Essays in Economic Geography: Spatial Frictions and Externalities

by

Daniel Velasquez

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2025

Doctoral Committee:

Professor Andrei Levchenko, Co-Chair
Professor Sebastian Sotelo, Co-Chair
Professor Sun Kyoung Lee
Professor Stephen Terry
Professor Dean Yang

Daniel Velasquez
danielvc@umich.edu
ORCID iD: 0000-0001-9709-2527

© Daniel Velasquez 2025

DEDICATION

To my wife, Lucia, and my son, Benjamin, whose love and support have been my anchor, and to my family and friends, who have always been there for me along this journey.

TABLE OF CONTENTS

DEDICATION ii

LIST OF FIGURES vi

LIST OF TABLES viii

LIST OF APPENDICES x

ABSTRACT xi

CHAPTER

1 Highways, Commuting and Trade: Unpacking Suburban Growth 1

1.1 Introduction 1
1.2 Background, Data and Stylized Facts 7
1.2.1 Background and Data 7
1.2.2 Stylized Facts 9
1.3 Reduced form: The Effect of Commuting and Trade Access 11
1.3.1 Measurement of commuting and trade access 12
1.3.2 Empirical strategy 16
1.3.3 Main results 20
1.3.4 Other results 23
1.3.5 Advantages relative to alternatives empirical strategies. 24
1.4 Theory 25
1.4.1 Environment 25
1.4.2 Preferences, commuting and migration 25
1.4.3 Production of tradable, non-tradable, and final goods 27
1.4.4 Trade 28
1.4.5 Floorspace and land 29
1.4.6 Scale externalities 30
1.4.7 Market clearing and equilibrium 31
1.4.8 Comparative statics in a simplified environment 32
1.5 Bridging Theory and Data 36
1.5.1 Model inversion 36
1.5.2 Calibration 38
1.5.3 Indirect inference 40

1.6 Counterfactuals: Unpacking Suburban Growth . . . 42
1.6.1 Aggregate effects: results . . . 43
1.6.2 Aggregate effects: discussion . . . 44
1.6.3 Heterogeneous effects . . . 46
1.7 Conclusions . . . 50

2 Transit Infrastructure and Couples' Commuting Choices in General Equilibrium . . . 52

2.1 Introduction . . . 52
2.2 Lima's Metro and Bus Rapid Transit System . . . 58
2.3 Data Sources, Stylized and Reduced-form Facts . . . 59
2.3.1 Data sources . . . 59
2.3.2 Stylized facts . . . 60
2.3.3 Reduced form evidence and gender inequality . . . 64
2.4 Theory . . . 67
2.4.1 Microfoundation of a household utility for the married . . . 68
2.4.2 Implications . . . 69
2.5 Bridging Theory and Data . . . 72
2.5.1 Quantitative model of city structure and couples . . . 72
2.5.2 Calibration and estimation from aggregate Data . . . 79
2.5.3 Estimation outside the model . . . 79
2.5.4 Estimation within the model . . . 79
2.5.5 Over-identification and untargeted data . . . 85
2.6 The Impact of Improving Commute Times . . . 87
2.6.1 Interdependent commute choices . . . 87
2.6.2 Independent commute choices . . . 89
2.6.3 Understanding the role of joint commuting within couples . . . 91
2.6.4 Policy: completing the Metro's network . . . 95
2.7 Conclusion . . . 96

3 Initial Conditions and the Biggest Push in U.S. History . . . 97

3.1 Introduction . . . 97
3.2 Estimation and Identification . . . 100
3.2.1 Historical Context and Data . . . 100
3.2.2 Empirical Model . . . 102
3.3 Empirical Results . . . 105
3.4 Quantitative Model and Structural Estimation . . . 111
3.4.1 Quantitative Model . . . 111
3.4.2 Estimation using indirect inference . . . 116
3.4.3 Results . . . 120
3.5 Counterfactual Analysis . . . 122
3.6 Conclusion . . . 124

APPENDICES . . . 126

BIBLIOGRAPHY 213

LIST OF FIGURES

FIGURE

1.1 New York and Hartford . . . 8
1.2 Population and employment share out of the nationwide total . . . 10
1.3 Share of workers by type of commute out of total number of workers in size bin . . . 11
1.4 Changes in commute and trade access between 1970 and 1960 . . . 14
1.5 Illustrative examples . . . 15
1.6 OLS estimates of the effect of commuting and trade access . . . 22
1.7 2SLS estimates of the effect of commuting and trade access . . . 23
1.8 Simplified geography before and after a highway construction . . . 33
1.9 Loss Function . . . 41
1.10 Population and employment growth across counties . . . 47
1.11 County growth across suburbs and urban cores . . . 48

2.1 Interdependence in Commuting Choices within Dual-earner Households . . . 63
2.2 Reduced Form Impact of Transit Infrastructure by Distance to CBD . . . 67
2.3 Matching the Geography of the Gender Earnings Gap . . . 86
2.4 Quantitative Impact of Transit Infrastructure on Earnings and the Earnings Gender Gap with Interdependent Commute Choices . . . 90
2.5 Impact with Interdependent Commute Choices by Distance to the CBD . . . 91

3.1 Federal Net Outlays as Percent of GDP . . . 101
3.2 Distribution of Log Contract Value Per 1930 Manufacturing Worker . . . 106
3.3 Coefficients with Varying Cutoffs of Treatment and Initial Conditions . . . 111
3.4 Unrestricted Loss Function . . . 119
3.5 Restricted Loss Function . . . 121
3.6 Long-term Impact of WWII Manufacturing Contracts on Manufacturing Employment Across U.S. Counties . . . 123
3.7 Long-term Impact of WWII Manufacturing Contracts on Manufacturing Employment Across U.S. Counties . . . 124

A.1 Distribution across counties of the share of people working on the same county they live . . . 126
A.2 Highway plans, historical routes, and actual highways . . . 127
A.3 OLS estimates of the effect of commuting and trade access . . . 128
A.4 2SLS estimates of the effect of commuting and trade access . . . 129
A.5 OLS estimates of the effect of commute and trade access on log population by commuting zone and county NCHS classification . . . 130

A.6 OLS estimates of the effect of commute and trade access on log employment by commuting zone and county NCHS classification . . . 131
A.7 OLS estimates of the effect of commuting and trade access and their interaction . . 132
A.8 OLS estimates of the effect of commuting and trade access conditioning on CZ times year-by-wave FE . . . 133
A.9 2SLS estimates of the effect of commuting and trade access conditioning on CZ times year-by-wave FE . . . 134
A.10 OLS estimates of the effect of commuting and trade access on census-tracts' population . . . 135
A.11 Dynamic Gravity Equation of Commuting . . . 148

B.1 The Line 1 of Lima's Metro . . . 150
B.2 The Bus Rapid Transit System . . . 151
B.3 Gravity in Commuting Flows among Singles and Married . . . 152
B.4 Constructed and Planned Lines . . . 153
B.5 Timing Assumptions . . . 154
B.6 Commuting and Over-identification . . . 155
B.7 Interdependence in Commuting Choices within Dual-earner Households . . . 156
B.8 Impact with Independent Commute Choices by Distance to the CBD . . . 157
B.9 Reduced Form Impact of Transit Infrastructure by Distance to CBD . . . 158
B.10 Constructed and Planned Lines . . . 169
B.11 The Reduced Form Impact of Transit Infrastructure on Rents and Expenditure by Distance to the CBD . . . 171
B.12 The Reduced Form Impact of Transit Infrastructure on Rents, Income, and Extreme Poverty . . . 172
B.13 The Reduced Form Impact of Transit Infrastructure on Census Outcomes by Distance to CBD, Part 1 . . . 173
B.14 The Reduced Form Impact of Transit Infrastructure on Census Outcomes by Distance to CBD, Part 2 . . . 174
B.15 The Reduced-form Impact of Transit Infrastructure, Data and Model . . . 177
B.16 Reduced-form Impact by Distance to the CBD, Data and Model . . . 178
B.17 Reduced-form Impact with Independent Commute Choices by Distance to the CBD 179
B.18 Decomposition of the Aggregate Impacts . . . 180
B.19 Incremental Impact of Metro Lines 2 and 4 by Distance to the CBD . . . 182
B.20 Imputed versus reported commuting times . . . 201
B.21 Property Values versus Predicted Rents by Municipality . . . 202

C.1 Cumulative Probability Distribution of Initial Manufacturing Employment . . . 208

LIST OF TABLES

TABLE

1.1 Commute and trade access effects on log population after 40 years 21
1.2 Commute and trade access effects on log employment after 40 years 21
1.3 The effect of Interstate Highways on population across county aggregates 43
1.4 The effect of Interstate Highways on determinants of workers' utility 46
1.5 Determinants of population growth in urban cores 49
1.6 Determinants of population growth in suburbs 50

2.1 Commuting Elasticity 61
2.2 Estimation from Aggregate Data, Calibration, and Gravity Equation 82
2.3 Estimation of Parameters 83
2.4 The Aggregate Impact of the Line 1 of the Metro and the BRT 93
2.5 Decomposition of the Aggregate Impacts 93

3.1 Pretrend Tests of Outcomes in 1900-1930 107
3.2 Long-Term Impact on Growth Rate of Outcomes 108
3.3 Long-Term Impact on Growth Rate of Outcomes by Treatment Size 110
3.4 Summary of calibration 118
3.5 Indirect Inference 120
3.6 Aggregate Differences in Manufacturing Growth in 2000 123

B.1 The Aggregate Impact of Line 1 of the Metro and the BRT 149
B.2 Commuting Elasticity by Presence of Kids 160
B.3 Commuting Elasticity and Origin-Destination Controls 161
B.4 Commuting and Wages 162
B.5 Reduced Form Impact of New Transit Infrastructure on Rents and Income 168
B.6 Time-Event Study of New Transit Infrastructure on Census Outcomes 169
B.7 Differences-in-Differences of New Transit Infrastructure on Mobility 170
B.8 The Aggregate Impact of the Line 1 of the Metro and the BRT - Mechanisms 176
B.9 The Aggregate Impact of Completing the Metro Network 181
B.10 Manually imputed speeds 200

C.1 Summary Statistics of Pre-WWII Regional Characteristics in 1930 209
C.2 Contract Allocation and Initial Manufacturing Conditions 210
C.3 Robustness Check Controlling for Industry Composition 210
C.4 Reduced-Form Estimation Based on Employment Shares in War-Related Sectors . . . 211
C.5 Long-Term Impact on Growth Rate of Outcomes by Contract Types 211

C.6 Relationship Between Post-WWII Military Spending and War Supply Contracts 212

LIST OF APPENDICES

A Appendix to Chapter 1 126

B Appendix to Chapter 2 149

C Appendix to Chapter 3 204

ABSTRACT

This dissertation contains three chapters related to the study of spatial frictions and externalities in economic geography. Chapter 1 examines how highways influence suburban growth by reducing both commuting and trade costs. Chapter 2 analyzes the impact of transit infrastructure on couples' commuting choices and its broader implications for gender disparities in the labor market. Chapter 3 explores the long-term effects of World War II manufacturing investments on the economic geography of U.S. counties.

In Chapter 1, titled "Highways, Commuting, and Trade: Unpacking Suburban Growth", I investigate the role of the U.S. Interstate Highway System in reshaping suburban and urban economic dynamics. Using historical highway construction data, census data, and a spatial general equilibrium model, I show that trade cost reductions, rather than commuting cost reductions alone, were a key driver of suburbanization. Highways accounted for 15% of suburban growth and 33% of urban core decline. Most of the effect on urban core decline is driven by the reductions in trade costs.

In Chapter 2, titled "Transit Infrastructure and Couples' Commuting Choices in General Equilibrium", I study how improvements in public transportation affect household commuting decisions. Using data from Lima's Metro and Bus Rapid Transit system, I provide empirical evidence that transit expansion influences labor market outcomes differently for men and women. By incorporating these insights into a general equilibrium model, I show that the interaction between transit accessibility and joint commuting decisions has significant implications for earnings and gender disparities.

In Chapter 3, titled "Initial Conditions and the Biggest Push in U.S. History", co-authored with Dominick Bartelme and Zhimin Li, I examine the persistent effects of World War II manufacturing contracts on local economic development. Empirically, we find that initial manufacturing employment levels determined whether wartime investments had long-lasting impacts. While some regions experienced sustained manufacturing growth, others saw only temporary gains, suggesting that coordination failures limited the persistence of economic shocks. Motivated by these findings, we develop a quantitative model incorporating non-constant elasticity external economies of scale. Counterfactual simulations suggest that while WWII spending reshaped local economic trajectories, its aggregate impact on national manufacturing employment and welfare was minimal.

CHAPTER 1

Highways, Commuting and Trade: Unpacking Suburban Growth

Abstract

I quantify the effect of U.S. Interstate Highway System on suburbanization and urban core decline via trade costs reductions beyond just commute costs reductions. I leverage variation in highway construction dates and driving time reductions to show that new highways affect both trade costs and commute 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 suburbanization and urban core decline were driven not only by reductions in commute times but also by trade cost reductions, which shifted economic activity away from urban cores.

1.1 Introduction

Since 1950, suburban counties have grown into major population centers, while urban cores have experienced population decline. Alongside these shifts, suburban counties have also become key employment hubs, suggesting a role that extends beyond merely serving as residential areas. This spatial reorganization of population and employment has had far-reaching consequences, contributing to environmental degradation, reduced density, and increased city fiscal and maintenance burdens.[1] However, the forces driving the reorganization of population and employment remain

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

debated.[2] Much of the debate centers on how the Interstate Highway System reduced commuting costs to urban centers, facilitating suburban living by improving access to urban jobs (Baum-Snow, 2007). Yet, as I argue in this paper, by reducing trade costs, highways also contributed to shifting jobs away from urban cores to suburban employment hubs. In this context, the paper asks: 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.? How much of the effect of highways can be attributed to reduced trade costs beyond 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 core decline. To this end, I highlight how trade linkages influence the effects of highways on population and employment heterogeneously across urban categories (e.g., urban cores, suburbs, rural, etc). For example, from Baltimore's perspective, shorter driving times to New York after the introduction of I-95 benefit firms in suburban areas along this route more than firms along I-70 to Pittsburgh. This advantage stems from the relative sizes of New York and Pittsburgh: New York's sheer size provides firms along I-95 with access to a much larger market, making it easier for them to sell their products. In contrast, firms along I-70 to Pittsburgh are closer to a smaller market. Additionally, I-95 attracts new residents to Baltimore's suburbs as their access to goods produced in New York increases. This thought experiment illustrates how, as trade costs to New York decline, both jobs and residents increasingly gravitate toward Baltimore's suburbs along I-95.

By comparing the effects of trade cost reductions to those of commute cost reductions, I also highlight a key conceptual distinction: trade cost reductions can influence bilateral relationships over long distances, whereas commute cost reductions cannot. For instance, the significant distance between Baltimore and New York makes commuting between the two cities impractical. As a result, any potential influence New York might have on Baltimore's suburban growth would occur through changes in market access rather than commuting access. Put it differently, trade cost reductions create 'pulling forces' that extend over greater distances (e.g., population and jobs from Baltimore's core gravitating toward its suburbs in the direction of New York), whereas commute cost reductions generate pulling forces that operate over more limited distances.

To study these issues and answer my research questions, I proceed in two steps. First, I show empirically that highways substantially reduced trade costs and commute costs between counties. Second, I use the results of the empirical exercise to discipline a spatial general equilibrium model via indirect inference. I quantify the role of highways in driving suburban growth and the decline of urban cores, focusing on the reduction in trade costs, in addition to the reductions in commuting

[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.

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 allow me to compute each decade's bilateral driving times between counties. 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.[3] In other words, on average across counties—and regardless of whether they are suburbs or not—trade cost reductions have a greater role than commute cost reductions in driving population and employment growth. This result holds across a battery of robustness tests, including concerns about the endogeneity of highway placement.

While trade cost reductions, on average, have a greater impact on population and employment growth than commute cost reductions, the question remains whether trade cost reductions have driven suburbanization and the decline of urban cores. To address this, in the second step, I use a quantitative spatial model to separately identify the contributions of trade and commute costs to suburbanization. This model incorporates trade, commuting, migration, and externalities in production and amenities, with tradable, non-tradable, and construction sectors.

Before discussing the model's results, I first outline the calibration procedure. I focus on the indirect inference approach I employed to match the model's population and employment responses to reductions in commute and trade costs with the corresponding responses observed in the data. Central to this calibration are the amenities externalities embedded in the model—specifically, scale externalities arising from the agglomeration of people leading to attractiveness gains. I argue that amenity externalities amplify the effects of reduced trade and commute costs. For instance, when commuting costs between Baltimore's suburbs and its core decrease, people relocate to suburbs. This migration, in turn, boosts suburban amenities through scale effects. As amenities improve, the attractiveness of the suburbs increases. Thus, further reductions in trade costs, such as to New York, create even stronger incentives for people to move to the suburbs, where the amenities have already been enhanced.

I estimate the parameters of the amenity externality function that allow the model to generate population and employment responses as close as possible to those observed in the data. To model amenities, I follow Bartelme, Li, and Velasquez (2024) and introduce a more flexible specification

[3]Standard deviations were calculated based on the distribution of trade and commute access improvements, respectively.

of scale externalities that nests the standard specification. I refer to this framework as heterogeneous amenity externalities, which contrasts with the conventional constant elasticity approach. Specifically, I allow amenity externalities to be active only within a middle range of population sizes. By matching the reduced-form effects of trade and commute access, I estimate the threshold at which these externalities become inactive.[4] Including heterogeneous amenity externalities gives the model greater flexibility in matching the reduced-form effects of trade and commute access. While constant amenity externalities can match population responses to either trade or commute access, they cannot account for both simultaneously. The model can only match the population response to commute and trade cost reductions with heterogeneous amenity externalities. The intuition is as follows: raw data reveals that the reduction in commute costs was most significant in denser suburban areas. In contrast, trade cost reductions were more prominent in less dense areas. To match the observed empirical fact that improvements in commute access generate minimal population responses on average, amenity externalities must not be active in denser areas, where commute cost reductions were substantial. However, to match the sizable population responses to reductions in trade costs, amenity externalities must be active in less dense areas, where trade cost reductions were more pronounced. Empirically, I find that amenity externalities become inactive in counties with populations above approximately 200,000. Furthermore, 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 understand separately the contributions of trade and commute costs to suburbanization while exploring their interaction. My calibration emphasizes that the observed patterns of suburban growth and urban core decline can only be understood by considering both trade and commute cost reductions. I simulate a counterfactual scenario without the Interstate Highway System, holding constant all other shocks between 1960 and 2020. I evaluate three cases: highways reduce only commute costs, only trade costs, and both simultaneously. Comparing observed data to the scenario where highways reduce trade and commute costs, I find that highways account for 15% of suburban growth and 33% of urban core decline. When highways reduce only commute costs, the model shows that these reductions account for 88% of the highways' impact on suburban growth. However, commute cost reductions alone do not explain the decline in urban cores, as suburban expansion in this scenario is primarily driven by migration from rural areas. In contrast, when trade cost reductions are added to the reductions in commute costs, they explain the remaining 12% of the impact on suburban

[4]One interpretation of this approach is that congestion forces dominate in larger locations, while smaller locations lack sufficient population to generate significant endogenous amenities. Meanwhile, TFP externalities, calibrated from existing literature, remain constant, preserving the productivity advantages of large cities.

growth and account for 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 constant trade costs, 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 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 trade and commute access. The interaction between cost 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.

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 most studies focus on transport infrastructure's effects on either trade or commute costs, I analyze both dimensions jointly, using a combination of reduced-form and structural approaches. In the reduced-form analysis, I use my new dataset of bilateral driving times to measure changes in market and commuting access across U.S. counties. I show that, on average, improvements in

[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.

[6]Trade: Michaels, 2008; Duranton et al., 2014; Faber, 2014; Allen and Arkolakis, 2014, 2023; Donaldson and Hornbeck, 2016; 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; Warnes, 2021; Brinkman and Lin, 2022; Tsivanidis, 2023; Weiwu, 2023, 2024; Velasquez, 2024.

commute access have a negligible impact on population and employment growth across all counties, while improvements in trade access drive significant growth in both outcomes. However, these aggregate results mask important heterogeneity across locations. The structural analysis uncovers this heterogeneity by separately characterizing the contributions of trade and commute access for each county, while matching the average effects estimated in the reduced-form analysis. I find that improvements in commute access are particularly influential in suburban areas, where they significantly drive population and employment growth. Moreover, improvements in trade access amplify these effects, interacting with commute access improvements to accelerate suburban expansion and the decline of urban cores.

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 productivity and amenity externalities.[8] Some of this literature focuses on amenity externalities.[9] I provide new estimates of dynamic amenity externalities and evidence of heterogeneous amenity 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 productivity 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 calibrations 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.

[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).

1.2 Background, Data and Stylized Facts

1.2.1 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 (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; Michaels, 2008; Duranton and Turner 2012; Frye, 2023; Brinkman and Lin, 2024).

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"

[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.

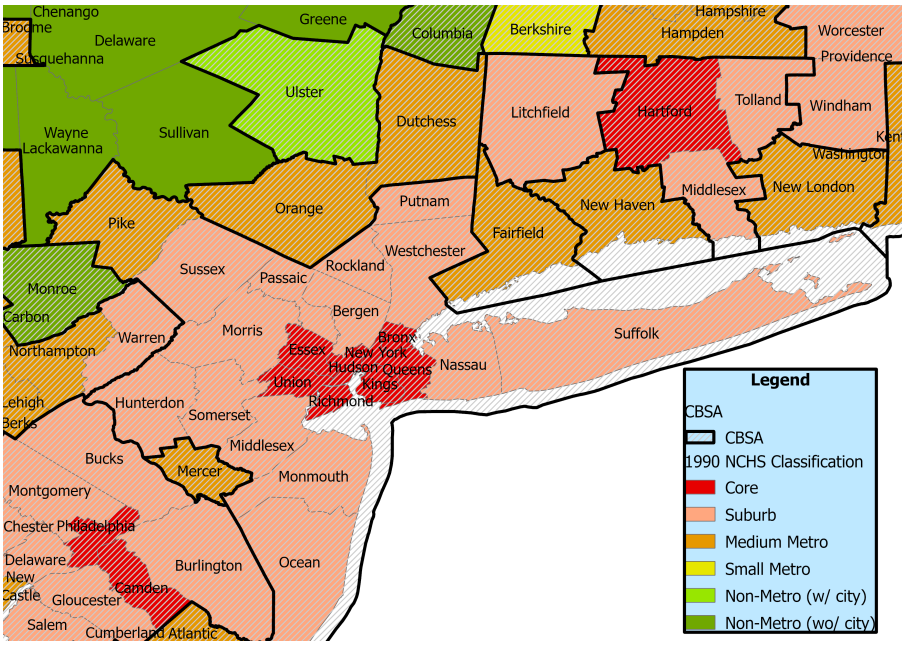
[11] The PR-511 data set identifies the year of construction of each highway segment (Baum-Snow, 2007; Baum-Snow, 2020; Weiwu, 2024).

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

metro counties (suburbs).[13] 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 1.1: New York and Hartford



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

Figure 1.1 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 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.

[13] 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.

1.2.2 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.

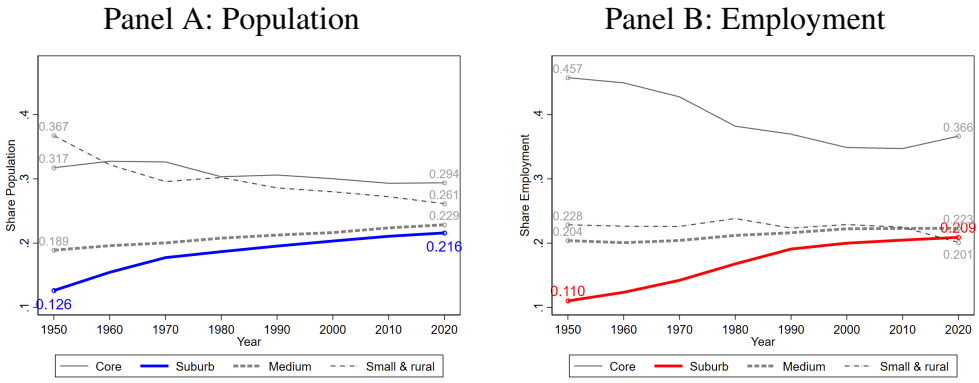
Fact 1: Suburbs emerged as population and employment centers at the expense of rural and core counties, respectively.

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 1.2 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 1.2: 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 1.3 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.

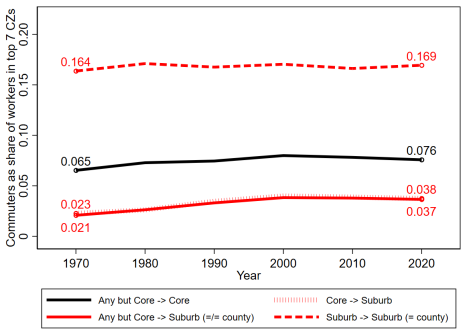
This paper argues that the growing prominence of suburbs as commuting destinations stems in part from highway construction, which reduced trade costs and shifted employment away from urban cores. While commuting access also plays a role, its effect is distinct: mid-sized suburban zones primarily draw 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 these facts serves as a motivation to set up a model with a realistic geography that

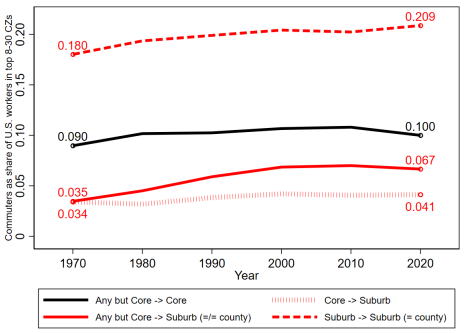
captures commute and trade (and migration) linkages.

Figure 1.3: Share of workers by type of commute out of total number of workers in size bin

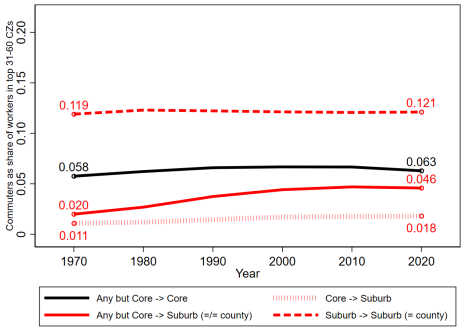
Panel A: Commuting in Top 7 Commuting Zones



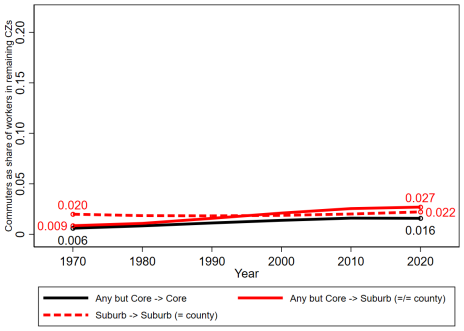
Panel B: Commuting in Top 8-30 Commuting Zones



Panel C: Commuting in Top 31-60 Commuting Zones



Panel D: Commuting in remaining Commuting Zones



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.

1.3 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).

