree. To address this, I define a treated group for each wave by focusing on locations with access increases above a specific threshold. For each treatment wave, the threshold is set at the 50th percentile of decade-by-decade changes in access measures.[15] For each highway expansion wave s (e.g., 1970, 1980, 1990, or 2000), counties j that experience substantial access increases are classified as treated. Control groups consist of counties that will be treated in future waves or remain untreated throughout the period. This ensures that units heavily treated in previous periods are not used as controls in later waves.

To estimate the impact of highway construction on population and employment growth, I use a series of two-period panel regressions centered around the construction event. For each highway wave s, the baseline period is $s-10$, and the post-treatment period is $s+h$, where h is the horizon of interest (e.g., $h=40$). Let $q \in \{s-10, s+h\}$ index pre and post periods. The regression specification is:

$$\underbrace{\ln y_{j,s,q}}_{\text{population and employment}} = \rho_h \underbrace{\log \Phi^W_{j,s,q}}_{\text{commuting}} + \varphi_h \underbrace{\log \Psi^T_{j,s,q}}_{\text{trade}} + \underbrace{\gamma_{s,q,h} + \gamma_{j,s,h}}_{\text{FE}} + \underbrace{\varpi'_h X_{j,s,q}}_{\text{controls}} + \tilde{\mu}_{j,s,h} \tag{3}$$

In the regression model, the terms $\log \Phi^W_{j,s,q}$ and $\log \Psi^T_{j,s,q}$ represent the commuting and trade market access, respectively, for county j during wave s, either before the construction of the highway $s-10$, or just after the construction of the highway $s+0$. Specifically, the access terms are defined as:

$$\log \text{Access}_{j,s,q} = \begin{cases} \log \text{Access}_{j,s-10} & \text{if } q = s-10, \\ \log \text{Access}_{j,s+0} & \text{if } q = s+h. \end{cases}$$

where Access can be either commute, Φ , or trade access Ψ^T. The key idea is to compute the *first change in access* due to highway expansion, $s+0$ versus $s-10$, and use it to estimate the impulse

[15] Sensitivity analyses confirm the robustness of results to different threshold choices.

response for $s-20, s+0, s+10, ..., s+40$. By using these definitions in the two-period panel regressions, I measure how the initial change in market access (e.g., $\log \Phi^W_{j,s+0} - \log \Phi^W_{j,s-10}$) affects population and employment growth. Put differently, this setup captures the cumulative dynamic effects of highways over time, allowing me to study how improvements in access propagate over multiple decades. I demean $\log \Phi$ and $\log \Psi$, and re-scale with standard deviation of their change between 1970 and 1960.

In all regressions, I include the following set of controls interacted with wave s and horizon h fixed effects: population in 1950, change in log population between 1960 and 1950, income in 1950, share of the population with college education in 1950, employment in 1950, change in log employment between 1960 and 1950, ratio of tradable to non-tradable employment, urban dummies in 1950, and state fixed effects. These controls address potential biases from baseline economic conditions, and pre-existing trends. State fixed effects capture unobserved policies or shocks at the state level, while the employment ratio accounts for economic structure. All this, helps ensure that the growth in unobservables such as TFP or amenities do not confound the relationship between access measures and economic outcomes.

To estimate the average effect of commute and trade access, it is important to include their interaction term. Omitting it would result in omitted variable bias if the interaction term is correlated with either commute or trade access and the interaction effect is nonzero. The interaction term is likely to correlate with one or both of the trade access measures. The interaction effect is also expected to be nonzero. When both trade and commute access improve simultaneously, their interaction can be negative, meaning the two channels may act as substitutes. For example, better trade access makes it easier for firms to move goods across regions, but if commute access also improves, workers may choose to commute to jobs elsewhere rather than work in the same county where trade access has improved. This reduces the local employment benefits that improved trade access alone would have generated. To capture such interactions, in some specifications I add the interaction term:

$$\ln y_{j,s,q} = \rho_h \log \Phi^{W,0}_{j,s,q} + \varphi_h \log \Psi^{T,0}_{j,s,q} + \zeta_h \underbrace{\left(\log \Phi^{W,0}_{j,s,q} \cdot \log \Psi^{T,0}_{j,s,q}\right)}_{\text{interaction}} + \gamma_{s,q,h} + \gamma_{j,s,h} + \varpi'_h X_{j,s,q} + \tilde{\mu}_{j,s,q} \quad (4)$$

If the interaction coefficient ζ_h is negative, it suggests that the improvement in commute access reduces the marginal benefit of trade access. By taking a partial derivative, I predict the average commute access effect as $\rho_h + \zeta_h \log \bar{\Psi}^{T,0}_{j,s,q}$, and the average trade access effect as $\varphi_h + \zeta_h \log \bar{\Phi}^{W,0}_{j,s,q}$.

2SLS Local projection differences in differences. Highway placement may not be random, which requires the use of instruments to address potential endogeneity. To this end, I employ four well-established instruments from the literature: the 1947 Federal Plan, the 1920 Pershing Plan, railroads constructed by 1898, and exploration routes from the 16th to 19th centuries. I assume that highway placement would follow the routes defined by these instruments and calculate driving times accordingly. To construct these instruments, I determine the time required to traverse the designated routes between various locations. I then define commuting and trade access measures as follows:

$$\tilde{\Psi}_d^k \propto -\sum_o \tilde{\lambda}_{od}^0 \tau_{od,t}^k \quad \text{and} \quad \tilde{\Phi}_{i,t}^k \propto -\sum_j \pi_{ij}^0 \tau_{ij,t}^k \tag{5}$$

where k indexes each of these four instruments. I interact each of these variables with time trends and use them as instruments for commuting access, trade access, and their interaction.

To be valid, the instruments must satisfy both exogeneity and relevance conditions. First, I will discuss the exogeneity condition. The first instrument is the 1947 Plan, developed by the Bureau of Public Roads to promote intercity trade and national defense (Baum-Snow, 2007, 2019; Duranton and Turner, 2012). Since the focus was on national defense and intercity trade, the argument is that this plan may not correlate with fundamentals at the neighborhood level within cities (i.e. counties within MSAs). However, it could still be related to fundamentals between cities (i.e. MSAs). To further address this concern, I also use the 1921 Pershing Plan, created under General John J. Pershing to prioritize military needs over trade (Michaels et al., 2019; Frye, 2023). Additionally, I utilize historical routes as instruments based on the premise that they are unlikely to correlate with county fundamentals from 1950 to 2020 while still predicting highway location. Specifically, I follow Duranton and Turner (2012) by using exploration routes from the 16th to 19th centuries (shapefiles from Brinkman and Lin, 2024) and historical railroads from 1898 (shapefiles from Atack, 2015).

Second, regarding the relevance condition, Appendix A.2 compares each of these instruments with highway placement by the year 2000. Visual inspection confirms that these instruments effectively predict the location of highways. While these instruments predict highway placement, they must also be correlated with the timing of highway construction without being directly related to the outcome of interest. By interacting these instruments with time trends, the underlying assumption is that areas predicted to have lower driving times tended to have highways built first. I do not have construction dates for these counterfactual roads, but if such data were available, the advantage of interacting market and commuting access measures with time trends is clear: using actual construction dates could introduce bias because unobservable factors may be correlated with both population and employment growth and construction dates. By relying on time trends instead of actual dates, I argue that it is plausible to claim exogeneity. However, this approach may lead to a weak first stage if the interacted time trends with market and trade access measures do not effectively predict changes in market and commute access. Fortunately, this is not the case, as the first stage is usually strong.

3.3 Main results

This section establishes that commute access alone does not drive population and employment growth; trade access plays an important role. Once trade access is accounted for, the effect of commute access disappears, highlighting the importance of trade networks in shaping counties' growth. These effects are essential inputs for the quantitative analysis conducted in the second step.

OLS results. I first present the results of increasing commuting access without controlling for trade access. Then, I add the controls for trade. Finally, I add the interaction of trade and commuting. I repeat this exercise with the variable for trade access. I use log population and log employment as outcomes.

I begin by looking at the impacts of changes in commuting access on log population and log employment. Table 1 displays the coefficients for the estimation when $h = 40$, illustrating the effects observed after 40 years, specifically for log population. Column (1) shows that a one standard deviation increase in commute access results in a 5.9% increase in log population. In Column (2), we see that a one standard deviation increase in trade access leads to a 10.9% increase in population. However, when both terms are included together in the model, each coefficient decreases in magnitude. Finally, the interaction term reveals a negative coefficient, suggesting that the simultaneous increases in trade and commute access may dampen the individual effects of each. We observe similar trends for log employment in Table 2. Panel A of Figure 6 presents the average effect of commute access after including the interaction term and computing average effects for the whole time horizon. It also displays a pre-trend test by estimating effects before highway construction. Controlling for trade access and its interaction with commute access yields commute effects of 0.03% for population and

1.5% for employment, both statistically indistinguishable from zero. I find no evidence of differential pre-trends.

Table 1: Commute and trade access effects on log population after 40 years

	(1)	(2)	(3)	(4)
	Log Population	Log Population	Log Population	Log Population
Commute Access	0.0588***		0.0408***	0.00235
	(0.0121)		(0.0119)	(0.0140)
Trade Access		0.109***	0.0884***	0.0958***
		(0.0134)	(0.0147)	(0.0146)
Interaction Term				-0.00192***
				(0.000330)
Observations	13336	13258	12248	12248

Notes: Standard errors clustered at the county-level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 2: Commute and trade access effects on log employment after 40 years

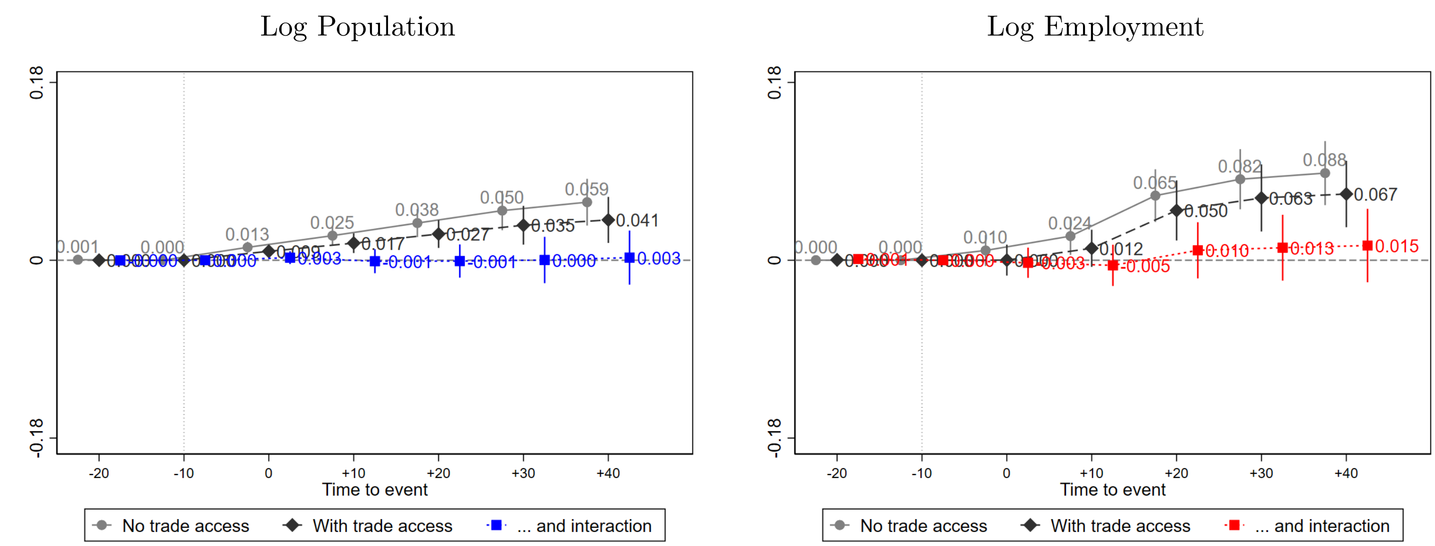
	(1)	(2)	(3)	(4)
	Log Employment	Log Employment	Log Employment	Log Employment
Commute Access	0.0882***		0.0671***	0.0145
	(0.0165)		(0.0172)	(0.0190)
Trade Access		0.137***	0.108***	0.118***
		(0.0195)	(0.0211)	(0.0209)
Interaction Term				-0.00263***
				(0.000474)
Observations	13336	13258	12248	12248

Notes: Standard errors clustered at the county-level in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

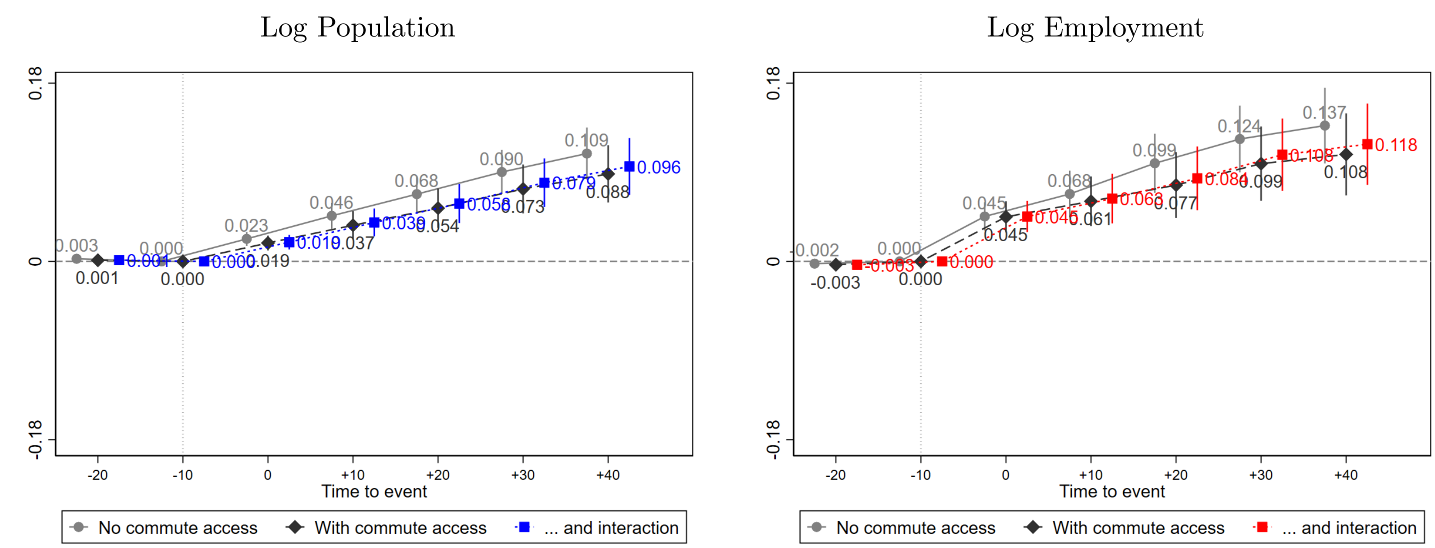
In Column (2), trade access has a larger effect than commute access, with a one standard deviation increase leading to a 10.9% rise in population. When controlling for commute access in Column (3), the trade access effect decreases slightly to 8.8%, indicating that part of its impact overlaps with commute access. In Column (4), after accounting for the interaction between trade and commute access, the trade effect stabilizes at 9.6%. Panel B of Figure 6 traces traces the dynamic impact over time, including a pre-trend test before highway construction. Studying the dynamic response of highway construction is avenue for future research. Figure A.3 shows that the impact is greater for employment than for population.

Figure 6: OLS estimates of the effect of commuting and trade access

Panel A: Commute Access



Panel B: Trade Access



All in all, OLS regressions show that, on average, highways affect population and employment at the county level primarily through trade access, not commuting access. The effects are greater for employment than for population. Since highway placement may not be random, even when pre-trends are parallel, in the next section, I use a 2SLS estimator to confirm my OLS findings.

2SLS results. Figure 7 summarizes the main findings. Panel A shows the average effect of increasing commute access (when controlling for trade access and the interaction). Panel B does so for trade access. In each panel there are two outcomes, population and employment. Each figure depicts in blue and red the results from using all instruments at the same time. The figure also depicts in the background the results from excluding a particular instrument (e.g., the 1947 Plan) while keeping the remaining instruments in the estimation (e.g., "All-1947 Plan"). Kleibergen-Paap F-statistics are reported below each picture. First stage is strong in most specifications, except when excluding exploration routes from the 16th-19th centuries.

There are three key takeaways. First, 2SLS regressions confirm that, on average, highways impact population and employment through trade access, not commuting access. After 40 years, the effect of commute access is negative and insignificant, while trade access has a positive and significant impact.[16] Second, these results hold across different sets of instruments. Third, the 2SLS results reveal that OLS underestimates the effect (OLS is lower than 2SLS) of trade access and overestimates the effect of commute access (OLS is very close to zero, whereas 2SLS is negative but not significant).

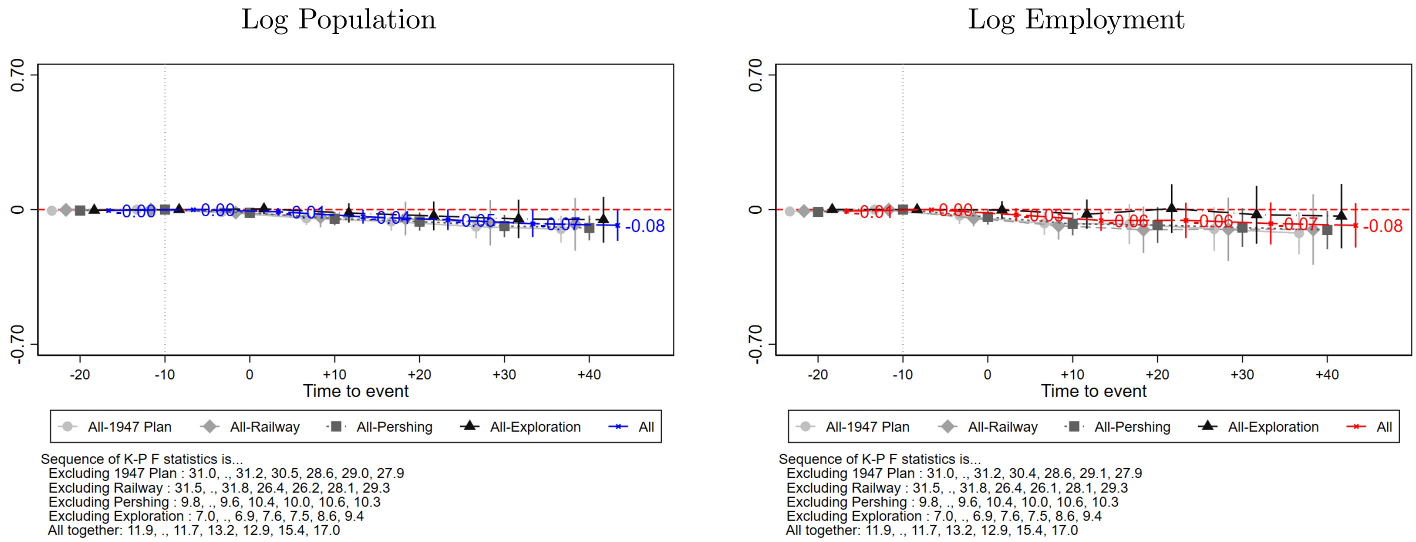
3.4 Other results

In Appendix A.4, I present robustness tests and additional results. First, in my baseline estimates, identification comes from a comparison of access measures between counties, even when they belong to different commuting zones but within the same state. In this appendix, as a robustness test, I introduce commuting zones fixed effects interacted with year-by-wave fixed effects. Thus, identification comes from comparing access measures between two or more counties within the same commuting zone. Qualitatively, results are similar: the average effect of trade access is positive, while commute access effects are closer to zero.

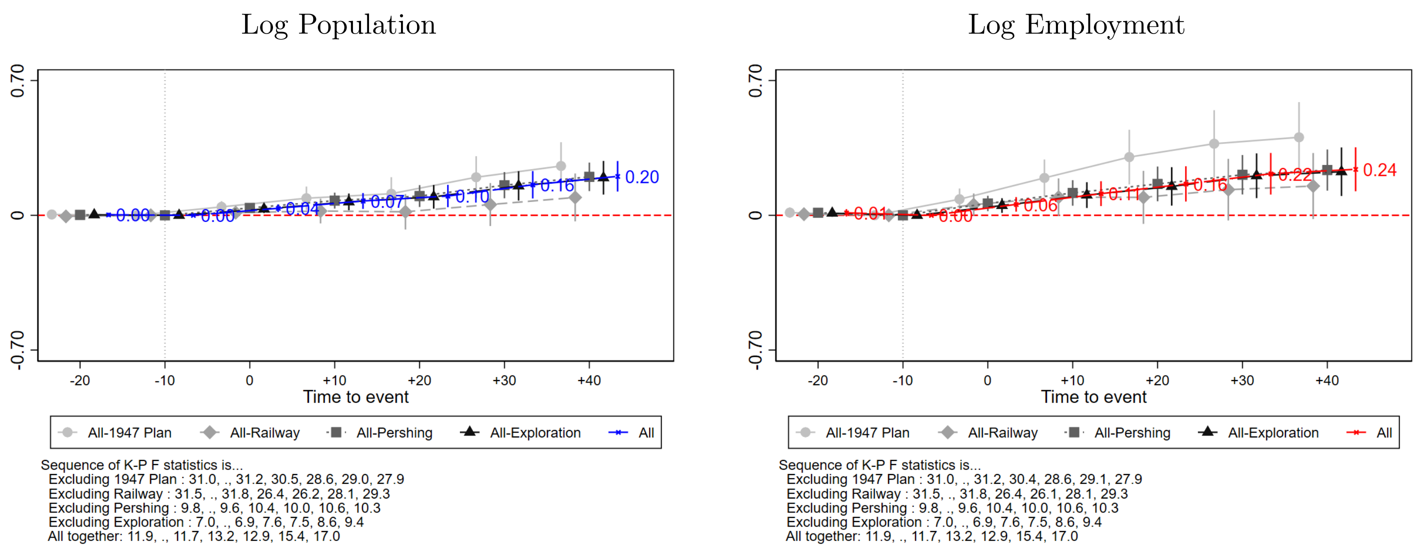
[16] Figure A.4 shows that the impact is greater for employment than for population, albeit sometimes the difference is not statistically significant.