1.3.1 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:

$$\Delta \log \Phi_{i,t}^{W} = -\sum_{j} \pi_{ij|i}^{0} \cdot \Delta \tau_{i,j,t} \tag{1.1}$$

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, $\Delta \log \Phi_{i,t}^{W}$, 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:

$$\Delta \log \Psi_{d}^{T} \propto -\sum_{o} \lambda_{od|o} \Delta \log \tau_{o,d,t}$$

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 $\lambda_{od|o} \approx \frac{L_d (\varsigma_{od})^{-\varepsilon}}{\Psi_o^T}$ where market access is approximated as $\Psi_o^T \approx \sum_{d=1}^{N} (\varsigma_{od})^{-\varepsilon} L_o$ and $\varsigma_{od} = \tau_{od}^{\varkappa}$ 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:

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

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

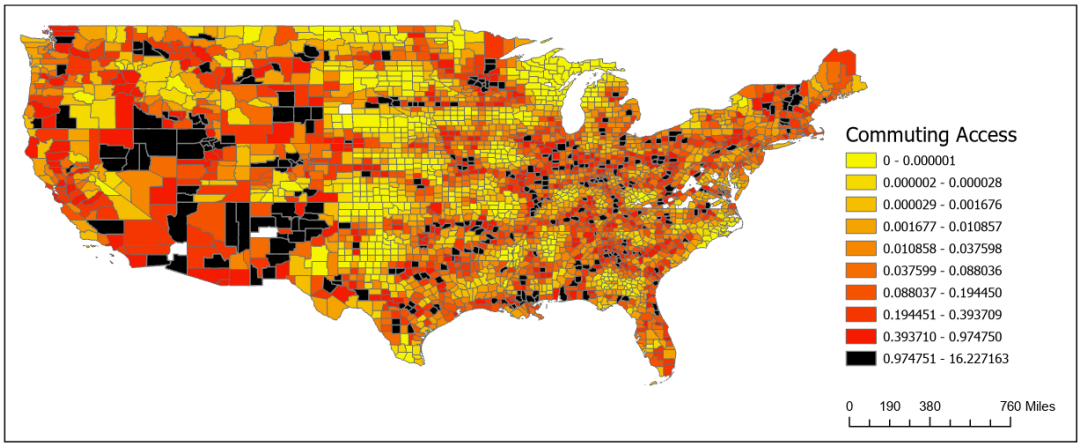
trade access and commute access, the reality is different. The correlation between these measures was only 23.3% from 1960 to 1970. The decade-by-decade correlation between these two was 32.8% from 2000 to 1960. This low correlation arises because trade access assigns positive weights to distant locations, 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 1.4, 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 1.5 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 1.4: Changes in commute and trade access between 1970 and 1960

Panel A: Commuting Access



Panel B: Trade Access

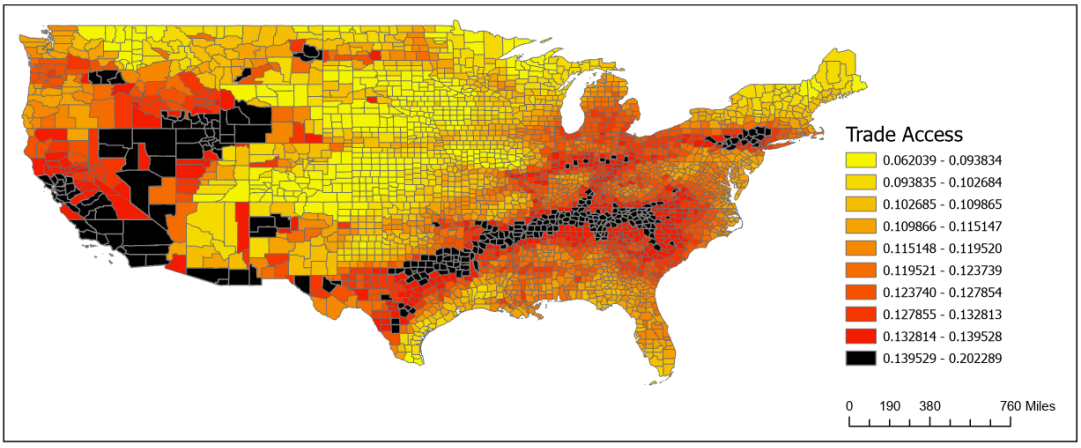
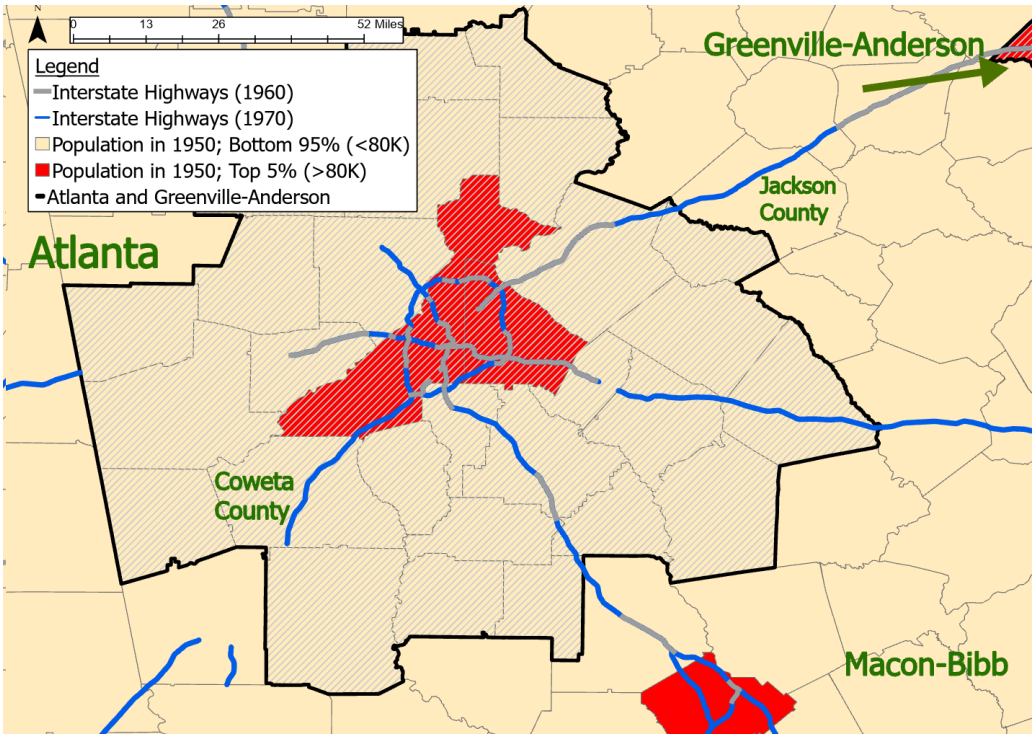
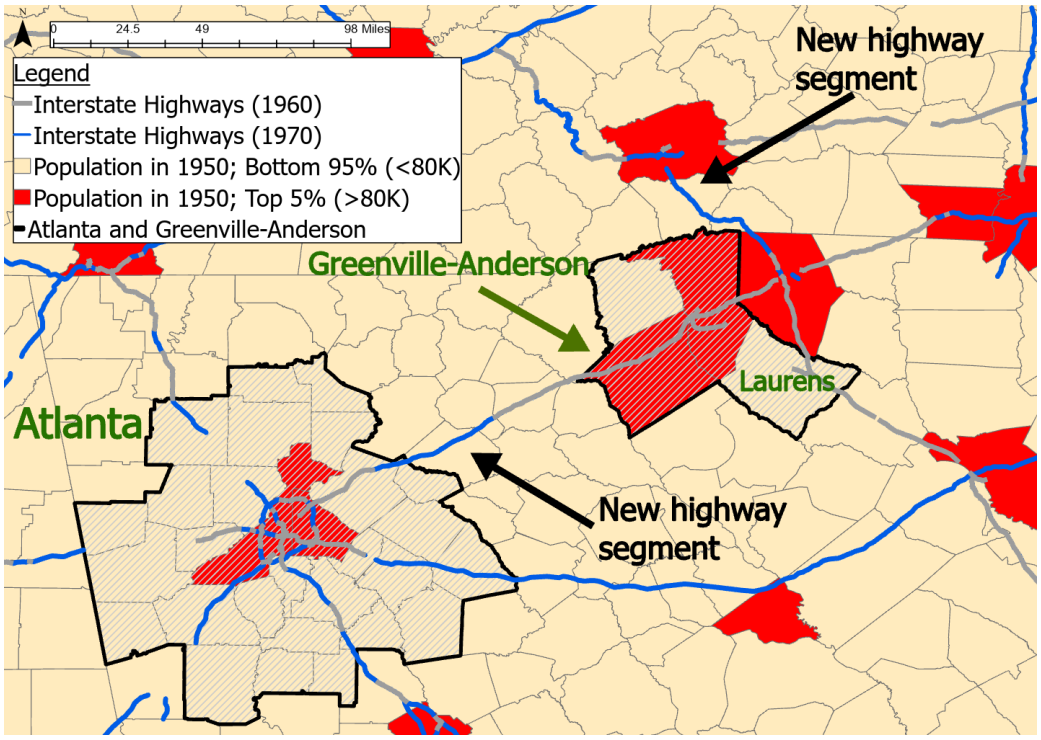


Figure 1.5: Illustrative examples

Panel A: Atlanta, Georgia



Panel B: Atlanta and Greenville-Anderson Metro areas



1.3.2 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:

$$y_{it} = \varrho\Phi_{it} + \varphi\Psi_{it}^{T} + FE_i + FE_t + \beta' X_{it} + \varepsilon_{it} \tag{1.2}$$

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; Michaels, 2008; Duranton and Turner 2012; Frye, 2024; 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.[14] 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:

$$\underbrace{\ln y_{j,s,q}}_{\text{population and employment}} = \rho_h \underbrace{\log \Phi^W_{j,s,q}}_{\text{commuting}} + \varphi_h \underbrace{\log \Psi^T_{j,s,q}}_{\text{trade}} + \underbrace{\gamma_{s,q,h} + \gamma_{j,s,h}}_{\text{FE}} + \underbrace{\varpi'_h X_{j,s,q}}_{\text{controls}} + \tilde{\mu}_{j,s,h} \quad (1.3)$$

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 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

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

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:

$$\begin{aligned} \ln y_{j,s,q} &= \rho_h \log \Phi^{W,0}_{j,s,q} + \varphi_h \log \Psi^{T,0}_{j,s,q} + \zeta_h \underbrace{\left(\log \Phi^{W,0}_{j,s,q} \cdot \log \Psi^{T,0}_{j,s,q}\right)}_{\text{interaction}} + \\ &+ \gamma_{s,q,h} + \gamma_{j,s,h} + \varpi_h' X_{j,s,q} + \tilde{\mu}_{j,s,q} \end{aligned} \tag{1.4}$$

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:

$$\tilde{\Psi}_d^k \propto -\sum_o \tilde{\lambda}_{od}^0 \tau_{od,t}^k \quad \text{and} \quad \tilde{\Phi}_{i,t}^k \propto -\sum_j \pi_{ij}^0 \tau_{ij,t}^k \tag{1.5}$$

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., 2008; Frye, 2024). 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.

1.3.3 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 1.1 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 1.2. Panel A of Figure 1.6 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 1.1: Commute and trade access effects on log population after 40 years

	(1) Log Population	(2) Log Population	(3) Log Population	(4) 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 1.2: Commute and trade access effects on log employment after 40 years

	(1) Log Employment	(2) Log Employment	(3) Log Employment	(4) 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 1.6 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 1.6: OLS estimates of the effect of commuting and 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.

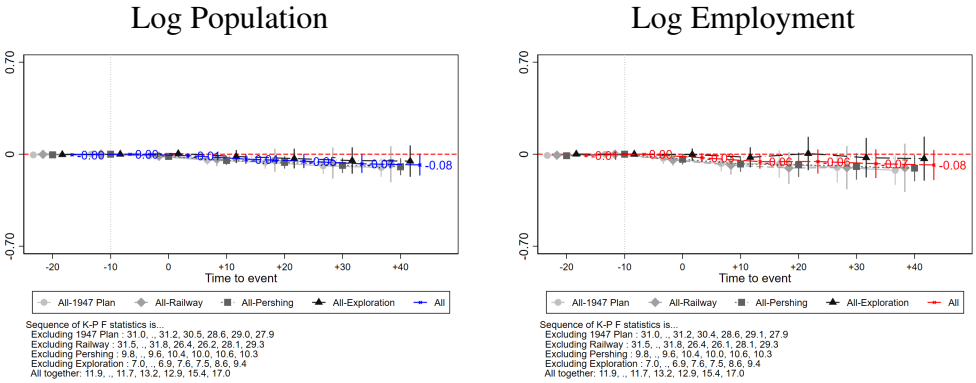
2SLS results. Figure 1.7 summarizes the main findings. Panel A shows the average effect of increasing commute access (when controlling for trade access and the interaction). Panel B does so for trade access. In each panel there are two outcomes, population and employment. Each figure depicts in blue 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

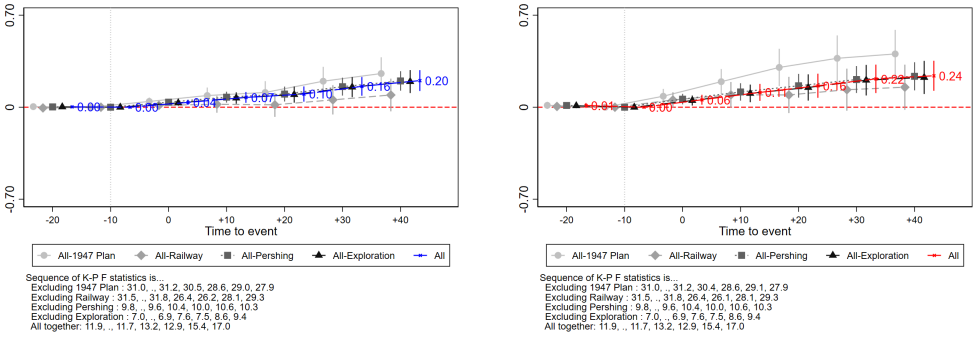
impact population and employment through trade access, not commuting access. After 40 years, the effect of commute access is negative and insignificant, while trade access has a positive and significant impact.[15] Second, these results hold across different sets of instruments. Third, the 2SLS results 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).

Figure 1.7: 2SLS estimates of the effect of commuting and trade access

Panel A: Commute Access



Panel B: Trade Access



1.3.4 Other results

In Appendix A.2, 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

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

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.

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

1.3.5 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.

1.4 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.

1.4.1 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.

1.4.2 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:

$$U_{n,i,j,t}(\omega) = \nu_i(\omega) \cdot \xi_{n,i,t} \cdot u_{i,t} \cdot \frac{\epsilon_j(\omega)}{d_{i,j,t}} \cdot C_{i,t}^{\alpha} H_{i,t}^{1-\alpha} \tag{1.6}$$

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:

$$y_{n,i,j,t}(\omega) \equiv w_{j,t} = P_{i,t} C_{i,t} + R_{i,t} H_{i,t} \tag{1.7}$$

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:

$$\pi_{ij|i,t} = \frac{\left(\frac{w_{j,t}}{d_{i,j,t}}\right)^{\theta}}{\Phi_{i,t}} \tag{1.8}$$

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:

$$\mathbb{E}[y|i] = \sum_{j} \pi_{ij|i,t} w_j \equiv \bar{y}_{i,t} \tag{1.9}$$

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

$$\mathbb{E}\left[U_{n,i,j,t}(\omega) \mid i, n\right] \propto \nu_i(\omega) \cdot \underbrace{\frac{\xi_{n,i,t} u_{i,t}}{P_{i,t}^{\alpha} R_{i,t}^{1-\alpha}} \cdot \Phi_{i,t}^{1/\theta}}_{\bar{U}_{n,i,t}} \equiv U_{n,i,t}^{\mathcal{E}}. \tag{1.10}$$

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 pa-

rameter 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:

$$L_{n,i,t} = \zeta_t \frac{\bar{U}^{\eta}_{n,i,t}}{\sum_{i \in N} \bar{U}^{\eta}_{n,i,t}} L_{n,t-1} \tag{1.11}$$

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.

1.4.3 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 $\eth$. The efficiency of producing good $\eth$ 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:

$$c^s_{o,t} = \left(\varrho^s_L w^{1-\sigma}_{o,t} + \varrho^s_H \tilde{R}^{1-\sigma}_{o,t} + (1 - \varrho^s_L - \varrho^s_H) P^{T\,1-\sigma}_{o,t}\right)^{\frac{1}{1-\sigma}} \tag{1.12}$$

where P_o^T is the CES price index intermediate goods, which is the same price index of tradable goods, and $\varrho_L^{NT} + \varrho_H^{NT} = 1$.

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:

$$P_{i,t}^{NT} = \left[\int_0^1 \left(p_{i,t}^{NT}(\eth)\right)^{1-\sigma_D} d\eth\right]^{\frac{1}{1-\sigma_D}} \quad \text{and} \quad P_{i,t}^{T} = \left[\int_0^1 \left(p_{i,t}^{T}(\eth)\right)^{1-\sigma_D} d\eth\right]^{\frac{1}{1-\sigma_D}} \tag{1.13}$$

$$P_{i,t} = \left[b\left(P_{i,t}^{T}\right)^{1-\sigma_D} + (1-b)\left(P_{i,t}^{NT}\right)^{1-\sigma_D}\right]^{\frac{1}{1-\sigma_D}} \tag{1.14}$$

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.

1.4.4 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, $p_{od}^s = \frac{c_o^s \varsigma_{od}^s}{a_o^s}$ 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:

$$\frac{X_{od}^s}{X_d^s} = \lambda_{od}^s = \frac{A_o^s \left(c_o^s \varsigma_{od}^s\right)^{-\varepsilon}}{\Psi_d^s} \tag{1.15}$$

Finally, to derive the price indexes, we plug the distribution of realized prices into the price indexes of equations 1.13.[16] Assuming $\sigma_D < 1 + \varepsilon$, the price indexes become:

$$P_d^{NT} = cons \cdot \left(\Psi_d^{NT}\right)^{-1/\varepsilon} \quad \text{and} \quad P_d^T = cons \cdot \left(\Psi_d^T\right)^{-1/\varepsilon} \tag{1.16}$$

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

1.4.5 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.[17]

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

$$c_o^{\text{Cons}} = \left(\beta w_o^{1-\nu} + (1-\beta) Q_o^{1-\nu}\right)^{\frac{1}{1-\nu}}, \tag{1.17}$$

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:

$$R_o = \frac{\left(\beta w_o^{1-\nu} + (1-\beta) Q_o^{1-\nu}\right)^{\frac{1}{1-\nu}}}{Z_o}, \tag{1.18}$$

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,

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

[17] 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.

own the land and use their land payments to consume the final good.

1.4.6 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:

$$
\begin{aligned}
A_{j,t}^{T} &= \bar{A}_{j,t}^{T} \cdot \left(\tilde{L}_{j,t-1}^{T}\right)^{\delta}, \qquad (1.19)\\
u_{i,t} &= \bar{u}_{i,t} \cdot \begin{cases} \gamma_0^{\phi} & \text{if } L_{i,t-1} < \gamma_0 \\ \left(L_{i,t-1}\right)^{\phi} & \text{if } \gamma_0 \leq L_{i,t-1} \leq \gamma_1 \\ \gamma_1^{\phi} & \text{if } L_{i,t-1} > \gamma_1 \end{cases}
\end{aligned}
$$

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 Ł$_{j,t-1}$. 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.

1.4.7 Market clearing and equilibrium

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

$$\underbrace{\tilde{L}_d w_d}_{\text{Commuting}} = \underbrace{\varrho_L^T \left(\frac{w_d}{c_d^T}\right)^{1-\sigma} X_d^T}_{\text{Tradable Sector}} + \underbrace{\varrho_L^{NT} \left(\frac{w_d}{c_d^{NT}}\right)^{1-\sigma} X_d^{NT}}_{\text{Non-Tradable Sector}} + \underbrace{\beta \left(\frac{w_o}{c_o^{Cons}}\right)^{1-\nu} X_d^{Cons}}_{\text{Construction Sector}} \tag{1.20}$$

$$\underbrace{\frac{Q_d T_d}{(1-\beta)\left(\frac{Q_d}{c_d^{Cons}}\right)^{1-\nu}}}_{=H_d R_d} = \underbrace{\varrho_H^T \left(\frac{\tilde{R}_d}{c_d^T}\right)^{1-\sigma} X_d^T}_{\text{Tradable Sector}} + \underbrace{\varrho_H^{NT} \left(\frac{\tilde{R}_d}{c_d^{NT}}\right)^{1-\sigma} X_d^{NT}}_{\text{Non-Tradable Sector}} + \underbrace{(1-\alpha)\bar{y}_d L_d}_{\text{By Workers}} \tag{1.21}$$

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:

$$X_d^T = \sum_o \frac{A_d^T \left(c_d^T \varsigma_{do}^T\right)^{-\varepsilon}}{\Psi_o^T} \left(\underbrace{P_o^T C_o^{W,T}}_{\text{Workers}} + \underbrace{P_o^T C_o^{H,T}}_{\text{Landlords}} + \underbrace{\frac{(1 - \varrho_L^T - \varrho_H^T)}{\varrho_L^T} \left(\frac{P_o^T}{w_o^T}\right)^{1-\sigma} w_o \tilde{L}_o^T}_{\text{Tradable}} \right) \quad (1.22)$$

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).[18]

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

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 $\left\{\varrho_L^T, \varrho_H^T, \varrho_L^{NT}, \mathrm{b}, \beta, \alpha, \theta, \eta, \varepsilon, \sigma, \sigma_D, \nu, \gamma, \phi, \zeta_t\right\}$, 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, $\sum_i L_{it} = \zeta_t \sum_i L_{it-1}$.

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:

$$\underbrace{w_d \sum_i \pi_{ij|i} L_i}_{\text{Labor Supply}} = \underbrace{\sum_o \lambda_{do|o} \left(\bar{y}_o L_o\right)}_{\text{Labor Demand}} \quad (1.23)$$

[18] 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.

where:

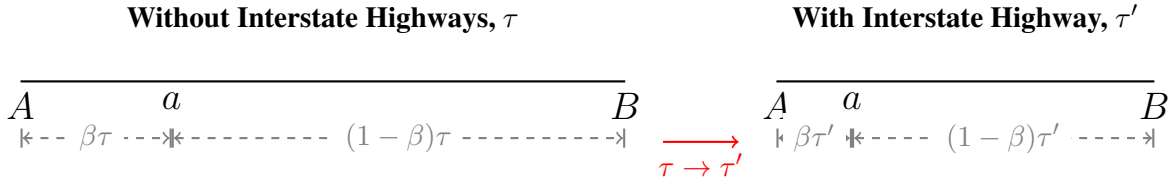
$$\lambda_{do|o} = \frac{A_d \left(c_d \varsigma_{do}\right)^{-\varepsilon}}{\Psi_o} \quad L_i = \frac{\bar{U}_i^{\eta} \bar{L}}{\sum_{i \in N} \bar{U}_i^{\eta}} \quad \bar{U}_{ni} = \frac{u_i}{P_i} \cdot \Phi_i^{1/\theta}$$

$$\bar{y}_i = \sum_j \pi_{ij|i} w_j \quad \pi_{ij|i} = \frac{\left(\frac{w_j}{d_{ij}}\right)^{\theta}}{\Phi_i} \quad \Phi_i = \sum_{l=1}^{S} \left(\frac{w_j}{d_{ij}}\right)^{\theta}$$

$$\Psi_i = \sum_j A_j \left(w_j \varsigma_{ji}\right)^{-\varepsilon} \quad P_i = \left(\Psi_i\right)^{-1/\varepsilon}$$

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 1.8 illustrates the geography before and after the highway construction, emphasizing the compression of driving times.

Figure 1.8: 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:

$$-\frac{1}{L_a}\frac{dL_a}{d\tau} = -\left(\Omega_{L_a/w_A} \cdot \frac{dw_A/d\tau}{w_A}\right) - \left(\Omega_{L_a/w_B} \cdot \frac{dw_B/d\tau}{w_B}\right) - \left(\Omega_{L_a/\varsigma} \cdot \varkappa \tau^{-1}\right) - \left(\Omega_{L_a/d} \cdot \kappa\right) \quad (1.24)$$

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 1.24. 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:

$$-\Omega_{L_a/d} = -\beta \varpi_A \left\{ \pi_{Aa} - \pi_{aA} \right\} - \beta \varpi_B \left\{ -\pi_{aA} \right\} > 0 \tag{1.25}$$

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:

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

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:

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

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

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

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:

$$\Omega_{L_a/w_B} = \varpi_A \left(\underbrace{\lambda_{BA|A} - \lambda_{Ba|a}}_{<0} \right) \eta - \varpi_B \underbrace{\left(1 - \lambda_{BB|B} + \lambda_{Ba|a}\right)}_{>0} \eta < 0 \tag{1.26}$$

$$\Omega_{L_a/w_A} = \left\{ \varpi_A \underbrace{\left(\left(\pi_{aA} - \lambda_{Aa|a}\right) - \left(\pi_{AA} - \lambda_{AA|A}\right) \right)}_{?} \eta + \varpi_B \underbrace{\left(\pi_{aA} + \lambda_{AB|B} - \lambda_{Aa|a}\right)}_{?} \eta \right\}$$

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.3.

1.5 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.

1.5.1 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 1.5.1: *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.4.1 □

Theorem 1.5.1 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 1.5.2: *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.4.2 □

Theorem 1.5.2 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 1.5.3: *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.*[20]

Proof. See Appendix A.4.3 □

Theorem 1.5.3 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 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.[21] 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.

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

[21] 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 ($\tilde{L}_d^{NT} + \tilde{L}_d^{T}$).

THEOREM 1.5.4: *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.4.4 □

Finally, Theorem 1.5.4 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, $L_{i,t}$.

1.5.2 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.5 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.5.[22] Then, I use estimates 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 $\varkappa_m\eta = 1.25$. Following the literature, I set $\eta = 4$. Population growth ζ_t is calibrated directly from the data.

[22]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 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.

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.[23]

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).[24] 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.[25] Regarding, the 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

[23] 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.

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

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

(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 1.5.1-1.5.4 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.

1.5.3 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.[26] I utilize an indirect inference approach and map the externality parameters to the reduced-form evidence I reported in Section 1.3.3. 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 1.5.1-1.5.4, I recover unobservable factors for each decade, hold them 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:

$$Loss = [\beta^{data} - \beta^{model}]I[\beta^{data} - \beta^{model}]' \tag{1.27}$$

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

Panel A of Figure 1.9 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

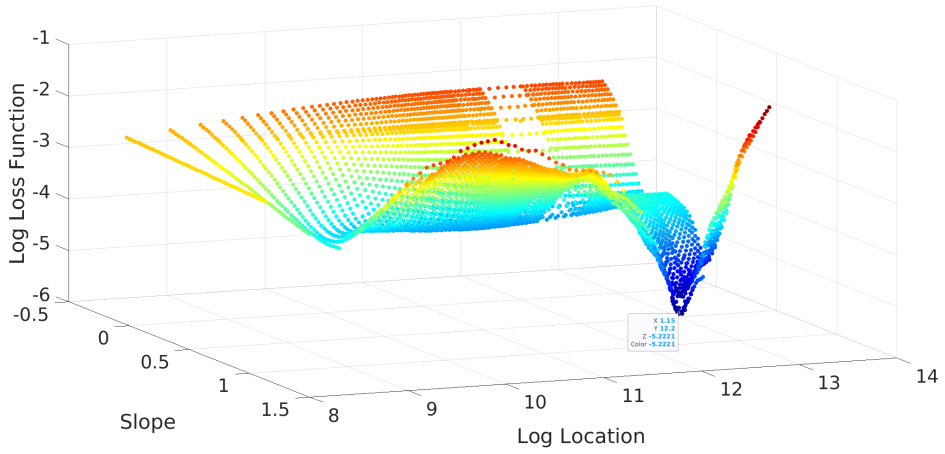
[26] To avoid estimating three parameters, I calibrate γ_0 using the average rural population of 3,000, meaning externalities are inactive below this threshold.

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.

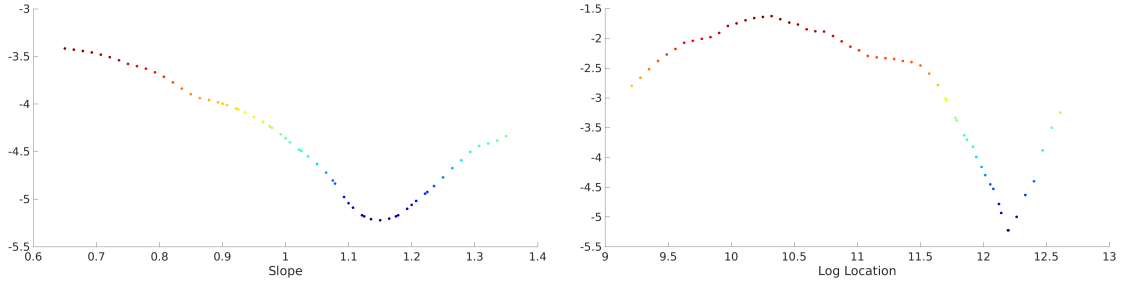
I estimate a slope parameter of $\phi = 1.15$ and a location threshold of $\gamma_1 = 198,790$. Panel B of Figure 1.9 shows sharp identification, especially for the location parameter. The estimated effects are: trade access increases population by 0.21 (data = 0.20) and employment by 0.17 (data = 0.24), 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 1.9: 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.

1.6 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.

1.6.1 Aggregate effects: results

Table 1.3 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 1.3: The effect of Interstate Highways on population across county aggregates

Panel A: Population share by NCHS classification				
	Core	**Suburbs**	**Medium**	**Small/Rural**
Data $\triangle$ (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 $\triangle$ (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.

1.6.2 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 1.6, 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,

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

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

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

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 1.4 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 1.3.), while reductions in trade costs encourage rural growth at the expense of cores (see "Trade Effect" in Table 1.3.).

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 1.3.) 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 1.3.) 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 1.3, 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,

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

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

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

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 1.4: 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

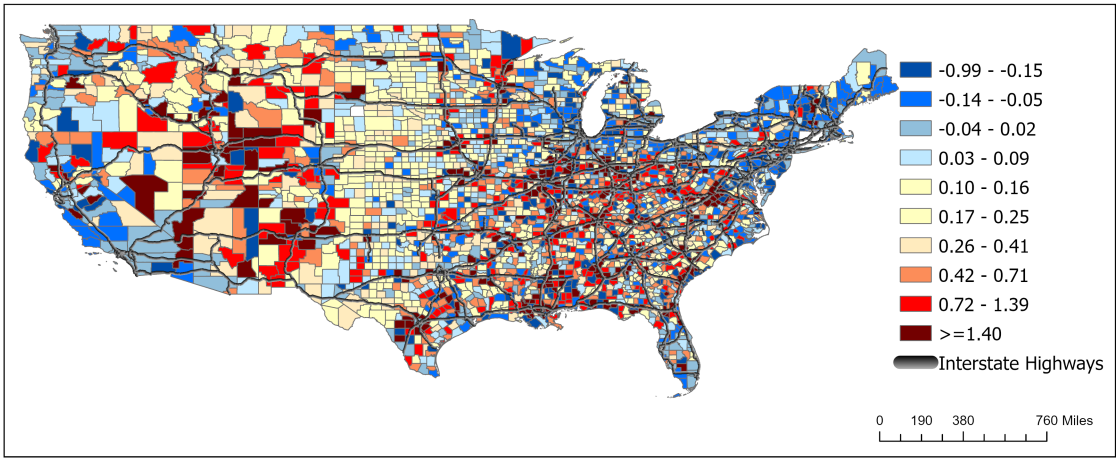
1.6.3 Heterogeneous effects

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 1.10 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 1.10: Population and employment growth across counties

Panel A: Population growth (highways - no highways)



Panel B: Employment growth (highways - no highways)

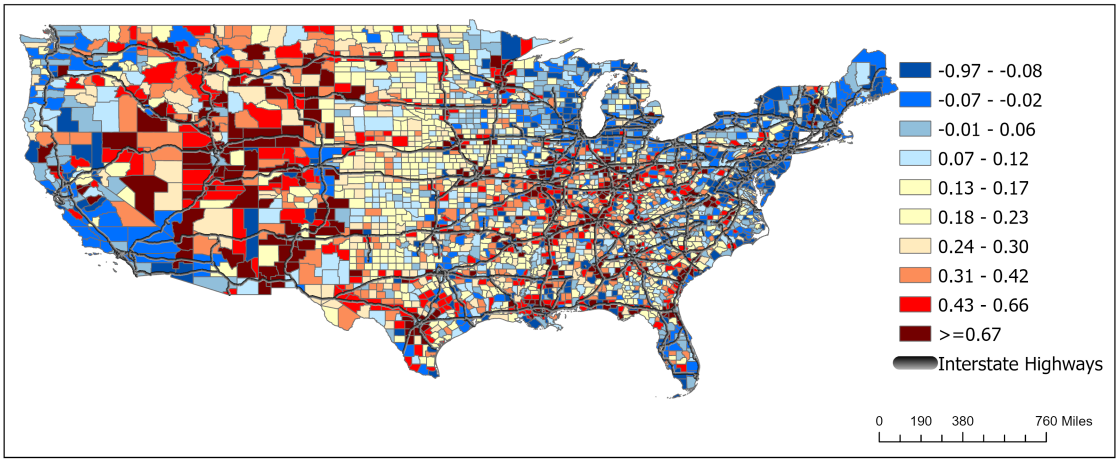


Figure 1.11: County growth across suburbs and urban cores

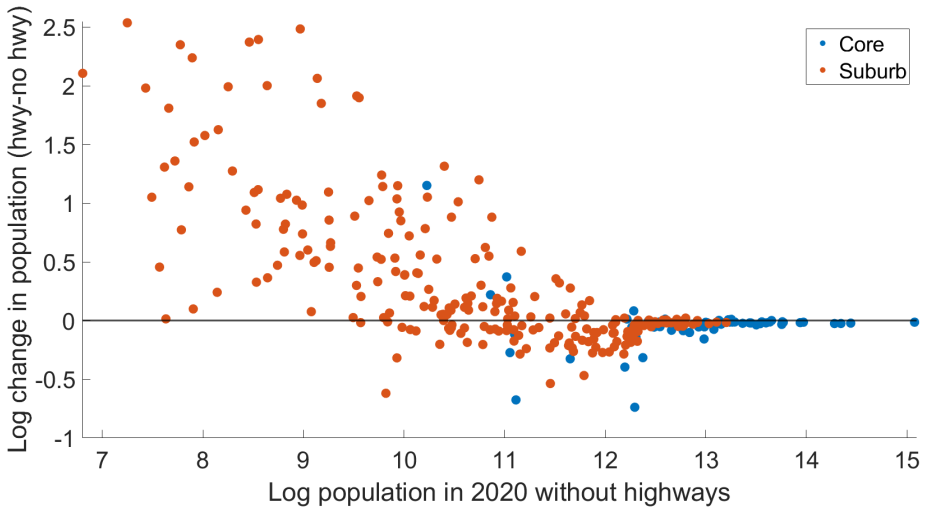


Figure 1.11 plots the log change in 2020 population between the world with highways and without highways against the 2020 population in the counterfactual world without highways, focusing on suburbs and cores. The results show that population growth in suburbs was concentrated in smaller suburbs, while 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 1.5 and 1.6 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 1.5 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 1.6 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 1.5: 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 1.6: 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)

1.7 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.

CHAPTER 2

Transit Infrastructure and Couples' Commuting Choices in General Equilibrium

Abstract

I study the impact of transit on commuting and gender earnings inequality among singles and couples. Using unique household and census data recording the commuting destinations of members, I show that (i) commuting is interdependent within couples and that (ii) women are more sensitive to commute time differences than men. I complement these findings by leveraging new transit infrastructure and a differences-in-differences approach. I find that gender inequality in earnings and labor participation decreased in areas where access increased the most. I set up a quantitative model of a city where households choose residence and commuting destinations, with interdependent commuting choices within couples. I simulate the effect of the new transit infrastructure on the earnings and commuting patterns of singles and couples. In areas with the most substantial reductions in commute times, the gender earnings gap among couples decreased by 12%, while remaining unchanged among singles. With interdependent commuting, a higher husband's earnings make the household more willing to sacrifice the wife's earnings to reduce total commuting costs. Absent interdependent commuting, the gender gap would have decreased even more in response to improved commute times.

2.1 Introduction

Urbanization brings forth advantages and challenges, some of which are specific to women. For example, cities' longer commutes disproportionately burden women as they try to balance their work and family responsibilities, limiting economic gains of urban areas.[1] With the global urban

[1]Black, Kolesnikova, and Taylor (2014) show that labor force participation rates of married women are negatively correlated with the metropolitan area commuting time in the U.S. More recently, Le Barbanchon, Rathelot, and Roulet (2021) show that gender differences in commute valuation can account for 14% of the residualized gender wage gap

population expected to increase by more than 2.5 billion by 2050, primarily in middle- and low-income countries, the challenges faced by women due to urbanization—and especially married women—are likely to become more pronounced as cities get more congested.

In the face of growing urbanization, governments worldwide have invested heavily in infrastructure projects to ease commuting in urban areas. To understand how these investments impact households, particularly women, it is essential to consider multiple-member households. In married households, labor supply and commuting decisions are made jointly, which means that when one spouse increases his earnings, the household may be willing to sacrifice a portion of the other spouse's earnings to reduce overall commuting costs. Thus, even when new transit infrastructure can improve households' total welfare, the interdependent commuting channel can offset the welfare gains experienced by married women, increasing gender earnings inequality. Yet, the literature about transit investments has paid limited attention to couples, even when they typically constitute more than 50% of households (OECD, 2023). Thus, in this paper, I ask: How does urban transit infrastructure affect labor supply and gender earnings inequality once we account for interdependent commuting decisions within households?

I provide new theory and evidence about how transit infrastructure differentially affects married households by studying the response to the Bus Rapid Transit system (BRT) and Line 1 subway construction in Lima, Peru. I assemble various georeferenced data sets at the neighborhood level, including commuting and household surveys, census, firm surveys, and infrastructure data. I start by documenting three facts. First, using novel census data recording the commuting destinations of workers, I find that commuting elasticities are heterogeneous across household groups. Specifically, women's commuting flows are more sensitive to improvements in commuting times than men's, especially when women cohabit with a partner (Fact 1). Second, by observing the commuting destinations of both spouses within households, I compute *conditional commuting flows*: a member's commuting probabilities given the partner's commuting destination. I show that the choice of where to work depends on the spouse's commuting choice, even after controlling for a destination's desirability. This implies that choices are interdependent within dual-earner households (Fact 2). Third, by leveraging transit infrastructure building plans, precisely georeferenced household and census data, and a differences-in-differences approach, I show that the introduction of the new transit infrastructure reduced gender earnings inequality and increased female labor force participation in locations closer to newly constructed stations relative to locations closer to planned stations, especially in the fringe of the city (Fact 3). These facts imply that transit infrastructure may have gendered impacts due to the heterogeneity in commuting elasticities and the interdependence in commuting choices. Both features regulate how more responsive women are to improved commute times than men. Therefore, these are two key dimensions that the model

in France.

should consider.

To gain a better understanding of the interaction between commute times and within household choices, I develop a general equilibrium model of commuting for married households. I start by microfounding the commuting model by assuming that married households achieve Pareto efficiency—as is standard in the family economics literature. This approach explicitly takes account of the fact that married households consist of several members sharing the same budget constraint. The main implication of the household model is that commuting choices are interdependent across spouses. Hence, there are two channels through which improved commute times can affect households. On the one hand, when times between origins and destinations improve, income at origins increases only if the new destinations are high-wage locations. Otherwise, if the new destinations are low-wage locations, improvements in commuting times may induce some workers to give up some of their earnings to have a shorter commute. I call this the *direct use channel* which is a mechanism present in the standard class of spatial models (e.g. Tsivanidis, 2023). On the other hand, when commute costs for one of the spouses decreases leading to higher household earnings, the household becomes more willing to sacrifice the other spouse's earnings in order to decrease her commute costs (and vice-versa). I call this the *indirect use channel* (or within-household spillovers). Thus, the final impact on the gender gap through these two channels depends on the geography of commute times improvements and the distribution of married and single households and male and female wages across locations.

Next, I embed the model of commuting choices for married households into a general equilibrium model of city structure in the spirit of Ahlfeldt et al. (2015) and Tsivanidis (2023). The model incorporates the following features. I consider both single and married households. I allow for commuting elasticities to be heterogeneous across genders and civil status, which is consistent with Fact 1. Moreover, the model includes a residential location choice with location elasticities that differ between single and married households. To account for the fact that spouses within households can endogenously specialize in labor or household production in response to improved commute times, I allow married households to choose to be either dual-earner or male-breadwinner households.[2] From the production side, the model includes multiple sectors that employ workers of each gender with a certain degree of substitution, allowing for gender gaps in the workplace (Acemoglu et al., 2004). Finally, the model allows for endogenous sectoral TFP to account for density effects in industries that employ men and women with different intensities (Rosenthal and Strange, 2012; Ahlfeldt, et al., 2015; Tsivanidis, 2023).

[2]I assume that married households are comprised of a male individual, i.e., the husband, and a female individual, i.e., the wife. A simplifying assumption usually made in the literature on family economics is that the husband always works. Moreover, in the data, female breadwinner households constitute just 1.3% of all households. This is why I turn off the husband's participation margin: it does not seem quantitatively important but would add greater complexity to the model.

I estimate the model using data from Lima, Peru. To estimate commuting elasticities, I exploit data on bilateral flows and times and the gravity structure of the model for commuting probabilities. I use Fact 1 to discipline these elasticities. Moreover, to estimate scale elasticities, I leverage the introduction of the new transit infrastructure to construct instruments. I weigh changes in commute times generated by the new infrastructure with commuting probabilities to produce Bartik-style instruments. Once parameters are estimated, I show that the model with interdependent commuting choices within couples fits better the data on (i) the gender earnings gap across municipalities and (ii) the bilateral commuting flows of dual-earner households than a model with independent choices. Moreover, although I do not use Fact 2 and 3 to discipline the model directly, the model with interdependent choices does a better job at approximating these facts than the model with independent choices, which validates my modeling decisions.

To understand the impact of improving commute times on gender earnings inequality, I study two counterfactuals. In the first one, commute times are set to pre-Metro and BRT implementation levels. In the second one, I predict what would happen if Lima's planned network of Metro lines is constructed. These exercises allow me to understand how improving commute times affects income by gender, the quantitative importance of the direct and indirect margins, and the relevance of general equilibrium effects.

I summarize the main findings of the first counterfactual. First, while earnings increased for both men and women due to the new infrastructure, I find that improving commute times decreased the gender gap in aggregate earnings among couples by more than 12% in the fringe of the city, that is, in the set of locations that gained access the most.[3] However, the gap remained constant among singles.

Second, to quantify the relevance of the direct and indirect channel, I perform a decomposition analysis, to a first-order. Income for each gender can increase in response to improved commute times due to the direct channel, the indirect channel, and general equilibrium effects. One the one hand, the approximation indicates that commuting times decreased more substantially along the routes used by married women than those used by married men. This, coupled with the higher elasticity of commuting with respect to commute time for women, contributed to a decrease in the gender gap through the direct use channel. On the other hand, the interdependence in commute choices within married households worsened gender earnings inequality through the indirect use channel. Shutting down the indirect use channel and considering all locations in the city, the gender gap in dual earner households would have decreased about 23% more. This is due to households willing to give up a portion of the wife's earnings to reduce the household's overall commuting

[3]The gap is defined as $\frac{\text{Male income}}{\text{Female income}}$ where in this example, males and females belong to married households. So, an effect of 12% means that female income grew 12 percentage points faster than male income: $\Delta \log(\text{Male income}) - \Delta \log(\text{Female income}) = -12\%$.

costs.

Finally, I re-do the analysis using a model with independent choices. This model would have led me to underestimate reductions in the gender earnings gap substantially. This is due to the model with independent choices substantially underestimating households' direct exposure to the new transit infrastructure, given my observed data.

In the face of growing urbanization, will future developments in the transit infrastructure reduce gender earnings inequality when we account for interdependent commuting within households? I perform a policy counterfactual exercise and predict what would happen if Lima's planned network of Metro lines is constructed. The purpose of this exercise is to quantify the importance of the Metro network in affecting gender inequality when a different set of locations experiences an improvement in commute times. My results show that the gap in aggregate earnings among married in remote locations would decrease by 16%.

This paper contributes to the literature on economic geography and on urban structure. By relying on models with a single income earner per household, this literature shows that improving commute times through a BRT (Tsivanidis, 2023), Metro systems (Severen, 2021; Zarate, 2021), steam railways (Heblich et al., 2020) or cable cars (Khanna et al., 2022) can have major impacts on many margins such as housing, employment, crime, informality, and welfare.[4] However, real households typically have multiple members, and this aspect is likely to be relevant when assessing the effects of new transit infrastructure. For example, Blundell et al., (2007) document that labor supply responses of couples depend on how they share resources within the household. This way, my paper provides a bridge between the field of family and household economics to the field of economic geography. I highlight that incorporating other members to households changes the way households are affected by reductions in commute costs. Since each member contributes to the household's total income, exposure to various shocks such as commute cost reductions depends on the degree of exposure and relative importance of each income source. Hence, reduced commute times can impact commuting choices through two channels: by affecting an individual's prospects and by affecting those of their spouse.[5]

Second, this paper relates to the literature that examines the trade-off between wages and non-wage job characteristics across genders. On the one hand, some papers find that reducing commute costs should decrease gender gaps. Le Barbanchon et al. (2021) use French administrative data on job search criteria and document that unemployed women have a shorter maximum acceptable

[4]See as well: Ahlfeldt et al. (2015), Allen et al. (2015), Monte et al. (2018), Owens III et al. (2020), and Brinkman and Lin (2024). See Redding and Rossi-Hansberg (2017) for a review of spatial quantitative models.

[5]Theoretically, a similar logic could be applied to reductions in trade costs and productivity shocks. When a household's income relies on multiple sources, the impact of a shock to any given source will be greater if that source contributes a larger portion of the household's total income. The more important a source of income is to the household, the more vulnerable they are to any disruptions or changes to that source. This is an avenue for future research.

commute than unemployed men. Using a job search model, they estimate that gender differences in commute valuation account for 14% of the residualized gender wage gap. Using U.S. data, Liu and Su (2022) show that the commuting and wage gaps are considerably smaller among workers living near city centers, especially for occupations with a high geographic concentration of high-wage jobs. They highlight that the geography of jobs amplifies the impact of commuting preferences on the gender wage gap.[6] On the other hand, Bütikofer, et al. (2022) exploit the opening of the Oresund Bridge as an exogenous change in access to a larger labor market for individuals residing in Malmo. They find that the bridge substantially increased commuting, wages, and gender inequality among Swedes living in Malmo. My paper contributes to the literature by explaining these apparently contradictory results. The effect of increasing market access on the gender gap depends on the type of jobs individuals are gaining access to. If jobs are predominantly male-intensive, as one can infer was the case in Bütikofer, et al. (2022), then it is likely that the gender gap will increase. Moreover, when two people are living together and sharing a budget, if the husband gains access to better jobs through a cross-spouse income effect, this reduces the value of longer commuting for the wife relative to a world without income pooling at the household level. This further reinforces the rise in gender earnings inequality. Consistent with this view, Bütikofer, et al. (2022) find that the increase in gender inequality was more pronounced within couples.

Related literature demonstrates that gender differences in the relationship between commuting and employment tend to be magnified among married individuals. For example, Black et al. (2014) document a substantial variation in the labor supply of married women across U.S. cities due to differences in commuting times.[7] My contribution to this literature is to incorporate the family and spatial dimensions together into a single tractable general equilibrium model, allowing me to quantify the aggregate impact of new transit infrastructure in the presence of married households. In this exercise, I account for commuting destinations, a kind of intensive margin understudied relative to others such as employment and working hours.

[6]Moreover, Kwon (2022) investigates whether the increases in connectivity across cities help reduce the gender gap in labor market outcomes in South Korea. In particular, she leverages a massive expansion of High-Speed Rail (HSR) beginning in 2004 and historical railroads constructed during the Japanese colonial era. She finds that the gender gap in the South Korean labor market decreased with the expansion of high-speed rail (HSR). Field and Vyborny (2022), through an experiment in Pakistan, show that reducing mobility constraints significantly impacts job searching for women.

[7]Kawabata and Abe (2018) examine the relationship between commute times and female labor force participation within Tokyo, showing that it differs markedly across households of different civil status. Carta and De Philippis (2018) explore the relationship between husbands' commuting times and wives' employment and family time allocation in Germany, revealing that a 1% increase in the husband's commuting distance reduces his wife's employment probability by 0.016 percentage points. Gu et al. (2021) use administrative records of home mortgages in Beijing and show that dual-income households systematically buy homes that are closer to the wife's workplace. Moreno-Maldonado (2023) builds a spatial model that reproduces the observed differences in labor supply between small and big cities in the U.S. In her model, spatial frictions act as a fixed to labor force participation. Removing spatial frictions reduces gender gaps in labor markets.

Finally, this paper contributes to the literature on family and household economics, which encompasses two main areas: the dynamics within existing unions and mate selection (Browning et al., 2011). My work mainly contributes to the first strand by showcasing how relaxing time constraints can impact earnings inequality within households through spouses' choices (see Chiappori, et al. (2022) for a review). This paper also illustrates how the family dynamic can influence decisions related to the spatial aspect of the data, such as choices about where to live and where to commute.

2.2 Lima's Metro and Bus Rapid Transit System

Lima is an ideal setting to perform this study for several reasons. First, as the political and economic hub of Peru, it comprises 30% of the population and 40% of the country's GDP, leading to a diverse range of economic activities across the region. Manufacturing, commerce, construction, and services are among the primary sectors, with the region of Lima contributing over 50% of their respective GDPs. These industries often exhibit gender-based disparities in employment opportunities, further accentuating the variation in labor prospects across different locations within the city. Second, Lima's significant traffic congestion is worth noting, with buses serving as the primary mode of transportation for 69% of work trips in 2010 (before the new transit infrastructure were introduced), followed by taxis, 'colectivos', or motorbike taxis (11%), and personal cars (9%). Third, two transportation infrastructures were implemented in recent years, generating variation in transportation access across different locations within the city. These were the Line 1 of the Metro and the Bus Rapid Transit System (BRT).

The Line 1 of the Metro is a 34.4 km long transportation system that serves over 500,000 daily passengers. Construction started by the end of the 80s during the first presidency of Alan Garcia. However, construction stopped short after Alberto Fujimori became president in 1990. Almost 20 years later, construction resumed during the second presidency of Alan Garcia. The first 22 km were completed in 2011, with the remaining 12.4 km completed in 2014. Although the Metro does not directly cross through the Central Business District, it runs through high-demand areas, including one of the most significant commercial centers in the city, the "Gamarra" market (see Figure B.1). The Metro has significantly improved commuting times, reducing travel times from 2 hours and 45 minutes to just one hour for end-to-end trips. On average, work trips on the Metro take around 24 minutes (JICA, 2013).

The introduction of the BRT system in Lima occurred around 2010. Typically, a BRT system consists of a dedicated system of roadways for buses and has been successfully implemented in cities such as Bogota and Buenos Aires. The goal of the BRT system is to provide faster speeds while maintaining the simplicity of a bus system. As shown in Figure B.2, Lima's BRT system

connects the fringe to the modern municipalities, as well as the Central Business District (CBD). The system spans 26 km and serves approximately 700,000 passengers daily. A World Bank study found that commuting times were reduced by approximately 25-45% as a result (Oviedo et al., 2019).

2.3 Data Sources, Stylized and Reduced-form Facts

I start by briefly summarizing the data sources. Then, I explain two stylized facts highlighting how commuting might be different for couples compared to singles. Finally, I perform a reduced-form analysis to study the impact of the new infrastructure on gender inequality measured through earnings and labor force participation.

2.3.1 Data sources

For this paper's analysis, I require several sources of data geographically identified at a fine scale. A zone is a set of 25 blocks approximately. A block is the finest available disaggregation used by the National Institute of Statistics and Informatics. Lima in 2017 was partitioned into more than 30,000 blocks and more than 1500 zones. My analysis is performed always at the zone level. Hence, throughout the paper, I refer to zones and blocks interchangeably.[8]

In particular, (i) I use the Population Census from 1993, 2007 and 2017 to get residence counts at the zone level for single men and women, and cohabiting couples. (ii) I also use the Economic Census of 2008 and an open administrative data for firms' location and employment figures in 2015 to calculate employment levels at the block-by-industry level. To compute commuting flows, I rely on several sources of data. (iii) I use the Population Census from 2017, which has data on commuting flows at the municipality level.[9] (iv) I complement this with several commuting surveys collected during 2010-2018 for the city of Lima. Then, (v) I use data on the road network from Open Street Maps and Google Maps API. Furthermore, (vi) I use the National Household Survey (ENAHO) from 2007-2017 waves, which is precisely georeferenced, to analyze rents at the zone level. (vii) Data on land use come from the Metropolitan Institute of Planning of Lima. Finally, (viii) I use the 2010 National Time Use survey to complement the data on commute times. Using my data set, I document two stylized facts that motivate my empirical design and modeling approach.

More information about the data is available in the Data Appendix, which can be found in the online version of the paper (here).

[8] I also use the word "location" to refer to zones. When referring to commuting, I use the word "origin" to refer to the zones where households live, and "destination" to refer to the zones where households work.

[9] There is 50 municipalities in Lima. Each municipality contains about 30 zones and 600 blocks.

2.3.2 Stylized facts

These stylized facts motivate incorporating married households to the model.

Fact 1: Commuting elasticity differs by civil status and gender. Using the Population Census, I classify households into five different groups. The first two consist of (i) *single females* and (ii) *single males*, which comprises 31.9% and 35.9% of households in 2017. Then I consider (iii) *male breadwinner* households, that is, households where the wife stays at home while the husband goes to work (13.3% of households). For the sake of completeness, I also consider (iv) *female breadwinner* households, although they just constitute 1.3% of households. Finally, I consider (v) *dual-earner* households, which represent 17.6% of households.[10] For each of these groups, I run the following reduced form regression:

$$\log \pi_{j|i,k} = \beta_k time_{ij} + FE_{i,k} + FE_{j,k} + \varepsilon_{ij} \tag{2.1}$$

where $\pi_{j|i,k}$ is the share of members of civil status-by-gender k that live in origin i that commute to j. $time_{ij}$ is the time that takes to commute from i to j. Moreover, $FE_{i,k}$ and $FE_{j,k}$ are fixed effects by origin i and destination j. Finally, β_k is the *commuting elasticity*. It measures how the probability of commuting to j would change in relative terms after increasing the commuting time by one minute.

To compute commuting probabilities $\pi_{j|i,k}$ I rely on the 2017 population Census, and aggregate the data at the municipality level.[11] I measure commuting times by zones using two data sources: Open Street Maps (OSM) and the 2010-2018 waves of Lima's Commuting Survey. I assign speeds to match documented speed times in Lima. I then aggregate these times to the municipality level by taking the median value across zones within municipalities.[12]

Table 2.1 shows the result of estimating this equation by PPML. Columns (1) and (2) show commuting elasticities for single males and females. These estimates suggest that single females are 20% more sensitive to changes in commuting times than single men. Columns (3) and (4) exhibit these estimates for breadwinner households. Comparing men and women in breadwinner households leads to a similar conclusion. Women are more sensitive than men in breadwinner households and the gap widens to 36%. Finally, columns (5) and (6) deliver the estimates for men and women in dual-earner households, showing that women are 40% more sensitive than men. The difference is always statistically significant at the 1% level. Figure B.3 depicts the bin scatter between the residual of log commuting probabilities, net of origin and destination fixed effects,

[10] The precise definition is available in the Data Appendix of the online version of the paper.

[11] While I can locate the origin of households at the block level, the data only records the destination municipality, and so I perform this analysis at the municipality level.

[12] In the Data Appendix available in the online version of the paper, I show that these times are well correlated with times reported in commuting surveys.

and the residual of commuting times. Panel A does it for singles, while Panel B for married in men and women breadwinner households. The graph confirms the log-linear relationship between commuting probabilities and commuting times, and the fact that the gap in elasticities is wider among the married.

All in all, these results show that women are more sensitive to improvements in commuting times than men especially when cohabiting with a partner. Hence, the model should allow for heterogeneity in commuting elasticities.