Figure 7: 2SLS estimates of the effect of commuting and trade access

Panel A: Commute Access



Panel B: Trade Access



Second, this paper examines whether gains in commuting and trade access from highways explain the suburbanization of population and employment at the county level. To explore suburbanization within counties, that is, at the census tract level, I construct a panel of census tracts from 1950. My findings suggest that trade access concentrates population within counties, while commute access decentralizes it. Although my model does not explicitly account for intra-county geography, these findings help bridge the gap between existing literature on the effects of highways on neighborhood-level population and my analysis.

Third, there may be concerns based on the prior that most commuting occurs within counties and changes in driving times between counties may have little impact. In this appendix, I exploit changes in driving times and commute flows to study the relationship directly. I estimate a dynamic gravity equation for commuting flows using data from 1970 to 2020, leveraging changes in commute times caused by staggered highway construction. I show that commuting flows between counties do respond to declines in commute times. However, recall that, on average, commute access is not increasing total population nor employment. These two outcomes are entirely possible: highways can reorganize commuting flows without significantly impacting total population and employment. Furthermore, these findings align with prior research on commuting semi-elasticities but offer a novel contribution by separately identifying short-and long-run elasticities

3.5 Advantages relative to alternatives empirical strategies.

There is some advantages of my empirical strategy relative to some alternatives. First, to study the effect of highways one alternative would be to use dummies or perhaps count the number of miles of highways per county. The problem with this idea is that it would be hard to disentangle the trade effect from the commute effect. By relying on the commute and trade access measures I can use a richer variation in the data (i.e. the location of a county within the road network), and even test for interaction effects. Second, I could use highways dummies and a measure of distance to the closest central business district to try to pin down the commute access effect. This is because one could speculate that the commute access effect to be greater on the fringe than on the central business district. However, the commuting access measure I utilize already handles this idea directly by using commuting flows as shares. My approach also do not rely on assumptions of what is a central business district and how far it is from a given county. Third, I could use long differences instead of decade-by-decade differences, thus avoiding the bias from persistent effects without relying on LP-DiD estimates. I argue that in the context of highways, by relying on decade by decade, and estimating dynamic impulse responses, I can leverage finer variation giving me a better chance of differentiating trade access effects from commute access effects.

4 Theory

I develop a Quantitative Spatial Model of the U.S., where locations are connected through trade, commuting, migration. Trade and commute cost reductions trigger migration responses which are amplified by externalities.

4.1 Environment

The model considers a finite and discrete set of locations, indexed by $i, j, o, d, n \in \mathcal{S}$, and an infinite sequence of discrete periods, $t \in \mathcal{T}$. Each location contains three sectors: tradable, non-tradable, and construction.

4.2 Preferences, commuting and migration

Preferences. There is a continuum of individuals indexed by ω, each living for two periods. In the first period ("childhood"), they consume what their parents consume. In the second period ("adulthood"), they choose where to live and work, supply labor inelastically, consume a final good and housing, and give birth to $\zeta_t \in (0, \infty)$ children. The utility of an individual ω born in n, living in i, and commuting to j is:

$$U_{n,i,j,t}(\omega) = \nu_i(\omega) \cdot \xi_{n,i,t} \cdot u_{i,t} \cdot \frac{\epsilon_j(\omega)}{d_{i,j,t}} \cdot C_{i,t}^{\alpha} H_{i,t}^{1-\alpha} \tag{6}$$

where $\nu_i(\omega)$ and $\epsilon_j(\omega)$ are idiosyncratic preferences for living in i and commuting to j. $d_{i,j,t} = \exp(\kappa \tau_{i,j,t})$ is the commute iceberg cost between origin i and destination j, and $\xi_{n,i,t}$ is the migration iceberg cost between origin n and destination i, both of which depend on driving times. $u_{i,t}$ are amenities, $C_{i,t}$ is final good consumption, and $H_{i,t}$ is residential floorspace consumption. Households maximize this utility subject to the budget:

$$y_{n,i,j,t}(\omega) \equiv w_{j,t} = P_{i,t} C_{i,t} + R_{i,t} H_{i,t} \tag{7}$$

where $w_{j,t}$ is the wage at j, $P_{i,t}$ is the price of the final good, $R_{i,t}$ is the rental price of residential floorspace.

Commuting. Conditional on their place of residence i, individuals draw workplace-specific preferences $\epsilon_j(\omega)$ from a Frechet distribution: $P(\epsilon_j \leq \epsilon) = \exp\left(-K\epsilon^{-\theta}\right)$, where θ is the shape parameter. They seek the location with the highest utility, which is equivalent to seeking the location with the highest net income: $\tilde{y}_{i,j,t}(\omega) = \frac{w_j \epsilon_j(\omega)}{d_{i,j,t}}$. In other words, they look for the location that provides the highest income after accounting for commuting costs and commuting preferences. Dispersion in commuting preferences affects sensitivity to commute costs. A higher θ (i.e., lower dispersion) implies greater sensitivity. To understand why, consider the case where preferences over workplace destinations, ϵ_j, are homogeneous (i.e., $\theta \to \infty$). In this scenario, even a small reduction in commute costs triggers a large reaction. On average, the probability of choosing location j given residence i is:

$$\pi_{ij|i,t} = \frac{\left(\frac{w_{j,t}}{d_{i,j,t}}\right)^{\theta}}{\Phi_{i,t}} \tag{8}$$

where commuting access is $\Phi_{i,t} = \sum_{l=1}^{S}\left(\frac{w_{l,t}}{d_{i,l,t}}\right)^{\theta}$. This is the gravity equation for commuting: it predicts the share of workers commuting from i to j based on the relative utility derived from working in j. The numerator reflects the utility associated with working in j, where higher wages $w_{j,t}$ and lower commuting costs $d_{i,j,t}$ increase the likelihood of choosing j. The denominator, $\Phi_{i,t}$, is a location-specific term that captures the total utility across all possible workplace destinations, ensuring the probabilities across all j sum to one. This equation highlights the trade-off between wages and commuting costs, with the parameter θ governing how sensitively workers respond to differences in these factors.

The expected income conditional on residence i is:

$$\mathbb{E}[y|i] = \sum_j \pi_{ij|i,t} w_j \equiv \bar{y}_{i,t} \tag{9}$$

The expected utility, conditional on residence i and birthplace n, is:

$$\mathbb{E}\left[U_{n,i,j,t}(\omega) \mid i, n\right] \propto \nu_i(\omega) \cdot \underbrace{\frac{\xi_{n,i,t} u_{i,t}}{P_{i,t}^{\alpha} R_{i,t}^{1-\alpha}} \cdot \Phi_{i,t}^{1/\theta}}_{\bar{U}_{n,i,t}} \equiv U_{n,i,t}^{\mathcal{E}}. \tag{10}$$

While expected income $\mathbb{E}[y|i]$ does not directly enter expected utility, the model incorporates the commuting access term, $\Phi^{1/\theta}$, which represents average income net of commuting costs and adjusted for idiosyncratic commuting preferences. Specifically, $\mathbb{E}\left[\frac{w_j \epsilon_j(\omega)}{d_{i,j,t}}\right] \propto \Phi_{i,t}^{1/\theta}$.

Migration. Assuming the idiosyncratic term $\nu_i(\omega)$ follows a Frechet distribution with shape parameter η, workers' utility is also distributed Frechet: $P(U^{\mathcal{E}}_{n,i,t} \leq x) = \exp\left[-\left(U^{\mathcal{E}}_{n,i,t}\right)^{\eta} x^{-\eta}\right]$. Workers choose the location offering the highest utility in the period they reach adulthood. The probability of choosing location i is $\pi^R_{i|n,t} = P\left(U^{\mathcal{E}}_{n,i,t} \geq \max_{l\neq i} U^{\mathcal{E}}_{n,l,t}\right)$. Thus, the number of individuals moving from location n to i is:

$$L_{n,i,t} \quad = \quad \zeta_t \frac{\bar{U}^{\eta}_{n,i,t}}{\sum_{i\in N} \bar{U}^{\eta}_{n,i,t}} L_{n,t-1} \tag{11}$$

where ζ_t is the number of children per adult. This equation says migration flows increase towards locations with higher indirect utility $\bar{U}_{n,i,t}$ relative to outside options $\sum_{i\in N} \bar{U}^{\eta}_{n,i,t}$, and from origins with more residents $L_{n,t-1}$. From these migration flows I can compute total population per location-time period. Once I know population per location, I can compute expenditure on final goods (i.e. tradable and non-tradable goods), and residential floorspace.

This completes the discussion of the household side of the model. Next, we turn to the production of tradable, non-tradable goods, and final goods, as well as the costs associated with shipping tradable goods across locations.

4.3 Production of tradable, non-tradable, and final goods

Tradable and non-tradable goods. This model is an application of the Eaton-Kortum framework. Each location has two sectors: sector s can be either tradable (T) or non-tradable (NT), each producing a continuum of goods indexed by ð. The efficiency of producing good ð in each sector-location pair $\{s, o\}$ is a realization of a random variable a^s_o, drawn from a Frechet distribution $F^s_o(a) = \exp\left(-A^s_o a^{-\varepsilon}\right)$, where ε is the shape parameter (governing comparative advantage, with lower ε implying more heterogeneity and stronger comparative advantage), and A^s_o represents absolute advantage. The cost of purchasing a good from location o in location d in sector s is given by the random variable $p^s_{od} = c^s_o \varsigma^s_{od}/a_o$, where c^s_o is the unit cost in o, and ς^s_{od} are the trade costs between o and d. In the non-tradeable sector we have $\varsigma^{NT}_{od} \to \infty$ for all $o \neq d$. In the tradeable sector we have that $\varsigma^T_{od} \geq 1$ for all o, d. For both sectors, ς_{oo} is normalized to one.

Firms in the tradable and non-tradable sectors employ structures $\tilde{H}_o$ and labor $\tilde{L}_o$ as inputs. For simplicity, only in the tradable sector, which is the primary focus of this paper, firms additionally use intermediate inputs. Structures, labor and intermediate inputs are aggregated with an elasticity of substitution σ. In principle, σ can be below one, which would imply that labor, floorspace, and intermediate inputs are gross complements. The cost bundle in sector $s \in \{T, NT\}$ is:

$$c^s_{o,t} \quad = \quad \left(\varrho^s_L w^{1-\sigma}_{o,t} + \varrho^s_H \tilde{R}^{1-\sigma}_{o,t} + (1 - \varrho^s_L - \varrho^s_H) P^{T\,1-\sigma}_{o,t}\right)^{\frac{1}{1-\sigma}} \tag{12}$$

where P^T_o is the CES price index intermediate goods, which is the same price index of tradable goods, and $\varrho^{NT}_L + \varrho^{NT}_H = 1$.

Final goods. The final good is a CES aggregate of the tradable and non-tradable goods with elasticity σ_D. The corresponding price indexes are:

$$P^{NT}_{i,t} = \left[\int_0^1 \left(p^{NT}_{i,t}(\eth)\right)^{1-\sigma_D} d\eth\right]^{\frac{1}{1-\sigma_D}} \quad \text{and} \quad P^{T}_{i,t} = \left[\int_0^1 \left(p^{T}_{i,t}(\eth)\right)^{1-\sigma_D} d\eth\right]^{\frac{1}{1-\sigma_D}} \tag{13}$$

$$P_{i,t} \quad = \quad \left[b\left(P^T_{i,t}\right)^{1-\sigma_D} + (1-b)\left(P^{NT}_{i,t}\right)^{1-\sigma_D}\right]^{\frac{1}{1-\sigma_D}} \tag{14}$$

From quantitative point of view, including the non-tradable sector in the model is important because it grew considerably between 1950-2020 as a result of structural change. From a qualitative point of view, it is important because the non-tradable sector interacts closely with trade and commuting. If tradable and non-tradable goods are complements, a decrease in trade costs may shift the household expenditure share from tradable to non-tradable goods, causing a reallocation of employment between sectors. Moreover, a reduction in commuting costs (i.e. a labor supply shock to a first order approximation) primarily affect prices in the labor-intensive sector, decreasing household expenditure share in that sector and increasing it elsewhere.

4.4 Trade

The likelihood that county o supplies a particular good to county d is the probability λ_{od} that o's price turns out to be the lowest. Recall that due perfect competition, $p^s_{od} = \frac{c^s_o \varsigma^s_{od}}{a^s_o}$ for both sectors. Thus, the lowest price realization in county d is $p_d = \min\{p_{od} : o = 1, \ldots, N\}$. Given a_o follows a Frechet distribution $P(a_o \leq a) = \exp(-A_o a^{-\varepsilon})$, we have that the probability that county o provides a good at the lowest price to county d is:

$$\lambda_{od|d} = \frac{A_o \left(c_o \varsigma_{od}\right)^{-\varepsilon}}{\Psi_d}$$

where $\Psi_d = \sum_{o=1}^{N} A_o \left(c_o \varsigma_{od}\right)^{-\varepsilon}$ summarizes states of technology, input costs, and geographic barriers, and their influence on prices in each county d. Notice that for the non-tradable sector $\Psi_d^{NT} = A_d^{NT} c_d^{NT-\varepsilon}$.

Due to the Frechet assumption, goods purchased by destination counties have a price distribution that does not depend on the source. A source county with better technology, lower input costs, or fewer barriers will sell a wider variety of goods, until the price distribution of the goods sold by the source county matches the overall price distribution in the destination. Since county's d average expenditure per good does not vary by source, the fraction of goods that county d buys from county o is also the fraction of its expenditure on goods from county o:

$$\frac{X_{od}^s}{X_d^s} = \lambda_{od}^s = \frac{A_o^s \left(c_o^s \varsigma_{od}^s\right)^{-\varepsilon}}{\Psi_d^s} \tag{15}$$

Finally, to derive the price indexes, we plug the distribution of realized prices into the price indexes of equations 13.[17] Assuming $\sigma_D < 1 + \varepsilon$, the price indexes become:

$$P_d^{NT} = cons \cdot \left(\Psi_d^{NT}\right)^{-1/\varepsilon} \quad \text{and} \quad P_d^{T} = cons \cdot \left(\Psi_d^{T}\right)^{-1/\varepsilon} \tag{16}$$

with $const = \left[\Gamma\left(\frac{\varepsilon+1-\sigma_D}{\varepsilon}\right)\right]$.

4.5 Floorspace and land

Each location is endowed with K_o units of land, which, combined with labor under an elasticity of substitution ν, produces floorspace H_o.[18]

Solving the problem of the firm yields the following cost bundle for floorspace:

$$c_o^{\text{Cons}} = \left(\beta w_o^{1-\nu} + (1-\beta) Q_o^{1-\nu}\right)^{\frac{1}{1-\nu}}, \tag{17}$$

[17] $G_d\left(p\right) = 1 - \exp\left(-\Psi_d p^{\varepsilon}\right)$.

[18] If labor and land are substitutes, differences in land prices can affect the production mix. For instance, in suburban areas where land prices are typically lower due to its abundance, firms tend to use more land than other inputs. This explains why factories, warehouses, and homes are generally larger in suburban and rural areas compared to central locations.

where w_o are wages, Q_o is the price of land, and β represents labor-specific productivity and $1-\beta$ represents land-specific productivity, normalized to sum to 1. Due perfect competition, the rental price of floorspace R_o is:

$$R_o = \frac{\left(\beta w_o^{1-\nu} + (1-\beta) Q_o^{1-\nu}\right)^{\frac{1}{1-\nu}}}{Z_o}, \tag{18}$$

where Z_o denotes the total factor productivity of the construction sector.

Total floorspace H is allocated between residential and commercial uses. This allocation is determined by the total demand from residential uses (from households) and commercial uses (the sum of demand from the tradable and non-tradable sectors).

Finally, revenue from floorspace production, $H_o R_o$, is distributed between labor and land payments. Specifically, the share of revenue allocated to labor payments is $\beta \frac{w_o^{1-\nu}}{c_o^{1-\nu}}$, and the share allocated to land payments is $(1-\beta)\frac{Q_o^{1-\nu}}{c_o^{1-\nu}}$. Developers, who are responsible for producing floorspace, own the land and use their land payments to consume the final good.

4.6 Scale externalities

The model incorporates externalities in two ways: through total factor productivity (TFP) and amenities. For TFP, I use a constant scale elasticity formulation (Allen and Donaldson, 2023), while for amenities, I adopt a heterogeneous scale elasticity (Bartelme, Li and Velasquez, 2024). The respective specifications are:

$$\begin{aligned} A^T_{j,t} &= \bar{A}^T_{j,t} \cdot \left(\tilde{L}^T_{j,t-1}\right)^{\delta}, \qquad (19) \\ u_{i,t} &= \bar{u}_{i,t} \cdot \begin{cases} \gamma_0^{\phi} & \text{if } L_{i,t-1} < \gamma_0 \\ (L_{i,t-1})^{\phi} & \text{if } \gamma_0 \le L_{i,t-1} \le \gamma_1 \\ \gamma_1^{\phi} & \text{if } L_{i,t-1} > \gamma_1 \end{cases} \end{aligned}$$

The key difference between these formulations lies in how externalities respond to changes in scale. The constant scale elasticity used for TFP assumes proportional productivity gains with increases in the labor force $\tilde{L}_{j,t-1}$. In contrast, the heterogeneous scale elasticity for amenities allows the strength of externalities to vary across regions according to population size $\mathrm{L}_{j,t-1}$. This specification captures the following logic in a reduced-form way: large locations face congestion effects that diminish the benefits of additional people, while small locations lack the critical mass required to generate endogenous amenities. Mid-sized regions—such as suburbs—strike a balance, avoiding excessive congestion while benefiting from a sufficiently large population.

I use heterogeneous externalities instead of a constant scale elasticity for these reasons. First, this approach provides the model with a better chance of matching the reduced-form evidence. Increases in trade access and commute access on population (and employment) constitute at least two key moments (four if employment is also considered). Thus, incorporating more parameters, rather than just one, offers greater flexibility. Second, in practice, I find that amenity externalities play a larger role in explaining the observed effects than TFP externalities, making it more practical to allow for heterogeneity where it matters most. Third, while it is theoretically possible to use heterogeneous scale elasticities for both TFP and amenities, doing so would substantially expand the parameter space. In practice, I employ an indirect inference approach to align the effects of commuting and trade access in the data with those in the model. Expanding the model to estimate six parameters would make this process computationally intractable.

The relationship between trade and commute access effects and amenity externalities is as follows. Both lower trade costs and lower commute costs trigger migration responses that are amplified by amenity externalities. Trade-induced migration occurs as households seek greater market access, while commute-induced migration results from households seeking better access to jobs. According to the data, the reduction in commuting costs was particularly significant in dense suburban areas, while reductions in trade costs were more pronounced in less dense areas. To reconcile the empirical observation that improvements in commute access generate minimal population responses on average, amenity externalities must be inactive in the denser areas where commute cost reductions were substantial. Conversely, to explain the large population responses to improvements in trade access, amenity externalities need to be active in the less dense areas where trade cost reductions were significant.

Finally, I assume dynamic rather than static externalities in the model. While both static and dynamic externality models can exhibit multiplicity when the slope of the externality function is large enough, this particular dynamic formulation ensures a unique equilibrium path, conditional on initial conditions, even when multiplicity may arise in the steady state. Thus, it provides a natural mechanism for equilibrium selection within the model, even when the estimated slope is large, allowing for the possibility of multiplicity.

4.7 Market clearing and equilibrium

Market clearing. The model reduces to market clearing conditions for employment and floorspace, where I have omitted the t subscript:

$$\underbrace{\tilde{L}_d w_d}_{\text{Commuting}} = \underbrace{\varrho_L^T \left(\frac{w_d}{c_d^T}\right)^{1-\sigma} X_d^T}_{\text{Tradable Sector}} + \underbrace{\varrho_L^{NT} \left(\frac{w_d}{c_d^{NT}}\right)^{1-\sigma} X_d^{NT}}_{\text{Non-Tradable Sector}} + \underbrace{\beta \left(\frac{w_o}{c_o^{Cons}}\right)^{1-\nu} X_d^{Cons}}_{\text{Construction Sector}} \tag{20}$$

$$\underbrace{\frac{Q_d T_d}{(1-\beta)\left(\frac{Q_d}{c_d^{Cons}}\right)^{1-\nu}}}_{=H_d R_d} = \underbrace{\varrho_H^T \left(\frac{\tilde{R}_d}{c_d^T}\right)^{1-\sigma} X_d^T}_{\text{Tradable Sector}} + \underbrace{\varrho_H^{NT} \left(\frac{\tilde{R}_d}{c_d^{NT}}\right)^{1-\sigma} X_d^{NT}}_{\text{Non-Tradable Sector}} + \underbrace{(1-\alpha)\bar{y}_d L_d}_{\text{By Workers}} \tag{21}$$

where note that $\tilde{L}_d$ is employment and L_d is population.

The left-hand side of the employment equation represents labor supply, which depend on commuting costs. The right hand side is total labor demand which depends on trade linkages. Moreover, the left-hand side of the floorspace equation, which is the supply of floorspace, already imposes land market clearing.

To close the model, I use the fact that exports from county i are equal to total income of the tradable sector $X_i^T = \sum_d X_{id}^T = \sum_d \lambda_{id|d} X_d^T$, with balanced trade ensuring exports equal imports: $X_i^T = P_i^T C_i^{W,T} + P_i^T C_i^{H,T} + (1 - \varrho_L^T - \varrho_H^T)\left(\frac{P_i^T}{c_i^T}\right)^{1-\sigma} X_i^T$, with $w_i \tilde{L}_i^T = \varrho_L^T \left(\frac{w_i^T}{c_i^T}\right)^{1-\sigma} X_i^T$. From these, I obtain an expression for X_d^T which I plug into the employment and floorspace market clearing conditions:

$$X_d^T = \sum_o \frac{A_d^T \left(c_d^T \varsigma_{do}^T\right)^{-\varepsilon}}{\Psi_o^T} \left(\underbrace{P_o^T C_o^{W,T}}_{\text{By Workers}} + \underbrace{P_o^T C_o^{H,T}}_{\text{By Landlords}} + \underbrace{\frac{(1 - \varrho_L^T - \varrho_H^T)}{\varrho_L^T} \left(\frac{P_o^T}{w_o^T}\right)^{1-\sigma} w_o \tilde{L}_o^T}_{\text{By Tradable Firms}} \right) \tag{22}$$

where $P_d^T C_d^{W,T} = b\left(\frac{P_d^T}{P_d}\right)^{1-\sigma} \alpha \bar{y}_d L_d$ is the consumption by workers, and $P_d^T C_d^{H,T} = b\left(\frac{P_d^T}{P_d}\right)^{1-\sigma} Q_d T_d$ is the consumption by landowners. The system of equations for employment and floorspace solves for wages (w_d) and land prices (Q_d).[19]

Equilibrium. We are ready to define the equilibrium of the model.