Table 2.1: Commuting Elasticity

	Single HH		Breadwinner HH		Dual-Earner HH	
	Males	Females	Males	Females	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time	-0.0514	-0.0617	-0.0466	-0.0634	-0.0480	-0.0673
	(0.0034)***	(0.0038)***	(0.0033)***	(0.0041)***	(0.0032)***	(0.0041)***
Origin FE	X	X	X	X	X	X
Destination FE	X	X	X	X	X	X
Gap (Female/Male-1)		20%		36%		40%
N	2500	2500	2500	2500	2500	2500

Notes: Observations are at the municipality-by-municipality level. The independent variable in all regressions is the median travel time between origin i and destination j is t, with origin and destination fixed effects included. Column (1) of the regression analysis uses the share of single male workers who reside in origin municipality i and work in destination municipality j as dependent variable. Column (2) uses the corresponding share for single female workers. Columns (3) and (4) focus on households where only the male or female breadwinner works, respectively. In columns (5) and (6), the shares are defined for dual-earner households, specifically for males and females separately. The difference between coefficients in column (1) and (2), between column (3) and (4), and between column (5) and (6) is statistically significant at the 1 percent level.

Robustness and heterogeneity. In Online Appendix B.2, I re-run equation 2.1 but incorporating a set of controls at the origin-destination level such as the average education of people taking that route, average age, etc. Conclusions are the same. Moreover, in Online Appendix B.2, I also study the relationship between commute times and commute flows conditional on having a young kid at home. Results are as expected: having young kids at home make women more sensitive to improvements in commute times relative to men.

Relation to literature. Previous papers featuring quantitative spatial model have relied on data about commuting flows and commuting times to discipline the trade-off between the two. While I am using a similar strategy, the richness of my Census data allows me to use commuting flows to characterize this trade-off across genders and civil status for the first time in the gravity literature. My estimates are within a similar range as those reported elsewhere.[13] Moreover, there is a related

[13]To name a few, point estimates of the semi-elasticity between commute times and commute flows were: [-0.0706,-0.0697] in Ahlfeldt, et al. (2015); [-0.042,-0.028] in Zarate (2021); [-0.038,-0.029] in Warnes (2021); [-0.046; -0.036] in Tsivanidis 2023; [-0.068,-0.039] in Khanna et al., 2022.

literature on compensating differentials and commute valuation under the lens of a job search model. In this research, wage data is regressed against commute times, leveraging some quasi-experimental variation, to identify the trade-off between wages and commute time at the origin level (e.g., Mulalic et al., 2013). The first to bring the gender dimension to this valuation problem was Manning (2003). More recently, Le Barbanchon et al. (2021) used data on the joint distribution of reservation job attributes and realized job bundles to discipline this trade-off across genders. In my case, I regress bilateral commute probabilities to commute times, allowing me to control for both origin and destination fixed effects. I do this for single men and women and married men and women. While the trade-off studied by the literature on compensating differentials is not the same object as the one I estimate, it is related because if you can commute longer distances, it is more likely that you will be able to find a location featuring higher wages.

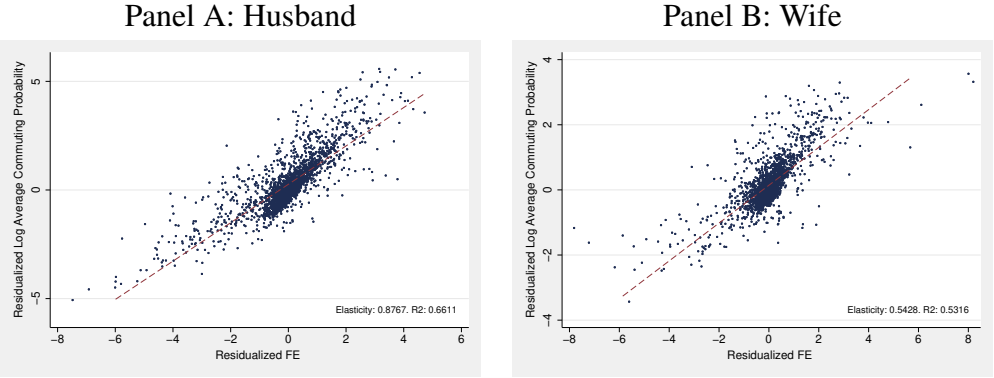
Fact 2: Choice of where to work depends on spouse's choice. In dual-earner households, the choice of where to work depends on the spouse's destination choice. To see why, first, note that if the household cares about some weighted average of the members' utility, the household will dislike members commuting to locations farther away. Second, conditioning on the first spouse's destination, the utility that the household derives from sending the other spouse to a destination depends on the income the first spouse is already bringing to the household. For example, if one of the spouses is already commuting to a high wage location, increasing household income, the other spouse's supply may be less sensitive to further increases in her own wage. This is akin to a result in labor economics where an income effect reduces the labor supply elasticity (Blundell et al., 2007).

I can test whether commuting probabilities of one spouse depend on those of the other spouse by using data about commuting choices in dual-earner households. Since I observe the complete household's roster, this allows me to compute *conditional* commuting probabilities. I denote the probability of member g of working in location j given that the spouse commutes to j' and lives in i as $\pi_{j|ij',g}$. I compute commuting probabilities at the municipality level. Given the city has 50 municipalitys, this object has $50 \times 50 \times 50$ cells. I estimate the following regression for each member in dual-earner households:

$$\log \pi_{j|ij',g} = \beta time_{ij} + FE_{i,j',g} + FE_{j,j',g} + \varepsilon_{ijj',g} \tag{2.2}$$

where $FE_{i,j',g}$ are a set of origin-spouse's destination fixed effects, which control for the fact that certain origin-spouse's destination pairs may provide greater access to labor markets to member g. Special focus should be given to $FE_{j,j',g}$, which are a set of destination pairs fixed effects controlling for the fact that certain destination pairs might be more desirable than others.

Figure 2.1: Interdependence in Commuting Choices within Dual-earner Households



Notes: I estimate $\log \pi_{j|ij',g} = \beta time_{ij} + FE_{i,j',g} + FE_{j,j',g} + \varepsilon_{ijj',g}$ for males and females in dual-earner households. Then, I recover the estimates of $FE_{j,j',g}$. I also control for the time it takes to go from workplace in j to workplace in j' to take into account that couples may decide destinations partially based on the distance between jobs. Finally, I estimate the coefficient ψ from $\log \pi_{j|j',g} = \psi_g \hat{FE}_{j,j'} + FE_j + FE_{j'}$. This figure depicts a scatter plot between residualized commuting probabilities and the residuals of $FE_{j,j'}$. Panel A depicts this plot for husbands while Panel B, for wives.

My procedure goes as follows. First, I estimate the previous equation. Then, I recover the estimates of $FE_{j,j',g}$. In a world with interdependence, the interaction of destinations should be highly predictive of conditional commuting probabilities once that I control for each members' independent reasons to work in a particular destination pair. In other words, the residual of $FE_{j,j',g}$—after controlling for fixed effects at one's own location j and the spouse's location j'—should be predictive of commuting probabilities.[14] In terms of the regression, the coefficient ψ from $\log \pi_{j|j',g} = \psi_g \hat{FE}_{j,j'} + FE_j + FE_{j'}$ should be different than zero. The greater the value of ψ, the lesser that choices are independent.[15] I show that this is in fact true. Figure 2.1 depicts a scatter plot between residualized commuting probabilities and the residuals of $FE_{j,j'}$. Panel A depicts this plot for husbands while Panel B, for wives. Residualized $FE_{j,j'}$ are highly predictive of residualized commuting probabilities for both. In fact, $\hat{\psi} = 0.87$ for men, while $\hat{\psi} = 0.54$ for women. The R-squared are 0.66 and 0.53, respectively.

To sum up, commuting choices are interlinked in dual-earner households. In the structural section of the paper I quantify how important is this interdependence when evaluating the impact of new transit infrastructure.

Robustness. One could argue that this procedure is capturing anything that varies at the destination-pair level, but that is unrelated to the joint decision that partners make regarding their commuting. In Online Appendix B.3, I provide suggestive evidence of a cross-spouse income

[14] I also control for the time it takes to go from workplace in j to workplace in j' to take into account that couples may decide destinations partially based on the distance between jobs.

[15] For simplicity note that I have averaged $\pi_{ij|j',g,k}$ over origins, and so $\pi_{j|j',g}$ is the average probability than one commutes to j given that his or her spouse commutes to j'.

effect by relying on a firm-level survey with data on wages at workplaces.

Relation to literature. To the best of my knowledge, the only paper exploring how one spouse's commuting affect the other spouse's employment prospects is Carta and De Phillips (2018). They show that increasing the husband's commute time, due to office or plant relocations, decreases his wife's employment probability. The effects are stronger for couples with children and for highly educated husbands. My approach differs from theirs in some dimensions. First, rather than studying the effect of increasing commute times on the wife's *employment probability*, I am studying the relationship between commute times and spouses' *commuting destinations*. Second, to do so, I am relying on a gravity approach that draws from a spatial quantitative model that I develop later in the paper. In particular, I am relying on data about *conditional commuting probabilities* as implied by the model. The richness of my Census data allows me to compute conditional commuting probabilities. This data has never been used before due to the limited size of commonly-used commuting surveys.

2.3.3 Reduced form evidence and gender inequality

Planned stations. I perform a reduced-form analysis by exploiting the planned network of Metro lines, that is, Line 2 and Line 4 of the Metro.[16] I compare locations closer to Line 1 of the Metro or the BRT (treatment locations) against locations closer to Line 2 and Line 4 (control locations). The basic idea is that we can account for location selection effects by comparing treated locations to locations closer to planned-but-not-constructed stations.

To illustrate how I define the treatment and control groups, Figure B.4 depicts the map of the constructed and planned stations, in addition to buffers around each station. In practice, I define the treatment group as all locations that are within 1.5 kilometers of any Line 1 or BRT station. Then, I define the control group as all locations that are within 3.2 kilometers from any Line 2 or Line 4 station. To avoid contamination from the treated group, I add the restriction that all locations in the control group should also be at least 2 kilometers away from all Line 1 and BRT stations. I use a bigger radius to define the control group because some locations are lost due to contamination which hurts statistical power. However, results are robust to alternative definitions.

Empirical strategy. I utilize two data sources: the 1993, 2007 and 2017 Population Censuses, and the 2007-2017 waves of the National Household Survey.

The 2007 and 2017 Population Censuses. Let us consider the following equation:

$$\begin{aligned} \log y_{lt} &= \beta^{Metro}\left[Treat_l^{Metro} \cdot After_t\right] + \beta^{BRT}\left[Treat_l^{BRT} \cdot After_t\right] \\ &+ FE_l + FE_t + \gamma' X_{lt} + \epsilon_{it} \end{aligned} \tag{2.3}$$

[16] I leave out Line 3 because it will run over the same locations as the BRT.

where y_{lt} is an outcome constructed from Census data, $Treat_l^{Metro}$ and $Treat_l^{BRT}$ are treatment dummies taking the value of one if the location l is closest to the Metro or the BRT station, respectively. I interact these variables with an after dummy which takes the value of one if t is 2017, and zero otherwise. β measures the effect of the new transit infrastructure on locations that are closer to the stations, relative to locations that are closer to yet-to-be-constructed stations. I introduce a set of location fixed effects FE_l, in addition to year fixed effects FE_t. To control for differential trends, I include a set X_{lt} of block-level controls, i.e. the physical size of the block (i.e. the area), the elevation of the block, the median slope, and a continuous measure of distance to the closest station (either constructed or planned) each interacted with year dummies FE_t.

I argue that being closer to a station does not guarantee greater access to relevant destinations. Gains in access should be higher in the fringe of the city since these areas were more isolated at the beginning of the sample. Conversely, gains in access in locations closer to the CBD may be less pronounced since workers were already close to valuable destinations. Thus, I allow for the treatment effects to vary by the distance of households to the CBD. I construct five dummies indicating if the household is located within 0-5, 5-7.5, 7.5-10, 10-12.5 kilometers, and more than 12.5 kilometers.

The 2007-2017 waves of the National Household Survey. I use 2007-2017 rounds of the Peruvian National Household Survey, georeferenced at the block level, to precisely locate households and classify them into treatment and control groups. Then, I consider the following equation:

$$\begin{aligned}\log y_{il(i)t} &= \beta^{Metro}\left[Treat_{l(i)}^{Metro}\cdot After_t\right]+\beta^{BRT}\left[Treat_{l(i)}^{BRT}\cdot After_t\right]\\ &+ \beta_F^{Metro}\left[Treat_{l(i)}^{Metro}\cdot After_t\cdot Female_i\right]+\beta_F^{BRT}\left[Treat_{l(i)}^{BRT}\cdot After_t\cdot Female_i\right]\\ &+ +\alpha Female_i+\gamma' X_{il(i)t}+FE_{l(i)}+FE_t+\epsilon_{it} \qquad (2.4)\end{aligned}$$

where y_{it} is a individual-level outcome such as labor earnings.[17] $Treat_{l(i)}^{Metro}$ and $Treat_{l(i)}^{BRT}$ are treatment dummies taking the value of one if the location l is closest to either the Metro or the BRT station. $After_t$ is a dummy indicating if the data comes from any year after the introduction of the new infrastructure i.e. 2010 or after. X_{1it} is a set of individual-level controls comprising age, years of education, civil status, mother's language, and moving history, each interacted with a female dummy. Finally, X_{2lt} is a set of block-level controls, i.e. the area of the block, the elevation of the block, the median slope, and a continuous measure of distance to the closest station (either constructed or planned) each interacted with year dummies FE_t and a female dummy. The parameters of interest are β_F^{Metro} and β_F^{BRT}. They measure the evolution of a women relative to men, as a result of being closer to an operative station rather than a yet-to-be-constructed station.

[17] Unfortunately, the size of the sample does not allow me to have enough statistical power to explore the heterogeneity in the impact across different types of households.

A positive value means that the gap between men and women is shrinking. Once again, I construct five dummies indicating if the household is located within 0-5, 5-7.5, 7.5-10, 10-12.5 kilometers, and more than 12.5 kilometers.

Fact 3. Within the fringe, gender inequity shrank. Panel A of Figure 2.2 shows the results of estimating equation 2.3 by PPML, while allowing for the heterogeneity of the effect based on households' location. I use the ratio of dual-earner households to male breadwinner households as dependent variable. This is a measure of female labor force participation. Roughly, as one gets closer to the fringe, the effect of either the Metro or the BRT increases. For instance, in the most remote area of the city (¿12.5 km), after the introduction of the new infrastructure, the number of dual-earner households relative to male breadwinner households rose by 9.4% in locations closer to Metro stations relative to locations closer to yet-to-be-constructed stations.

Next, in Panel B I depict the estimation of equation 2.4 by PPML, and in particular, of parameters β_F^{Metro} and β_F^{BRT}. I use labor earnings as dependent variable. Overall, women's earnings increased more than those of men in locations closer to stations relative to locations closer to yet-to-be-constructed stations. Since my household survey has a limited sample size, estimation is noisy. However, patterns are strikingly similar to those in Panel A. For example, the zigzagging profile of the Metro impacts can be observed with both outcomes, even though they belong to two entirely different datasets. Furthermore, the almost monotonic impact of the BRT on Panel A is once again observed in Panel B.

In conclusion, this section shows that gender inequality tended to improve in response to transit investments, especially in remote areas. This information will be useful later, when evaluating the empirical success of my structural model.

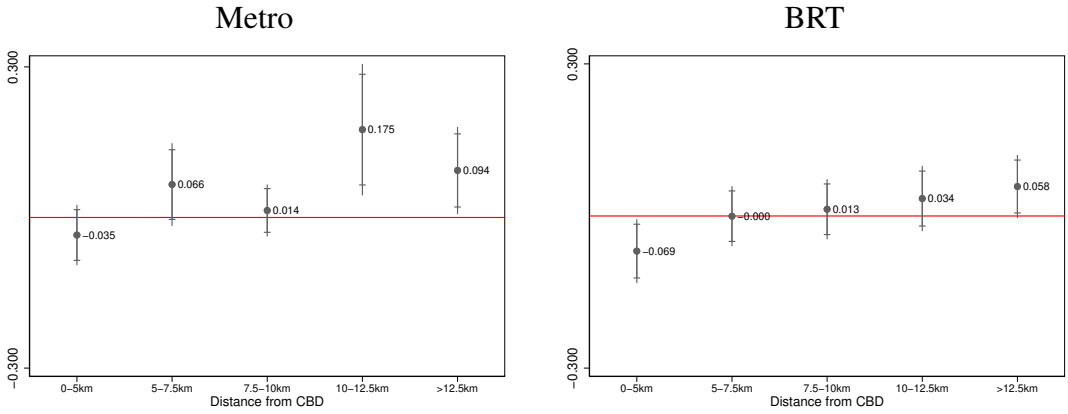
Pre-trends and other outcomes. In Online Appendix B.4, I perform a similar regression analysis exploring the impacts on other outcomes available in the population census and the household survey. Moreover, by estimating time-event studies, I show that pre-trends are parallel for a large set of outcomes, which suggests that comparing stations to yet-to-be-constructed stations is taking care of endogeneity concerns.

Relation to literature. Some have relied on quasi-experimental variation induced by new infrastructure to show that increased market access to *other cities* reduces gender inequality (Kwon, 2022). Others have found the opposite (Bütikofer, et al., 2022). Relative to these papers, I exploit variation in market access induced by new transit infrastructure *within the same city*. I argue that commutes within the same city are more representative of a typical commute, as most people in the world live in big cities, where work trips tend to happen within the same metropolitan area. Note that papers that rely on within-city variation induced by new transit infrastructure focus on different outcomes.[18]

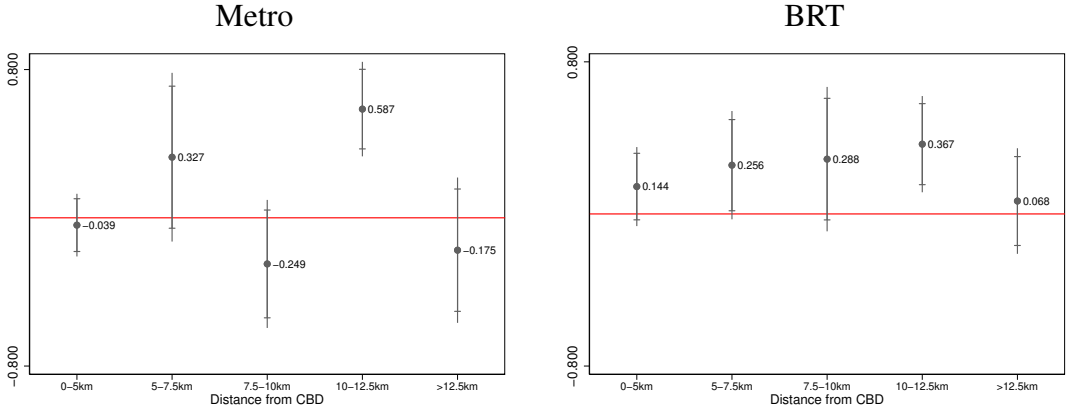
[18]See Tsivanidis, 2023; Warnes, 2021, Severen, 2021; Zarate, 2021; Khanna et al., 2022.

Figure 2.2: Reduced Form Impact of Transit Infrastructure by Distance to CBD

Panel A: Dual-Earners to Male Breadwinner



Panel B: Female Labor Earnings relative to Male Labor Earnings



Notes: For Panel A, I interact treatment dummies in equation 2.3 with a complete set of dummies indicating the distance to the CBD. Then, I plot β^{Metro} and β^{BRT} by distance to CBD. For Panel B, I interact treatment dummies in equation 2.4 with distance dummies. I also add the interaction of distance dummies and the female dummy . Then, I plot β_F^{Metro} and β_F^{BRT} by distance to CBD. 5% and 10% confidence intervals depicted.

2.4 Theory

To build intuition, I set up a simple model of commuting only with married households. Then, I study the model's implications. In Section 2.5, I embed this choice structure into a spatial model of city structure that accounts for general equilibrium.

2.4.1 Microfoundation of a household utility for the married

Environment. The city consists of N married households distributed in J locations. Locations are indexed by i, j or j'. Spouses are indexed by k or k' (or h and w when indicated).

Preferences for each spouse are defined by V_{ij}^k, which is the utility of spouse k living in i and working in j, such that:

$$V_{ij}^k = \log \frac{C_{k,i}^{\beta^k} H_{R_i}^{1-\beta^k}}{d_{ij}^k} \tag{2.5}$$

where $C_{k,i}$ is the consumption of family member k of the final good, H_{R_i} is the household consumption of residential floorspace. Note that residential floorspace is a public good within the couple, which explains why I drop the k subscript. d_{ij}^k is an iceberg commuting cost $d_{ij}^k = \exp(\kappa_k t_{ij})$. t_{ij} is the time it takes to commute to j from i, and κ_k is the rate at which commute times are transformed into commute costs. The household budget constraint is: $PC_{k,i} + PC_{k',i} + r_{R_i} H_{R_i} = w_{k,j} + w_{k',j'}$, where P is the price of the final good, which does not depend on the origin location, because in the quantitative exercise I assume that the final good is freely traded within the city. r_{R_i} is the rental value of the residential floorspace. Finally, $w_{k,j}$ is the wage received by member k when working at j, whereas $w_{k',j'}$ is the wage received by the other member at j'.

Optimal consumption of the final good and floorspace. Following the literature on family economics, I assume households only choose efficient outcomes (Browning et al., 2011). That is, conditional on residential and workplace locations, for any given vector of prices and wages $(P, r_{R_i}, w_{k,j}, w_{k',j'})$ and location characteristics $(d_{ij}^k, d_{ij'}^{k'})$, an allocation $(C_{k,i}, H_{R_i})$ is efficient if there exists a feasible $\bar{V_{ij}}^{k'}(P, r_{R_i}, w_{k,j}, w_{k',j'})$ such that $(C_{k,i}, H_{R_i})$ solves the problem:

$$\begin{aligned} \max_{C_{k,i}, H_{R_i}} \quad & V_{ij'}^k \\ \text{s.t.} \quad & PC_{k,i} + PC_{k',i} + r_{R_i} H_{R_i} = w_{k,j} + w_{k',j'} \\ \text{and} \quad & V_{ij}^{k'} \geq \bar{V_{ij}}^{k'}(P_i, r_{R_i}, w_{k,j}, w_{k',j'}) \end{aligned}$$

This maximization problem can be broken into two stages. In the first stage, households choose the optimal level of residential floorspace H_{R_i} and the disposable income allocated to each spouse, $x_{k,i}, x_{k',i}$. In the second, each spouse allocates their income into the consumption of the final good. Then, the first stage becomes:

$$\begin{aligned} \max_{x_{k,i}, H_{R_i}} \quad & W_{ijj'} \equiv \lambda V_{ij'}^{k'} + V_{ij}^k \\ \text{s.t.} \quad & x_{k,i} + x_{k',i} + r_{R_i} H_{R_i} = w_{k,j} + w_{k',j'} \equiv y_{ijj'} \end{aligned}$$

The first order conditions yield two sets of equations. The first set equalizes the marginal utility derived from the consumption of one spouse to that of the other spouse: $\lambda \frac{\partial V_{ij}^{k'}}{\partial x_{k',i}} = \frac{\partial V_{ij'}^{w}}{\partial x_{k,i}}$. The second set, also known as the Bowe-Lindahl-Samuelson conditions, says that members' marginal willingness to pay should be equalized to the marginal price of residential floorspace: $\frac{\partial V_{ij}^{k}/\partial H_{R_i}}{\partial V_{ij}^{k}/\partial x_{k,i}} + \frac{\partial V_{ij'}^{k'}/\partial H_{R_i}}{\partial V_{ij'}^{k'}/\partial x_{k',i}} = r_{R_i}$.

In the quantitative exercise, I abstract away from bargaining considerations within households and only consider the feature that spouses share the same budget. For that reason, I assume that λ, typically thought as a measure of bargaining power, is exogenous and equal to one so that the household weights spouses' utility equally. Under this assumption, households behave as a single entity that makes all choices, i.e. households behave as unitary households. Moreover, I assume that spouses have the same preferences over consumption and housing and so $\beta^k = \beta^{k'} = \beta$. These assumptions and the first order conditions yield the following solutions to the maximization problem: $x_{ijj'}^{k} = x_{ijj'}^{k'} = \frac{\beta}{2} y_{ijj'}$ and $r_{R_i} H_{R_i} = (1-\beta) y_{ijj'}$.

Household's indirect utility. Plugging solutions back into the household utility function, I get that:[19]

$$\tilde{W}_{ijj'} \propto \left(d_{ij}^{k} d_{ij'}^{k'}\right)^{-1} \left(\frac{w_{k,j} + w_{k',j'}}{P^{\beta} r_{R_i}^{(1-\beta)}}\right) \tag{2.6}$$

This household utility postulates that the household dislikes the commuting performed by both members, which is a natural outcome from the fact that the household cares about the average utility of members. Moreover, greater income net of prices provides greater utility.

2.4.2 Implications

Marginal rates of substitution. In this section, I explain what are the implications of this choice structure. Specifically, I ask two questions. First, say that the times from origin i to destination j' improve for the wife by one minute. How much can the husband's commuting times increase while keeping household utility $\tilde{W}_{ijj'}$ constant? Second, keeping husband's commuting times constant, what is the wage of the wife that the household would be willing to give up in order to improve her commute time by one minute?

To answer the first question, I take a total derivative of the utility function with respect to commute times of spouses and equalize it to zero. This yields $dt_{ij}^{h} = -\frac{\kappa_h}{\kappa_w} dt_{ij}^{w}$ where h and w index the husband and the wife. This result means that the household can substitute the time

[19] I have taken the exponential function to both sides, taken the root square and redefined the meaning of κ^k. These preferences should represent the same preferences ordering since all operations were monotonic transformations.

freed up from the wife's commuting with $\frac{\kappa_h}{\kappa_w} dt^w_{ij'}$ minutes in husband's commuting. $\frac{\kappa_h}{\kappa_w}$ acts as an exchange rate at which the time of the wife can be substituted with the time of the husband.

To answer the second question, I follow a similar strategy. To decrease the commute times of the wife by one minute, the household would be willing to give up income $w_{h,j} + w_{w,j'}$ by κ_w log points: $d\log(w_{h,j} + w_{w,j'}) = -\kappa_w dt^w_{ij'}$. This is a feature also present in the class of spatial models with iceberg costs and constant returns to scale utility functions. The difference in this setting is that I generalize it: income depends on the wage of the two spouses, and so each spouse's commuting is responsible of sourcing a share of the household's income. Noting that $d\log(w_{h,j} + w_{w,j'}) = \frac{1}{w_{h,j}+w_{w,j'}} dw_{w,j'}$, then $\frac{dw_{w',j}}{w_{w,j'}} = -\kappa_w \frac{w_{h,j}+w_{w,j'}}{w_{w,j'}} dt^w_{ij'}$. Thus, as $w_{h,j}$ increases, households' willingness to give up the wife's wage in exchange of reduced commute times rises.

Gravity equations. I now assume that there is a continuum of households indexed by ω. Moreover, I assume that individuals' utility function (equation 2.5) is heterogeneous over their commute destinations, $\epsilon_{k,j}(\omega)$, and $\epsilon_{k',j'}(\omega)$. Solving the maximization problem yields:

$$W_{ijj'}(\omega) \propto \frac{\epsilon_{k,j}(\omega)\epsilon_{k',j'}(\omega)}{d^k_{ij} d^{k'}_{ij'}} \left(\frac{w_{k,j} + w_{k',j'}}{P^\beta r^{(1-\beta)}_{R_i}} \right) \tag{2.7}$$

I assume preference shocks are Frechet distributed with shape parameters θ_k and $\theta_{k'}$. Thus, conditional on spouse k''s destination j', the household evaluates the utility it would get from sending k to each possible destination and chooses the one that maximizes it. Similarly, conditional on spouse k's destination j, the household evaluates all possible destinations of member k'. Then, I get that:

$$\pi^k_{j|ij'} = \frac{W^{\theta_k}_{ijj'}}{\sum_l W^{\theta_k}_{ilj'}} \quad \text{and} \quad \pi^{k'}_{j'|ij} = \frac{W^{\theta_{k'}}_{ijj'}}{\sum_l W^{\theta_{k'}}_{ijl}}$$

The first expression says that the share of spouses k commuting to location j conditional on their residence i and their spouses commuting to j' depends on total income $w_{h,j} + w_{w,j'}$ net of commute costs $d^k_{ij} d^{k'}_{ij'}$, relative to available options $\sum_l W^{\theta_k}_{ilj'}$. For simplicity in this exposition, I assume that $\theta_k = \theta_{k'} = \theta$. In the quantitative exercise, I allow for the shape parameters to vary across household members which is consistent with Fact 1. Finally, unconditional commuting probabilities are:

$$\pi^k_{j|i} = \frac{\sum_l W^\theta_{ijl}}{\sum_{l'} \sum_l W^\theta_{ill'}} \quad \text{and} \quad \pi^{k'}_{j'|i} = \frac{\sum_l W^\theta_{ilj}}{\sum_{l'} \sum_l W^\theta_{ill'}}$$

Comparative statics. I compute the impact of improving the commuting times of spouse k' on the commuting probabilities of spouse k. First, note that improving k''s commuting times does not

affect k's conditional commuting probabilities. This is because when k''s commuting times to j' improve, the utility received in every destination of k changes proportionally to the initial utility: $\frac{dW_{ijj'}}{dt_{ij'}^{k'}} = -\kappa_{k'} W_{ijj'}, \forall j$.

However, k's unconditional probabilities are affected as the relative ranking of utility $W_{ijj'}$ across destinations j and j' is reordered thanks to the improvement in commuting times: $\pi_{j|i}^{k} = \frac{\sum_l (W_{ijl}+1[l=j']\kappa_{k'}W_{ijj'})^{\theta}}{\sum_{l'}\sum_l (W_{ill'}+1[l'=j']\kappa_{k'}W_{ilj'})^{\theta}}$. This is as if households, by having more time at their disposal, reallocate spouses across destinations such that utility is maximized. Commuting to location j will increase if the resulting numerator $\sum_l \left(W_{ijl} + 1[l = j']\kappa_{k'} W_{ijj'}\right)^{\theta}$ is greater than the numerator of the remaining options. Intuitively, this will happen when the location j features high wages or low commute costs, as the impact on $W_{ijj'}$ is proportional. If wages in j are too low or commute costs to j too high, utility maximization leads households to decrease the probability at which they send spouses k to such destination.

Mathematically, how do k's unconditional probabilities change when k''s times improve? Taking a derivative with respect to $t_{ij'}^{k'}$, I get:

$$\frac{\partial \pi_{j|i}^{k}}{\partial t_{ij'}^{k'}} = -\theta\kappa_{k'}\pi_{j'|i}^{k'}\left(\pi_{j|ij'}^{k} - \pi_{j|i}^{k}\right) = -\theta\kappa_{k'}\pi_{j|i}^{k}\left(\pi_{j'|ij}^{k'} - \pi_{j'|i}^{k'}\right) \tag{2.8}$$

Spouses k will increase their probability of commuting to location j if the probability that spouses k use the route j given that their spouses k' use the improved route j' is higher than the unconditional probability that spouses k use the route j. In other words, the probability that k commutes to j will increase when initial conditional commuting probabilities $\pi_{j|ij'}^{k} = (\frac{w_{k,j}+w_{k',j'}}{d_{ij}^{k}})^{\theta} / \sum_l (\frac{w_{k,l}+w_{k',j'}}{d_{il}^{k}})^{\theta}$ are high enough.

How improvements in commuting time affect income? Defining spouses k's income at origin i as the sum of wages across destinations $y_i^k = \sum_l w_{k,l}\pi_{k,l|i}N_i^k$, where N_i^k is the number of spouses k living in origin i. Differentiating it with respect to commute times between origin i and a destination j' of both spouses k and k', I get:

$$-\frac{\partial}{\partial t_{ij'}}\log y_i^k = \underbrace{\theta\kappa_k \frac{w_{k,j'}\pi_{j'|i}^{k}}{\sum_l w_{k,l}\pi_{l|i}^{k}}\left(1 - \frac{\bar{y}_i^k}{w_{k,j'}}\right)}_{\text{direct use}} + \underbrace{\theta\kappa_{k'}\sum_l \frac{w_{k,l}\pi_{l|i}^{k}}{\sum_l w_{k,l}\pi_{l|i}^{k}}\left(\pi_{j'|il}^{k'} - \pi_{j'|i}^{k'}\right)}_{\text{indirect use}} \tag{2.9}$$

The effect on total income in origin i for spouse k depends on two terms: the direct use channel, and the indirect use channel (or within household spillovers). The direct use channel will be positive if $w_{k,j'} > \bar{y}_i^k$, where $\bar{y}_i^k = \sum_j \left(w_{k,j}\pi_{j|i}^{k}\right)$. In other words, when times between origin i and destination j' improve, aggregate income in origin i increases if wages at the destination that is now more accessible are higher than the average income in origin i. The intuition is that

income should increase only if the new destination j' is a high-wage location. Otherwise, if the new destination is a low-wage location, improvements in commuting times may induce some workers to give up some of their earnings to have a shorter commute.[20] How much would it increase depends on three terms. First, it depends on the share of income that destination j' accounts for, $\frac{w_{k,j'}\pi^k_{j'|i}}{\sum_j\left(w_{k,j}\pi^k_{j|i}\right)}$. Second, it depends on the difference between wages $w_{k,j'}$ and average income $\bar{y}^k_i$. Finally, it depends on the commuting elasticity $\theta\kappa^k$.

The indirect use channel indicates that when commute costs for spouses k' decrease, this affects the households' allocation of spouses k across destinations as explained above. When this happens, the origin i stops receiving the income from locations where the spouse k is no longer working (i.e. when $\pi^{k'}_{j'|il} < \pi^{k'}_{j'|i}$), but starts receiving the income from locations where spouses k are commuting to (when $\pi^{k'}_{j'|il} > \pi^{k'}_{j'|i}$). We need to sum across all locations l to account for all spillovers generated by spouses k' to spouses k working in locations l. Finally, note that if choices are independent, as in standard spatial models, conditional commuting probabilities become equal to unconditional probabilities, collapsing the indirect channel to zero.

In this section I have shown the microfoundation of a utility function tailored for married households. I have explained the major implications of this choice structure in terms of commuting and have shown how improvements in commute times may impact income across origins.

2.5 Bridging Theory and Data

This section starts by summarizing the main components of the quantitative model. Afterwards, it describes the calibration and estimation procedures. Finally, it shows the model's performance when matching untargeted moments.

2.5.1 Quantitative model of city structure and couples

2.5.1.1 Environment

This is a model in which the internal structure of the city is driven by a tension between agglomeration forces (production externalities) and dispersion forces (commuting costs and an inelastic supply of land). The city consists of a set of discrete locations indexed by $i = 1, ..., S$. These locations differ in their commute times to other locations, their housing floorspace, households' productivity (in household production), amenities, and industries' productivity as I explain later.

[20] Note, however, that new metro lines improve commuting times between a set of origin and destination locations. The derivative I have taken only considers an improvement between a particular origin and a particular destination, assuming all other times remain the same which is physically impossible. If if its quicker to reach j' from i, it will also be quicker to reach a neighboring location j''.

Moreover, the city is populated by a fixed measure of households N_k, where $k \in \{m, f, mf\}$. If $k = m$ or $k = f$ the household is comprised of a single male worker or a single female worker, respectively. If $k = mf$ the household is a couple. For simplicity, singles and husbands always work in this model.[21]

Furthermore, each location houses multiple industries which produce a consumption good using labor and commercial floorspace. Each location-by-industry's consumption good is aggregated into a final good with CES preferences. Males and females are imperfect substitutes in the production function. Landowners choose how to allocate a fixed amount of floorspace across residential and commercial use.

2.5.1.2 Households

Single households

A. Preferences in single households. Singles $g \in \{m, f\}$ derive Cobb-Douglas utility from the consumption of a freely traded numeraire good ($C^g(\omega)$) with parameter β^g, from the consumption of residential floorspace $H_{R^g}(\omega)$ with parameter $1 - \beta^g$, and from an amenity accounting for the average preference of each group to live in location i (u_i^g). Moreover, individuals experience disutility from commuting in an iceberg fashion, $d_{ij}^g \geq 1$, where $d_{ij}^g = \exp(\kappa^g t_{ij})$, and t_{ij} is the time it takes to commute between locations i and j. The parameter κ^g governs the size of these commuting costs for each gender. Finally, workers are heterogeneous in their preferences for working in location j, $\epsilon_j^g(\omega)$, and for living in residence i, $\varphi_i^g(\omega)$. Hence, singles need to decide where to work based on the trade-off between higher wages but longer commutes. Similarly, they need to decide where to live by balancing out the expected income against housing costs and amenities. Concretely, indirect utility is given by:

$$V_{ij}^g(\omega) \quad \propto \quad u_i^g \varphi_i^g(\omega) P^{-\beta^g} r_{R_i}^{\beta^g - 1} \left[\frac{w_{jg} \epsilon_j^g(\omega)}{d_{ij}^g} \right] \tag{2.10}$$

where P is the price of the freely traded final good, and r_{R_i} is the rental value of residential floorspace. Panel A of Figure B.5 depicts the timing assumptions.

B. Commuting. Conditional on where to live, singles (indexed by ω) draw a vector of location-specific preferences for work across the city, $\epsilon_j^g(\omega)$, iid from a Frechet distribution $F(\epsilon_j) =$

[21] To simplify, in this paper couples are comprised of a male and a female. Another simplifying assumption usually done in the literature on family economics is that the husband always works. In my data, female breadwinner households constitute just 1.3% of all households. This is why I turn off this margin: it does not seem to be quantitatively important but would add greater complexity to the model.

$\exp\left(-T^g \epsilon_j^{-\theta^g}\right)$. From standard Frechet operations, we know that:

$$\pi^g_{j|i} = \frac{\left(\frac{w_{jg}}{d^g_{ij}}\right)^{\theta^g}}{\Psi_{R^g_i}} \tag{2.11}$$

where the probability singles g choose destination j given origin i is: $\pi^g_{j|i}$. Moreover, $\Psi_{R^g_i} = \sum_{l=1}^{S}\left(\frac{w_{lg}}{d^g_{ij}}\right)^{\theta^g}$ is a market access term indicating how close are jobs to people. Thus, one key trade-off is that individuals are attracted to locations paying a high wage but dislike commuting long distances. That is, they compare the net wage they would get from commuting to particular destination with other available options embodied in $\Psi_{R^g_i}$. The dispersion of preferences is specific for each gender, and it determines how sensitive singles are to changes in commute costs. For example, when the dispersion of preferences shocks is low (i.e. high θ^g), choices are more sensitive to commute costs.

C. Residence Location. Before singles draw a vector of location-specific preferences, they draw a value for residence preferences, $\varphi^g_i(\omega)$, and decide where to live. $\varphi^g_i(\omega)$ is distributed Frechet with shape parameter $\eta^g > 1$ and average E^g. So, singles choose to live in i only if they attain the highest expected utility across all locations. Then, the probability that a single lives in location i is:

$$\pi^g_i = \left(\frac{r^{1-\beta^g}_{R_i}}{u^g_i \Psi^{1/\theta^g}_{R^g_i}}\right)^{-\eta^g}\left(\sum_l \left(\frac{r^{1-\beta^g}_{R_l}}{u^g_l \Psi^{1/\theta^g}_{R^g_l}}\right)^{-\eta^g}\right)^{-1} \tag{2.12}$$

This expression says that, when deciding where to live, singles balance out locations with better amenities and greater access to jobs against the cost of housing.

Married households

A. Preferences in married households. Individuals in married households derive utility from the same components as singles. Thus, preferences for married households can be microfounded following a similar procedure as in Section 2.4.1. Note that I denote husbands and wives using the subscripts h and w, respectively.

I additionally incorporate the labor participation choice of women. To do so, I assume that households can engage in household production with location-specific productivity $\xi^{mf}_{i\ell}$, which depends on whether both spouses work ($\ell = 1$) or just one ($\ell = 0$).[22] There could also be id-

[22]This term represents location-specific amenities that increases productivity of household production. For example, if household production is related to raising kids, one could think that it is easier to raise kids in an environment with low crime rates, and with availability of good schools and childcare.

iosyncratic reasons why a household prefers to have both spouses working in the labor market, which are embodied in the preference shock $\alpha_{\ell}^{mf}(\omega)$. I further assume that males and females have the same preference over residential location φ_i^{mf} and derive the same utility from local amenities u_i^{mf}. They only differ on their preferences over destinations for work, ϵ_j^k, and commuting time d_{ij}^k. Thus, indirect utility of a couple ω living in i, with the husband working at j and the wife in j' (if she works) is given by:

$$W_{i\ell jj'}^{mf}(\omega) \propto u_i^{mf}\varphi_i^{mf}(\omega)\,\xi_{i\ell}^{mf}\alpha_{\ell}^{mf}(\omega)\,\frac{\epsilon_j^h(\omega)}{d_{ij}^h}\left(\frac{\epsilon_{j'}^w(\omega)}{d_{ij'}^w}\right)^{\ell} P^{-\beta^{mf}} r_{R_i}^{\beta^{mf}-1} y_{\ell jj'}^{mf}(\omega) \quad (2.13)$$

where household ω's income is $y_{\ell jj'}^{mf}(\omega) = w_{jh} + w_{j'w}\ell(\omega)$. Notice that a couple ω performs the same choices as singles. Additionally, they have to choose whether the wife is going to work ($\ell = 1$ if she works, $\ell = 0$ if she stays at home), and the location where she is going to work (indexed by j'). Thus, married couples face additional trade-offs. First, if the wife decides to work, households will dispose of more income to spend in housing and the final good. However, if the wife stays, households will use her unit of labor to produce a household good with location-specific productivity $\xi_{i\ell=0}^{mf}$. Second, there is a potential trade-off between commute time and income, as explained in the previous section. Panel B of Figure B.5 depicts the timing assumptions.

B. Husbands' Commuting Choice in Dual-Earner Households. I assume for simplicity that households choose the husband's workplace after knowing the workplace of the wife. Households compare the utility they would receive if the husband works in a particular location, against the utility they would derive if he works in some other location, conditionally on where the wife works. Hence, I compute the *conditional* probability that married men living i commute to j given that their wives work and commute to j', i.e. $\pi_{j|i,\ell=1,j'}^h = P\left[W_{i,\ell=1,j,j'}(\omega) \geq \max_{l\neq j} W_{i,\ell=1,l,j'}(\omega)\right]$. Simple manipulations result in the following expression:

$$\pi_{j|i,\ell=1,j'}^h = \frac{\left(\frac{w_{jh}+w_{j'w}}{d_{ij}^h}\right)^{\theta^h}}{\Phi_{R_{i,\ell=1,j'}}^h} \quad (2.14)$$

where $\Phi_{R_{i,\ell=1,j'}}^h = \sum_l \left(\frac{w_{lh}+w_{j'w}}{d_{il}^h}\right)^{\theta^h}$ is a market access term for husbands. It indicates how close are jobs to husbands but conditional on the wife's wage. It can also be thought as the expected utility of households when the wife works at j', but before households have decided where to send the husband. As in single households, married men are more likely to commute to a location when it pays a high income net of commute cost. However, now income includes the wage earned by the wife when she works at j'.

C. Wives' Commuting Choice in Dual-Earner Households. To decide where the wife works, the household compares the expected utility it would receive across all destinations. So,

$$\pi^{w}_{j'|i,\ell=1} = \frac{\left(\frac{\left(\Phi^{h}_{R_{i,\ell=1,j'}}\right)^{1/\theta^h}}{d^{w}_{ij'}}\right)^{\theta^w}}{\Phi^{w}_{R_{i,\ell=1}}} \tag{2.15}$$

where $\Phi^{w}_{R_{i,\ell=1}} = \sum_l \left(\frac{\left(\Phi^{h}_{R_{i,\ell=1,l}}\right)^{1/\theta^h}}{d^{w}_{il}}\right)^{\theta^w}$ is a market access term for dual-earner households. These households can have greater access to jobs either through the husband or the wife, and so $\Phi^{w}_{R_{i,\ell=1}}$ acts as an average of the two. This equation says that wives will commute to locations where they expect a higher household income, $\Phi_{R^{h}_{i,\ell=1,j'}}$, net of their own disutility of commuting, $d^{w}_{ij'}$.

D. Husbands' Commuting Choice in Male Breadwinner Households. In the case of male breadwinner households, a similar logic as in the case of single households delivers the following expression for commuting probabilities:

$$\pi^{h}_{j|i,\ell=0} = \frac{\left(\frac{w_{jh}}{d^{h}_{ij}}\right)^{\theta^h}}{\Phi^{h}_{R_{i,\ell=0}}} \tag{2.16}$$

So, in terms of commuting, married males in breadwinner households and singles behave similarly, although with potentially different commuting elasticities.

E. Labor Force Participation. Conditional on their residence, households compare the expected utility they would obtain if the wife works or if she stays at home performing household production. Since households are heterogeneous in their preference for a stay-at-home wife, and assuming this heterogeneity is distributed Frechet with shape parameter ν, the probability that a couple decides to be a dual-earner household is:

$$\mu_i = \frac{\left[\left(\Phi^{w}_{R_{i,\ell=1}}\right)^{1/\theta^w}\right]^{\nu}}{\Psi^{mf}_{R_i}} \tag{2.17}$$

where $\Psi^{mf}_{R_i} = \left[\left(\Phi^{w}_{R_{i,\ell=1}}\right)^{1/\theta^w}\right]^{\nu} + \left[\xi_{i\ell=0}\left(\Phi^{h}_{R_{i,\ell=0}}\right)^{1/\theta^h}\right]^{\nu}$. Household productivity in domestic goods $\xi_{i\ell=1}$ is normalized to one, and hence $\xi^{mf}_{i\ell=0}$ indicates how much more productive are households on domestic production in location i when the wife stays at home relative to the situation in which she

participates in the labor market.[23] This expression says that if both spouses have opportunities to work by having access to nearby jobs, then the likelihood of both working increases relative to the situation where only males have greater access to jobs.

F. Residence Location. Couples choose to live in i only if they attain a higher expected utility than in all other locations. Then the probability household ω in location i is:

$$\pi_i^{mf} = \left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_i^{mf} \Psi_{R_i^{mf}}^{1/\nu}} \right)^{-\eta^{mf}} \left(\sum_l \left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_l^{mf} \Psi_{R_l^{mf}}^{1/\nu}} \right)^{-\eta^{mf}} \right)^{-1} \tag{2.18}$$

Hence, not only the location elasticity η is different in single households than in married couples, but also the married have other considerations when deciding where to live. That is, they consider locations where on average both spouses have better labor opportunities. If only males have good opportunities then they look for locations with greater household productivity. Thereby, this directly links female's and male's market conditions.

2.5.1.3 Production

I model the production side using standard assumptions in the literature. Following an Armington assumption, there are $s \in 1, ..., K$ industries which produce varieties differentiated by location under perfect competition. These goods are aggregated with CES preferences by consumers with an elasticity of substitution of $\sigma_D > 1$. Following Tsivanidis (2023), firms produce using a Cobb-Douglas technology over labor and commercial floorspace: $Y_{js} = A_{js} N_{js}^{\alpha_s} H_{F_{js}}^{1-\alpha_s}$ where $N_{js} = \left(\sum_g \alpha_{sg} \tilde{L}_{F_{jgs}}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$. Industries aggregate workers of different genders using a CES aggregator with an elasticity of substitution of σ.[24] Note that the total labor share is the sum of each gender's labor share: $\alpha_s = \sum_g \alpha_{sg}$. Finally, A_{js} is the productivity of location j for firms in industry s. Importantly, industries differ in the intensity in which they use different type of workers α_{sg}.

2.5.1.4 Housing

Market clearing

Each location is characterized by a fixed amount of floorspace H_i. A fraction ϑ_i is allocated to residential use and $1 - \vartheta_i$ to commercial use. Market clearing requires the supply of residential

[23] Since $\left(T^h\right)^{1/\theta^h} \Gamma\left(\frac{\theta^h - 1}{\theta^h}\right)$ only affects the scale of welfare I have normalized this constant to one hereafter.

[24] Hence, in equilibrium, a gender gap between males and females wages may arise in each location. Also, I assume that single and married men are perfect substitutes. Single and married women are perfect substitutes too. Hence, conditional on a particular destination, married and single individuals of a particular gender will receive the same wage. However, income may vary across these groups because they may commute to different destinations.

floorspace, $\vartheta_i H_i r_{R_i}$, to equal demand, $H_{R_i} = \sum_k H_{R_i^k}$, where k indexes household groups. This is the same for commercial floorspace. The supply of commercial floorspace, $(1-\vartheta_j)H_j r_{F_j}$, must equal the demand from the production-side, $H_{F_j} = \sum_s H_{F_{js}}$.[25]

Floorspace use allocation.

As in Ahlfeldt et al. (2015) landowners allocate floorspace to its most profitable use. When convenient they allocate ϑ_i to residential. They receive r_{R_i} per unit. For commercial use, land regulations limit the return to each unit, hence they receive $(1-\tau_i)r_{F_i}$. Then:

$$\begin{aligned} \vartheta_i = 1, \quad & \text{if } r_{R_i} > (1-\tau_i)r_{F_i} \\ (1-\tau_i)r_{F_i} = r_{R_i}, \quad & \forall \{i : \vartheta_i \in (0,1)\} \\ \vartheta_i = 0, \quad & \text{if } r_{R_i} < (1-\tau_i)r_{F_i} \end{aligned} \tag{2.19}$$

2.5.1.5 Agglomeration spillovers

Previous literature has found that agglomeration spillovers might shape the aggregate impact of new transit infrastructure (Tsivanidis, 2023). Productivity in each location depends on an exogenous component $\bar{A}_{js}$ that reflects the location's fundamentals (such as slope of the land, access to roads, etc.), and the endogenous employment in industry s in that location. I model endogenous productivity using a constant scale elasticity varying across sectors:

$$A_{js} \quad = \bar{A}_{js} \left(\tilde{L}_{F_{js}} \right)^{\phi_{F_s}} \tag{2.20}$$

where ϕ_{F_s} measures the overall effect of agglomeration forces in sector s.

2.5.1.6 Definition of Equilibrium

Definition 2. Given the model's parameters $\left\{\beta^k, \kappa^k, \theta^k, T^k, E^g, \eta^k, \alpha_{sg}, \sigma, \sigma_D, \phi_{Fs}, , \nu\right\}$, city population by household-group $\left\{N^k\right\}$, and exogenous location-specific characteristics $\left\{H_i, \xi_{i\ell}, \bar{u}_i^k, \bar{A}_{js}, t_{ij}, \tau_i\right\}$, the general equilibrium of the model is given by the vector $\left\{N_{Ri}^k, \mu_{i\ell}, L_{F_{jg}}, w_{jg}, r_{R_i}, r_{F_j}, \vartheta_i, \right\}$ such that labor markets clear, floorspace markets clear, and the population adds up to the city total.

The solution algorithm is in section B.8 of the online version of the paper.

[25] To avoid the inefficiency generated by absentee landlords, I am assuming that landlords spend all their income in consuming the final good.

2.5.2 Calibration and estimation from aggregate Data

I estimate industry-gender labor input shares and housing expenditures from surveys (Panel A of Table 2.2), and calibrate the production function's elasticity of substitution across genders, demand elasticity of substitution, and industry labor input share (Panel B).

2.5.3 Estimation outside the model

By taking logs to commuting probabilities in single and breadwinner households I can estimate the parameters of interest by PPML. From equations 2.11 and 2.16 I get the following:

$$\begin{aligned}
\log \frac{L_{R_{ij}^g}}{N_{R_i^g}} &= -\theta^g \kappa^g \cdot t_{ij} + \underbrace{\theta^g \cdot \log\left(w_{jg}\right)}_{\text{FE}_j} - \underbrace{\log \Psi_{R_i^g}}_{\text{FE}_i} \\
\log \frac{L_{R_{ij,\ell=0}^h}}{N_{R_{i,0}^{mf}}} &= -\theta^k \kappa^k \cdot t_{ij} + \underbrace{\theta^h \cdot \log\left(w_{jh}\right)}_{\text{FE}_j} - \underbrace{\log \Phi^h_{R_{i,\ell=0}}}_{\text{FE}_i}
\end{aligned}$$

which are exactly the same equations I estimate for Fact 1 reported in columns 1 to 4 of Table 2.1. I plug into the model the commuting elasticities I estimate for singles. For married households, I plug in the commuting elasticities of breadwinner households.[26] The underlying assumption is that among married of a given gender there is no heterogeneity in the shape parameter associated to their commuting preference shocks.[27] Finally, to get an estimate for the Frechet shape parameters in the commuting preferences, I follow the growing consensus in the literature (Ahlfeldt et al., 2015; Tsivanidis, 2023; Zarate, 2021) and set the iceberg disutility parameter κ^k to $\kappa^k = 0.01$, from which I get estimates of $\hat{\theta}^m = 5.14$, $\hat{\theta}^f = 6.17$, $\hat{\theta}^h = 4.66$, and $\hat{\theta}^w = 6.34$, as shown in Panel C of Table 2.2.

2.5.4 Estimation within the model

Model inversion. I recover unobservables such as location productivities, household productivities and amenities that rationalize the observed data as a model equilibrium. From the production-side I need to observe workplace employment levels per industry. Intuitively, given male and female intensities per industry, greater employment in a sector is informative of employment by gender. The propositions in the Appendix B.9 formalize these ideas.

[26] While in the model I am not considering female breadwinner households, I can still use data on them to estimate the commuting elasticity of married women.

[27] I make this assumption because equation 2.15 implies that conditioning on origin and destination fixed effects may not be enough to consistently estimate the commuting elasticities for dual-earner households. In any case, columns 5 to 6 show very similar elasticities as those shown in column 3 and 4.