[19] We can substitute $R_d = \frac{\left(\beta w_d^{1-\nu} + (1-\beta) Q_d^{1-\nu}\right)^{\frac{1}{1-\nu}}}{Z_d}$ into the system to express R_d in terms of Q_d.

Definition 1. Given geography of bilateral costs $\left\{\varsigma_{od,t}^{NT}, \varsigma_{od,t}^{T}, d_{od,t}, \xi_{n,i,t}\right\}_{o,d,t}$, set of characteristics and initial conditions $\left\{A_{d,t}^{T}, A_{d,t}^{NT}, u_{d,t}, T_d, Z_{d,t}, \bar{L}_0\right\}_{d,t}$, and parameters $\{\varrho_L^T, \varrho_H^T, \varrho_L^{NT}, \mathrm{b}, \beta, \alpha, \theta, \eta, \varepsilon, \sigma, \sigma_D, \nu, \gamma, \phi, \zeta_t\}$, the equilibrium is a path of prices, $\{w_{d,t}, Q_{d,t}\}$, and allocations, $\left\{H_{d,t}, \tilde{L}_{d,t}\right\}_{d,t}$, such that there is floorspace market clearing, labor market clearing, total population adds up, $\sum_i L_{it} = \zeta_t \sum_i L_{it-1}$.

4.8 Comparative statics in a simplified environment

To build intuition about how changes in commuting and trade access affect outcomes, I simplify the framework and perform comparative statics. My focus is on understanding the impact of reduced driving times on suburbs and urban cores.

First, I streamline the production and consumption structure. Each location contains a single tradable sector that relies exclusively on labor. While migration decisions remain, I assume no migration frictions. I also disregard externalities, reducing the model to a framework with one sector, one input, and explicit commuting and trade interactions across locations. With these assumptions, the labor market clearing condition simplifies to:

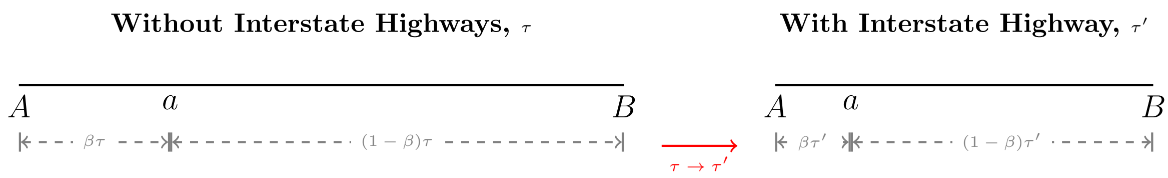
$$\underbrace{w_d \sum_i \pi_{ij|i} L_i}_{\text{Labor Supply}} = \underbrace{\sum_o \lambda_{do|o} \left(\bar{y}_o L_o\right)}_{\text{Labor Demand}} \tag{23}$$

where:

$$\lambda_{do|o} = \frac{A_d \left(c_d \varsigma_{do}\right)^{-\varepsilon}}{\Psi_o} \quad L_i = \frac{\bar{U}_i^{\eta} \bar{L}}{\sum_{i\in N} \bar{U}_i^{\eta}} \quad \bar{U}_{ni} = \frac{u_i}{P_i} \cdot \Phi_i^{1/\theta}$$
$$\bar{y}_i = \sum_j \pi_{ij|i} w_j \quad \pi_{ij|i} = \frac{\left(\frac{w_j}{d_{ij}}\right)^{\theta}}{\Phi_i} \quad \Phi_i = \sum_{l=1}^{S} \left(\frac{w_j}{d_{ij}}\right)^{\theta}$$
$$\Psi_i = \sum_j A_j \left(w_j \varsigma_{ji}\right)^{-\varepsilon} \quad P_i = \left(\Psi_i\right)^{-1/\varepsilon}$$

Second, I streamline the geography to focus on three linearly arranged locations: a core city A, its suburb a, and another core city B. The initial travel time between A and B is τ minutes, with the suburb a positioned $\beta\tau$ minutes from A and $(1-\beta)\tau$ minutes from B. I analyze how constructing a highway between A and B, which reduces travel time (τ), affects the spatial distribution of economic activity. Figure 8 illustrates the geography before and after the highway construction, emphasizing the compression of driving times.

Figure 8: Simplified geography before and after a highway construction



In response to the construction of the highway, the population in suburb a increases according to the following expression:

$$-\frac{1}{L_a}\frac{dL_a}{d\tau} = -\left(\Omega_{L_a/w_A} \cdot \frac{dw_A/d\tau}{w_A}\right) - \left(\Omega_{L_a/w_B} \cdot \frac{dw_B/d\tau}{w_B}\right) - \left(\Omega_{L_a/\varsigma} \cdot \varkappa\tau^{-1}\right) - \left(\Omega_{L_a/d} \cdot \kappa\right) \qquad (24)$$

Here, the first term, $-\left(\Omega_{L_a/w_A} \cdot \frac{dw_A/d\tau}{w_A}\right)$, represents migration to a driven by changes in the equilibrium wage in A. The second term, $-\left(\Omega_{L_a/w_B} \cdot \frac{dw_B/d\tau}{w_B}\right)$, captures migration to a resulting from changes in the equilibrium wage in B. The third term, $-\left(\Omega_{L_a/\varsigma} \cdot \varkappa_T\tau^{-1}\right)$, reflects migration to a due to reductions in trade costs enabled by the highway. Finally, the fourth term, $-\left(\Omega_{L_a/d} \cdot \kappa\right)$, accounts for migration to a prompted by reductions in commute costs generated by the highway. For simplicity, I assume wages in a serve as the numeraire, implying $\frac{dw_a/d\tau}{w_a} = 0$.

First-order approximations. I start by studying the first-order effects of highway construction which are given by the last two terms of equation 24. The response of the population to commute cost reductions is governed by the elasticity $-\Omega_{L_a/d}$:

$$\Omega_{L_a/d} = \eta\left\{-\varpi_A\left(\beta\pi_{Aa} + \pi_{AB}\right) + \left(1-\varpi_a\right)\left(\beta\pi_{aA} + \left(1-\beta\right)\pi_{aB}\right) - \varpi_B\left(\pi_{BA} + \left(1-\beta\right)\pi_{Ba}\right)\right\}$$

where ϖ_i is the share of households residing in i.

Reducing commuting costs affects location choices, as individuals consider commuting access when deciding where to live. Ceteris paribus, on average, more people will choose the location offering better commuting access. Consequently, when commuting times change across all locations, it is essential to identify where commuting access increases the most. For example, commuting access in A increases by $\beta\pi_{Aa} + \pi_{AB}$, while commuting access in B rises by $\pi_{BA} + (1-\beta)\pi_{Ba}$. Similarly, commuting access in a increases by $\beta\pi_{aA} + (1-\beta)\pi_{aB}$.

To a first-order approximation, the greater the commuting flows from a location to other regions, the larger the impact of reducing commute costs on that location's commuting access. How these changes in commuting access influence the decision to live in a depends on their effect on the utility of living in a (which receives a weight of 1) relative to their effect on the average utility, where each of these changes is weighted by ϖ_i.

Now, let us assume that no one commutes to or from B, effectively treating B as being 'too far away'. In this scenario, commuting flows exist only between A and a. Furthermore, suppose the share of people commuting to A from a is initially higher than the share commuting in the opposite direction. In other words, I assume that $\pi_{AB} = \pi_{BA} = \pi_{Ba} = \pi_{aB} = 0$ and $\pi_{aA} > \pi_{Aa}$. Under these assumptions, the following relationship holds:

$$-\Omega_{L_a/d} = -\beta\varpi_A \left\{\pi_{Aa} - \pi_{aA}\right\} - \beta\varpi_B \left\{-\pi_{aA}\right\} > 0 \tag{25}$$

To a first-order approximation, if workers in suburb a commute to other locations more intensively than workers in A or B, then suburb a will expand in response to reductions in commuting costs (note the double negative).

Now, let's turn to trade cost effects. The response of population in a given trade cost reductions is:

$$-\Omega_{L_a/\varsigma} \quad = \eta \left\{-\varpi_A \left(\lambda_{aA|A} + \lambda_{BA|A}\right) + (1 - \varpi_a) \left(\lambda_{Aa|a} + \lambda_{Ba|a}\right) - \varpi_B \left(\lambda_{AB|B} + \lambda_{aB|B}\right)\right\}$$

where ϖ_i is the share of the population living in i.

Trade cost reductions, to a first order, affects location choices because households aim to minimize the price of tradable goods. Since trade costs are decreasing everywhere, it is important to identify where the reductions are the most pronounced. This depends on which locations, ex-ante, are more dependent on trade to supply the goods that households consume.

We assume that suburb a relies more on core city A than A relies on suburb a, and that suburb a also relies more on A than core B relies on A. Specifically, $\lambda_{aA|A} < \lambda_{AB|B} < \lambda_{Aa|a}$. Moreover, suburb a relies more on core B than A relies on B, and more than core B relies on suburb a, i.e.,

$\lambda_{aB|B} < \lambda_{BA|A} < \lambda_{Ba|a}$. Under these assumptions, we can show that:

$$-\Omega_{L_a/\varsigma} = -\left(\underbrace{\lambda_{aA|A} - \lambda_{Aa|a}}_{<0} + \underbrace{\lambda_{BA|A} - \lambda_{Ba|a}}_{<0}\right) - \varpi_B\left(\underbrace{\lambda_{AB|B} - \lambda_{Aa|a}}_{<0} + \underbrace{\lambda_{aB|B} - \lambda_{Ba|a}}_{<0}\right) > 0. \quad (26)$$

By imposing $\lambda_{aA|A} < \lambda_{AB|B} < \lambda_{Aa|a}$ and $\lambda_{aB|B} < \lambda_{BA|A} < \lambda_{Ba|a}$, we assume that suburb a is highly dependent on both A and B. Thus, constructing a highway connecting A and B particularly benefits suburb a, since it lies between the two cores.

If we had instead imposed $\lambda_{aA|A} < \lambda_{Aa|a} < \lambda_{AB|B}$ and $\lambda_{aB|B} < \lambda_{Ba|a} < \lambda_{BA|A}$, the effect could reverse, leading to a decrease in population in suburb a despite its reliance on both A and B (more than what each core relies on the suburb). This is because, under this alternative assumption, the cores are highly interdependent, and constructing highways connecting them would cause them to amplify their growth, and expand at the expense of the suburb.

Accounting for general equilibrium responses in wages. While I do not fully characterize the derivatives $\frac{dw_A/d\tau}{w_A}$ and $\frac{dw_B/d\tau}{w_B}$ in the main text, I do explain how changes in wages affect suburban growth. The impact of wages on suburban population is given by:

$$\begin{aligned}
\Omega_{L_a/w_B} &= \eta\left\{-\varpi_A\left(\pi_{AB} - \lambda_{BA|A}\right) + (1-\varpi_a)\left(\pi_{aB} - \lambda_{Ba|a}\right) - \varpi_B\left(\pi_{BB} - \lambda_{BB|B}\right)\right\} \\
\Omega_{L_a/w_A} &= \eta\left\{-\varpi_A\left(\pi_{AA} - \lambda_{AA|A}\right) + (1-\varpi_a)\left(\pi_{aA} - \lambda_{Aa|a}\right) - \varpi_B\left(\pi_{BA} - \lambda_{AB|B}\right)\right\}
\end{aligned}$$

The difference $\pi_{ij} - \lambda_{ji|i}$ indicates that when wages in location j increase, commute access to location i improves because more people commute to j, but market access to i decreases as goods from j become more expensive. If $\pi_{ij} > \lambda_{ji|i}$, commute access increases more than market access decreases, prompting migration to location i; otherwise, market access dominates, leading to outmigration. This demonstrates how trade and commuting interact in shaping population responses to wage changes.[20] For example, if a location i primarily sends workers to producing locations without sourcing goods from them, wage increases in the producing locations will raise commute access more than what market access will decrease, leading to migration to the location i. Conversely, if a location i relies on goods from the producing location more than it sends workers, wage increases will reduce market access

[20] Trade and commuting also interact in shaping the wage responses to highway construction, $\frac{dw_A/d\tau}{w_A}$ and $\frac{dw_B/d\tau}{w_B}$.

more than they improve commute access, resulting in outmigration from location i.

By applying the restrictions $\lambda_{aA|A} < \lambda_{AB|B} < \lambda_{Aa|a}$ and $\lambda_{aB|B} < \lambda_{BA|A} < \lambda_{Ba|a}$, we can show that:

$$\begin{aligned} \Omega_{L_a/w_B} &= \varpi_A \left(\underbrace{\lambda_{BA|A} - \lambda_{Ba|a}}_{<0} \right) \eta - \varpi_B \underbrace{\left(1 - \lambda_{BB|B} + \lambda_{Ba|a}\right)}_{>0} \eta < 0 \\ \Omega_{L_a/w_A} &= \left\{ \varpi_A \underbrace{\left(\left(\pi_{aA} - \lambda_{Aa|a}\right) - \left(\pi_{AA} - \lambda_{AA|A}\right)\right)}_{?} \eta + \varpi_B \underbrace{\left(\pi_{aA} + \lambda_{AB|B} - \lambda_{Aa|a}\right)}_{?} \eta \right\} \end{aligned} \tag{27}$$

The first equation indicates that if wages in B increase, migration to suburb a will decrease. This is because the loss of competitiveness in location B causes workers to move from B to A rather than to a, as suburb a depends heavily on B, and the reduction in market access is significant. In the alternative scenario where the cores were highly interdependent, particularly where core A relies on B for goods, then a rise in wages in core B could potentially increase the population in suburb a. The second equation shows that, under the current assumptions, the effect of a wage increase in A on suburb a is ambiguous. It depends on whether the suburb relies on A for goods more than as a commuting destination. If the net reliance of the suburb on A (reliance on goods relative to employment) exceeds the net reliance of core B, then the effect could favor migration to a.

Ultimately, whether suburbs expand in response to highways affecting wages depends on several factors discussed here. These include the elasticity of migration with respect to wage changes, as well as the effects of highways on wages, $\frac{dw_A}{d\tau}/w_A$ and $\frac{dw_B}{d\tau}/w_B$, which I discuss in Appendix A.5.

5 Bridging Theory and Data

In this section, I describe the process of matching the model to real data, estimating key parameters, and calibrating the remaining ones.

5.1 Model inversion

Based on the knowledge on some parameters, and availability of some data, I can invert the model and recover all the model's unobservables, which are required to simulate counterfactual scenarios. A key feature of the model is that we can invert for unobservables, and then decompose them into endogenous (i.e. due externalities) and exogenous unobservables. The following theorems formalize these ideas.

THEOREM 5.1: *Bilateral commuting flows. Given parameter* $\{\theta\}$*, data* $\{w_d\}$*, and bilateral commute costs* $\{d_{od}\}$*, there exists a unique set of commuting preferences* $\{D_{od}\}$ *that rationalize bilateral commute flows* $\{\pi_{od|o}\}$*.*

Proof. See Appendix A.1 □

Theorem 5.1 uses data on wages and commute costs to predict commuting flows. The theorem's logic relies on identifying commuting preferences D_{od} such that $\frac{D_{od}(w_d/d_{od})^\theta}{\Phi_o} = \pi_{od|o}$, where $\Phi_o = \sum_d D_{od}(w_d/d_{od})^\theta$.

THEOREM 5.2: *Construction labor market clearing. Given parameters* $\{\nu, \beta\}$ *and data* $\{\tilde{L}_d, w_d, H_d, R_d\}$*, there is a unique set of construction TFP,* $\{Z_d\}_d$*, and land supply,* $\{K_d\}_d$*, that rationalize the data as an equilibrium of the model.*

Proof. See Appendix A.2 □

Theorem 5.2 uses labor and labor market clearing in the construction sector plus perfect competition, along with data on floorspace, wages, and sectoral employment, to infer land supply and construction productivity.

THEOREM 5.3: *Non-tradable and tradable labor market clearing. If* $\alpha\bar{y}_{|d}L_d + (1-\beta)\left(\frac{Q_d}{c_d^{Cons}}\right)^{1-\nu} R_d H_d \geq \tilde{L}_d^{NT} w_d, \forall d$*, given parameters* $\{\sigma, \sigma_D, \varepsilon, \theta, \nu, \varrho^T, \varrho_L^{NT}, b, \alpha, \beta\}$*, data* $\{L_d, \tilde{L}_d^s, w_d, H_d, R_d,\}$*, and bilateral costs* $\{\varsigma_{od}^{NT}, \varsigma_{od}^T, d_{od}\}$*, there is a unique set of TFP in the tradable and non-tradable sectors* $\{A_d^T, A_d^{NT}\}_d$ *that rationalize the data as an equilibrium of the model.*[21]

Proof. See Appendix A.3 □

Theorem 5.3 leverages labor market clearing in both the tradable and non-tradable sectors, alongside perfect competition, to recover unobserved sectoral productivity using data on floorspace units

[21] $L_{d,-1}$ indicates population in the previous period.

and prices, wages, and sectoral employment. A sufficient condition is that the maximum income available for spending within a location is at least as large as the wage bill in the non-tradable sector, which we observe from the data. In models without commuting, this condition is trivially satisfied.[22] Once this condition holds, we can recover the total factor productivity (TFP) in the non-tradable sector by ensuring the income spent on workers in the non-tradable sector matches the wage bill implied by the data.

THEOREM 5.4: *Migration. If* $\alpha \bar{y}_{|d} L_d + (1-\beta)\left(\frac{Q_d}{c_d^{Cons}}\right)^{1-\nu} R_d H_d \geq \tilde{L}_d^{NT} w_d, \forall d$*, given parameters* $\{\sigma, \sigma_D, \varepsilon, \theta, \nu, \varrho^T, \varrho_L^{NT}, b, \alpha, \beta, \eta\}$*, data* $\{L_d, L_{d,-1}\tilde{L}_d^s, w_d, H_d, R_d, \}$*, and bilateral costs* $\{\varsigma_{od,t}^{NT}, \varsigma_{od}^{T}, d_{od}, \xi_{od}\}$*, there is a unique set of amenities* $\{u_d\}$ *that rationalize the data as an equilibrium of the model.*

Proof. See Appendix A.4 □

Finally, Theorem 5.4 uses all previous theorems, and hence requires the same data and parameters, plus migration elasticities η and bilateral migration frictions $\xi_{ni,t}$. It relies on the migration equations derived from the model to rationalize the time series of population across counties, $L_{i,t}$.

5.2 Calibration

This section clarifies which parameter values I estimate from the data, and which I calibrate from existing literature.

Externalities. Scale externalities for TFP, δ, are calibrated following updated estimates in the literature (Bartelme et al., 2024). I choose a value of $\delta = 0.1$, which is close to the median across sectors. I estimate amenity externalities following an indirect inference approach, which is explained in detail in the following section.

Commuting parameters. I assume an exponential representation of commute costs following papers in the literature: $d_{ij} = \exp(\kappa \tau_{ij})$. Appendix A.7 estimates the relationship between commute times and commute flows in the short- and long-run. For the model, I assume an average semi-elasticity of $\kappa\theta$=0.0375 which is the average of all the coefficients shown in Appendix A.7.[23] Then, I use estimates

[22] For instance, in a model without commuting or a floorspace market, the condition simplifies to $\bar{y}_d L_d \geq \tilde{L}_d^{NT} w_d$, where income and population equal wages and total employment ($\tilde{L}_d^{NT} + \tilde{L}_d^{T}$).

[23] This paper presents a static model of commuting, despite the dynamic gravity equation indicating differences between short- and long-run elasticities. I am currently developing a paper that explores the implications of dynamic commuting. Since commuting is static in this paper, I simply assume an average elasticity across time. If I instead a calibrate

of θ from the literature to calibrate θ and κ. In particular, Monte et al. (2018) estimates suggest a value of $\theta = 3.3$. From, there I get a value of $\kappa = 0.0375/3.3 = 0.0114$.

Migration parameters. I calibrate migration costs as done typically in the literature and assume: $\xi_{ij} = \tau_{ij}^{-\varkappa_m}$. For the sake of computing migration costs, I fix 1960 driving times to represent county distances. This assumes highways influence migration only indirectly through prices, commuting access, and amenities, but not directly through reductions in driving times. From 2000–2010 migration data, I estimate a gravity equation and find $\varkappa_m \eta = 1.25$ (see Appendix A.8 for details). Following the literature, I set $\eta = 4$. Population growth ζ_t is calibrated directly from the data.

Trade parameters. I calibrate trade costs as $\varsigma_{ji} = B_{s(i),s(j)} \tau_{ji}^{\varkappa_T}$. Using Monte et al. (2018), I set the elasticity of driving times to trade flows as $\varkappa_T \varepsilon = 1.5$, with $\varepsilon = 4$ from Broda and Weinstein (2006). The term $B_{s(i),s(j)}$ representing home bias equals one for intra-state trade and exceeds one for inter-state trade. I calibrate $B_{s(i),s(j)}$ to match 1992 data on average intra-state expenditure from the Freight Analysis Framework.

Floorspace production parameters. Based on Albouy et al., (2018), the share of land in floorspace production, $1 - \beta$, is about 0.4. The elasticity of substitution between land and labor, drawn from Epple et al., (2010), Ahlfeldt and McMillen (2018, 2020), and Combes et al., (2021), ranges from 0.75 to 1.4. I select the value of $\nu = 1.20$ to reflect increased substitution toward land as its price falls.[24]

Tradable and non-tradable production parameters. For the production of tradables, intensities for labor, floorspace, and intermediate sum to one ($\sum_k \varrho_k^T = 1$). For non-tradables, the only factors are floorspace and labor, and their intensities sum to one ($\sum_k \varrho_k^{NT} = 1$). I assume labor and land intensities of $\varrho_L^T = \varrho_H^T = 0.25$ for tradables, with intermediates at $1 - \varrho_L^T - \varrho_H^T = 0.50$, aligning with input-output data and Valentinyi and Herrendorf (2008).[25] For non-tradables, I assume labor and land intensities of $\varrho_L^{NT} = 2/3$ and $\varrho_L^T = 1/3$, following Valentinyi and Herrendorf.[26] Regarding, the

the elasticity in this paper using the value after 40 years ($\theta\kappa = 0.060$), every decade's commuting will be excessively responsive to commuting times.

[24] 1.20 is the middle value that allows for substitution (i.e. above one) and that is within the range of estimates (below 1.4). Note that estimates in the literature concern land and capital, not labor as in this paper. Future versions may include intermediate inputs as a proxy of capital.

[25] See Table 7, Valentinyi and Herrendorf (2008), with re-normalized values excluding non-tradable intermediates, and capital relabeled as floorspace.

[26] See Table 7, Valentinyi and Herrendorf (2008), with re-normalized values excluding intermediates, and capital relabeled as floorspace.

elasticity of substitution between floorspace and labor, I assume a value of $\sigma = 0.3$, the midpoint value in Behrens, et al., (2022) range of 0.2-0.4.

Household consumption parameters. From the Bureau of Labor Statistics Expenditure Surveys, the share of household consumption allocated to land is approximately $1 - \alpha = 1/3$. The intensity of tradables in the final goods bundle, denoted by b, is directly derived from the data and calibrated to ensure that the sum of exports is equal to the sum of imports:

$$\sum_i X_i^T = \sum_i \left(P_i^T C_i^{W,T} + P_i^T C_i^{H,T} + (1 - \varrho_L^T - \varrho_H^T) \left(\frac{P_i^T}{c_i^T} \right)^{1-\sigma} X_i^T \right)$$

where $P_d^T C_d^{W,T} = b \left(\frac{P_d^T}{P_d} \right)^{1-\sigma} \alpha \bar{y}_d L_d$ is the consumption by workers, and $P_d^T C_d^{H,T} = b \left(\frac{P_d^T}{P_d} \right)^{1-\sigma} Q_d T_d$ is the consumption by landowners. I assume an elasticity of substitution between tradables and non-tradables of $\sigma_D = 0.5$, roughly consistent with findings in the literature (e.g. 0.44 in Stockmand and Tesar, 1995; 0.74 in Mendoza, 1991).

Theorems and unobservables. With these parameters (except for externalities), I apply Theorems 5.1-5.4 to back out unobservables for each county-by-decade. This allows me to match county-level data by decade on population, employment, bilateral commuting, wages, floorspace prices, and floorspace supply. I then use the recovered amenities to perform the indirect inference approach outlined in the next section.

5.3 Indirect inference

The goal of this section is to estimate the amenity externality function, which includes three parameters: the slope of the externality, ϕ, and two location thresholds, γ_0 and γ_1.[27] I utilize an indirect inference approach and map the externality parameters to the reduced-form evidence I reported in Section 3.3. The primary advantage of employing an indirect inference approach is that it allows us to leverage the general equilibrium responses generated by the model and compare them to those observed in the data.

I estimate the effects of trade and commute access on population and employment using model-generated data. Using Theorems 5.1-5.4, I recover unobservable factors for each decade, hold them

[27] To avoid estimating three parameters, I calibrate γ_0 using the average rural population of 3,000, meaning externalities are inactive below this threshold.

fixed at their 1960 levels, and solve the model under two scenarios: one without highways and one incorporating highways from the Pershing Plan. This approach assumes that the assignment of highways under the Pershing Plan is effectively random. By comparing population and employment growth in the scenario with highways to the counterfactual scenario without highways, and relating those changes to the shifts in commute and trade access induced by the highways, I estimate the effects of trade and commute access on population and employment after 40 years.

By comparing the model's moments to the data, I search for the amenity elasticity ϕ and the upper threshold γ_1 that minimizes the difference between the model's reduced-form regressions and the data's reduced-form evidence. I use the IV estimates as target moments. That is, I target four moments: the trade and commute access effect on population and employment. I minimize the quadratic loss function:

$$Loss = [\beta^{data} - \beta^{model}]I[\beta^{data} - \beta^{model}]' \tag{28}$$

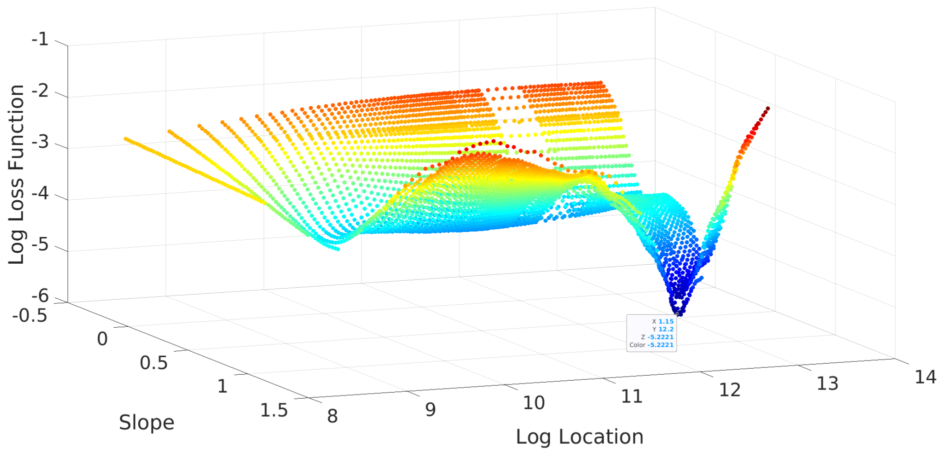
where β^{data} and β^{model} are row vectors, and I is the identity matrix.

Panel A of Figure 9 presents the resulting loss function. As the location parameter approaches infinity, the loss function increases, indicating a poorer fit between the model and the data. This serves as evidence against the constant scale formulation. Similarly, as the slope parameter approaches zero, the loss function increases for every location parameter, providing evidence against the no externality formulation.

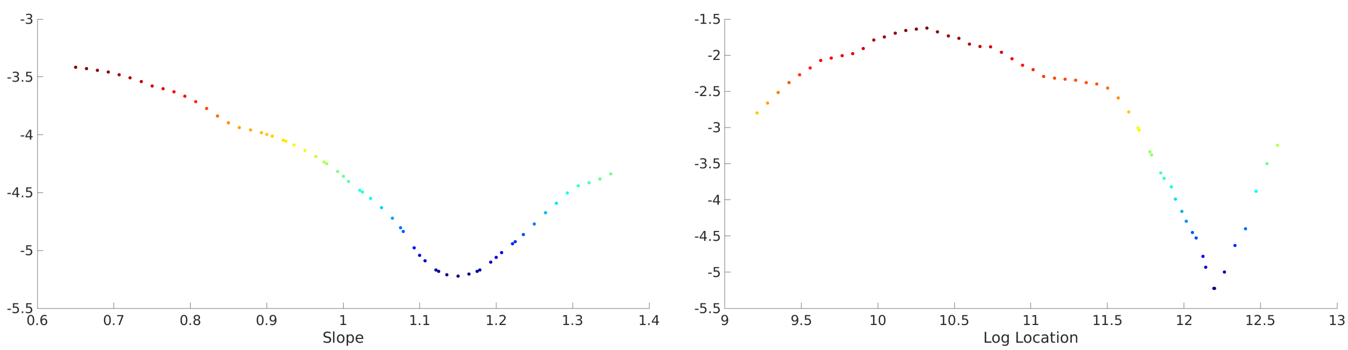
I estimate a slope parameter of $\phi = 1.15$ and a location threshold of $\gamma_1 = 198,790$. Panel B of Figure 9 shows sharp identification, especially for the location parameter. The estimated effects are: trade access increases population by 0.21 (data = 0.20) and employment by 0.17 (data = 0.24), while commute access decreases population by -0.10 (data = -0.08) and employment by -0.10 (data = -0.08). The model closely matches population changes but underestimates trade access effects on employment.

Figure 9: Loss Function

Panel A: Log Loss vs Slope and Log Location



Panel B: Fixing one parameter to its estimated value while varying the remaining one



Regarding the magnitude of these parameter estimates, some comments are in order. While the slope is larger than typical values in the literature, it remains reasonable for two reasons. First, in contrast to constant scale elasticities—where large locations experience unbounded improvements as externalities never diminish—heterogeneous scale elasticities limit the amplification of externalities to the range between γ_0 and γ_1. This ensures that beyond γ_1, population growth stabilizes, yielding a unique steady state in partial equilibrium for large locations even for large slopes. Second, some studies assume negative amenity externalities to capture congestion effects associated with fixed land. While my model includes variable floorspace, land remains fixed, which introduces an endogenous congestion force. My calibration already accounts for these features, allowing the estimated amenity externalities to be higher, and even positive, as seen here. Moreover, my calibration also considers imperfect migration given the Frechet assumption for migration preferences, where the shape parameter acts as another dispersion force.

With regards the location parameters, their value highlights how congestion offsets amenity externalities in the largest counties, and the lack of critical mass in the smallest counties. By 1960, approximately 1,800 out of 3,092 counties lie within the interval $\gamma_0 = 3,000$ to $\gamma_1 = 198,790$, with the 35 largest counties falling outside the upper threshold. By 2020, approximately 2400 counties lie in this interval, with 120 counties above the upper threshold. This suggests that in these large counties, congestion forces are strong enough to offset any positive amenity externalities, which is a plausible finding. Importantly, this does not imply that living in large cities offers no advantages. My calibration assumes TFP externalities follow a constant scale formulation, meaning productivity gains from agglomeration in large cities remain intact.

6 Counterfactuals: Unpacking Suburban Growth

In this section, I use a model calibrated to the U.S. economy. To calibrate the model, I match county-level data on population, employment, commuting flows, wages, floorspace prices, and floorspace supply from 1950 to 2020. This allows me to back out unobservable shocks. I then simulate a counterfactual scenario without the Interstate Highway System, holding all other shocks between 1960 and 2020 constant. I evaluate three cases: highways reduce only commute costs, only trade costs, and both simultaneously.

First, I discuss the aggregate effects of constructing highways, decomposing the overall impact into three components: the trade effect, the commute effect, and the interaction between trade and commute. Second, to better understand the aggregate results on suburbanization and urban core decline, I examine which components of the utility function were most affected by trade and commute cost reductions. Finally, I explore the heterogeneity across suburbs and urban cores. I identify which suburban characteristics predict suburban growth and urban core decline.

6.1 Aggregate effects: results

Table 3 shows my main findings. Panel A reports changes in population share by NCHS classification, while Panel B shows changes in employment share. Each row presents changes in percentage points. The first row in each panel shows the observed change in shares between 1960 and 2020 (Data Δ).

The second row displays the total effect of highways by comparing the simulated economy with and without highways. The next two rows decompose this total effect into the commute effect (only reduced commuting costs) and the trade effect (only reduced trade costs). Finally, the last row shows the interaction effect, which captures the additional impact when both commuting and trade cost reductions are introduced simultaneously. The interaction is computed as the difference between the total effect and the sum of the commute and trade effect. Each column represents a category in the NCHS classification of counties (Core, Suburbs, Medium, and Small/Rural areas).

Highways account for 0.94 percentage points of the 6.14 percentage point increase in population share attributed to suburbs (15.3%), while they explain -0.88 percentage points of the -3.33 percentage point decline in core areas (26.6%). In the data, medium cities grew by 3.27 percentage points, while small and rural areas shrank by 6.10 percentage points. In contrast, the model predicts medium cities shrinking by 0.60 percentage points and small and rural places growing by 0.53 percentage points.

Table 3: The effect of Interstate Highways on population across county aggregates

Panel A: Population share by NCHS classification				
	Core	**Suburbs**	**Medium**	**Small/Rural**
Data Δ (2020-1960)	-0.0333	0.0614	0.0327	-0.0610
Total Effect (hwy - no hwy)	-0.0088	0.0094	-0.0060	0.0053
Commute Effect	0.0013	0.0083	0.0036	-0.0131
Trade Effect	-0.0100	-0.0035	-0.0125	0.0260
Interaction Effect	-0.0001	0.0047	0.0030	-0.0075
Panel B: Employment share by NCHS classification				
	Core	**Suburbs**	**Medium**	**Small/Rural**
Data Δ (2020-1960)	-0.0157	0.0504	0.0319	-0.0669
Total Effect (hwy - no hwy)	-0.0027	0.0030	-0.0041	0.0040
Commute Effect	0.0033	0.0054	0.0019	-0.0103
Trade Effect	-0.0074	-0.0047	-0.0086	0.0206
Interaction Effect	0.0015	0.0022	0.0026	-0.0063

I unpack the growth driven by highways across three terms: commute-only, trade-only, and interaction effects. The reduction in commute costs contributed to the relative decline of small and rural areas compared to medium cities, suburbs, and core counties, with suburbs experiencing the most growth. Regarding trade access, the reduction in trade costs shifted populations away from core and medium cities toward small and rural areas. While suburbs also contracted, they did so to a lesser extent than cores, so suburbs grew relative to cores, even when both shrunk in absolute terms.

The interaction effect shows that simultaneous reductions in trade and commute costs further enhanced suburban growth, particularly at the expense of rural areas. Notably, the interaction effect accounts for 50% (0.0047/0.0094) of the total increase in suburbs, underscoring the significance of trade in the suburbanization process, which is characterized by the decline of cores and the rise of suburbs.

For employment, as shown in Panel B, highways similarly explain about 17.2% of the decline in core areas and 6% of the rise in suburbs. Qualitatively the commute, trade and interaction effects show similar patterns.

6.2 Aggregate effects: discussion

I now explore the factors driving the observed patterns of suburban growth and urban decline. First, I examine why reductions in commute costs alone promote suburban growth, while reductions in trade costs alone support growth in rural areas. Then, I analyze the interaction between trade and commute cost reductions.

Starting from equation 6, locations with larger increases in job access (i.e., commute access) and consumer access to goods (i.e., trade access) become more desirable places to live. Other factors that influence location choice include access to non-tradable goods, rental prices of floorspace, and amenities, though I focus here on commute and trade access.

(i) First-order changes in commute and trade access. I calculate the average change in commute access due to changes in driving times,

$$\Delta \log \Phi_i = \sum_j \pi_{ij,2020,\text{no hwy}} \times \left(\log d_{ij,\text{hwy}} - \log d_{ij,\text{no hwy}}\right),$$

and the average change in trade access due to changes in driving times,

$$\Delta \log \Psi_i = \sum_j \lambda_{ji,2020,\text{no hwy}} \times \left(\log \varsigma_{ji,\text{hwy}} - \log \varsigma_{ji,\text{no hwy}}\right),$$

across cores, suburbs, medium metros, and rural areas. I use 2020 commute and trade shares from the counterfactual scenario without highways as weights. Growth in commute and trade access is normalized relative to cores, for example, $\Delta \log \Phi_{\text{suburb}} - \Delta \log \Phi_{\text{core}}$. Panel A of Table 4 presents the average values of commute and trade access changes across counties, aggregated into cores, suburbs, medium metros, and rural areas. The results indicate that, to a first order, commute access increased most in suburban locations, while trade access rose most in rural areas. This aligns with findings from the previous section: reductions in commute costs alone favor suburban growth (see "Commute Effect" in Table 3.), while reductions in trade costs encourage rural growth at the expense of cores (see "Trade Effect" in Table 3.).

However, some questions still remain. If commute cost reductions increased commute access in all locations more than in core areas, why do rural areas shrink when commute costs are reduced? (See "Commute Effect" in Table 3.) Similarly, if trade access rises in suburban areas relative to core when trade costs are lowered, why do these areas not grow accordingly? (see "Trade Effect" in Table 3.) To address these questions, next I explore how the general equilibrium effects on prices might explain these patterns.

(ii) Interaction effects between reductions in trade and commute costs. As per Table 3, the interaction between trade cost and commute cost reductions led to further growth in suburbs and medium cities, at the expense primarily of rural areas. Why is that? In Panel B, I calculate changes in commute access based on shifts in wages,

$$\Delta \log \Phi_i = \sum_j \pi_{ij,2020,\text{no hwy}} \times \left(\log w_{j,\text{hwy}} - \log w_{j,\text{no hwy}}\right),$$

and changes in trade access based on shifts in unit costs,

$$\Delta \log \Psi_i^T = -\sum_j \lambda_{ji,2020,\text{no hwy}} \times \left(\log uc^T_{ji,\text{hwy}} - \log uc^T_{ji,\text{no hwy}}\right),$$

This time shifts are computed by comparing the observed data where highways reduce both trade and commute costs against a counterfactual scenario where highway were never built. I also compute shifts in other components of the utility function such as access to non-tradable goods (i.e. inverse of non-tradable prices), floorspace rental prices, and amenities.

The results show that equilibrium price responses increased suburban population through commute access, trade access, access to non-tradable goods, and lower floorspace prices. Meanwhile, amenities primarily grew in medium cities and suburbs. In other words, endogenous amenities play a key role in shaping the interaction between commute and trade cost reductions. Commute cost reductions spurred suburban population growth. Once suburbs reached a critical mass, trade cost reductions—though to a first order favoring rural areas more than suburban areas—further boosted suburban growth, since it triggered a higher accumulation of endogenous amenities than in rural areas.

To sum up, reductions in commute costs primarily benefited suburban areas, where increased job access spurred migration and amplified endogenous amenities, creating a reinforcing growth loop. Trade cost reductions, however, favored rural growth by improving access to consumer goods, drawing residents outward. When both costs declined simultaneously, these forces interacted: suburban areas became focal points for growth as they absorbed both commuter-driven and trade-driven migration.

Table 4: The effect of Interstate Highways on determinants of workers' utility

Panel A: First-order changes in commute and trade access				
	Core	**Suburbs**	**Medium**	**Small/Rural**
Commute access ($\Delta \log \Phi$)	0	0.593	0.340	0.355
Trade access ($\Delta \log \Psi^T$)	0	0.045	0.051	0.075
Panel B: Simultaneous commute and trade cost reductions and their effect on components of workers' utility				
	Core	**Suburbs**	**Medium**	**Small/Rural**
Commute access ($\Delta \log \Phi$)	0	-0.003	0.023	0.060
Trade access ($\Delta \log \Psi^T$)	0	-0.006	-0.007	-0.006
Access to non-tradable goods ($\Delta \log \Psi^{NT}$)	0	-0.084	-0.164	-0.323
Floorspace prices ($\Delta \log R$)	0	0.402	0.454	0.435
Endogenous amenities ($\Delta \log u$)	0	0.213	0.213	0.143

6.3 Heterogeneous effects

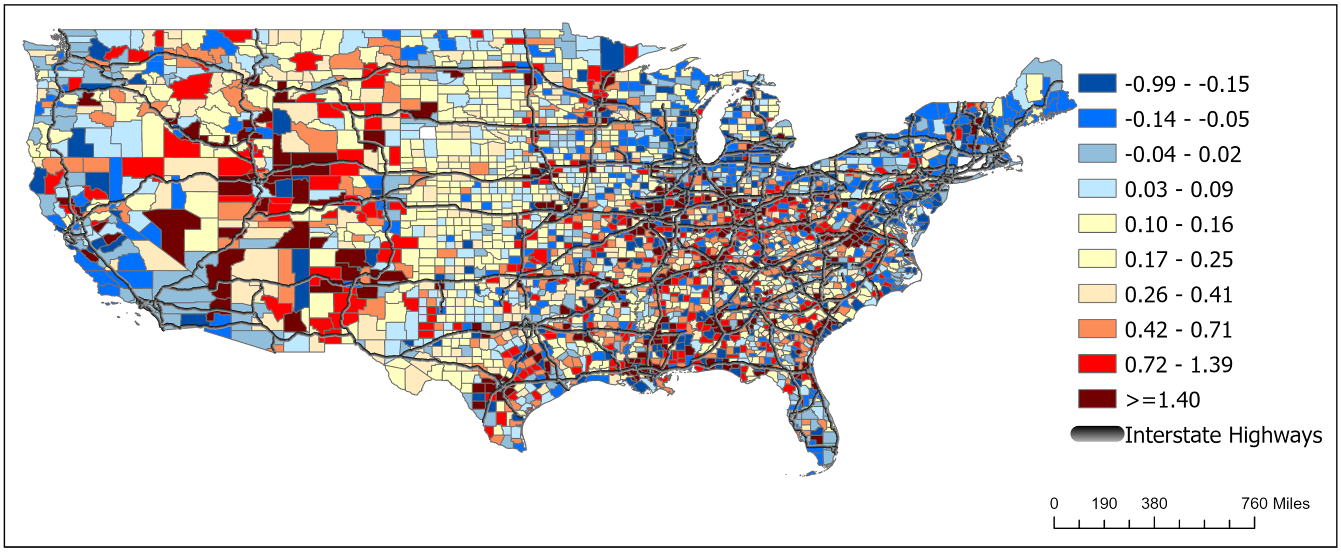
So far, I have focused on the aggregate impacts of highways. However, there is significant heterogeneity across suburbs. In this section, I identify which suburbs expanded due to highways, particularly as a result of trade cost reductions.

Figure 10 highlights the heterogeneity in population and employment growth across U.S. counties. First, the maps reveal that growth disproportionately occurred inland, away from the coasts. Since many of the largest counties are located along the coasts, this suggests a trend toward decentralization.

Second, while population and employment growth are correlated, they do not align perfectly. For example, the employment decline was more pronounced in coastal areas, particularly in the southwest (e.g., Los Angeles), than population decline. Conversely, the central west region experienced stronger employment growth relative to population growth.