Agglomeration economies. I can recover productivity from $A_{js} = W_{js}^{\alpha_s} r_{F_j}^{1-\alpha_s} X_{js}^{1/(\sigma_D-1)} X$. Hence, productivity is the residual that ensures the definition for firm sales X_{js} to hold. This way, we know that locations characterized by high employment levels relative to the price of commercial floorspace and wages are also locations with high productivity. I recover productivity in 2007, before transit improvements, and in 2017. Dividing by the geometric mean, taking logs, and taking first differences, we can express TFP in equation 2.20 as:

$$\Delta \ln \frac{A_{js}}{\tilde{A}_s} = \phi_{F_s} \Delta \ln \frac{\tilde{L}_{F_{js}}}{\tilde{\tilde{L}}_F} + \underbrace{\Delta \ln \frac{\bar{A}_{js}}{\tilde{\tilde{A}}_s}}_{\gamma' X_{js} + \varepsilon_{js}^A}$$

where $\Delta \ln \frac{\bar{A}_{js}}{\tilde{\tilde{A}}_s}$ is approximated by a set of controls, $\gamma' X_{js}$ and an error term, ε_{js}^A.[28]

Estimating the strength of agglomeration economies is a challenging endeavor as locations may be more productive because more people work there, or because locations with high productivity attract more workers. To identify this parameter, we need to use a shock to labor supply. Therefore, we can exploit the model and the introduction of new transit infrastructure to generate such an instrument. By taking a full differentiation to $\left\{\sum_k \sum_i \pi_{j|i,\ell}^k N_{R_i^k}\right\}$ with respect to commuting times from origin locations (while keeping market access terms constant) we get the following Bartik-style instruments:

$$Z_{Fj}^m = -\theta^m \kappa^m \sum_{i'} \frac{\pi_{j|i'}^m N_{R_{i'}^m}}{\sum_i \pi_{j|i}^m N_{R_i^m}} \cdot dt_{i'j} \qquad Z_{Fj}^f = -\theta^f \kappa^f \sum_{i'} \frac{\pi_{j|i'}^f N_{R_{i'}^f}}{\sum_i \pi_{j|i}^f N_{R_i^f}} \cdot dt_{i'j}$$

$$Z_{Fj}^w = -\theta^w \kappa^w \sum_{i'} \frac{\pi_{j|i',\ell=1}^w N_{R_{i',\ell=1}^{mf}}}{\sum_i \pi_{j|i,\ell=1}^w N_{R_{i,\ell=1}^{mf}}} \cdot dt_{i'j} \qquad Z_{Fj\ell=0}^h = -\theta^h \kappa^h \sum_{i'} \frac{\pi_{j|i',\ell=0}^h N_{R_{i',\ell=0}^h}}{\sum_i \pi_{j|i,\ell=0}^h N_{R_{i,\ell=0}^h}} \cdot dt_{i'j}$$

$$Z_{Fj\ell=1}^h = -\theta^h \kappa^h \sum_{i'} \frac{\pi_{j|i',\ell=1}^h N_{R_{i',\ell=1}^{mf}}}{\sum_i \pi_{j|i,\ell=1}^h N_{R_{i,\ell=1}^{mf}}} \cdot dt_{i'j} \quad - \quad \theta^w \kappa^w \sum_{i'} \frac{\pi_{j|i',\ell=1,j}^h \pi_{j|i',\ell=1}^w N_{R_{i',\ell=1}^{mf}}}{\sum_i \pi_{j|i,\ell=1}^h N_{R_{i,\ell=1}^{mf}}} \cdot dt_{i'j}$$

For each destination j, the Bartik-instruments weigh the change in commuting times with the share of workers from each origin. I use inferred shares $\pi_{j|i}$ from 2007 data. Since I have many groups of households, this calculation is made for each, providing five possible instruments to exploit. The identifying assumption is that the growth of exogenous productivities is uncorrelated with the

[28] I control for a set of block-level covariates to account for differential growth in locations of such characteristics. In particular, I include the distance to the closest BRT station, the distance to the closest Metro station, and the distance to the Central Business District and its square. This allows me to control for anything that might have changed in places near stations but that are not related to improvement in commuting times. I also control for the physical size of the block and its square, its size in terms of the number of households in 2007, its slope, its elevation, and the number of dual-earner households as a share of married couples in 2007. Moreover, I include the average market access across household-types as measured in 2007. Finally, I control for the dependent variable at its value in 2007.

Bartik instruments conditional on controls: $\mathcal{E}\left[\varepsilon^{A}_{js} \cdot Z^{k}_{Fj\ell} | X_{js}\right] = 0$.

To construct sector-specific instruments, I use the model to compute the share of female and male workers in each location-by-industry in 2007, which I use to average the five instruments presented above in the following way: $Z_{Fjs} = \frac{L_{F_{jfs}}}{L_{F_{js}}}(0.5 \cdot Z^{f}_{Fj} + 0.5 \cdot Z^{w}_{Fj}) + \frac{L_{F_{jms}}}{L_{F_{js}}}(0.33 \cdot Z^{m}_{Fj} + 0.33 \cdot Z^{h}_{Fj\ell=1} + 0.33 \cdot Z^{h}_{Fj\ell=0})$. So, location-by-sectors employing a greater share of women, my sector-specific instrument gives more weight to female Bartik shocks such as Z^{f}_{Fj} and Z^{w}_{Fj}.

Table 2.2: Estimation from Aggregate Data, Calibration, and Gravity Equation

Panel A: Estimation from Aggregate Data			
Parameter	Description	Value	Source
$[\alpha_{f1}, \alpha_{f2}, \alpha_{f3}, \alpha_{f4}, \alpha_{f5}, \alpha_{f6}, \alpha_{f7}]$	Female input share by industry	$[0.21, 0.42, 0.41, 0.35, 0.29, 0.29, 0.20]$	ENAHO (earnings data)
$[1-\beta_m, 1-\beta_f, 1-\beta_{mf}]$	Household expenditure in housing	$[0.212, 0.208, 0.171]$	ENAHO (expenditure data)
Panel B: Calibration			
Parameter	Description	Value	Source
σ	Male-female elasticity of substitution	2	Johnson and Keane (2013)
σ_D	Elasticity of substitution of demand	5	Freenstra et al. (2018)
α_s	Labor input share by industry	1-0.2 $\forall_s$	Ahlfeldt et al. (2015)
Panel C: Commuting Elasticity			
Parameter	Description	Value	Source
$[\kappa^m, \kappa^f, \kappa^h, \kappa^w]$	Iceberg disutility	$[0.01, 0.01, 0.01, 0.01]$	Tsivanidis (2023)
$[\hat{\theta}^m, \hat{\theta}^f, \hat{\theta}^h, \hat{\theta}^w]$	Shape parameter in commuting preferences	$[5.14, 6.17, 4.66, 6.34]$	Gravity Equation. See Section 2.5.3.

Notes: Considered industries are (in order): Manufacture (w/o textiles), Textiles, Services, Business Services, Wholesale Trade, Retail Trade, Transportation. ENAHO stands for the 2007-2018 waves of the national household survey.

Table 2.3: Estimation of Parameters

Panel A: Agglomeration Externalities (ϕ_{F_s})				Panel B: Labor Supply and Location Elasticities			
	OLS	2SLS	F - Weak Id		OLS	2SLS	F - Weak Id
	(1)	(2)	(3)		(1)	(2)	(3)
Manufacture (w/o textiles)	0.6230	0.1057	43.54	Labor Supply Elasticity (ν)	0.3644	1.2809	57.84
	(0.0146)***	(0.0877)			(0.0776)***	(0.2418)***	
Manufacture (textiles)	0.4836	0.1668	19.02	Location Single Females (η^f)	0.1801	2.0458	30.47
	(0.0184)***	(0.0968)*			(0.0707)***	(0.4363)***	
Services	0.7255	0.0590	19.72	Location Single Males (η^m)	0.2046	1.8154	31.13
	(0.0126)***	(0.1561)			(0.0733)**	(0.4121)***	
Business Services	0.5991	0.1837	30.98	Location Married (η^{mf})	0.1106	1.0155	42.71
	(0.0200)***	(0.0888)**			(0.0519)***	(0.1853)***	
Wholesale Trade	0.6261	0.1591	39.43				
	(0.0158)***	(0.0814)*					
Retail Trade	0.8204	0.2435	30.09				
	(0.0126)***	(0.0924)***					
Transportation	0.6112	0.1573	42.37				
	(0.0138)***	(0.0798)**					

Notes: Panel A of the analysis estimates agglomeration externalities for each industry at the zone level. The study focuses on Lima, which is made up of 1500 zones. The dependent variable is the log change in TFP between 2017 and 2007. TFP is the productivity residual that ensures the definition for industry sales holds given observable data. The independent variable is the log change in sectoral employment. To address endogeneity concerns, the analysis employs IV estimation using Bartik-style instruments that leverage the change in commuting times. I include the euclidean distance to the closest BRT station, the euclidean distance to the closest Metro station, and the euclidean distance to the Central Business District and its square, the physical size of the block and its square, its size in terms of the number of households in 2007, its slope, its elevation, and the number of dual-earner households as a share of married couples in 2007, the average market access across household-types as measured in 2007, and the dependent variable at its value in 2007. Panel B estimates labor supply and location elasticity. For the labor supply elasticity, the dependent variable is the log difference in the odds ratio in female labor force participation in each zone, while the independent variable is the log difference between the utility when participating and when staying at home. The analysis leverages changes in TFP to construct instruments and controls for a complete set of euclidean distances (to the BRT, to the Metro, to the CBD). Additionally, the analysis controls for the physical size of the block, its size in terms of the number of households in 2007, total employment in 2007, the number of dual-earner households as a share of married couples in 2007, and the block's slope and area. For the location elasticity, the dependent variable in the final regression is the log difference in the share of workers of a specific gender-civil status living in a particular zone, while the independent variable is a measure of indirect utility backed out from the model. The analysis exploits changes in TFP to construct instruments and includes controls such as the euclidean distance to the Metro, to the BRT, and to the CBD. The study also includes the size of the block, its slope and elevation, 2007 rents, and the dependent variable at its value in 2007.

Results are shown in Panel A of Table 2.3. Two observations are in order. First there is some heterogeneity in the strength of agglomeration economies across industries (see column 2 of Panel A). We have sectors such as Services with a scale elasticity statistically indistinguishable from zero, and Retail Trade where the scale elasticity is 0.24. Second, this heterogeneity might affect the extent to which the new infrastructure affects the gender gap. If the new transit infrastructure provides women with access to female-intensive industries that have high agglomeration economies (such as textiles), the gender gap might close compared to a counterfactual city where the new transit infrastructure is placed elsewhere.

Labor supply elasticity. I recover wages and compute market access $\Phi^{w}_{R_{i\ell=1}}$ and $\Phi^{h}_{R_{i\ell=0}}$. By taking logs and first differences to the odds ratio of labor participation in equation 2.17:

$$\Delta \log \frac{\mu_i^{mf}}{1-\mu_i^{mf}} = \nu \left[\frac{\Delta \log \Phi_{R^w_{i,\ell=1}}}{\theta^w} - \frac{\Delta \log \Phi_{R^h_{i,\ell=0}}}{\theta^h} \right] + \underbrace{\nu \left(\Delta \log \xi_{i\ell=1} - \Delta \log \xi_{i\ell=0} \right)}_{\gamma' X_i + \varepsilon_i^{\xi}}$$

where the term in bracket measures how the accessibility to jobs for dual-earner households improved relative to job accessibility for male breadwinner households. If changes in job access are correlated with changes in household productivity, then estimating this equation by OLS will lead to biased estimates of ν. To bypass this endogeneity concern I use the model to produce instruments. Since this is a supply-side parameter, we require variation from the demand. I use the TFP terms I recovered to define the following instruments: $Z_{R^{mf}_{is}} = \sum_j \left(d^w_{ij}\right)^{-\theta^w} d \ln A_{js}$. The identifying assumption is that $\mathbb{E}_i \left[\xi_i \cdot Z_{R^{mf}_{is}} | X_i \right] = 0$. Results are shown in Panel B of Table 2.3.[29] I estimate a labor supply elasticity ν of 1.28.

Location elasticity. I recover wages and household productivity and compute market access $\Psi_{R_i^g}$ and $\Psi_{R_i^{mf}}$. Manipulating the equations for location choices in 2.12 and 2.18, I get:

$$\begin{aligned}
\Delta \log \frac{N_{R_i^g}}{N^g} &= \chi_g + \eta^g \left[\left(\beta^g - 1\right) \Delta \log r_{R_i} + \frac{1}{\theta^g} \Delta \log \Psi_{R_i^g} \right] + \underbrace{\eta^g \Delta \log u_i^g}_{\gamma' X_i + \varepsilon_i^g} \\
\Delta \log \frac{N_{R_i^{mf}}}{N^{mf}} &= \chi_{mf} + \eta^{mf} \left[\left(\beta^{mf} - 1\right) \Delta \log r_{R_i} + \frac{1}{\nu} \Delta \log \Psi_{R_i^{mf}} \right] + \underbrace{\eta^{mf} \Delta \log u_i^{mf}}_{\gamma' X_i + \varepsilon_i^{mf}}
\end{aligned}$$

If growth in exogenous amenities is correlated with the growth in market access measures $\Psi_{R_i^g}$ and $\Psi_{R_i^{mf}}$ then the estimation of the location elasticities η^g and η^{mf} would be biased. Once again, I

[29] Once again I control for the complete set of euclidean distances (to the BRT, to the Metro, to the CBD). Moreover, I control for the physical size of the block, for its size in terms of the number of households in 2007, total employment in 2007, the number of dual-earner households as a share of married couples in 2007, and the block's slope and its area.

exploit changes in TFP to construct instruments for the growth in market access measures. Results are shown in Panel B of Table 2.3.[30] My results indicate that singles ($\hat{\eta}^f \approx \hat{\eta}^m \approx 2$) are much more mobile than married couples ($\hat{\eta}^{mf} \approx 1$).

2.5.5 Over-identification and untargeted data

In this subsection, I examine the extent to which the model can account for the observed variation in two outcomes not targeted for calibration, i.e., the geography of the gender earnings gap and commuting probabilities. In both cases, I compare predictions from the model with interdependent commuting against an alternative model where commuting is independent within couples. To give the alternative model the best chance to replicate the data, I back out exogenous characteristics, assuming this is the true model.[31]

Gender earnings gap. I use the 2011-2018 waves of the national household survey, which consists of the period after the introduction of the new transit infrastructure. I compute the average income at the municipality level for both men and women in dual-earner households. With these measures in hand, I calculate the gender gap by dividing male income by female income.[32] Next, I solve the two models using their backed-out 2017 exogenous characteristics. This yields a measure of income for men and women in dual-earner households at the zone level, which I then aggregate at the municipality level. Finally, I compute the gender earnings gap by dividing the two.

Panel A of Figure 2.3 shows the results for the model with interdependent commuting. The y-axis and x-axis showsthe gender gap (in logs) as measured in the data and the model, respectively. Note that I normalize the gap in both cases by setting the municipality of Surquillo's log gap to zero. For example, let us focus on Ventanilla, which is at the northeast side of the graph. In the model, the gap in Ventanilla is 47% higher than in Surquillo. In the data, the gap in Ventanilla is 54% higher than in Surquillo.

The message of Panel A is that even though we are only using earnings data to calibrate the gender share in each industry—the α_s in the model—we are quite successful in matching the geographical variation in the gender earnings gap in dual-earner households. In contrast, a model where commuting is independent within couples yields a strikingly different picture: Panel B

[30] I include controls such as the euclidean distance to the Metro, to the BRT and to the CBD. I also include the size of the block, its slope and its elevation, the 2007 rents, and dependent variable at its value in 2007.

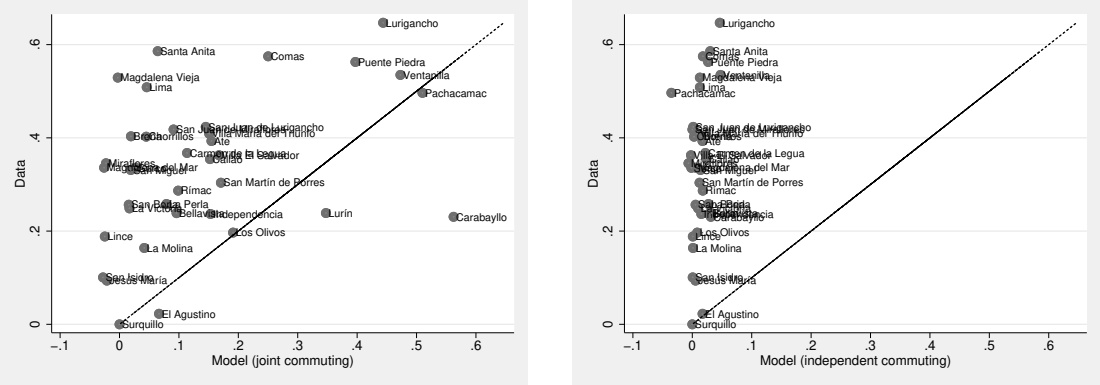
[31] Maximizing utility function: $U^{mf}_{i\ell jj'}(\omega) = u^{mf}_i \nu^{mf}_i(\omega) \cdot \xi^{mf}_{i\ell} \alpha^{mf}_\ell(\omega) \cdot \frac{\epsilon^h_j(\omega)}{d^h_{ij}} C^h_{i\ell j}(\omega)^{\beta^{mf}} H_{R^h_{i\ell j}}(\omega)^{1-\beta^{mf}} \cdot \left[\frac{\epsilon^w_{j'}(\omega)}{d^w_{ij'}} C^w_{i\ell j'}(\omega)^{\beta^{mf}} H_{R^w_{i\ell j'}}(\omega)^{1-\beta^{mf}}\right]^\ell$ subject to budget constraints $PC^h_{i\ell j}(\omega) + r_{R_i} H_{R^h_{i\ell j}}(\omega) = w_{jm}$ and $PC^w_{i\ell j'}(\omega) + r_{R_i} H_{R^w_{i\ell j'}}(\omega) = w_{j'f}\ell_i$ delivers a model with independent commuting probabilities. Another way of setting up a model with independent commuting is by assuming both that (i) iceberg costs in commuting are productivity costs, and (ii) idiosyncratic shocks to commuting are productivity shocks rather than preference shocks. Welfare in both models is isomorphic.

[32] I only keep municipalities for which I had at least 30 observations.

shows almost no geographical variation in the gender earnings gap.[33]

Figure 2.3: Matching the Geography of the Gender Earnings Gap

Panel A: Interdependent Commute Choices Panel B: Independent Commute Choices



The y-axis shows the gender earnings gap (in logs) computed from data on labor earnings at the household level and aggregated at the municipality level. The x-axis shows the gender earnings gap as predicted by the model (in logs) at the municipality level. In Panel A, I use the model with interdependent commute choices. In Panel B, I use the model with independent commute choices. For the y- and x-axis, I normalize the gender gap by setting it to zero in the municipality of Surquillo. The black line in both figures is the 45 degrees line.

Commuting probabilities. I use the Census data to compute the probability that a worker commutes between any of the 49 municipalities of Lima in 2017, yielding 2401 pairs of bilateral commuting probabilities. I solve the model using 2017 exogenous characteristics and commuting times under two models. Then, I compare the predicted commuting probabilities with those observed in the 2017 Census.

Figure B.6 plots a bin scatter of commuting probabilities in the model and in the data. Panel A focuses on females in dual-earner households. The model with joint commuting performs better when predicting commuting flows. The left hand side of Panel A shows that the model with independent commuting is able to predict around 40% of the variation in the data (correlation of 0.63). On the right hand side, Panel A of Figure B.6 reveals that the model with interdependent commuting is able to explain 56% of the variation in the data (correlation of 0.75). Panel B exhibits the results for males in dual-earner households. Both models have similar performances (R^2 of 0.39, and correlation of approximately 0.63).

Finally, Figure B.7, reproduces the estimation procedure for Fact 2 and tests whether commuting probabilities of one spouse depend on those of the other spouse (see Section 2.3.2). To do so,

[33]The model with interdependent commuting is more reliable than the model with independent choices in predicting income data for men, but especially for women in dual-earner households. The correlation of the average income of women in dual-earner households (in logs) between the model with interdependent commuting choices and the data is 0.72. However, for the model with independent commuting choices, the correlation is 0.52. For men, the model with interdependent commuting choices yields a correlation of 0.61, while the model with independent commuting choices, 0.56.

I use data on conditional commuting probabilities generated by the model. The main takeaway is that even though I did not target this moment in the calibration, the R-squared is very far from zero as observed in the data.[34]

Therefore, although the model with interdependent choices is an abstraction, it is able to capture critical features of couples' commuting patterns that the model with independent choices is unable to capture, especially in the case of women in dual-earner households.

2.6 The Impact of Improving Commute Times

This section quantifies how the new transit infrastructure affected the gender earnings gap and aggregate real earnings through improved commute times under a model with interdependent commuting (Section 2.6.1), and a model with independent choices (Section 2.6.2). Then, I explain how and why quantitative results would change if commuting choices are independent within couples (Section 2.6.3). Finally, I perform a policy counterfactual where the Metro network is completed according to the city's plans (Section 2.6.4).

Outcomes of interest. For each gender in each type of household I construct a measure of real income across locations $RY_i^k = \frac{y_i^k}{P^{\beta^k} r_R^{1-\beta^k}}$ where k indexes a gender-by-household group. Then, I compute aggregate real income RY^k simply by summing up RY^k over all locations. Since I am interested not only on aggregate income but on average earnings, I construct a measure of real income per worker $\bar{RY}^k = RY^k/N_{R^k}$, where N_{R^k} is the number of workers that belong to gender-household type k. Finally, I also compute the number of dual-earner households as a share of married households.

Counterfactual exercises. I remove the Metro and the BRT from the city, while keeping exogenous characteristics fixed at their 2017 values. I perform this counterfactual exercise for two cases, the model with interdependent commuting, and the model with independent commuting. I use observed data to recover unobservables under the model with interdependent commuting. Then, I solve each model conditionally on this set of unobservables.

2.6.1 Interdependent commute choices

Impact on the gender earnings gap across locations. Figure 2.4 maps the evolution of real income RY_i and gender gap $\text{Gap}_i^k = RY_i^{k,m}/RY_i^{k,f}$ across origins i caused by the improved times. Each color accounts for 20% of observations. The map in the left hand side of Panel A shows that real income increased mostly in the northern and southern fringe of the city, while it shrunk

[34] By construction, with independent commuting the R-squared would be zero.

in central areas of the city and on the eastern fringe. This is the consequence of two features. (i) Households in the northern and southern fringe of the city were the main beneficiaries of the infrastructure, as it allowed them to commute faster to central areas. (ii) Increased labor supply coming from the northern and southern fringe pushed down wages, especially in destinations in central areas. On the right hand side of Panel A, I exhibit the impact on the gender gap in aggregate real income. The gap decreased the least in central areas, and the most in the fringe. However, impacts vary by household-type and location. Panel B maps the impact on the gender gap in single households and in married households. Two differences between singles and married arise. First, gap reductions tend to be larger for married. Second, the decrease in the gender gap among married tend to be concentrated in the northern and southern fringe. This is true for singles but to a lesser extent.

In Figure 2.5, I estimate average impacts in different parts of the city by comparing the gap when the Metro and BRT are available against the gap without them. I compute these impacts for locations within 0 to 5 km from the CBD, from 5 to 7.5 km, from 7.5 to 10 km, from 10 to 12.25 km, and from more than 12.5 km. In Panel A, I show the impact on the aggregate real earnings gap per location. Since the number of households across locations is evolving endogenously, it is helpful to look at the increase on average earnings. In Panel B, I show the impact on the average real earnings gap. Both gaps decreased mainly in the fringe. This effect is almost entirely driven by married households for which aggregate gaps decreased by more than 12% and per worker gaps by 6.9%. Furthermore, the aggregate earnings gap increased for singles in some locations, but the average gap remained the same.

In sum, the improvement in commute times reduced the gender gap among married households according to the model with interdependent choices. This effect was greater in the fringe of the city. Next, I summarize aggregate impacts.

Aggregate impact. Table 2.4 reports my aggregate findings.[35] It reports the percentage growth in the outcomes described above for singles, married, and all grouped together.

First, both single males and single females see their income grow modestly by 0.60% and 0.39%. Since males aggregate income grew faster than females', the gender gap in aggregate earnings increased by 0.48%. When looking at average earnings, the gender gap increased by 0.56%. Second, turning to married households, aggregate income from male breadwinner households decreased by 1.61%, but average income increased by 0.66%. This is because many breadwinner households endogenously became dual-earner households. In fact, the new transit infrastructure led to an increase in female labor force participation of 1.74%, representing 9.48% of the total growth in labor force participation in the data. Since the number of dual-earner households is in-

[35] I compute the gender earnings gap by dividing male real income by female real income and subtracting one: $\text{Gap} = {}_{RY^{k,m}}/_{RY^{k,f}} - 1$ or $\bar{\text{Gap}} = {}_{\bar{RY}^{k,m}}/_{\bar{RY}^{k,f}} - 1$, where k indexes singles and couples.

creasing endogenously, it is helpful to look at the increase on average earnings. Males' average earnings in dual-earner households grew by 1.17%, whereas females' average earnings rose by 2.19%. These impacts led to a reduction of the gender gap in aggregate and average earnings, which decreased by 4.13% and 2.54%, respectively. Finally, regarding the overall impact of the new transit infrastructure, aggregate and real income per worker grew by 1.07% and 0.70%, while the aggregate gap and the gap per worker decreased by 2.15% and 1.05%.

In Online Appendix B.5, I turn off mechanisms from the full model to assess their quantitative relevance. Without externalities, the endogenous LFP choice, and location choices, the reduction in the aggregate earnings gap becomes 1.14% instead of 2.15%.

2.6.2 Independent commute choices

I re-do the counterfactual exercise assuming independent commuting.[36]

Impact across locations. This is shown in Figure B.8. Panel A reveals that married households experienced a decrease in the gap in aggregate earnings of about 8.6%. However, Panel B shows that this is not due to a decrease in the gap in earnings per worker. Rather, the decrease shown in Panel A is driven by the fact that the improvement in commute times induced male breadwinner households to become dual-earner households, thereby increasing the mass of income accumulated by married women relative to married men.

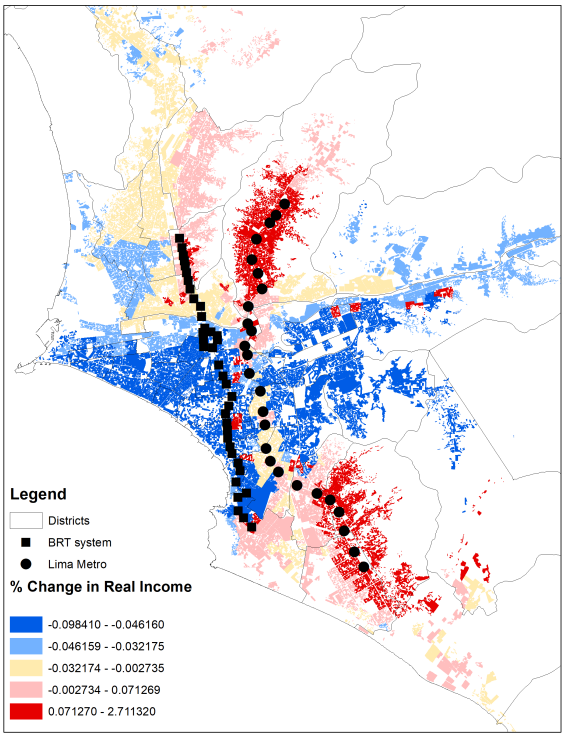
Aggregate impact. Table B.1 shows the aggregate impact when using a model with independent commute choices. Three observations are in order. First, this model underestimates the impact of the new transit infrastructure on aggregate real income by 1 - 0.0066/0.0107 = 38% and on real income per worker by 1 - 0.0036/0.0070 = 49%. Second, while the gender gap in aggregate earnings in married households closes by 2.47%, this is mainly due to the labor force participation margin leading to an overall increase in the mass of income accumulated by females. When looking at earnings per worker, this model suggests that the gap in married households decreased by 0.13% (instead of shrinking by 2.54% as calculated for the model with interdependent commuting). Third, this model underestimates the impact on the gender gap in aggregate earnings by 1 - 0.0150/0.0215 = 30%, and on the gender gap in earnings per worker by 1 - (0.0045/0.0105) = 57%.

[36] Maximizing utility function: $U^{mf}_{i\ell jj'}(\omega) = u_i^{mf}\nu_i^{mf}(\omega)\cdot\xi^{mf}_{i\ell}\alpha^{mf}_{\ell}(\omega)\cdot\frac{\epsilon^h_j(\omega)}{d^h_{ij}}C^h_{i\ell j}(\omega)^{\beta^{mf}}H_{R^h_{i\ell j}}(\omega)^{1-\beta^{mf}}\cdot\left[\frac{\epsilon^w_{j'}(\omega)}{d^w_{ij'}}C^w_{i\ell j'}(\omega)^{\beta^{mf}}H_{R^w_{i\ell j'}}(\omega)^{1-\beta^{mf}}\right]^{\ell}$ subject to budget constraints $PC^h_{i\ell j}(\omega)+r_{R_i}H_{R^h_{i\ell j}}(\omega)=w_{jm}$ and $PC^w_{i\ell j'}(\omega)+r_{R_i}H_{R^w_{i\ell j'}}(\omega)=w_{j'f}\ell_i$ delivers a model with independent commuting probabilities. Another way of setting up a model with independent commuting is by assuming both that (i) iceberg costs in commuting are productivity costs, and (ii) idiosyncratic shocks to commuting are productivity shocks rather than preference shocks. Welfare in both models is isomorphic. In any case, I keep the property that married men and women locate together to give it the best chance to replicate the data. Although, it is fairly simple to set up an alternative model where preferences are defined for males and females irrespective of their marital status. By definition, in this alternative model, commuting choices would be independent since there would not be any notion of married households.