Figure 10: Population and employment growth across counties

Panel A: Population growth (highways - no highways)



Panel B: Employment growth (highways - no highways)

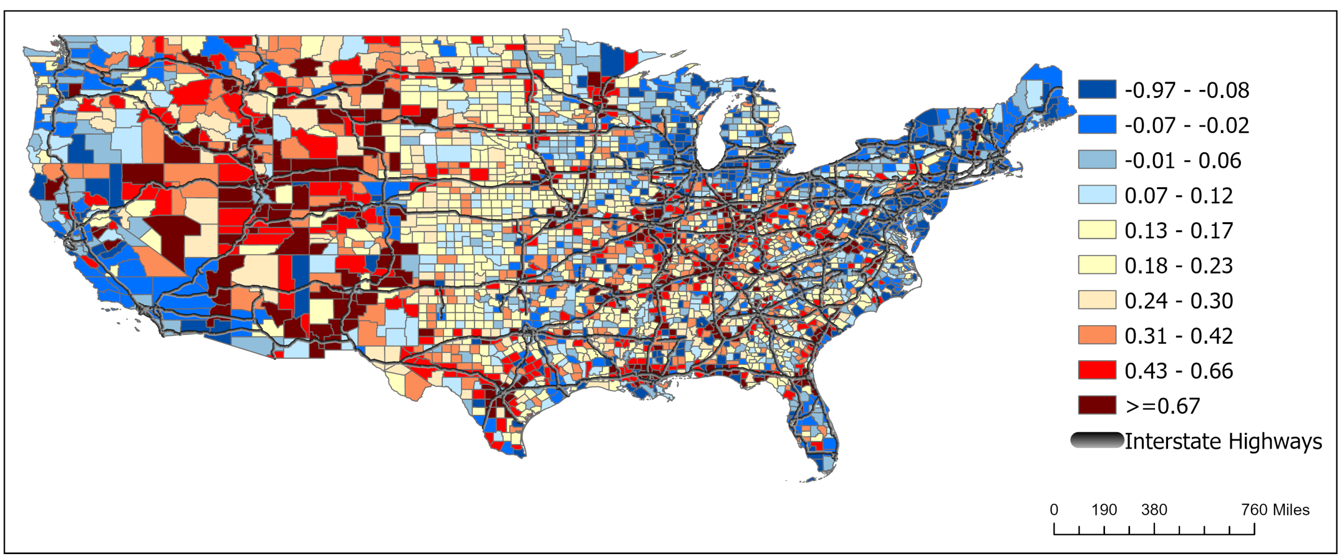


Figure 11: County growth across suburbs and urban cores

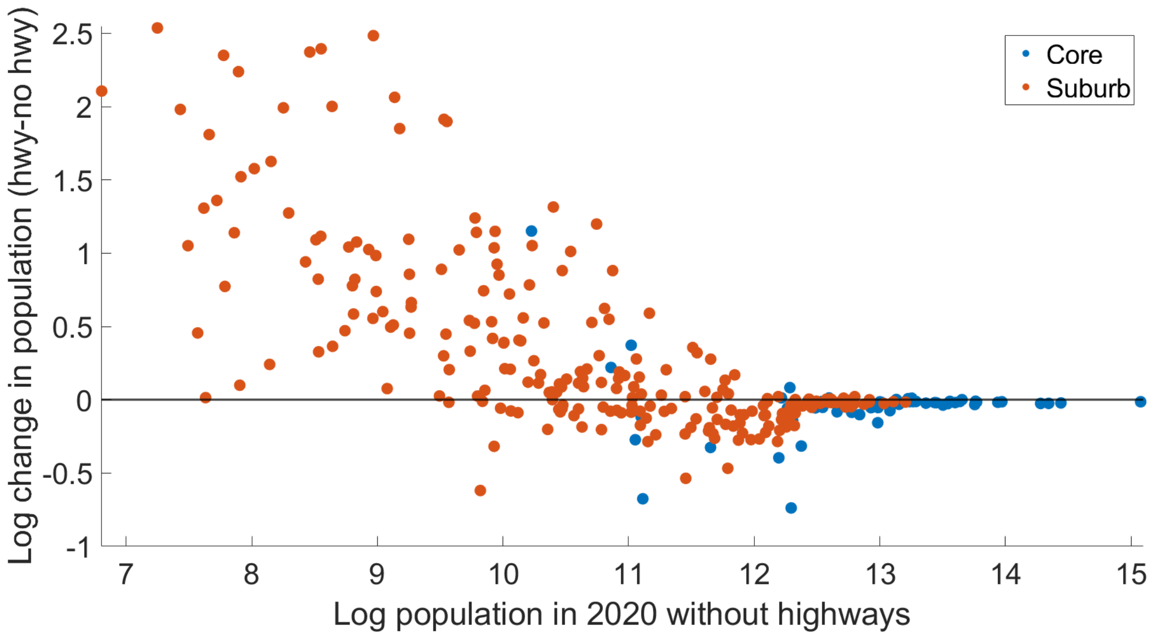


Figure 11 plots the log change in 2020 population between the world with highways and without highways against the 2020 population in the counterfactual world without highways, focusing on suburbs and cores. The results show that population growth in suburbs was concentrated in smaller suburbs, while larger suburbs and most core areas experienced population decline.

I then use regression analysis to characterize population growth in suburbs and cores based on their trade and commute linkages. In Tables 5 and 6 I relate population growth to several factors for cores and suburbs separately. Column 1 includes population in the counterfactual scenario without highways and its square. Column 2 adds the share of commuters and the share of expenditures within the same county (base category is commuters and expenditures elsewhere). Column 3 includes the share of commuters to core and suburban areas, while Column 4 adds the share of expenditures in core and suburban areas (base category is commuters and expenditures in medium cities pooled with rural areas).

Table 5 shows that smaller cores experience the most growth or the least decline. Cores are more likely to grow when residents spend a larger share of their expenditures on goods produced in other cores, rather than locally or in suburbs and rural areas (Column 4). This pattern suggests that stronger trade links between core areas support core growth. For suburbs, Table 6 shows that smaller suburbs grow faster in response to highway construction. Suburban growth is also influenced by a higher share of local workers, as well as workers commuting to other suburbs and cores. Moreover, spending within the suburb itself is an important driver of growth. Together, these results indicate that highways foster suburban growth when suburbs also function as employment hubs, not just as residential locations for core commuters.

Table 5: Determinants of population growth in urban cores

Variable	Population growth			
	(1)	(2)	(3)	(4)
Log Initial Population	-1.7426***	-1.7170***	-1.5911***	-2.1576***
	(0.5142)	(0.5124)	(0.5267)	(0.5154)
Log Initial Population Squared	0.0681***	0.0656***	0.0598***	0.0777***
	(0.0203)	(0.0202)	(0.0210)	(0.0200)
Share Working Locally		-0.1446	-0.4221	0.0083
		(0.3660)	(0.3956)	(0.3783)
Share Spent Locally		0.2620	0.3537*	0.0548
		(0.1931)	(0.1984)	(0.2014)
Share Working in Core			-1.3961	-0.6863
			(2.1317)	(1.9727)
Share Working in Suburb			-1.9858	-0.7371
			(2.1683)	(2.0716)
Share Spent in Core				2.5160***
				(0.8969)
Share Spent in Suburb				1.6563
				(1.0751)

Table 6: Determinants of population growth in suburbs

Variable	Population growth			
	(1)	(2)	(3)	(4)
Log Initial Population	-1.9216***	-1.4113***	-1.3716***	-1.4383***
	(0.2783)	(0.2917)	(0.2770)	(0.2869)
Log Initial Population Squared	0.0766***	0.0470***	0.0440***	0.0480***
	(0.0134)	(0.0144)	(0.0137)	(0.0145)
Share Working Locally		0.5004	0.9637***	0.7788**
		(0.3206)	(0.3219)	(0.3506)
Share Spent Locally		1.3863***	1.2194***	1.6723***
		(0.3425)	(0.3265)	(0.4791)
Share Working in Core			2.1007***	2.1980***
			(0.6409)	(0.6760)
Share Working in Suburb			1.0649	1.3788*
			(0.6705)	(0.7216)
Share Spent in Core				-0.2700
				(0.5860)
Share Spent in Suburb				-0.8159
				(0.7320)

7 Conclusions

In this paper, I examine the U.S. Interstate Highway System's role in reshaping population and employment patterns, showing that highways significantly reduced trade costs as well as commuting costs. This reduction in trade costs, alongside commuting improvements, was crucial in driving suburban growth and urban core decline.

First, using variation in highway construction dates and driving time reductions from 1950 to 2020, I estimate the impacts of commuting and trade access on county-level population and employment. I find that a one standard deviation increase in trade access raises population and employment by 9.6% and 11.8%, respectively, while a similar increase in commuting access results in only a 0.03% and 1.5% rise. This suggests that trade cost reductions have a more substantial impact on growth than commuting cost reductions.

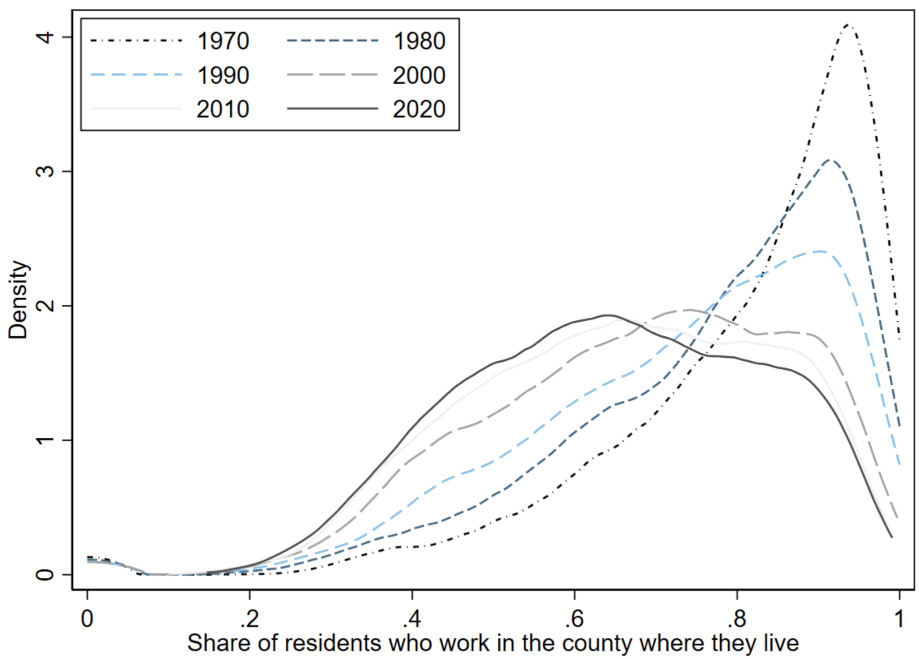
Second, I develop a quantitative spatial model incorporating trade, commuting, migration, production, and amenity externalities. Using indirect inference, I calibrate this model to the U.S. economy and simulate a counterfactual without the Interstate Highway System. When highways reduce both trade and commute costs, they account for 15% of suburban growth and 33% of urban core decline. Commute cost reductions alone explain 88% of suburban growth, primarily at the expense of rural areas, while trade cost reductions account for the remaining suburban growth and all of the decline in urban cores.

My analysis highlights that commute cost reductions largely benefit suburban areas by improving job access and boosting local amenities, while trade cost reductions encourage rural growth by improving access to goods. When both costs decrease, suburban areas absorb growth from both commuters and trade-related migration, reinforcing their role as residential and commercial hubs.

A Appendix

A.1 Figures and Tables

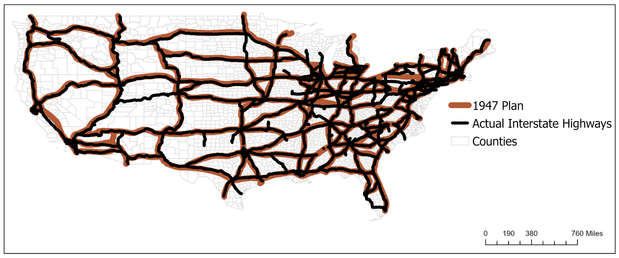
Figure A.1: Distribution across counties of the share of people working on the same county they live



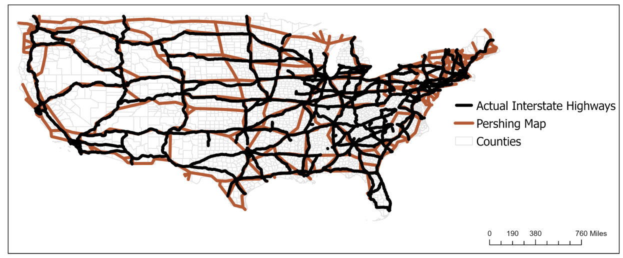
To make sure counties definition over time are consistent, I interpolate commuting flows using Eckert et al. (2020) crosswalks. Commuting flow data comes from the Journey To Work Database (1970-2020).

Figure A.2: Highway plans, historical routes, and actual highways

1947 plan

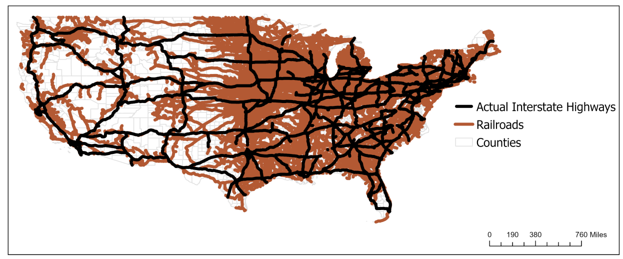


Pershing plan



16th-19th centuries explorations

Railroads by 1898

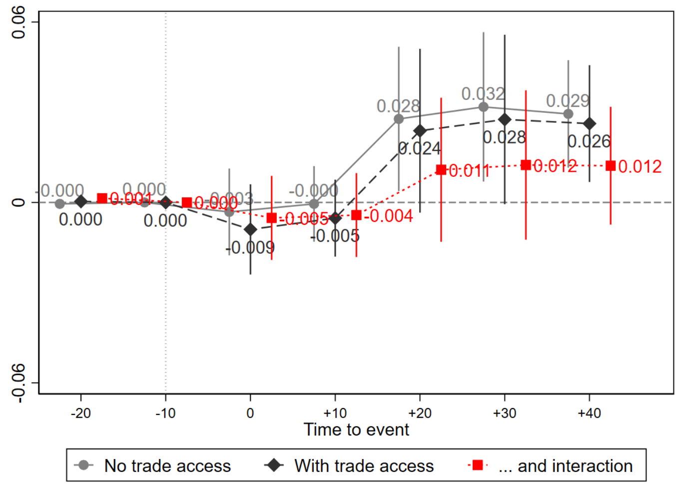


Shapefile sources: 1947 Plan (Baum-Snow, 2019); Pershing Plan (Frye, 2023); Exploration Routes (Brinkman and Lin, 2024); Railrods (Atack, 2015), Interstate Highways (Highway Performance Monitoring System, 2005)

Figure A.3: OLS estimates of the effect of commuting and trade access

Panel A: Commute Access

Log Employment - Log Population



Panel B: Trade Access

Log Employment - Log Population

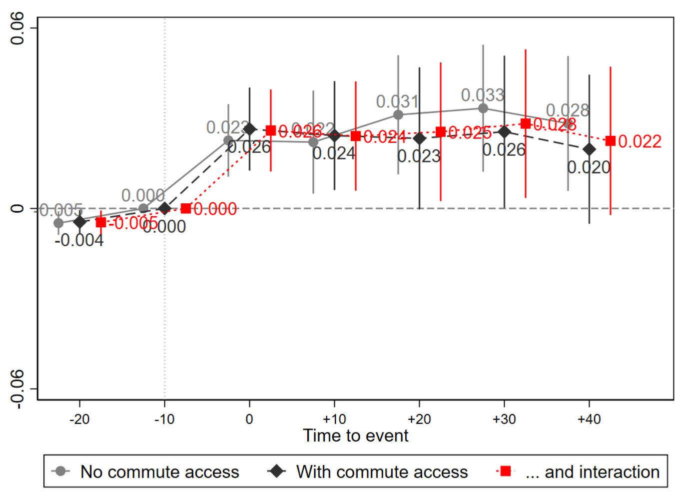
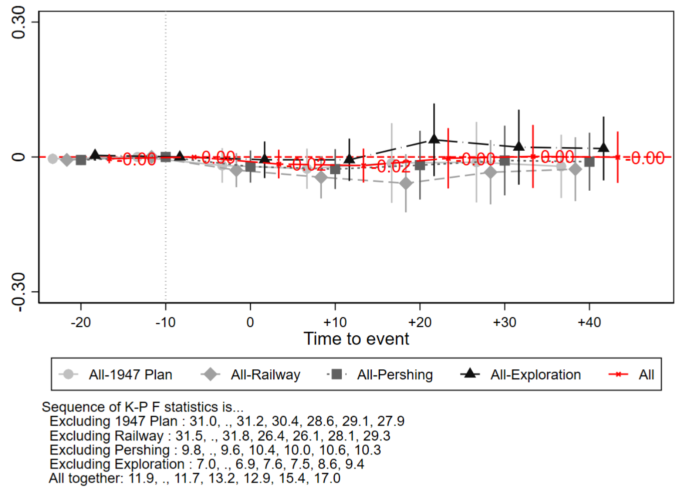


Figure A.4: 2SLS estimates of the effect of commuting and trade access

Panel A: Commute Access

Log Employment - Log Population



Panel B: Trade Access

Log Employment - Log Population

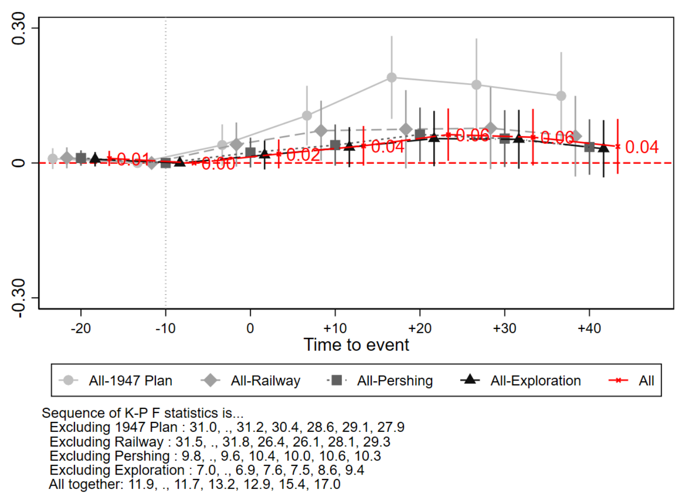
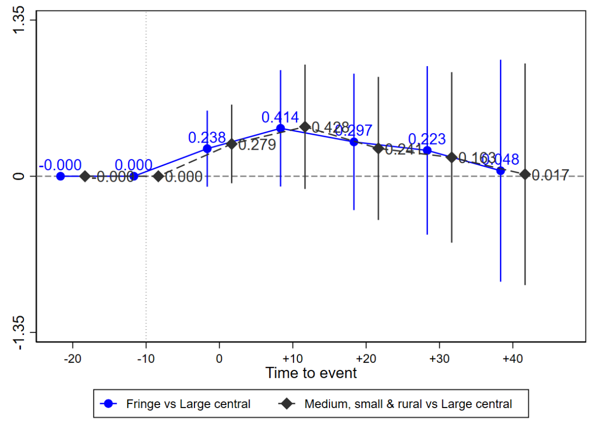


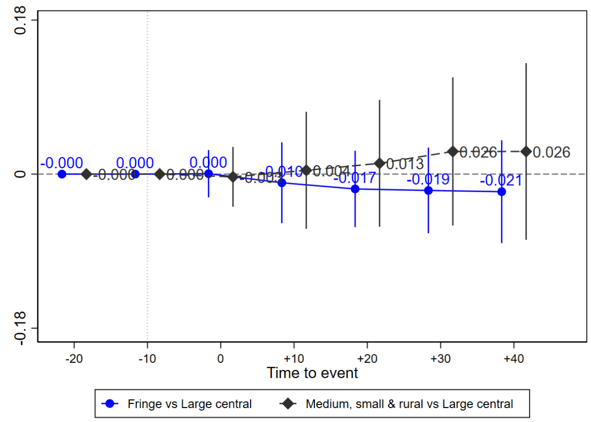
Figure A.5: OLS estimates of the effect of commute and trade access on log population by commuting zone and county NCHS classification

Panel A: Top 7 commuting zones

Commute

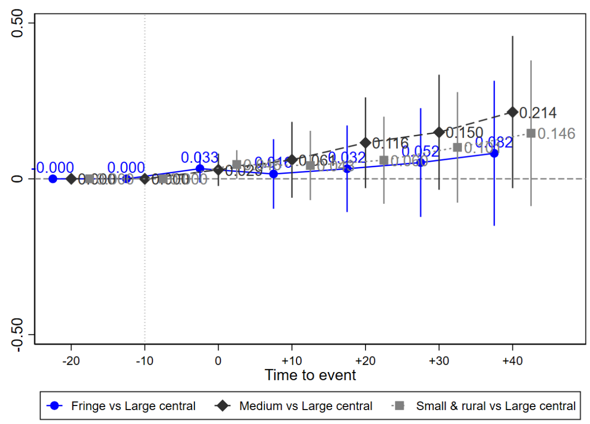


Trade

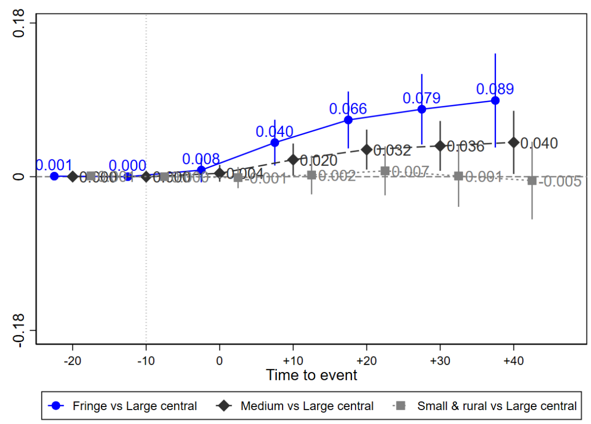


Panel B: Top 8-60 commuting zones

Commute

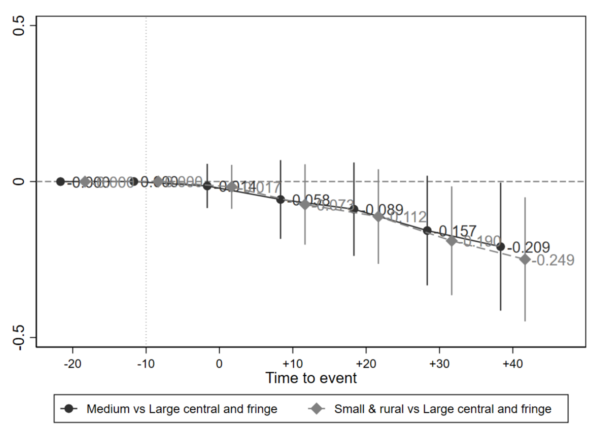


Trade



Panel C: Remaining commuting zones

Commute



Trade

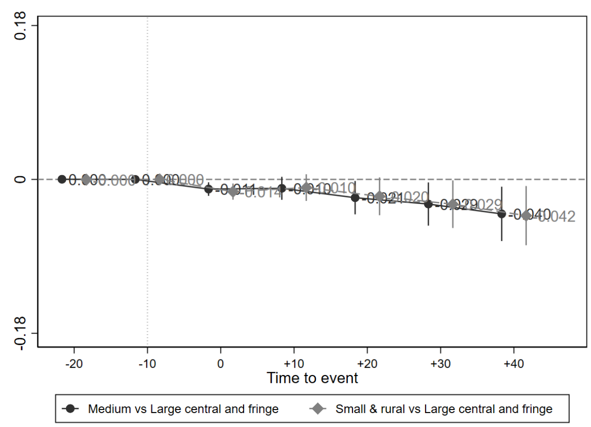
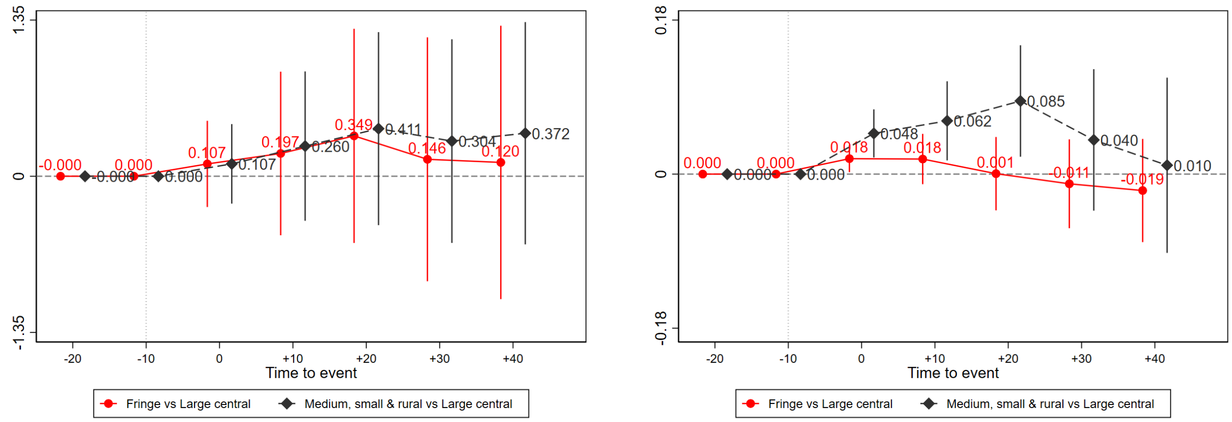