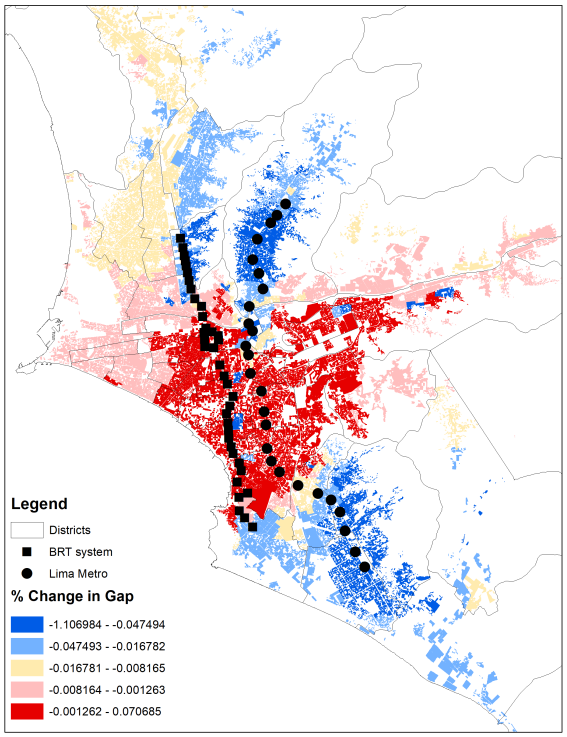
Figure 2.4: Quantitative Impact of Transit Infrastructure on Earnings and the Earnings Gender Gap with Interdependent Commute Choices

Panel A:

Real Income

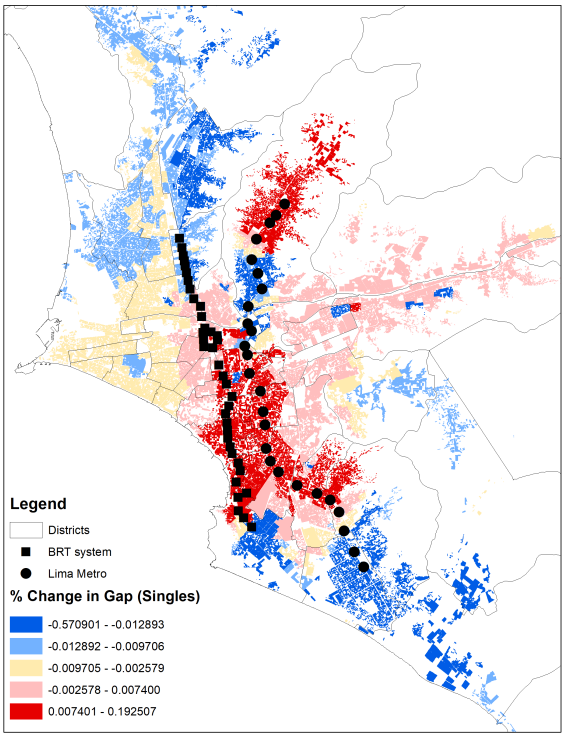


Gender Gap in Real Income

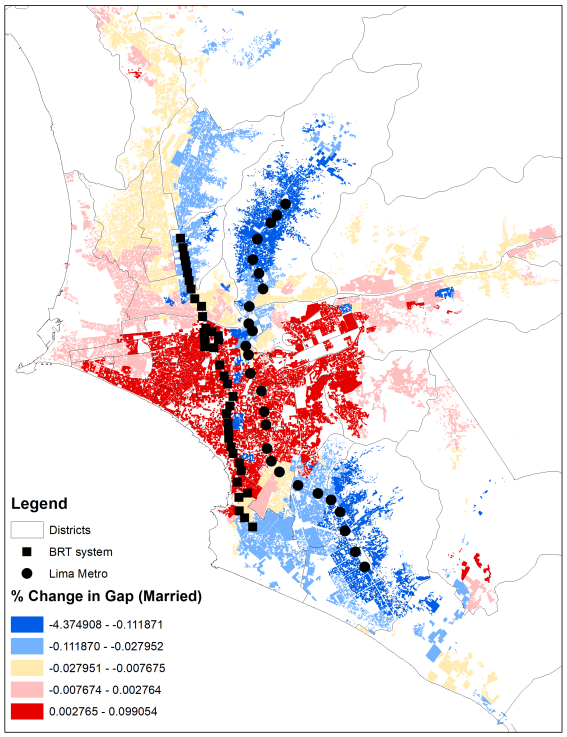


Panel B:

Gender Gap (singles)



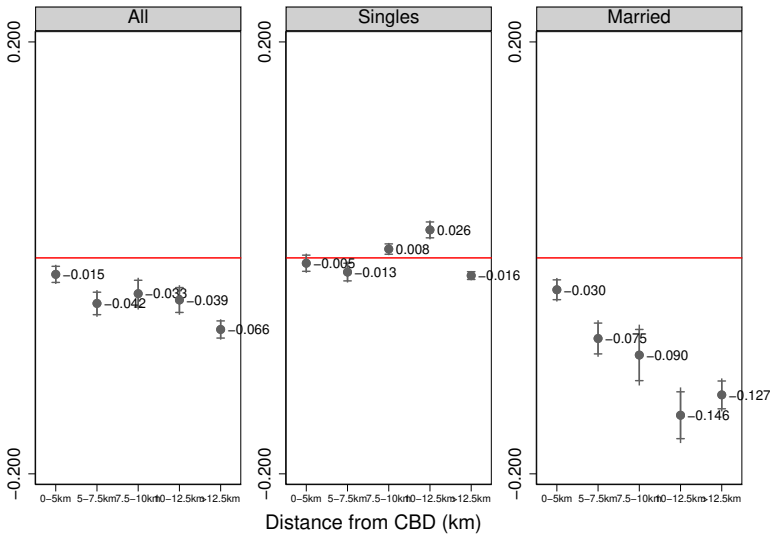
Gender Gap (married)



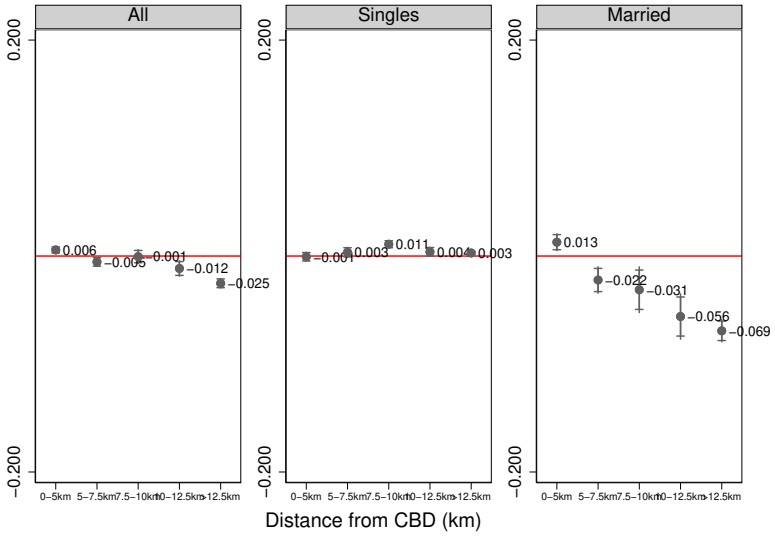
See Section 2.6.1.

Figure 2.5: Impact with Interdependent Commute Choices by Distance to the CBD

Panel A: Gap in Aggregate Real Earnings



Panel B: Gap in Real Earnings per Worker



See Section 2.6.1.

2.6.3 Understanding the role of joint commuting within couples

First order decomposition. I have established that removing interdependent commuting leads to a lower reduction of the gender gap in response to improved commute times. As a first exercise aiming to understand the differences between the models with interdependent and independent commuting, starting from the model with interdependent commuting, I decompose the impact of

improving commute times into several margins to a first order.

For a set of commute times changes $dt_{ij'}, \forall i, j'$, I can use equation 2.9 to decompose the aggregate impact, ΔGap, to a first order, into three pieces: the direct use channel, the indirect use channel, and the general equilibrium impact (which is the difference between the total impact and the sum of the direct and the indirect use channels):

$$\Delta Gap = \Delta Direct + \Delta Indirect + \Delta GE \tag{2.21}$$

Panel A of Table 2.5 depicts how the impact on the aggregate earnings gap, ΔGap, can be decomposed into different margins for singles, dual-earner households, and women in dual-earner households against men in male-breadwinner households.

First, among singles, it reveals that from a direct use perspective, improvements in commute times marginally reduced the gap (-15% out of total impact of 0.0033). In other words, single women tended to use the routes that were improved more than single men. However, general equilibrium effects are responsible for 115% of impacts on the gap. This means that general equilibrium effects favored single men. Why general equilibrium effects favored men relative to women? In part this is because the increased labor supply of women led to a decrease in average wages at workplaces. On average, wages decreased by 1.6% for men, and decreased by 2.3% for women, so the decrease was more pronounced for women.

Panel A of Table 2.5 also shows the impacts among spouses in dual-earner households. It reveals that the direct use channel contributed the most in reducing gaps (131%). This means that married women were initially more exposed to the routes that were improved. However, the indirect use channel (-23%) and the general equilibrium effects (-8%) tended to widen the gaps. Recall that according to equation 2.9, the probability that wives commute to low-wage locations increases through the indirect use channel when the initial conditional commuting probabilities to these locations are high enough. Since, ex-ante, women tended to commute to low-wage locations, the indirect channel led to a widening of the gap.

Finally, Panel A of Table 2.5 shows the comparison of women in dual-earner households against men in male breadwinner households. The direct use and the general equilibrium channel tended to decrease the gap in aggregate earnings, explaining 42% and 69% of the total impact, respectively. This is because the share of male breadwinner households decreased, and so the mass of income they generated also decreased relative to dual-earner households. The indirect channel tended to widen the gap, explaining -11% of the total impact.

In sum, this exercise indicates that, to a first order, the gender gap decreased for married households because the new infrastructure improved the routes that married women were using the most. However, the indirect use channel reinforced existing gaps. If we set this channel to zero while

keeping everything else constant, the gap among married households would have decreased even more.

Table 2.4: The Aggregate Impact of the Line 1 of the Metro and the BRT

Interdependent Commuting Choices								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0060	0.0039	-0.0161	0.0292	0.0396	0.0070	0.0213	0.0107
Real Income per Worker	0.0060	0.0039	0.0066	0.0117	0.0219	0.0070	0.0122	0.0070
Gender Gap	0.0048			-0.0413		-0.0215		
Gender Gap (per Worker)	0.0056			-0.0254		-0.0105		
# D.E. / # Married				0.0174				

Notes: I conduct a counterfactual exercise where I remove the Metro and the BRT from the city, while keeping exogenous characteristics fixed at their 2017 values. I perform this counterfactual exercise for two cases, one in which one spouse's choice is dependent on the other spouse's choice, and one in which spouses perform their commuting choices independently from each other. To perform these counterfactual exercises, I use observed data to recover unobservables under the model with interdependent choices. Then, I solve each model conditionally on this set of unobservables.

Table 2.5: Decomposition of the Aggregate Impacts

Panel A: Interdependet Commuting Choices							
	ΔGap	$\Delta Direct$		$\Delta Indirect$		ΔGE	
	Total	Absolute	$\frac{Absolute}{Total}$	Absolute	$\frac{Absolute}{Total}$	Absolute	$\frac{Absolute}{Total}$
Singles	0.0033	-0.0005	-15%	0	0%	0.0038	115%
Dual-Earners	-0.0108	-0.0141	131%	0.0025	-23%	0.0008	-8%
Women D-E vs Men B-W	-0.0477	-0.0199	42%	0.0051	-11%	-0.0328	69%

Panel B: Interdependet vs Independent Commuting Choices								
	$\Delta^2 Gap$	ΔGap	$\Delta^2 Direct$		$\Delta^2 Indirect$		$\Delta^2 GE$	
	Total	(indep.)	Absolute	$\frac{Absolute}{Total}$	Absolute	$\frac{Absolute}{Total}$	Absolute	$\frac{Absolute}{Total}$
Singles	0.0028	0.0005	0.0032	114%	0	0%	-0.0004	-14%
Dual-Earners	-0.0124	0.0017	-0.0115	93%	0.0025	-20%	-0.0034	27%
Women D-E vs Men B-W	-0.0250	-0.0227	-0.0190	76%	0.0051	-20%	-0.0110	44%

Notes: I conduct a counterfactual exercise where I remove the Metro and the BRT from the city, while keeping exogenous characteristics fixed at their 2017 values. I perform this counterfactual exercise for two cases, one in which spouses' commuting choices are interdependent, and one in which spouses perform their commuting choices independently. To perform these counterfactual exercises, I use observed data to recover unobservables under the model with interdependent choices, and solve each model conditionally on this set of unobservables. For Panel A, I decompose the total impact on the aggregate gender earnings gap, $\Delta Gap = \log \frac{y_{Male}}{y_{Female}}$, into three margins, $\Delta Direct$, $\Delta Indirect$ and ΔGE. For Panel B, I decompose the total impact on $\Delta^2 Gap \equiv \Delta Gap(Interdependent) - \Delta Gap(Independent)$ into $\Delta^2 Direct$, $\Delta^2 Indirect$, and $\Delta^2 GE$.

Complete decomposition. In the previous subsection, I decomposed the impact of improving commute times into three margins: the direct and indirect use channels and general equilibrium. Then, to understand the incidence of interdependent commuting within couples, I quantified the indirect use channel in Panel A of Table 2.5. However, this exercise implicitly assumed that removing interdependent commuting would only shape the impact of improving commute times through

the indirect use channel since the direct channel is kept fixed. This is reasonable if enough data is available to feed the direct channel into the model, in practice, fixing this channel. However, this is not a reasonable assumption in most cases because the data required is not so readily available. Removing interdependent commuting would likely affect the model's measurement of the direct channel and general equilibrium since everything is interrelated and interacts with each other. For example, if the geography of wages at workplaces is different under the model with interdependent commuting relative to the model with independent commuting, then commuting responses will also be different. Thus, the incidence of the direct channel would likely differ. Similarly, wage responses might be different between the two models, leading to different general equilibrium effects.

To understand the effect of having interdependent commuting within couples, and how this shapes the impact of improving commute times, I extend the decomposition of the previous section in the following way, akin to a differences-in-differences framework. First, I solve the model with interdependent commuting and decompose the impact of improving commute times into the three margins explained above (Scenario A). Second, I solve the model with independent commuting and perform the decomposition, this time only into two margins: the direct use channel, and general equilibrium (Scenario B). This is because, by construction, the indirect use channel is zero. Finally, I compare the effect of improving commute times on the gender gap under scenario A, $\Delta Gap(A)$, against scenario B, $\Delta Gap(B)$:

$$\begin{aligned} \Delta Gap(A) - \Delta Gap(B) &= \big[\Delta Direct(A) - \Delta Direct(B)\big] + \big[\Delta Indirect(A) - \Delta Indirect(B)\big] \\ &+ \big[\Delta GE(A) - \Delta GE(B)\big] \end{aligned} \tag{2.22}$$

where $\Delta Indirect(B) = 0$ by construction. This decomposition is shown in Panel B of Table 2.5. I define $\Delta^2 Gap \equiv \Delta Gap(A) - \Delta Gap(B)$. I also show for ease of reading the impact of improving commute times on the gender gap under the model with independent commuting; see $\Delta Gap(indep)$. For example, among singles, the gap increased by 0.0028 log points more than it would have increased in a model with independent commuting (0.0005 log points). Among dual-earners, the gap decreased by 0.0124 log points more than in a model with independent commuting (0.0017 log points). Finally, the gap between women in dual-earner households and men in male breadwinner households decreased by 0.0250 log points more than in a model with independent commuting (-0.0227 log points). This merely revisits what I already established in Section 2.6.2: the model with independent commuting underestimates the decrease in the gender earnings gap. Next, I explain why.

I decompose $\Delta^2 Gap$ into $\Delta^2 Direct$, $\Delta^2 Indirect$, and $\Delta^2 GE$. The main takeaway is that most of the differences between the models with interdependent and independent choices are explained

by differences in the direct channel. This channel constitutes 114% of $\Delta^2 Gap$ for singles, 93% for dual-earners, and 76% for women in dual-earner households vs men in breadwinner households. Why is that? By equation 2.9, the direct channel for $k \in \{\text{Singles, D-E, D-E vs B-W}\}$ is:

$$\Delta Direct^k = \theta_M \kappa \frac{w_{M,j'} \pi^M_{j'|i}}{\sum_l w_{M,l} \pi^M_{l|i}} \left(1 - \frac{\bar{y}^M_i}{w_{M,j'}}\right) \Delta t_{ij} - \theta_F \kappa \frac{w_{F,j'} \pi^F_{j'|i}}{\sum_l w_{F,l} \pi^F_{l|i}} \left(1 - \frac{\bar{y}^F_i}{w_{F,j'}}\right) \Delta t_{ij} \quad (2.23)$$

where M and F stand for males and females in group k. This means that to the extent that the geography of unobserved wages w, commuting probabilities, π, and household income $\bar{y}$ inferred by the two models differ, the contribution of the direct channel will differ, too. Intuitively, these terms explain how exposed a particular household is to improvements in commute times. Consistent with this, in Section 2.5.5, I established that the model with interdependent commuting did a much better job in predicting the geography of the gender earnings gap and commuting probabilities, especially for the married.[37]

Comparing data- and model-based reduced-form responses. In this subsection, I compare the reduced-form response to improved infrastructure against the predictions by the model. I solve for unobservables for both models to give them the best chance in replicating the data. I re-estimate the same reduced form equation using data generated by the model. I do it for the odds ratio in female labor force participation and the gender earnings gap. Results are shown in Panels A and B of Figure B.9, respectively. In summary, the model with interdependent commuting is more consistent with responses observed in the data than the model with independent commuting (compare Figures B.9 and Figures 2.2). This is especially true for responses in the gender earnings gap.[38]

2.6.4 Policy: completing the Metro's network

I use the model with joint commuting to evaluate the effect of completing Lines 2 and 4 of the Metro. Overall, real income would grow an additional 0.95%, and the gap would have decreased an additional 1.87%. The complete discussion is in Online Appendix B.7.

[37] The indirect channel stills explain a substantial portion of the total difference between the two models, but it goes in the opposite direction (-20%). General equilibrium effects also explain a substantial portion of the differences due to wages responding very differently in the two models as labor supply responses are obviously different.

[38] A comparison of data- and model-based responses to the new transit infrastructure using other outcomes is available in Online Appendix B.6.

2.7 Conclusion

Commuting choices are interdependent across spouses. As the wage of one of the spouses increases, households are willing to give up more of the other spouses' income to decrease commute costs. Thus, improving commute times can affect households through two channels: the direct and the indirect use channels. Incorporating the indirect channel has substantial implications about how households react to improvements in commute times, and fundamentally, on the measurement of the exposure to improved commute times.

My results show that the model with interdependent commuting choices fits better the data on the geography of the gender earnings gap and bilateral commuting flows of dual-earner households than the model without interdependence. They also show that reduced commute times shrunk the gender gap, primarily where access increased the most. This is due to, relative to men, married women were more intensively using improved routes. Finally, using a model with independent choices to perform this analysis would have led me to conclude that reductions in the gender earnings gap were substantially lower, which is counterfactual given the reduced-form responses observed in the data. This is because the model with independent choices would have miscalculated households' exposure to improved commute times (i.e. through the direct and indirect channels).

This analysis exemplifies how is that improving commute times can affect the gender gap, even if it was not a direct policy objective. However, if earnings inequality is part of policymakers' objective function, they should take into account the interdependence of commuting choices within households when designing transit infrastructure projects. This is because through the indirect use channel in commuting choices, ceteris paribus, improving one spouse's prospects could negatively affect the other spouse's earnings.

CHAPTER 3

Initial Conditions and the Biggest Push in U.S. History

Abstract[1]

This paper examines the impact of WWII on U.S. economic geography. We approach this question empirically, theoretically, and quantitatively. Our empirical results show positive long run effects of WWII spending on overall population and manufacturing activity only in counties with initially low manufacturing employment. Motivated by these findings, we develop a multisector economic geography model that features external economies of scale in manufacturing with non-constant elasticity. We estimate the model and show that an S-shaped scale function in which sizable scale economies are limited to locations below about 1500 manufacturing workers matches the data well. We use the estimated model to re-examine how WWII shaped the evolution of U.S. postwar economic geography. Despite the presence of "big push" dynamics and multiple steady states, the aggregate effects of WWII spending on long run outcomes are quite limited due to the small domain over which coordination failures operate as well as the untargeted nature of the spending.

3.1 Introduction

During World War II, the United States government undertook the largest and most abrupt public spending program in U.S. history. Federal net outlays went from roughly 10% of GDP to over 40% in a few years, then subsequently back down to less than 15% a few years later (Figure 3.1). Much of this increased spending consisted of supply contracts awarded to domestic firms for the production of materials and equipment for the war effort. What were the effects, if any, of this massive but temporary program on the long run evolution of U.S. economic geography? This question has drawn interest in light of the broader literature on historical persistence, path dependence and place-based policies in economic geography, as well as the specific historical

[1] Co-authored with Dominick Bartelme and Zhimin Li

experience of the post-war United States with the decline of the Manufacturing Belt and the spread of industry to the Sun Belt.[2]

This paper revisits the long run impact of WWII spending on U.S. economic geography empirically, theoretically, and quantitatively. We begin by linking a comprehensive dataset of war supply contracts during WWII with data on the long-run growth in population and manufacturing employment of U.S. counties from 1930 to 2000. Using a potential outcomes framework along with controls for initial conditions, we compare the long run outcomes of counties that received large war supply contracts relative to their manufacturing employment to those that received small or no contracts. Consistent with the prior literature, we find that treated counties experienced higher long-run population growth and manufacturing employment growth relative to untreated counties. Our primary empirical contribution is to show that these positive treatment effects are driven almost entirely by counties with initially small manufacturing employment, which experience economically large and statistically significant long run increases in population and manufacturing employment. Treated counties with initially large employment in manufacturing show strong short-run gains in population and manufacturing employment, but these gains are almost completely reversed in subsequent decades.

We interpret our results as providing support for "big push" or "poverty trap" theories of regional specialization and development [Rosenstein-Rodan, 1943, Murphy et al., 1989, Azariadis and Stachurski, 2005]. In these theories, coordination failures can result in a region being stuck in an equilibrium that features a low level of economic activity relative to its potential. A large enough temporary shock induces a higher level of activity, which then becomes self-sustaining and permanent. However, as suggested by the term "poverty trap," most versions of the theory posit some scale of economic activity beyond which coordination failures diminish in importance and temporary shocks no longer have permanent effects.[3] While these theories have motivated a sizable empirical literature on historical persistence and path dependence in a variety of contexts,[4] the theoretical prediction of heterogeneity in the long-run response based on initial conditions has been largely neglected. It is precisely this prediction of the limited relevance of coordination failures beyond a certain scale of economic activity that we find support for in this paper.

Motivated by our findings, we develop a quantitative economic geography model with many regions that features the possibility of coordination failures leading to path dependence in long run equilibria, but parameterizes the scale of economic activity over which coordination failures

[2]Econometric studies include Fishback and Cullen [2013], Fishback and Jaworski [2016], Jaworski [2017] and Li and Koustas [2019]. A few examples from the vast historical literature include Wright [1986], and Hooks and Bloomquist [1992].

[3]If one accepts the premise that finite inputs however applied result in finite output, then it logically follows that there must some limitation to the scope of coordination failures in any finite economy.

[4]A few prominent examples in a vast literature include Rauch [1993], Bleakley and Lin [2012], Davis and Weinstein [2002], Kline and Moretti [2014], Juhász [2018].

can arise. We build on the recent framework of Allen and Donaldson [2023] by adding multiple sectors in a waythat allow for long run structural change, and more importantly the possibility of non-constant-elasticity local external economies of scale in the manufacturing sector that may be diminishing in the size of the local sector. We calibrate the model by choosing the scale function so that the response of the model to temporary demand shocks, stemming from temporary government purchases of manufactured goods, matches the heterogeneous responses that we found in the data. We find that the model matches the data well with a scale function that approximates an S-curve in manufacturing employment: initially flat up to 250 manufacturing workers in the county, increasing with elasticity of about 0.2 up to about 1500 manufacturing workers, then flat again thereafter. The calibrated model generates a large number of interior steady states and path dependence. In contrast, the constant elasticity version of the model, which is ubiquitous in the quantitative literature, has trouble matching the empirical moments precisely because it lacks a natural mechanism to limit the domain over which coordination failures apply.

We use the calibrated model to revisit how the WWII war supply effort shaped the evolution of U.S. postwar economic geography. We simulate the model with and without the WWII spending shocks, and compare the outcomes in the year 2000. We find that the WWII spending did have long run effects on U.S. economic geography, with a number of smaller counties experiencing large gains or losses in manufacturing employment. However, the long run effects on aggregate manufacturing employment, welfare, and regional reallocation are minimal. This is due to both the limited scope for reallocation to raise aggregate productivity (limited domain for coordination failures to operate) as well as the distribution of the WWII contracts, which were not designed to improve allocative efficiency. History did matter, but it mattered little in this instance.

Our paper contributes several literatures. The first is the extensive empirical literature on historical persistence and path dependence in economic geography (e.g. Rauch [1993], Bleakley and Lin [2012], Davis and Weinstein [2002], Kline and Moretti [2014], Juhász [2018]). Our main contribution is to focus on the heterogeneity of persistence with respect to initial conditions, a neglected topic in the literature. Our paper also relates to the literature on quantitative economic geography, recently surveyed in Redding and Rossi-Hansberg [2017]. Our primary contribution to this literature is to develop a tractable quantitative model that embeds "big push" dynamics and parameterizes their scope. Finally, we contribute to the historical literature on the impact of WWII on U.S. economic geography. Several scholars have posited that WWII played a significant role in the diffusion of manufacturing activity to U.S. regions beyond the manufacturing belt, such as the South (e.g., Wright [1986], Hooks and Bloomquist [1992]). However, others have pointed out that the manner in which spending was allocated, primarily towards counties with a large established manufacturing base, could have had the opposite effect of reinforcing existing patterns of specialization [Fishback and Cullen, 2013, Rhode et al., 2018, Fishback and Jaworski, 2016].

In an empirical study, Jaworski [2017] finds no evidence that Southern counties receiving more wartime investment had faster growth. Our findings support the view that the long run impact of the *temporary* component of WWII spending was rather limited. However, we do find that the Pacific coast showed the largest long run impact, consistent with Rhode [2003]. WWII also led to some permanent changes in the composition and scale of military spending, which we do not study but could have had significant geographic effects.

3.2 Estimation and Identification

This section briefly describes the historical context of federal military supply contracts during WWII and our data sources, then discusses our approach to estimation and identification.