Figure A.6: OLS estimates of the effect of commute and trade access on log employment by commuting zone and county NCHS classification

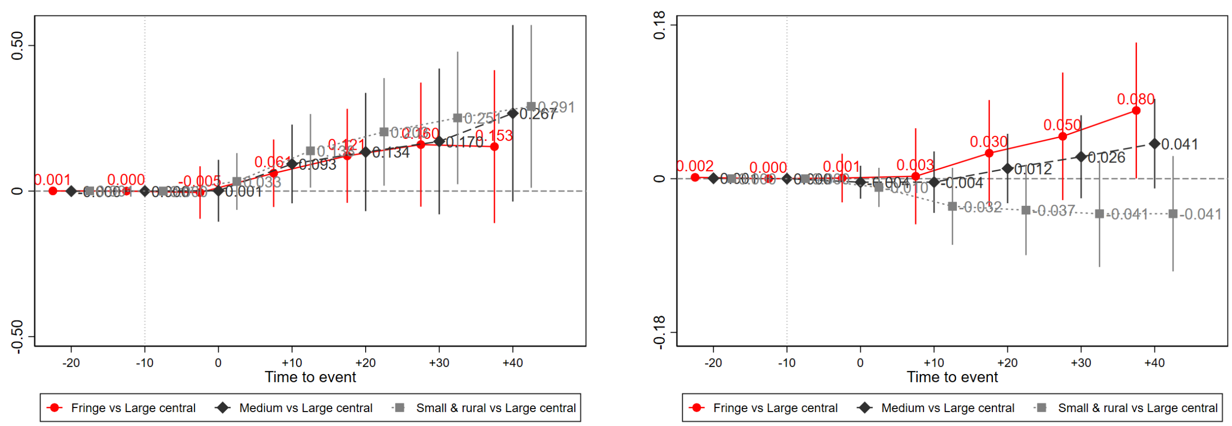
Panel A: Top 7 commuting zones

Commute Trade



Panel B: Top 8-60 commuting zones

Commute Trade



Panel C: Remaining commuting zones

Commute Trade

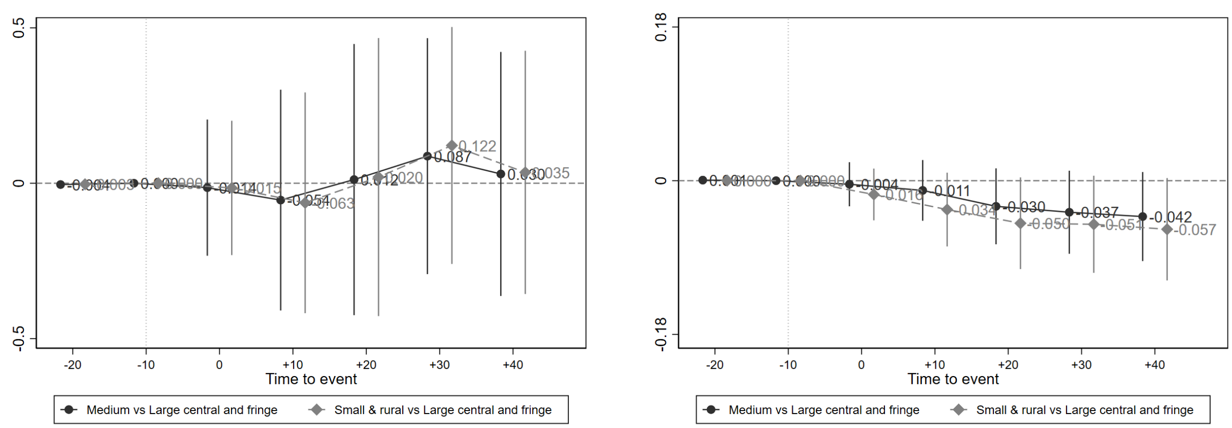
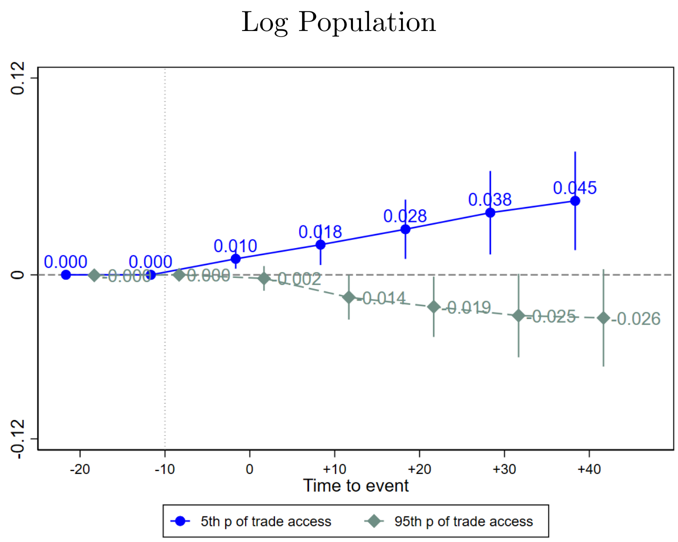
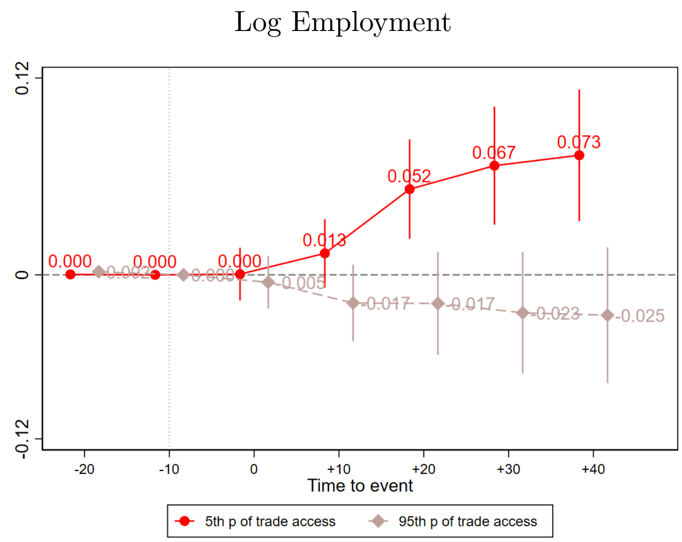


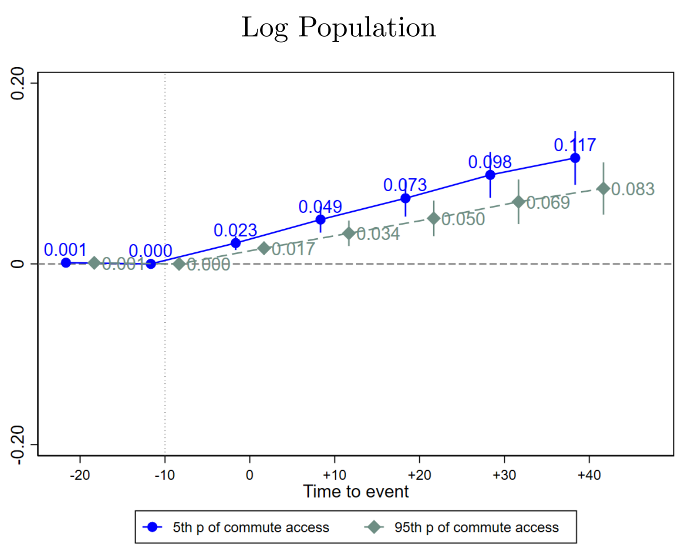
Figure A.7: OLS estimates of the effect of commuting and trade access and their interaction

Panel A: Commute Access





Panel B: Trade Access



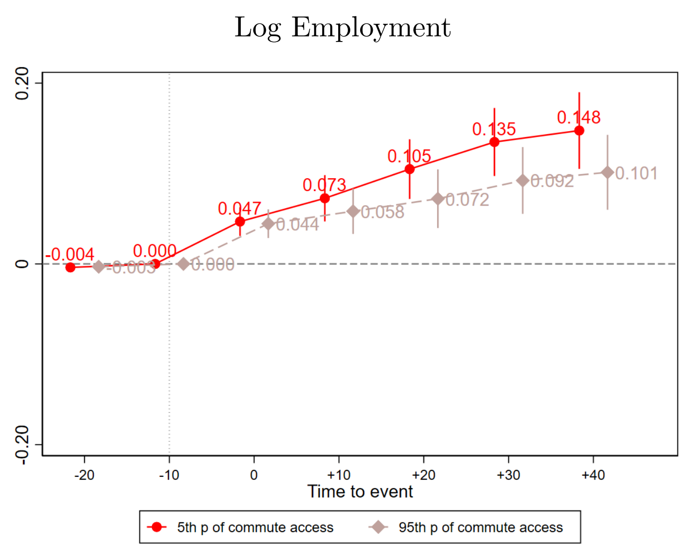
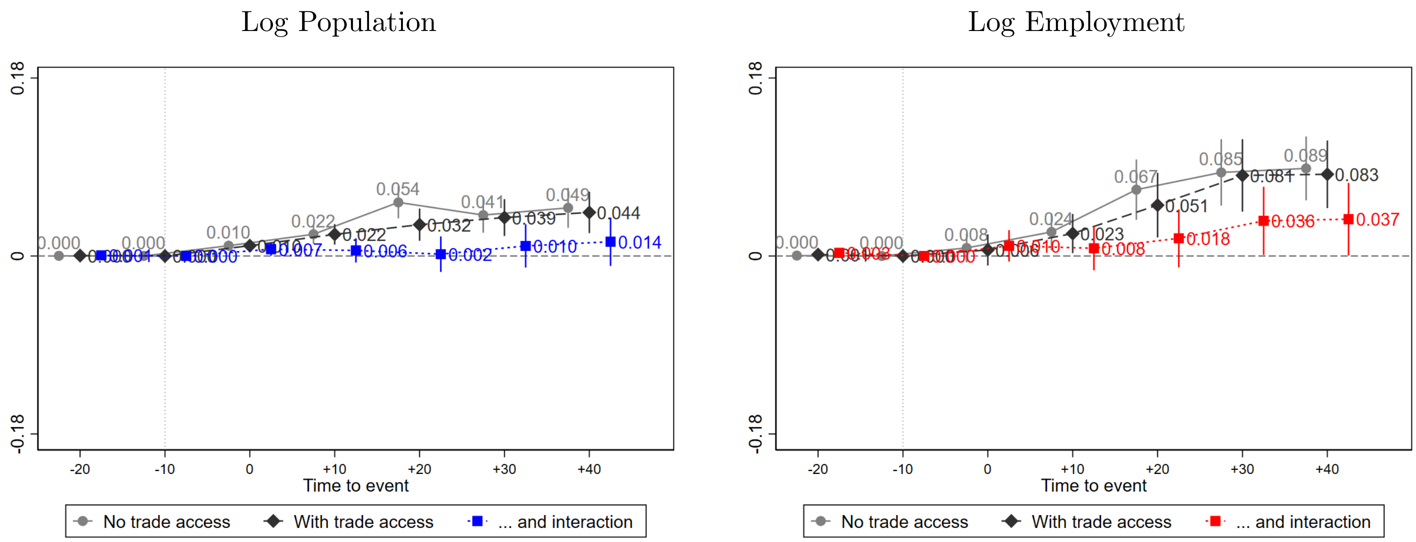


Figure A.8: OLS estimates of the effect of commuting and trade access conditioning on CZ times year-by-wave FE

Panel A: Commute Access



Panel B: Trade Access

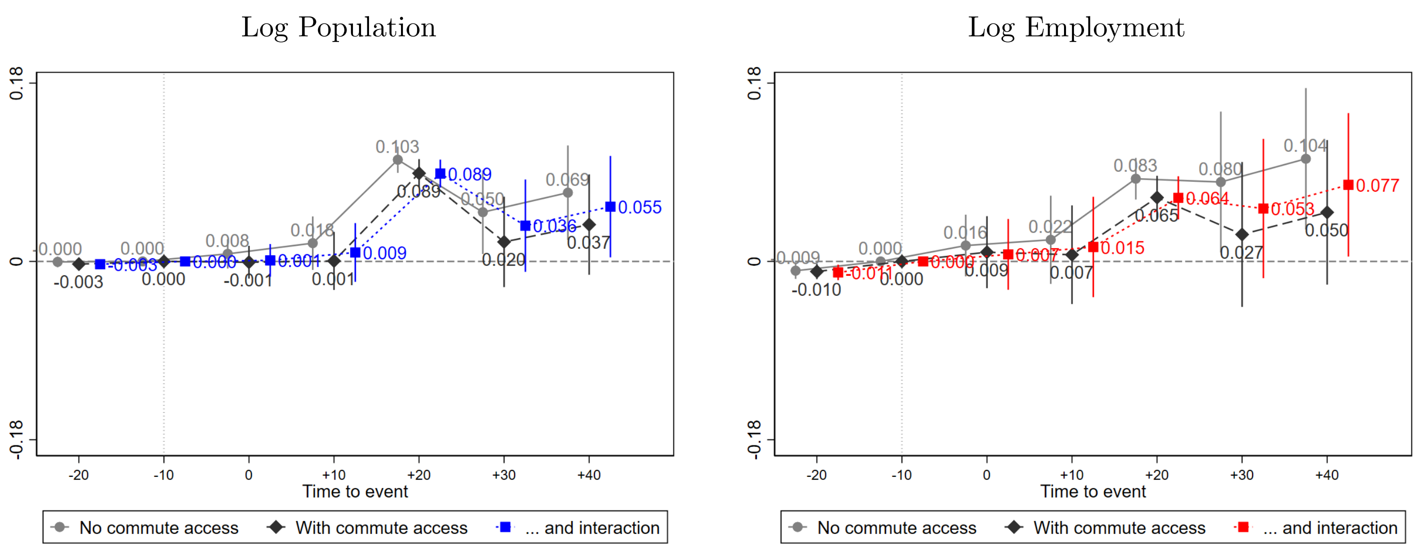
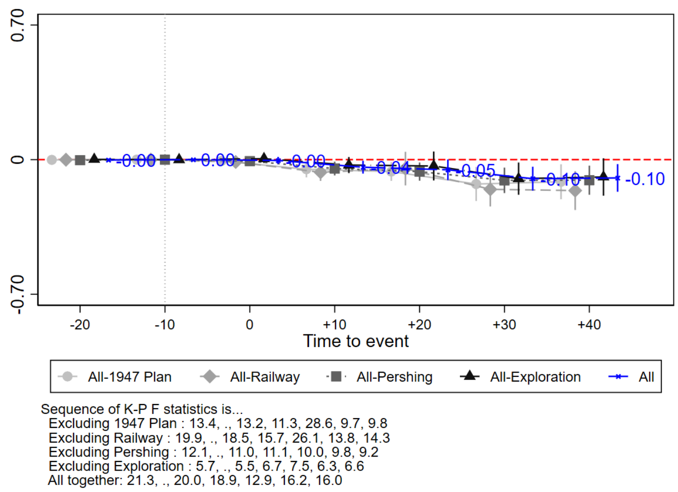


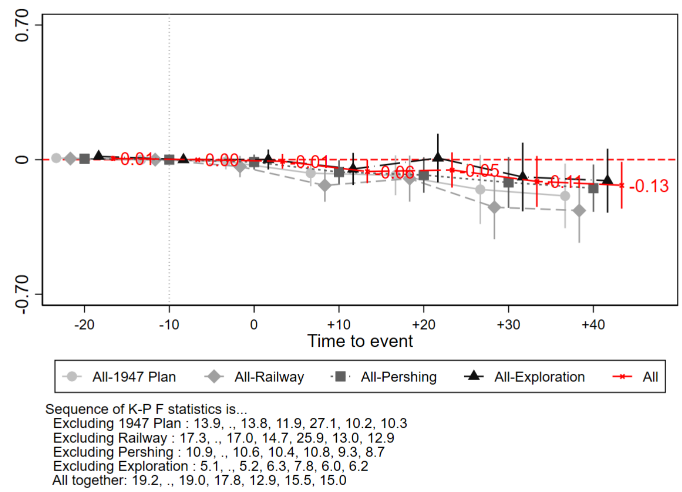
Figure A.9: 2SLS estimates of the effect of commuting and trade access conditioning on CZ times year-by-wave FE

Panel A: Commute Access

Log Population

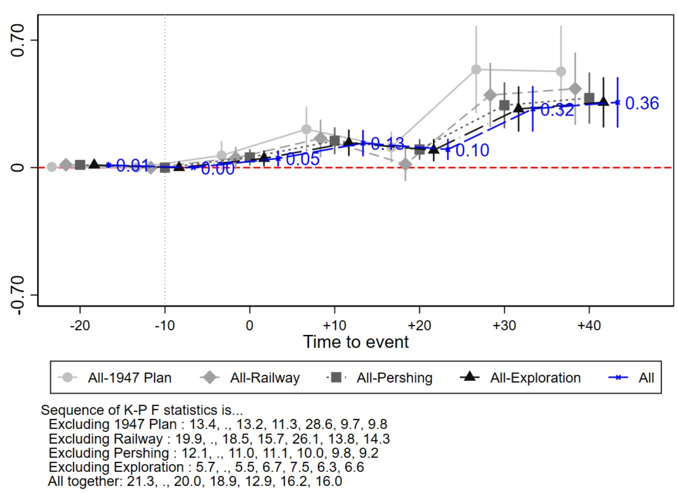


Log Employment



Panel B: Trade Access

Log Population



Log Employment

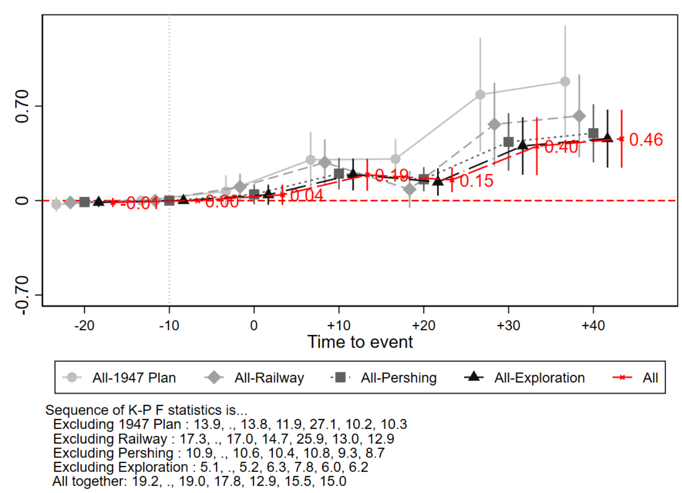
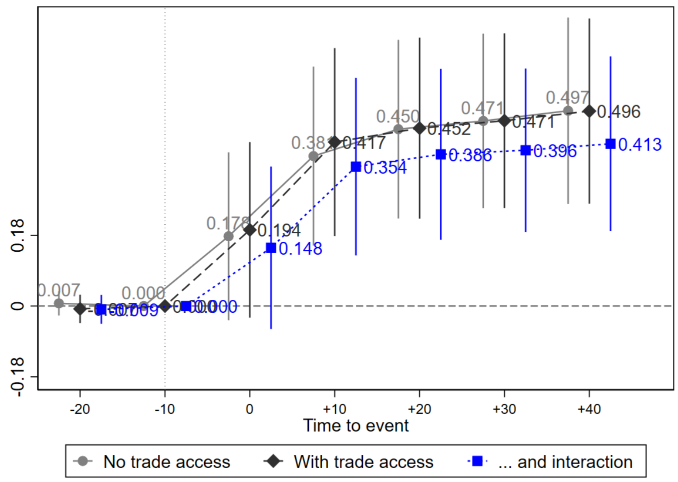


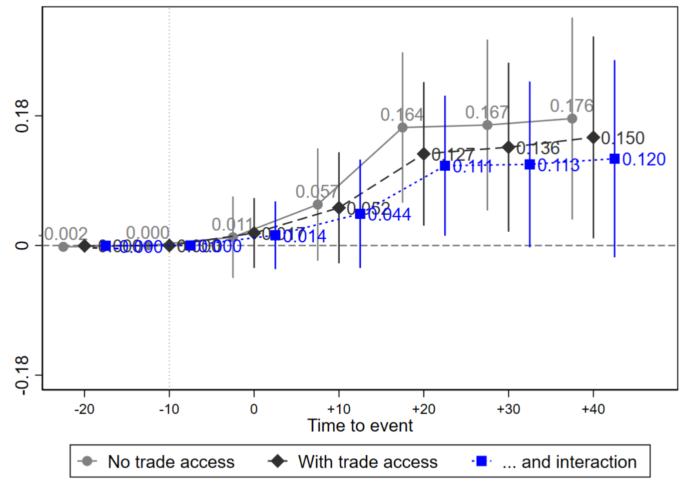
Figure A.10: OLS estimates of the effect of commuting and trade access on census-tracts' population

Panel A: Commute Access

Log Population (unweighted)

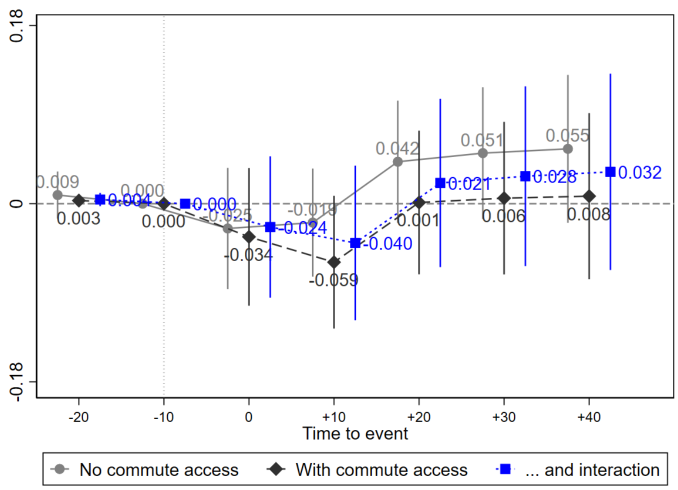


Log Population (weighted)

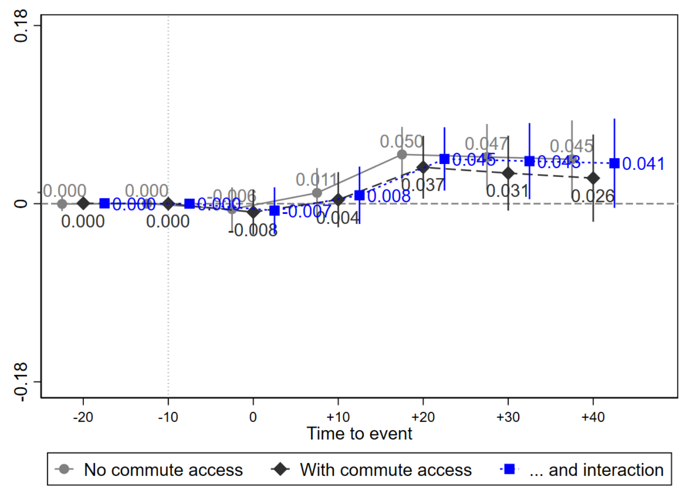


Panel A: Trade Access

Log Population (unweighted)



Log Population (weighted)



A.2 Solution algorithm

TBD.

A.3 Data Appendix

TBD.

A.4 Other results

Heterogeneity across commuting zones

Facts 1 and 2 in Section 2.2 established that suburbs, especially in mid-sized commuting zones, emerged as key employment hubs. Section 3.3 established that the average effect of trade access is positive, whereas the average effect of commute access is statistically not distinguishable from zero. This section explores the intersection of these two ideas. That is, it analyzes how commuting and trade access affect peripheral counties compared to central ones, across commuting zones of different sizes.

How? I restrict the sample by commuting zone size. Then, I re-run regressions, interacting commute and trade access with indicators for large central areas (i.e. core areas), suburbs, medium metros, and small metros pooled with rural areas (following the NCHS classification). In most cases, I choose large central counties as base category. Thus, coefficients are interpreted in relative terms.

Figure A.5 presents population results. Panel A focuses on the top 7 commuting zones (LA, NY, Chicago, Houston, SF, NJ, Boston).[28] The findings reveal that, after 40 years, commuting and trade access effects on smaller areas, relative to large central counties, are statistically insignificant. Employment results, shown in Figure A.6, follow a similar pattern.[29]

Panel B examines the top 8-60 commuting zones. The key finding is that trade access drives population and employment decentralization, while commuting access plays a smaller role. First, regarding trade access, Figure A.5 shows that it significantly affects population in fringe counties compared to large central counties. The same is true for medium metros relative to large central counties, though to a lesser extent. Figure A.6 show similar patterns for employment. Second, regarding commute access, it also has positive relative impacts but these are statistically insignificant[30]. Overall, these findings support Facts 1 and 2: increased trade access has made suburbs and intermediate cities important employment hubs.

[28] I group medium, small, and rural counties to ensure a larger sample. Within the biggest commuting zones there are just a few counties classified as either medium, small or rural.

[29] However, in the short-run (0-10 years), only commute access has relative impacts for population, while only trade access has relative impacts for employment.

[30] Although they are statistically significant for employment as per Figure A.6.

Panel C focuses on the remaining commuting zones. For this panel, I group counties classified as large central and fringe into a single category, again because of small sample considerations. I use this category as base category.[31] For population, both commute and trade access show negative coefficients across all areas, meaning that rises in commute and trade access have *larger* impacts on fringe counties relative to medium metros (and fringe counties relative to smaller metros and rural areas). For employment, only trade effects are statistically significant. In other words, both commute and trade access increase population density in the smallest commuting zones, though trade access only does so for employment.

In summary, suburbs have become major employment hubs due to trade access. In large commuting zones, decentralization produced by commute and trade effects is not persistent. However in mid-sized commuting zones decentralization is persistent. In the smallest commuting zones, commute access concentrates population and employment towards higher-density areas, while trade access only does so for employment.[32]

Interaction. In previous sections, I introduced the interaction between commute and trade access and treated it like a control, while also computing the average effect of either commute or trade access. In this section, I explore the sign and magnitude of this interaction term with greater detail.

Theoretically, when trade and commute access both improve simultaneously, the interaction effect can be negative, meaning the two channels substitute for each other. For instance, if trade access improves, making it easier for firms to move goods across regions, but commute access also increases, workers may prefer to commute to jobs in other regions instead of working locally. This weakens the local employment benefits that improved trade access would otherwise provide.

Figure A.7 shows the commute and trade effect for population and employment. I compute the effect of increasing commute access assuming that trade access increases at the 5th percentile and 95th percentile to illustrate the interaction's magnitude. The reverse is done for trade access. Since commute access effects decrease with trade access, there is a negative interaction. When trade access is low enough, commute access becomes positive and statistically significant, even when the average impact of commute access is zero in Figure 6. In contrast, the impact of trade access remains positive, even with high commute access changes.

[31] This is because within the smallest commuting zones there are just a few counties classified as either large central or fringe.

[32] 2SLS results are available in Appendix, but conclusions are similar despite weak first-stage results in some subsamples.

Controlling for commuting zones fixed effects interacted with year-by-wave fixed effects. In the baseline, I control for state fixed effects interacted with year-by-wave fixed effects. Identification comes from a comparison of access measures between counties, even when they belong to different commuting zones but within the same state.

An interesting comparison would be between counties with highways leading to different connections, but within the same commuting zone. For example, from the point of view of a city like Baltimore or Washington DC, finding more economic activity around a highway leading to Manhattan, NY, versus one leading to Charlottesville, VA would highlight the trade access effects of highways. To do a comparison similar to this example, I introduce commuting zones fixed effects interacted with year-by-wave fixed effects (and lose the state fixed effects). The commuting zone fixed effects interacted with year-by-wave fixed effects will be perfectly colinear with the access measures of commuting zones comprised by a single county. Thus, identification comes from comparing access measures between two or more counties within the same commuting zone.

OLS results are available in Appendix A.8. Overall results are similar to the baseline, but with a few nuances. First, commuting effects are positive and statistically significant for employment: a one standard deviation increase in commuting access rises employment by 3.7% (in Figure 6 the effect was of 1.5% but statistically, zero). For population, the effect is still indistinguishable from zero. Second, trade effects are almost halved: increasing trade access rises population by 5.5% (instead of 9.6%) and employment by 7.7% (instead of 11.8%).

While OLS estimates give commuting access a bigger role than in the baseline, the opposite is true for 2SLS estimates. Estimating the effect of trade and commute access by 2SLS, while also controlling for commuting zone times year-by-wave fixed effects strengthens the impact of trade (see Appendix A.9). Trade effects nearly double, with 40-year coefficients for population and employment increasing to 0.36 and 0.46, respectively (rather than 0.20 and 0.24). Commuting effects remain negative but they are now statistically significant.

Census tracts. This paper examines whether gains in commuting and trade access from highways explain the suburbanization of population and employment at the county level. To explore suburbanization within counties, that is, at the census tract level, I construct a panel of census tracts from 1950 to 2020. Only population data is available and not employment.[33] I rerun the regression using census-

[33] To ensure consistent census tract definitions, I use crosswalks from Lee and Lin (2018) and the Longitudinal Tract

tract population as the outcome and introduce census tract fixed effects to trace out how population changes in response to a county-wide increase in trade and commute access.[34] The analysis is done both with and without weighting by tract population in 1950. If trade access only affects population when weighted, it suggests that trade access increases population in the most populous tracts.

OLS results (Appendix A.10) show that without weights, commute access increases population, while trade access has no effect after 40 years. However, when population weights are applied, commute access has a smaller impact than without weights. In turn, trade access shows a statistically significant positive effect. This suggest that trade access effects are concentrated on the most populous tracts,[35] whereas commute access effects are so in the least populous tracts.[36]

In summary, trade access decentralizes population between counties but concentrates it within counties. Moreover, commuting access has no effect on suburbanization between counties but decentralizes population within counties.

Does commuting between counties respond to changes in commute times? There may be concerns that studying changes in commute access at the county level is unfructuous, under the prior that most commuting occurs within counties and changes in driving times between counties have little impact. First, Figure A.1 already suggests otherwise. It shows that many residents work outside their home county, and that this number has been increasing over time. Second, I exploit changes in driving times and commute flows to study the relationship directly.

In this section, I estimate a dynamic gravity equation for commuting flows using data from 1970 to 2020, leveraging changes in commute times caused by staggered highway construction. Results in Appendix A.7 show that a one-minute increase in commute time reduces the probability of commuting by 1.8% in the short run, but this effect grows to 6.0% after 40 years. These findings align with prior research on commuting semi-elasticities but offer a novel contribution by separately identifying short- and long-run elasticities.

Data Base. The sample includes only census tracts in MSAs with a positive population in 1950. If a tract was not part of an MSA in 1950 but appears later, its 1950 data is recorded as missing.

[34] I also lose the state-by-year-by-wave fixed effects to avoid over saturation of the regression with fixed effects.

[35] Or, at the very least, trade effects are somewhat constant across tracts of different sizes since coefficients are similar when the regression is weighted or unweighted.

[36] I do not report 2SLS estimates because instruments in this subsample are weak: the Kleibergen-Paap F-statistics are around 2.

A.5 Wage responses to highway construction in an environment with a simplified geography

Recall that in the simplified environment discussed in the main text, the market clearing conditions are given by:

$$\underbrace{w_d \sum_i \pi_{ij|i} L_i}_{\text{Labor Supply}} = \underbrace{\sum_o \lambda_{do|o} \left(\bar{y}_o L_o\right)}_{\text{Labor Demand}} \tag{29}$$

where:

$$\begin{aligned}
\lambda_{do|o} &= \frac{A_d \left(c_d \varsigma_{do}\right)^{-\varepsilon}}{\Psi_o} \quad & L_i &= \frac{\bar{U}_i^{\eta} \bar{L}}{\sum_{i \in N} \bar{U}_i^{\eta}} \quad & \bar{U}_{ni} &= \frac{u_i}{P_i} \cdot \Phi_i^{1/\theta} \\
\bar{y}_i &= \sum_j \pi_{ij|i} w_j \quad & \pi_{ij|i} &= \frac{\left(\frac{w_j}{d_{ij}}\right)^{\theta}}{\Phi_i} \quad & \Phi_i &= \sum_{l=1}^{S} \left(\frac{w_j}{d_{ij}}\right)^{\theta} \\
\Psi_i &= \sum_j A_j \left(w_j \varsigma_{ji}\right)^{-\varepsilon} \quad & P_i &= \left(\Psi_i\right)^{-1/\varepsilon}
\end{aligned}$$

By totally differentiating this system with respect to driving times τ, we can use the implicit function theorem to compute the wages response to highway construction.

In particular, and imposing wages in a as the numeraire, wage responses to highway construction are given by the solution of the following system of equations

$$\begin{bmatrix} \left(\mathit{\Psi}_{RHS_A/A} - \Sigma_{LHS_A/A}\right) & \left(\mathit{\Psi}_{RHS_A/B} - \Sigma_{LHS_A/B}\right) \\ \left(\mathit{\Psi}_{RHS_B/A} - \Sigma_{LHS_B/A}\right) & \left(\mathit{\Psi}_{RHS_B/B} - \Sigma_{LHS_B/B}\right) \end{bmatrix} \cdot \begin{bmatrix} \frac{dw_A/d\tau}{w_A} \\ \frac{dw_B/d\tau}{w_B} \end{bmatrix} =$$
$$\begin{bmatrix} \left(\Sigma_{LHS_A/d} - \mathit{\Psi}_{RHS_A/d}\right)\{\kappa\} + \left(\Sigma_{LHS_A/\varsigma} - \mathit{\Psi}_{RHS_A/\varsigma}\right)\left\{\varkappa_T \tau^{-1}\right\} \\ \left(\Sigma_{LHS_B/d} - \mathit{\Psi}_{RHS_B/d}\right)\{\kappa\} + \left(\Sigma_{LHS_B/\varsigma} - \mathit{\Psi}_{RHS_B/\varsigma}\right)\left\{\varkappa_T \tau^{-1}\right\} \end{bmatrix}$$

where $\Sigma_{LHS_i/j}$ terms are labor supply responses in location i to changes in wages in location j (or to changes in commute costs and trade costs). $\mathit{\Psi}_{RHS_i/j}$ terms are labor demand responses in location i to changes in wages in location j (or to changes in commute costs and trade costs).

The vector to the right group the first order effects of changing trade and commute costs on labor supply and labor demand. Whether highways increase labor supply more than labor demand depends whether Σ terms are higher than the Ψ terms.

As it can be inferred from equation 29, labor supply responses to highway construction and to equilibrium changes in wages depend on three factors: on direct wage responses, responses in commuting flows, and responses in location choices. Labor demand responses, in turn, depend on how trade flows respond to highway construction, and in the response in total income by location.

Thus, labor supply and demand responses Σ and Ψ depend on commute and trade linkages. Below, I characterize each:

$$
\begin{aligned}
\Sigma_{LHS_A/A} &= \left(1 + s_{A|A}\cdot\left[\left(1-\pi_{AA}\right)\theta + \frac{\Omega_{L_A/w_A}}{L_A}\right] + s_{a|A}\cdot\left[\left(1-\pi_{aA}\right)\theta + \frac{\Omega_{L_a/w_A}}{L_a}\right] + s_{B|A}\cdot\left[\left(1-\pi_{BA}\right)\theta + \frac{\Omega_{L_B/w}}{L_B}\right.\right.\\
\Sigma_{LHS_A/a} &= -\left(s_{A|A}\cdot\left[\pi_{Aa}\theta - \frac{\Omega_{L_A/w_a}}{L_A}\right] + s_{a|A}\cdot\left[\pi_{aa}\theta - \frac{\Omega_{L_a/w_a}}{L_a}\right] + s_{B|A}\cdot\left[\pi_{Ba}\theta - \frac{\Omega_{L_B/w_a}}{L_B}\right]\right)\\
\Sigma_{LHS_A/B} &= -\left(s_{A|A}\cdot\left[\pi_{AB}\theta - \frac{\Omega_{L_A/w_B}}{L_A}\right] + s_{a|A}\cdot\left[\pi_{aB}\theta - \frac{\Omega_{L_a/w_B}}{L_a}\right] + s_{B|A}\cdot\left[\pi_{BB}\theta - \frac{\Omega_{L_B/w_B}}{L_B}\right]\right)\\
\Sigma_{LHS_A/d} &= \left(s_{A|A}\cdot\left[\left(\beta\pi_{Aa} + \pi_{AB}\right)\theta + \frac{\Omega_{L_A/d}}{L_A}\right] + s_{a|A}\cdot\left[\left(\beta\left(\pi_{aA}-1\right) + \left(1-\beta\right)\pi_{aB}\right)\theta + \frac{\Omega_{L_a/d}}{L_a}\right] + s_{B|A}\cdot\left[\left(\left(\pi_{BA}\right.\right.\right.\right.\\
\Sigma_{LHS_A/\varsigma} &= \left(s_{A|A}\frac{\Omega_{L_A/\varsigma}}{L_A} + s_{a|A}\frac{\Omega_{L_a/\varsigma}}{L_a} + s_{B|A}\frac{\Omega_{L_B/\varsigma}}{L_B}\right)\\
\Sigma_{LHS_a/A} &= -\left(s_{A|a}\cdot\left[\pi_{AA}\theta - \frac{\Omega_{L_A/w_A}}{L_A}\right] + s_{a|a}\cdot\left[\pi_{aA}\theta - \frac{\Omega_{L_a/w_A}}{L_a}\right] + s_{B|a}\cdot\left[\pi_{BA}\theta - \frac{\Omega_{L_B/w_A}}{L_B}\right]\right)\\
\Sigma_{LHS_a/a} &= \left(1 + s_{A|a}\cdot\left[\left(1-\pi_{Aa}\right)\theta + \frac{\Omega_{L_A/w_a}}{L_A}\right] + s_{a|a}\cdot\left[\left(1-\pi_{aa}\right)\theta + \frac{\Omega_{L_a/w_a}}{L_a}\right] + s_{Ba}\cdot\left[\left(1-\pi_{Ba}\right)\theta + \frac{\Omega_{L_B/w_a}}{L_B}\right]\right)\\
\Sigma_{LHS_a/B} &= -\left(s_{A|a}\cdot\left[\pi_{AB}\theta - \frac{\Omega_{L_A/w_B}}{L_A}\right] + s_{a|a}\cdot\left[\pi_{aB}\theta - \frac{\Omega_{L_a/w_B}}{L_a}\right] + s_{B|a}\cdot\left[\pi_{BB}\theta - \frac{\Omega_{L_B/w_B}}{L_B}\right]\right)\\
\Sigma_{LHS_a/d} &= \left(s_{A|a}\cdot\left[\left(\beta\left(\pi_{Aa}-1\right) + \pi_{AB}\right)\theta + \frac{\Omega_{L_A/d}}{L_A}\right] + s_{a|a}\cdot\left[\left(\beta\pi_{aA} + \left(1-\beta\right)\pi_{aB}\right)\theta + \frac{\Omega_{L_a/d}}{L_a}\right] + s_{B|a}\cdot\left[\left(\pi_{BA} + \right.\right.\right.\\
\Sigma_{LHS_a/\varsigma} &= \left(s_{A|a}\frac{\Omega_{L_A/\varsigma}}{L_A} + s_{a|a}\frac{\Omega_{L_a/\varsigma}}{L_a} + s_{B|a}\frac{\Omega_{L_B/\varsigma}}{L_B}\right)\\
\Sigma_{LHS_B/A} &= -\left(s_{A|B}\cdot\left[\pi_{AA}\theta - \frac{\Omega_{L_A/w_A}}{L_A}\right] + s_{a|B}\cdot\left[\pi_{aA}\theta - \frac{\Omega_{L_a/w_A}}{L_a}\right] + s_{B|B}\cdot\left[\pi_{BA}\theta - \frac{\Omega_{L_B/w_A}}{L_B}\right]\right)\\
\Sigma_{LHS_B/a} &= -\left(s_{A|B}\cdot\left[\pi_{Aa}\theta - \frac{\Omega_{L_A/w_a}}{L_A}\right] + s_{a|B}\cdot\left[\pi_{aa}\theta - \frac{\Omega_{L_a/w_a}}{L_a}\right] + s_{B|B}\cdot\left[\pi_{Ba}\theta - \frac{\Omega_{L_B/w_a}}{L_B}\right]\right)\\
\Sigma_{LHS_B/B} &= \left(1 + s_{A|B}\cdot\left[\left(1-\pi_{AB}\right)\theta + \frac{\Omega_{L_A/w_B}}{L_A}\right] + s_{a|B}\cdot\left[\left(1-\pi_{aB}\right)\theta + \frac{\Omega_{L_a/w_B}}{L_a}\right] + s_{B|B}\cdot\left[\left(1-\pi_{BB}\right)\theta + \frac{\Omega_{L_B/}}{L_B}\right.\right.\\
\Sigma_{LHS_B/d} &= \left(s_{A|a}\cdot\left[\left(\beta\pi_{Aa} + \left(\pi_{AB}-1\right)\right)\theta + \frac{\Omega_{L_A/d}}{L_A}\right] + s_{a|a}\cdot\left[\left(\beta\pi_{aA} + \left(1-\beta\right)\left(\pi_{aB}-1\right)\right)\theta + \frac{\Omega_{L_a/d}}{L_a}\right] + s_{B|a}\cdot\left[\left(\pi\right.\right.\right.\\
\Sigma_{LHS_B/\varsigma} &= \left(s_{A|B}\frac{\Omega_{L_A/\varsigma}}{L_A} + s_{a|B}\frac{\Omega_{L_a/\varsigma}}{L_a} + s_{B|B}\frac{\Omega_{L_B/\varsigma}}{L_B}\right)
\end{aligned}
$$

$$\Psi_{RHS_{A/A}} = \lambda_{AA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_A}}{L_A}+\frac{\pi_{AA}w_A}{\bar{y}_A}-\left(1-\lambda_{AA|A}\right)\varepsilon\right)+\lambda_{Aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_A}}{L_a}+\frac{\pi_{aA}w_A}{\bar{y}_a}-\left(1-\lambda_{Aa|a}\right)\varepsilon\right)$$

$$\Psi_{RHS_{A/a}} = \lambda_{AA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_a}}{L_A}+\frac{\pi_{Aa}w_a}{\bar{y}_A}+\lambda_{aA|A}\varepsilon\right)+\lambda_{Aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_a}}{L_a}+\frac{\pi_{aa}w_a}{\bar{y}_a}+\lambda_{aa|a}\varepsilon\right)+\lambda_{AB|B}L_B\bar{y}_B$$

$$\Psi_{RHS_{A/B}} = \lambda_{AA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_B}}{L_A}+\frac{\pi_{AB}w_B}{\bar{y}_A}+\lambda_{BA|A}\varepsilon\right)+\lambda_{Aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_B}}{L_a}+\frac{\pi_{aB}w_B}{\bar{y}_a}+\lambda_{Ba|a}\varepsilon\right)+\lambda_{AB|B}L_B$$

$$\Psi_{RHS_{A/d}} = \lambda_{AA|A}L_A\bar{y}_A\cdot\left(\frac{\Omega_{L_A/d}}{L_A}\right)+\lambda_{Aa|a}L_a\bar{y}_a\cdot\left(\frac{\Omega_{L_a/d}}{L_a}\right)+\lambda_{AB|B}L_B\bar{y}_B\cdot\left(\frac{\Omega_{L_B/d}}{L_B}\right)$$

$$\Psi_{RHS_{A/\varsigma}} = \lambda_{AA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/\varsigma}}{L_A}+\left(\lambda_{aA|A}+\lambda_{BA|A}\right)\varepsilon\right)+\lambda_{Aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/\varsigma}}{L_a}+\left(\lambda_{Aa|a}+\lambda_{Ba|a}-1\right)\epsilon\right)+\lambda_{AB|B}L$$

$$\Psi_{RHS_{a/A}} = \lambda_{aA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_A}}{L_A}+\frac{\pi_{AA}w_A}{\bar{y}_A}-\left(1-\lambda_{AA|A}\right)\varepsilon\right)+\lambda_{aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_A}}{L_a}+\frac{\pi_{aA}w_A}{\bar{y}_a}-\left(1-\lambda_{Aa|a}\right)\varepsilon\right)$$

$$\Psi_{RHS_{a/a}} = \lambda_{aA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_a}}{L_A}+\frac{\pi_{Aa}w_a}{\bar{y}_A}+\lambda_{aA|A}\varepsilon\right)+\lambda_{aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_a}}{L_a}+\frac{\pi_{aa}w_a}{\bar{y}_a}+\lambda_{aa|a}\varepsilon\right)+\lambda_{aB|B}L_B\bar{y}_B\left(\right.$$

$$\Psi_{RHS_{a/B}} = \lambda_{aA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_B}}{L_A}+\frac{\pi_{AB}w_B}{\bar{y}_A}+\lambda_{BA|A}\varepsilon\right)+\lambda_{aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_B}}{L_a}+\frac{\pi_{aB}w_B}{\bar{y}_a}+\lambda_{Ba|a}\varepsilon\right)+\lambda_{aB|B}L_B\bar{y}$$

$$\Psi_{RHS_{A/d}} = \lambda_{aA|A}L_A\bar{y}_A\cdot\left(\frac{\Omega_{L_A/d}}{L_A}\right)+\lambda_{aa|a}L_a\bar{y}_a\cdot\left(\frac{\Omega_{L_a/d}}{L_a}\right)+\lambda_{aB|B}L_B\bar{y}_B\cdot\left(\frac{\Omega_{L_B/d}}{L_B}\right)$$

$$\Psi_{RHS_{a/\varsigma}} = \lambda_{aA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/\varsigma}}{L_A}+\left(\lambda_{aA|A}+\lambda_{BA|A}-1\right)\varepsilon\right)+\lambda_{aa|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/\varsigma}}{L_a}+\left(\lambda_{Aa|a}+\lambda_{Ba|a}\right)\epsilon\right)+\lambda_{aB|B}L_B\bar{y}$$

$$\Psi_{RHS_{B/A}} = \lambda_{BA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_A}}{L_A}+\frac{\pi_{AA}w_A}{\bar{y}_A}-\left(1-\lambda_{AA|A}\right)\varepsilon\right)+\lambda_{Ba|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_A}}{L_a}+\frac{\pi_{aA}w_A}{\bar{y}_a}-\left(1-\lambda_{Aa|a}\right)\varepsilon\right)$$

$$\Psi_{RHS_{B/a}} = \lambda_{BA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_a}}{L_A}+\frac{\pi_{Aa}w_a}{\bar{y}_A}+\lambda_{aA|A}\varepsilon\right)+\lambda_{Ba|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_a}}{L_a}+\frac{\pi_{aa}w_a}{\bar{y}_a}+\lambda_{aa|a}\varepsilon\right)+\lambda_{BB|B}L_B\bar{y}_B$$

$$\Psi_{RHS_{B/B}} = \lambda_{BA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/w_B}}{L_A}+\frac{\pi_{AB}w_B}{\bar{y}_A}+\lambda_{BA|A}\varepsilon\right)+\lambda_{Ba|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/w_B}}{L_a}+\frac{\pi_{aB}w_B}{\bar{y}_a}+\lambda_{Ba|a}\varepsilon\right)+\lambda_{BB|B}L$$

$$\Psi_{RHS_{B/d}} = \lambda_{BA|A}L_A\bar{y}_A\cdot\left(\frac{\Omega_{L_A/d}}{L_A}\right)+\lambda_{Ba|a}L_a\bar{y}_a\cdot\left(\frac{\Omega_{L_a/d}}{L_a}\right)+\lambda_{BB|B}L_B\bar{y}_B\cdot\left(\frac{\Omega_{L_B/d}}{L_B}\right)$$

$$\Psi_{RHS_{B/\varsigma}} = \lambda_{BA|A}L_A\bar{y}_A\left(\frac{\Omega_{L_A/\varsigma}}{L_A}+\left(\lambda_{aA|A}+\lambda_{BA|A}-1\right)\varepsilon\right)+\lambda_{Ba|a}L_a\bar{y}_a\left(\frac{\Omega_{L_a/\varsigma}}{L_a}+\left(\lambda_{Aa|a}+\lambda_{Ba|a}-1\right)\epsilon\right)+\lambda_{BB}$$

A.6 Theorems and propositions

Theorem A.1: *Given parameter $\{\theta\}$, data $\{\pi_{od}, w_d\}$, and bilateral commute costs $\{d_{od}\}$, there exists a unique set of commuting preferences $\{K_{od}\}$ that rationalize bilateral commute flows.*

Proof. From the formula for commute flows:

$$\pi_{ij|i} = \frac{\left(K_j\frac{w_j}{d_{i,j}}\right)^{\theta}}{\sum_l\left(K_l\frac{w_l}{d_{i,l}}\right)^{\theta}} \tag{60}$$

Since we observe $\pi_{ij|i}, d_{ij}$, and w_j, we solve for K_j for all j, given that the system of equations is homogeneous of degree one in K_j.

□

THEOREM A.2: *Given parameters $\{\nu, \beta\}$ and data $\{\tilde{L}_d, w_d, H_d, R_d\}$, there is a unique set of construction TFP $\{Z_d\}$ and land supply $\{T_d\}$ that rationalize the data as an equilibrium of the model.*

Proof. From labor market clearing in the construction sector, we infer unit costs as:

$$\tilde{L}_d^C w_d = \underbrace{\beta \left(\frac{w_d}{c_d^{Cons}}\right)^{1-\nu} R_d H_d}_{\text{by construction}} \quad \Rightarrow \quad c_d^{Cons} = \beta^{\frac{1}{1-\nu}} \left(\frac{\tilde{L}_d^C w_d^{\nu}}{R_d H_d}\right)^{-\frac{1}{1-\nu}}$$

We then recover productivities in the construction sector:

$$R_d = \frac{c_d^{Cons}}{Z_d}$$

Using the cost function definition $c_o^{Cons} = \left(\beta w_o^{1-\nu} + (1-\beta) Q_o^{1-\nu}\right)^{1/(1-\nu)}$, we can infer land prices Q_o:

$$(1-\beta) Q_d^{1-\nu} = \left(\frac{R_d H_d}{\tilde{L}_d^C} - w_d\right) \beta w_d^{-\nu}$$

Finally, from land market clearing:

$$(1-\beta) \left(\frac{Q_o}{c_o^{Cons}}\right)^{1-\nu} R_o H_o = Q_o T_o$$

we can recover land supply T_o. □

THEOREM A.3: *If $\alpha \bar{y}_{|d,t} L_{d,t} + (1-\beta) \left(\frac{Q_d}{c_d^{Cons}}\right)^{1-\nu} R_{d,t} H_{d,t} \geq \tilde{L}_{d,t}^{NT} w_{d,t}, \forall d$, given parameters $\{\sigma, \sigma_D, \varepsilon, \theta, \nu, \varrho^T, \varrho_L^{NT}, b, \alpha, \beta\}$, data $\{L_{d,t}, \tilde{L}_{d,t}^s, w_{d,t}, H_{d,t}, R_{d,t},\}$, and bilateral costs $\{\varsigma_{od,t}^{NT}, \varsigma_{od,t}^{T}, d_{od,t}\}$, there is a unique set of TFP in the tradable and non-tradable sectors $\{A_{dt}^T, A_{dt}^{NT}\}$ that rationalize the data as an equilibrium of the model.*

Proof. From previous theorems we have inferred prices and supply in the land market, plus commute flows and therefore average income. These will be inputs later in the proof.

We start by computing compute unit costs and the price index in the non-tradable sector:

$$c_o^{NT} = \left(\varrho^{NT} w_o^{1-\sigma} + \left(1-\varrho^{NT}\right)\tilde{R}_o^{1-\sigma}\right)^{1/(1-\sigma)} \quad \text{and} \quad P_o^{NT} = cons \cdot \left(A_d^{NT} c_o^{NT-\varepsilon}\right)^{-1/\varepsilon} \tag{61}$$

with $const = \left[\Gamma\left(\frac{\varepsilon+1-\sigma_D}{\varepsilon}\right)\right]$.

We use labor market clearing the in the non-tradable and tradable sectors to search for the set of TFPs that solves the following system of equations

$$\underbrace{\tilde{L}_d^T w_d}_{\text{Data}} = \underbrace{\varrho_L^T \left(\frac{w_d}{c_d^T}\right)^{1-\sigma} X_d^T}_{\text{Tradable Sector}} \tag{62}$$

$$\underbrace{\tilde{L}_d^{NT} w_d}_{\text{Data}} = \underbrace{\varrho_L^{NT} \left(\frac{w_d}{c_d^{NT}}\right)^{1-\sigma} X_d^{NT}}_{\text{Non-Tradable Sector}} \tag{63}$$

where:

$$X_d^T = \sum_o \frac{A_d^T \left(c_d^T \varsigma_{do}^T\right)^{-\varepsilon}}{\Psi_o^T} \left(\underbrace{P_o^T C_o^{W,T}}_{\text{By Workers}} + \underbrace{P_o^T C_o^{H,T}}_{\text{By Landlords}} + \underbrace{\frac{\left(1-\varrho_L^T-\varrho_H^T\right)}{\varrho_L^T}\left(\frac{P_o^T}{w_o^T}\right)^{1-\sigma} w_o \tilde{L}_o^T}_{\text{By Tradable Firms}} \right) \tag{64}$$

$$X_d^{NT} = \underbrace{(1-b)\left(\frac{P_d^{NT}}{P_d}\right)^{1-\sigma} \alpha \bar{y}_d L_d}_{\text{By Workers}} + \underbrace{(1-b)\left(\frac{P_d^{NT}}{P_d}\right)^{1-\sigma} Q_d T_d}_{\text{By Landlords}} \tag{65}$$

$$P_d^T = cons \cdot \left(\sum_{o=1}^{N} A_o^T \left(c_o^T \varsigma_{od}^T\right)^{-\varepsilon}\right)^{-1/\varepsilon} \tag{66}$$

and $P_d^T C_d^{W,T} = b\left(\frac{P_d^T}{P_d}\right)^{1-\sigma} \alpha \bar{y}_d L_d$ is the consumption by workers, and $P_d^T C_d^{H,T} = b\left(\frac{P_d^T}{P_d}\right)^{1-\sigma} Q_d T_d$ is the consumption by landowners. Every element in these equations, except for A_{NT} and A_T is either known directly from the data (e.g. $L_d^s w_d$), inferred from a previous theorem (e.g. $Q_d T_d$), or inferred in this theorem (e.g. P_d^T).

In fact, one can prove that the system is homogenous of degree 0, provided that parameters ϱ_x^T scale appropriately to a given normalization. Multiply TFP in the non-tradable and tradable sectors by a constant c. Notice that this term cancels out everywhere except in the demand of intermediate inputs by tradable firms, where we are left with:

$$c^{-(1-\sigma)/\epsilon}\frac{(1-\varrho_L^T-\varrho_H^T)}{\varrho_L^T}\left(\frac{P_o^T}{w_o^T}\right)^{1-\sigma}w_o\tilde{L}_o^T \tag{67}$$

Hence, we need to set the new $\frac{(1-\tilde{\varrho}_L^T-\tilde{\varrho}_H^T)}{\tilde{\varrho}_L^T}$ to equal to $(1/c^{-(1-\sigma)/\epsilon})\frac{(1-\varrho_L^T-\varrho_H^T)}{\varrho_L^T}$, so that the c term cancels out here as well.

Finally, notice that total income (by workers and landlords) in the right hand side of the non-tradables' labor market clearing equation is either observed directly in the data, or inferred from the data in a previous step. Then, the model can only match the wage bill in the non-tradable sector if the share of income spent on the non-tradable sector is large enough. This condition is trivially satisfied in a model without commuting but with a non-tradable sector. □

THEOREM A.4: *If* $\alpha\bar{y}_{|d,t}L_{d,t}+(1-\beta)\left(\frac{Q_d}{c_d^{Cons}}\right)^{1-\nu}R_{d,t}H_{d,t}\geq\tilde{L}_{d,t}^{NT}w_{d,t},\forall d$*, given parameters* $\{\sigma,\sigma_D,\varepsilon,\theta,\nu,\varrho^T,\varrho_L^{NT},b,\alpha,\beta,\zeta_t\}$*, data* $\{L_{d,t},\tilde{L}_{d,t}^s,w_{d,t},H_{d,t},R_{d,t},\}$*, and bilateral costs* $\{\varsigma_{od,t}^{NT},\varsigma_{od,t}^T,d_{od,t},\xi_{od,t}\}$*, there is a unique set of amenities* $\{u_d\}$ *that rationalize the data as an equilibrium of the model.*

Proof. From previous theorems we recover $P_{i,t}$, and $\bar{y}_{i,t}$. From migration equations we have the following system of equations which, in every period, is homogeneous of degree one in amenities $u_{i,t}$:

$$L_{i,t} = \sum_n \zeta_t\frac{\bar{U}_{n,i,t}^\eta}{\sum_{i\in N}\bar{U}_{n,i,t}^\eta}L_{n,t-1} \tag{68}$$

$$\bar{U}_{n,i,t} = \frac{\xi_{n,i,t}u_{i,t}}{P_{i,t}^\alpha R_{i,t}^{1-\alpha}}\cdot\bar{y}_{t,|i} \tag{69}$$

□

A.7 Dynamic Commuting Gravity Equation

In this section, I estimate a dynamic gravity equation using commute flows data and commute time changes from 1970-2020. Due to the staggered nature of highway construction, I exploit variation in changes in commute times produced at different dates to estimate a dynamic gravity equation of commuting. I follow Dube et al., (2023) and their local projection-based differences-in-differences (LP-DiD) approach. The local projection approach is a statistical technique introduced in Jorda (2005). The basic idea is exploit panel data to estimate dynamic impulse responses. The LP-DiD approach utilize local projections to estimate dynamic effects in addition to the so-called 'clean-control' condition. The clean control condition avoids bias by dropping 'unclean' observations.

What are the units of observation in this context? Commuting is a very local phenomenon. For example, even if a trip from San Francisco to New York would take 20 hours instead of 40 hours, still probably nobody would do this commute. In other words, flows between locations that are extremely distant between each other are obviously insensitive to changes in driving times. This way, I define my units of observation as all origin-destination pairs (i, j) that were at a distance of less than 5 hours in 1960. Results are robust to changes around the 5-hours threshold.

How do I build 'treatment' and 'control' groups? In this setting, treatment comes in many waves as highways were constructed in 1970, 1980, 1990, and 2000. For each treatment wave happening on date s, I define a 'treatment' and a 'control' group in the following way. A pair (i, j) is said to be 'treated' if commute times between locations i and j were reduced for the first time at date s. A pair (i, j) is said to be a 'control' pair if it is going to be treated in the future, but has not been treated yet; or, if it was never treated at all. In this way, for treatment wave s, I drop all pairs (i, j) that were treated for the first time before wave s. By this point, we already have a mapping between treatment wave s, and origin-destination pairs (i, j) that belong to either the treatment or the control group of treatment wave s. Finally, for each treatment wave s, I keep in the sample all observations at periods that are around a treatment window: $t \in \big(s - H, s + H\big)$.

I stack treatment and control groups for each treatment wave s and estimate the following regressions, clustering standard errors at county-pair level:

$$\Delta_h \log \pi_{i,j,t} = \theta_h \Delta_0 \tau_{i,j,t} + \underbrace{\phi_{i,t,h} + \phi_{j,t,h}}_{\text{FE}} + \epsilon_{i,j,t,h} \tag{70}$$

$$\Delta_h \tau_{i,j,t} = \alpha_h \Delta_0 \tau_{i,j,t} + \underbrace{\sigma_{i,t,h} + \sigma_{j,t,h}}_{\text{FE}} + \varepsilon_{i,j,t,h} \tag{71}$$

where the parameter of interest is $\tilde{\theta}_h = \theta_h / \alpha_h$. This is equivalent to using the first change in commute times $\Delta_0 \tau_{i,j,t}$ as an instrument for changes in commute times happening at any time after the first treatment: $\Delta_h \tau_{i,j,t}$.[37] The advantage in gravity settings is that it allows the researcher to control for for origin-decade fixed effects and destination-decade fixed effects. They account for idiosyncratic changes in the amenity value of origin locations, and for idiosyncratic changes in productivity of destination

[37] I estimate equation (70) by PPML, and equation (71) by OLS

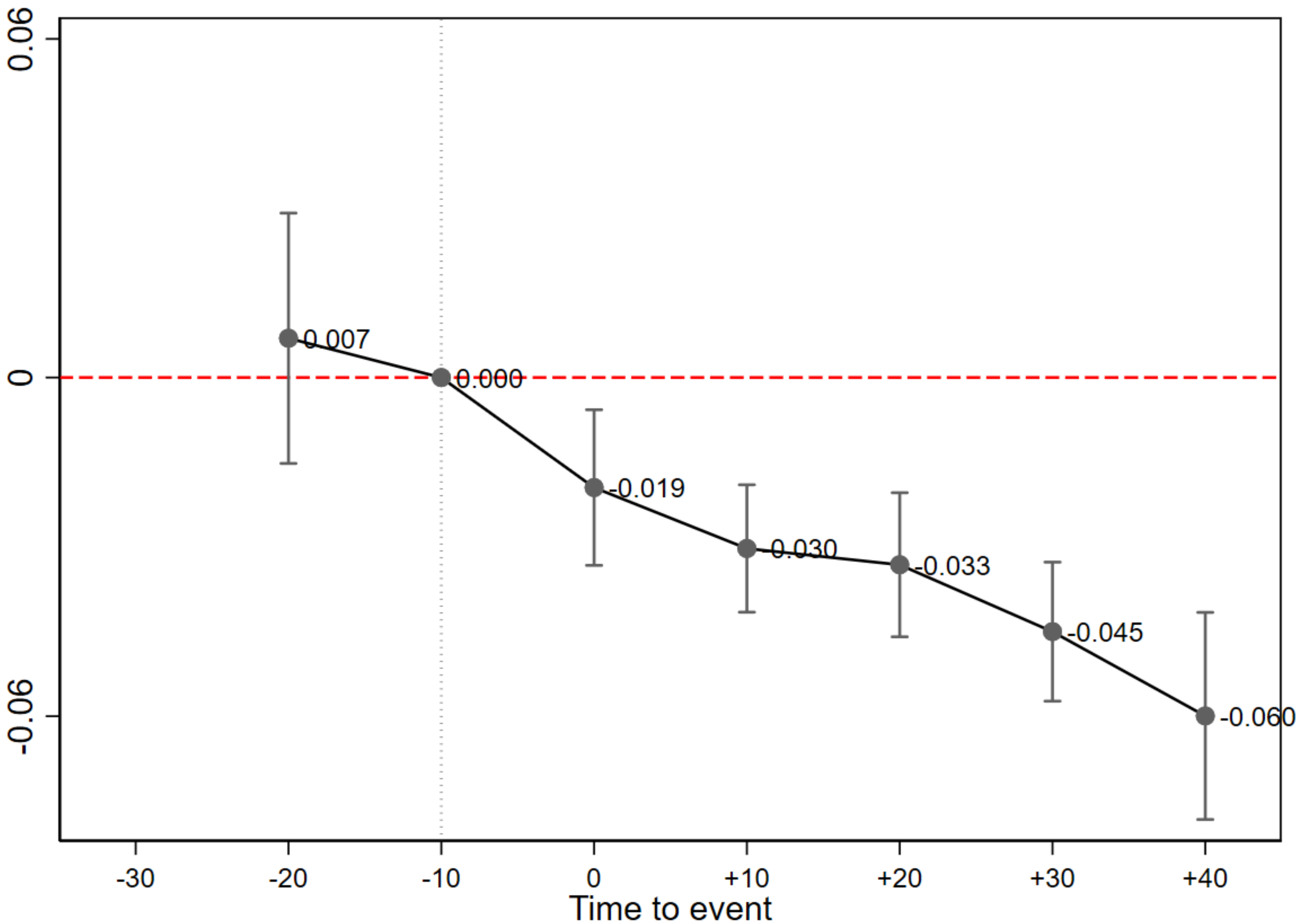
locations.

In Figure A.11, I show the results of estimating equation (70), which is akin to a reduced form regression. This figure reveals that when commute times rises by one minute between location i and location j, the probability of commuting to j decreases by 1.9%, approximately. However, this impact rises over time in absolute terms. After 40 years, the probability of commuting to j decreases by 6.0%. This semi-elasticity of -0.060 is statistically different at a 5% level than the -0.018 estimated for the short-run effect. Hence, changes in commute times have long-lived and growing effects on commuting flows as cities and structures are built. This is a novel fact. Importantly, I do not detect differential pre-trends in commute flows.

The values of these coefficients are consistent with previous estimates in the literature. For instance, depending on the specification, Tsivanidis (2023) finds an impact between -0.025 and -0.053 over a span of 20 years in a sample of urban planning zones within Bogotá, Colombia (an urban planning zone is a somewhat smaller area than a typical municipality). Ahlfeldt et al. (2015), find a semi-elasticity of -0.070 in a cross-section of origin-destination districts in Berlin. Zarate (2022) finds a semi-elasticity between -0.028 and -0.042 in a cross-section of origin-destination municipalities within Mexico city. Velasquez (2024) estimates a semi-elasticity between -0.047 and -0.067 in a cross-section of origin-destination municipalities within Lima, Perú.

The main takeaways from this section are that (i) changes in commute times between counties produce changes in commute flows, (ii) the effects are greater the longer the time horizon. Also, (iii) while previous estimates of the semi-elasticity between commute times and commute flows are in the same ballpark, in most studies, it is unclear whether estimates of the commuting elasticity is a long-run elasticity, an elasticity over a fixed time horizon, or a mix of short- and long-run elasticities. My study provides estimates of the short- and the long-run elasticities, as well as their path over time.

Figure A.11: Dynamic Gravity Equation of Commuting



A.8 Migration Gravity Equation

TBD.