3.2.1 Historical Context and Data

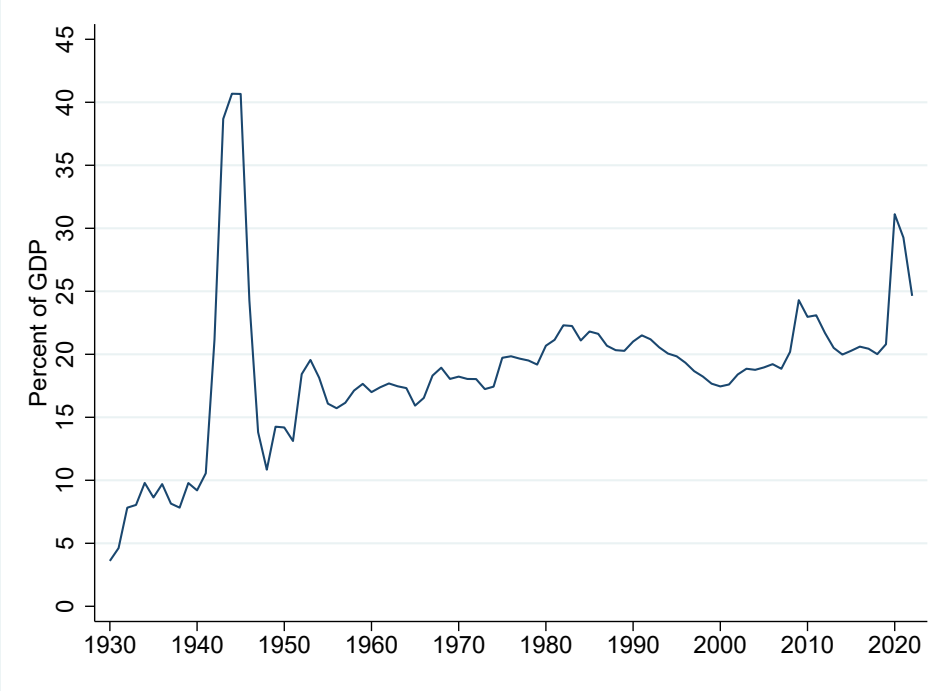
During World War II, the federal government assumed an unprecedented level of control over the U.S. economy. At their peak, federal government expenditures accounted for over 40 percent of the Gross Domestic Product, the highest proportion ever recorded (Figure 3.1). The federal government allocated a staggering $196 billion for military supply contracts (primarily for manufactured goods) and invested an additional $31 billion in new production facilities between June 1940 and June 1945.[5] This remarkable endeavor to turn the U.S. into the "Arsenal of Democracy" stands as the most significant economic intervention by the federal government in U.S. history.

The allocation of war supply contracts during World War II was primarily determined by economic considerations rather than political factors. Rhode et al. [2018] find that the strongest predictors of the assignment of supply contracts at the state level were existing general industrial capacity prior to the war in 1939 and specialized industrial capacity for aircraft production. Their findings suggest that the government prioritized regions with established industrial infrastructure and specialized capabilities, reflecting a strategic approach to efficiently meet the demands of wartime production [Fishback and Cullen, 2013]. The ability of the United States to rapidly procure and produce goods and services contributed significantly to the Allied victory in WWII [**?**].

The data for WWII supply contracts come from volumes originally issued by the Civilian Production Administration (formerly the War Production Board), which record all defense contracts greater than $50,000 issued between June 1940 and September 1945. Our measure of government defense spending is the aggregate monetary value of war supply contracts in each county during that period. The contract values across counties exhibit a large variation, with a minimum value of $26,200 and a maximum of $281 million. The mean contract value was $1.3 million with a

[5]For reference, U.S. GDP was about $102 billion in 1940.

Figure 3.1: Federal Net Outlays as Percent of GDP



Note: This figure shows the federal net outlays as percent of GDP over time. Source: Federal Reserve Bank of St. Louis, U.S. Office of Management and Budget; see the link here.

duration of 7.4 months. Data on county characteristics are the same as those in Kline and Moretti [2014] and come from a variety of sources including the Census of Population and Housing and Agricultural Census. Finally, to ensure comparability of county definitions over time, we use the crosswalk provided by Eckert et al. [2020b] to obtain a sample of harmonized counties between 1900 and 2000. Additional details can be found in Online Appendix C.1.

3.2.2 Empirical Model

Our goal for the empirical analysis is to estimate the impact of this massive but temporary increase in government spending on local economic outcomes in the long run, and in particular to understand how this impact varied with the initial size of the local manufacturing sector. Our motivation for exploring this potential heterogeneity comes from "big push" or "poverty trap" theories of regional specialization and development [Rosenstein-Rodan, 1943, Murphy et al., 1989, Azariadis and Stachurski, 2005], in which coordination failures can result in a region being stuck in an equilibrium that features a low level of economic activity relative to its potential. Large temporary shocks to underdeveloped regions can induce a higher level of activity, which then becomes self-sustaining and permanent. This implication has been born out in previous research on WWII spending as well as in many (though not all) other contexts.[6]

We focus on a second important implication: temporary shocks can only have permanent effects on regions that are stuck in a low-activity equilibrium. For regions that are already highly developed, even a very large temporary shock will not affect its long run trajectory. This theoretical prediction of heterogeneity in the long-run response based on initial conditions has seen little or no empirical investigation, and is the focus of our analysis in this paper. Following a long tradition in the literature, we think of manufacturing as the sector generating increasing returns during this time period, and hence the size of the manufacturing sector as being the key initial condition for understanding heterogeneous responses to shocks.[7]

We draw on several additional implications of the theory to inform our empirical analysis. Since an underdeveloped manufacturing sector is a stable equilibrium, small shocks will have no long run effect even in underdeveloped regions. On the other hand, once a shock is large enough to move the region onto a trajectory leading to the higher equilibrium, further increases in the size of the shock will not alter the long run outcome, although the trajectory leading to that long run outcome will be different. These considerations suggest dichotomizing the war spending variable based on a threshold, dividing counties into "treated" (large shocks) and "untreated" (small or no

[6]For the U.S. experience in WWII, see, for example, Fishback and Jaworski [2016], Li and Koustas [2019]. Examples of studies in other contexts include Eckert et al. [2020b], Kline and Moretti [2014], Juhász [2018].

[7]See Bartelme et al. [2019] and Lashkaripour and Lugovskyy [2023] for some recent evidence on scale economies in manufacturing.

shocks) categories for the analysis. Similar reasoning suggests dichotomizing the initial conditions, dividing counties into "high" and "low" initial manufacturing employment (our measure of sector size). We therefore divide observations (counties) into four quadrants based on these criteria, then compare long run outcomes in the treated quadrants (both "high" and "low") to their untreated counterparts.

To be clear, we expect the conditional expectation of the treatment effect to vary continuously with both the initial conditions and the treatment size.[8] Our discretization strategy is dictated by the need to maximize the relevant information that can be obtained from limited data. Rather than estimate the entire two-dimensional surface of treatment effects, which is infeasible to do with any precision, we average the treatment effects over the most relevant domains by choosing appropriate thresholds that define the quadrants. The choice of thresholds is guided by both substantive and statistical considerations; we discuss our choices in more detail and conduct robustness checks in the following section, but we stress that there is no right or wrong way to choose these thresholds, only more or less informative ones. We also note that in no case are the treatment effects to be interpreted as structural parameters of some underlying model. We will use the estimated treatment effects as moments to be targeted when estimating our structural model in Section 3.4.

We follow Kline and Moretti [2014] and adopt the potential outcomes framework as our empirical model, with the Average Treatment Effect on the Treated (ATT) as our primary parameter of interest.[9] If county i actually received treatment, the county i treatment effect is the difference between the actual county i outcome and the (unobserved) counterfactual outcome for county i had it not received treatment. The ATT is simply the average of these individual treatment effects taken over all counties that actually received treatment. Formally, let y_i^0 be the outcome of interest (e.g., the long run change in manufacturing employment) in county i *if it had not received treatment*. We assume that

$$y_i^0 = \alpha_i^0 + \theta_0 X_i + \epsilon_i^0, \tag{3.1}$$

where X_i is a set of observable control variables and ϵ_i is an unobserved term. If county i *were to*

[8] Continuity is expected because the treatment size and initial manufacturing employment are not sufficient to identify which counties will be deflected from their counterfactual (untreated) long run equilibrium by the treatment. For example, some counties with low initial manufacturing employment will already be headed towards a "high" equilibrium and will thus have a treatment effect of zero for even large treatment sizes, even though most counties with the exact same initial manufacturing employment will have positive treatment effects for most treatment sizes. Integrating over unobserved heterogeneity leads to continuity of the conditional expectation despite the discrete nature of the transition from a low to high equilibrium.

[9] See Imbens and Wooldridge [2009] for a review of the literature on estimation and inference in this model. The ATT can be consistently estimated under weaker assumptions than the average treatment effect (ATE). It also has an appealing interpretation in historical counterfactual analysis, especially when general equilibrium effects are important, since it avoids considering very different policies than those actually implemented.

recieve treatment, the outcome y_i^1 would given by

$$y_i^1 = \alpha_i^1 + \theta_1 X_i + \epsilon_i^1. \tag{3.2}$$

Identification is achieved by assuming that both residuals are uncorrelated with the treatment conditional on the controls,

$$E[\epsilon_i^0 | X_i, D_i] = E[\epsilon_i^1 | X_i, D_i] = 0, \ \forall i. \tag{3.3}$$

The treatment effect for county i is given by $\alpha_i^1 - \alpha_i^0 + (\theta_1 - \theta_0)X_i$, and the ATT for each subgroup is obtained by averaging across treated counties in each subgroup. We allow for the possibility that the treatment effect is correlated with the observables, $\theta_0 \neq \theta_1$; if $\theta_0 = \theta_1$, the model collapses to a standard difference-in-differences regression. Equations (3.1) and (3.2) are estimated separately on the untreated and treated populations, respectively, and the estimates then combined to compute the ATT.[10]

The model (3.1) - (3.3) can be regarded as a local approximation to a structural model, such as the one we develop in Section 3.4. The controls X_i should include observables from the initial state that summarize all aspects of the local initial conditions useful in predicting the future evolution of local productivity and amenities. The controls should also account for WWII spending in nearby counties, which will affect outcomes in county i through spatial spillovers (e.g. those generated by trade and migration frictions). Conditional on the appropriate set of controls, the residuals ϵ_i reflect the portion of future local productivity/amenity growth that is unforecastable based on the initial state. Our baseline identification assumption is that the local WWII supply spending is uncorrelated with this unforecastable component of local productivity/amenity growth.

This raises the question of where our identifying variation is coming from, given that treatment is non-randomly assigned and that we control extensively for observables in the initial state. We think of our identifying variation as coming from three primary sources. One is the composition of local industry; two places with the same level of manufacturing employment (which we control for in the regression) may have very different wartime demand depending on how useful the products they specialize in are for the war effort (see Rhode et al. [2018] for some evidence). Since the composition of wartime demand is very different from the composition of peacetime demand, this variation is plausibly exogenous with respect to the future (long run) evolution of local productivity and amenities. The second source of variation is political, social or informational connections between wartime decisionmakers and particular geographic locations or firms, which again is plausibly exogenous with respect to long run changes in fundamentals. The third source is the winner-take-all nature of most contracts, which implies that there might be very little

[10] This is known as the Oaxaca-Blinder estimator [Oaxaca, 1973, Blinder, 1973]. See also Kline [2011] for some robustness properties of the estimator and Kline and Moretti [2014] for an application in a similar context to ours.

difference between a county that recieved a large contract and one that did not. This last source could be particularly influential for counties with an initially small manufacturing base, for which the allocation of a single large contract could have a material effect on output and employment.

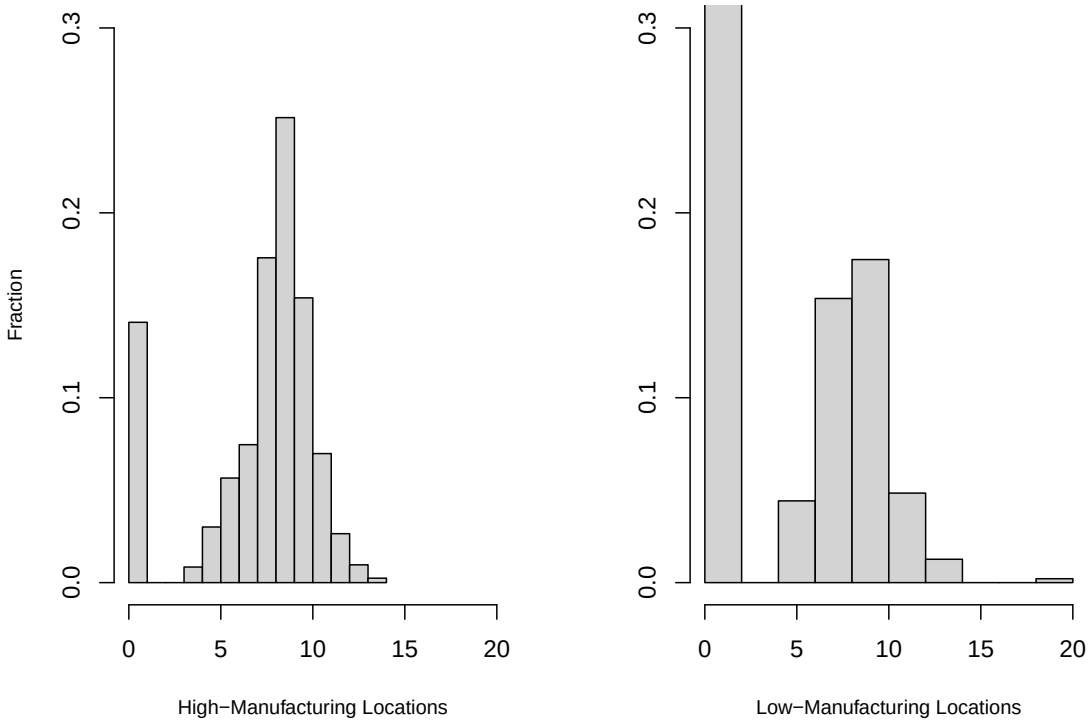
Admittedly, it is difficult to be completely confident that there is no confounding selection into treatment whenever the treatment is not explicitly randomly assigned. We address this concern, as best we are able, with two data exercises and a theoretical point. First, we use a placebo test to show that treatment status does not predict outcomes in the pre-treatment period. Second, we use data on industry composition to show that our results are robust to controlling for this source of variation. Finally, we note that our primary interest is in the *difference* in the ATT between locations with initially low and high manufacturing employment. Our identifying assumption (3.3) is sufficient to identify the levels of both treatment effects, but the difference in the ATTs is identified under the weaker assumption that selection bias equally affects both groups.

3.3 Empirical Results

Treatment and Initial Conditions We split the set of counties into those with an initially high manufacturing base—where log manufacturing employment in 1930 is above the 50th percentile nationally—and those with an initially low manufacturing base. We assign treatment status to each county based on the WWII war supply contract value per capita. Figure 3.2 plots the densities of log contract value per capita for locations with initially high and low manufacturing employment. The distributions differ primarily in the larger mass of initially low manufacturing base counties with small contract values; such allocations are relatively rare for counties with high initial manufacturing shares. The distributions are more comparable after conditioning on recieving some minimum contract value, although the right tail is somewhat thicker for the initially low manufacturing counties. Counties with a high level of war supply contracts were more densely populated and had higher manufacturing employment in the prewar period (see Table C.1 for summary statistics). This is not surprising because war supply contracts were allocated based primarily on existing production capacity [Fishback and Cullen, 2013].

However, these features do raise the question of how to choose the definition of treatment so as to balance the dual objectives of ensuring an adequate number of control units for both groups and having the treatment group assignment lead to similar distributions of contract value within the treatment and control groups, across initial conditions. Our baseline is to assign counties to the treatment group if their log contract value per 1930 manufacturing employee exceeds 8, a round number very close to the mean. We conduct extensive robustness checks on both cutoffs (initial conditions and treatment) and discuss them further below.

Figure 3.2: Distribution of Log Contract Value Per 1930 Manufacturing Worker



Notes: This figure shows the histograms of log average contract value per 1930 manufacturing worker across locations with initially high and low manufacturing employment. To ease comparison of the histograms, the fraction of locations with zero values is truncated at in the right panel.

Pretrend Test To adjust for pre-existing differences across counties before the WWII, we follow previous literature and control for a rich set of county-level variables related to prewar economic, social, demographic, and geographical characteristics in 1930 [Kline and Moretti, 2014, Li and Koustas, 2019]. The covariates include log population, log employment, log manufacturing employment, agricultural employment share, manufacturing employment share, log manufacturing wages, dummies for wages in manufacturing or trade being missing, farm values, urban population share, the share of white population, the share of the population age 5 or above that are illiterate, the share of whites who are foreign-born, the share of households with a radio, maximum elevation, elevation range, and area.[11] Since contract assignment is strongly correlated with

[11]To capture potential temporal correlation of shocks, we control for socioeconomic characteristics in 1930 and 1900 (except log population and log manufacturing employment in 1900 to prevent multicollinarity, as we also control for trends in these variables).

Table 3.1: Pretrend Tests of Outcomes in 1900-1930

	(1) Pop 1900-1930	(2) ManufEmp 1900-1930	(3) Pop 1900-1930	(4) ManufEmp 1900-1930	(5) Pop 1900-1930	(6) ManufEmp 1900-1930
Treat	−0.001	0.004	−0.004	−0.008	0.000	0.008
	(0.004)	(0.007)	(0.004)	(0.015)	(0.005)	(0.008)
Location	All	All	LowM	LowM	HighM	HighM
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows pretrend tests for the "treatment" effect on the growth rate between 1900 and 1930 (defined as the log difference between 1930 and 1900 levels divided by three) of the outcome variables. Columns (1)–(2) shows pretrend tests for the whole sample, columns (3)–(6) break them down by the initial levels of manufacturing employment: whether a county has manufacturing employment in 1930 below or above the 50th percentile. Treated counties correspond to those that received WWII war supply contract value per capita exceeds 8 in logarithmic terms. Block-bootstrapped standard errors at the state level are in parenthesis.

initial manufacturing conditions [Rhode et al., 2018], we also control for relevant variables in the manufacturing sector including log total manufacturing employment, manufacturing employment share, and log manufacturing wages. We also control for an indicator variable for whether a region is classified as having initially high manufacturing employment, as well as an interaction of this indicator with the log prewar population, log prewar manufacturing employment, andthe prewar manufacturing share. We also control for trends in population and manufacturing employment between 1900 and 1930. To account for potential spillover effects from neighboring counties, we control for the war supply contract value per capita of neighboring counties. To better balance the treated and control groups, we drop counties whose pre-WWII characteristics yield extreme predicated probability of treatment higher than 99% or lower than 1%.

We first assess whether the set of control variables perform well to adjust for observable differences across counties. To do so, we conduct a placebo test by estimating the "effects" of defense spending on 1900-1930 changes in population and manufacturing employment, for all counties together as well as separately for initially low and high manufacturing counties. This test attempts to detect if, conditional on the control variables, the outcome variables exhibit differential trends before WWII across treated and control counties. Because that time period predates the war, a finding of a significant relationship would signal a potential failure of our identification assumption. As shown in Table 3.1, no statistically or economically significant effects are discernable across all outcomes and subgroups.

Main Results Panel (a) of Table 3.2 shows the results of estimating equations (3.1) and (3.2) and computing the ATT across all observations, regardless of initial conditions. Columns (1)–(2) of Table 3.2 show the long-run effects of WWII spending on the growth of population and manu-

Table 3.2: Long-Term Impact on Growth Rate of Outcomes

Panel (a): Baseline						
	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat	0.010	0.016	0.023	0.060	0.002	−0.017
	(0.006)	(0.007)	(0.008)	(0.012)	(0.006)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Panel (b): Initial Condition						
	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat*LowM	0.029	0.038	0.029	0.087	0.028	−0.003
	(0.008)	(0.010)	(0.009)	(0.019)	(0.009)	(0.010)
Treat*HighM	0.005	0.010	0.021	0.053	−0.004	−0.021
	(0.007)	(0.008)	(0.009)	(0.013)	(0.007)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the impact of wartime spending on the growth rate of outcomes. Columns (1)–(2) show the treatment effect on the growth rate of outcome variables between 1930 and 2000 (defined as the log difference between 2000 and 1930 levels divided by seven). Columns (3)–(6) show the treatment effect on the growth rate of outcome variables by period: 1930-1960 and 1960-2000. Panel (a) shows the treatment effects for the whole sample, and panel (b) breaks them down by the initial levels of manufacturing employment: whether a county has manufacturing employment in 1930 below or above the 50th percentile. Treated counties correspond to those that received WWII war supply contract value per capita exceeds 8 in logarithmic terms. Block-bootstrapped standard errors at the state level are in parenthesis.

facturing employment between 1930 and 2000, with results expressed in terms of decadal growth rates. The impacts on the growth in population and manufacturing employment between 1930 and 2000, with the latter date being over 5 decades after the war ended and wartime spending subsided, are 1% and 1.6% per decade, respectively. These modest positive long run averge effects are consistent with the previous literature on the geographic impact of WWII spending [Fishback and Cullen, 2013, Fishback and Jaworski, 2016, Li and Koustas, 2019]. For the immediate post-war period—up until 1960—we see much larger differential growth rates of 2.3% and 6% for population and manufacturing employment in treated counties (columns (3)-(4). However, these sizable impacts die out or are partially reversed over the next several decades (columns (5)-(6)). This dynamic pattern of growth and (partial) reversal reflects the temporary nature of the WWII demand shock and the gradual adjustment of treating counties to the loss of wartime demand for their products.

Panel (b) of Table 3.2 reports the ATT separately for counties with high (*HighM*) and low (*LowM*) initial manufacturing bases. For regions with initially low manufacturing employment,

wartime defense spending increased the long term growth rate of population and manufacturing employment levels by 2.9% and 3.8%, respectively, between 1930 and 2000. For regions with initially high manufacturing employment, these effects are much smaller both economically and statistically: the long run effects are 0.5% and 1.0%, respectively, with both statistically indistinguishable from zero. Columns (3)-(6) report the dynamic patterns, which show that both treated groups experienced similar large positive effects on both population and manufacturing between 1930 and 1960. However, these positive impacts almost completely reverse over the next 40 years for the initially high manufacturing group, leaving the long run effects for this group close to zero. In contrast, the initially small treated locations experienced little or no reversal from 1960-2000, leaving their population and manufacturing employment significantly higher after 60 years, relative to the controls.

Our results show that the strong positive long run effects of WWII spending on population and manufacturing employment are found only for regions with initially small manufacturing sectors, as predicted by "big push" or "poverty trap" models. The dynamic pattern of the effects also supports the "big push" theory; it is quite reassuring that the short term effects are similar for both groups, since we would expect a large temporary expansion in employment due to the demand shock regardless of any endogenous productivity effect. However, as the demand shock recedes only the locations that have been pushed to a higher productivity state retain significant employment gains, while they largely fade away for locations that were not.

Additional Results and Robustness Checks Table 3.3 breaks the treatment size into two categories, with the log contract value per 1930 manufacturing employment above or below median, and re-estimates baseline regression equations (3.1) and (3.2). The results are as expected; larger treatments lead to larger short (reflecting the larger direct effect) and long (reflecting the higher probability of being pushed onto a higher productivity trajectory) term gains for all groups. The long run effects continue to be much larger for the initially small manufacturing group.

Figure 3.3 shows how varying the cutoffs for the initial high/low split as well as the treatment size affects the results. Classifying more locations as "small" and less as "large" results in a monotonic decline in both coefficients, with the effects for "large" locations turning negative as the category comes to include only the very largest locations. The spread between the estimates is fairly robust. These results suggest that it is indeed the smaller percentiles of the initial conditions that are driving the positive long run effect. The results are also not too sensitive to the definition of treatment.

As mentioned above, our identifying variation comes from 3 primary sources: the composition of local industry, locally specific political and informational connections, and true randomness from the lumpy and winner-take-all nature of the contracts. The only observable element is the

Table 3.3: Long-Term Impact on Growth Rate of Outcomes by Treatment Size

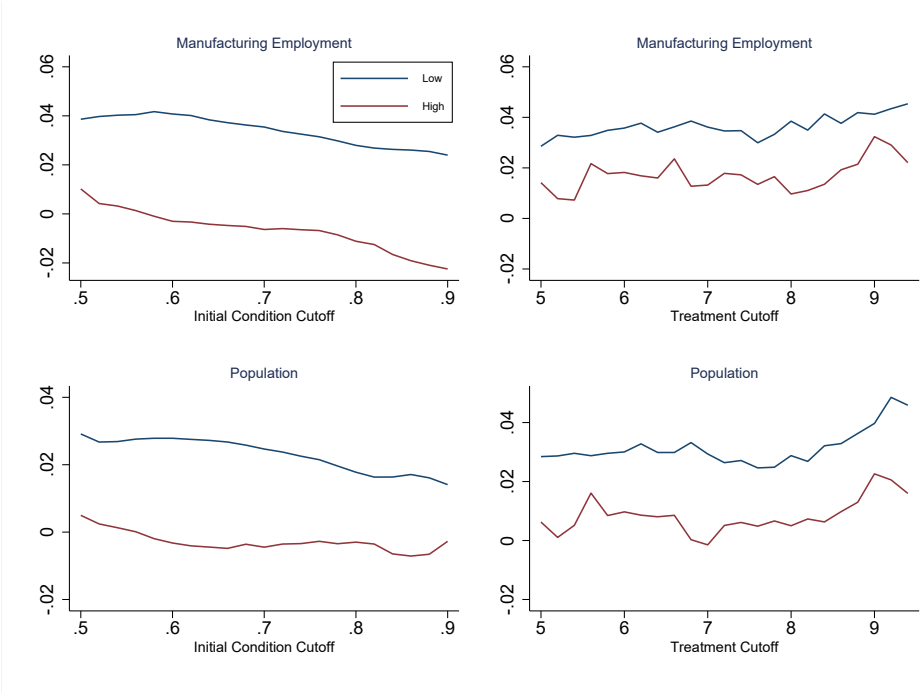
	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat Big*LowM	0.050	0.052	0.042	0.117	0.056	−0.011
	(0.017)	(0.016)	(0.019)	(0.022)	(0.015)	(0.017)
Treat Small*LowM	0.014	0.029	0.021	0.067	0.009	0.001
	(0.008)	(0.011)	(0.011)	(0.022)	(0.010)	(0.013)
Treat Big*HighM	0.014	0.022	0.038	0.091	−0.001	−0.025
	(0.009)	(0.013)	(0.013)	(0.022)	(0.008)	(0.011)
Treat Small*HighM	−0.000	0.001	0.009	0.027	−0.007	−0.017
	(0.007)	(0.008)	(0.008)	(0.011)	(0.007)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows the impact of wartime spending on the growth rate of outcomes by treatment size. Columns (1)–(2) show the treatment effect on the growth rate of outcome variables between 1930 and 2000 (defined as the log difference between 2000 and 1930 levels divided by seven). Columns (3)–(6) show the treatment effect on the growth rate of outcome variables by period: 1930-1960 and 1960-2000. Low and high initial manufacturing base is defined by being above or below the 50th percentile of manufacturing employment in 1930. Treated counties correspond to those that received WWII war supply contract value per capita exceeds 8 in logarithmic terms. The boundary between "small" and "large" treatment is defined by a per capita supply contract value of XXX, which is the median value in the treated group. Block-bootstrapped standard errors at the state level are in parenthesis.

composition of local industry, and we now explore how using (or not using) this information affects our estimates. We divide manfuacturing sectors into those highly related to the war effort (e.g. auto, ship and aircraft manufacturing) and all others. Table C.2 shows that the share of prewar manufacturing employment in these sectors is highly correlated with both treatment status and contract value. It is plausible that this industry composition is also related to the future evolution of productivity and amenities, despite our other controls, which would violate our exclusion restriction. Therefore, as a robustness check Table C.3 reports the results of re-running our basic specification while controlling for industry composition. Re-assuringly, the results are very similar to those reported in Table 3.2.

Finally, we discuss the concern that our results are driven by possible serial correlation of military spending. While military spending dropped precipitously at the end of the war, it did not drop all the way to the prewar level and it could be that locations with high wartime spending disproportionately benefited from this continued spending. We conduct two tests to address the concern. First, we show that long-term effects, especially for locations with initially low manufacturing, are present in across different categories of contract spending including those meant for civilian use. Specifically, we break contract spending data into three coarse categories as follows: "food, textile, chemicals, miscellaneous," "metal," and "machinery, electronics, transportation." As shown in Table C.5, many of the key coefficients are statistically significant for all categories of contract

Figure 3.3: Coefficients with Varying Cutoffs of Treatment and Initial Conditions



Notes: This figure shows the coefficients as the cutoffs of treatment and initial condition definitions vary. The left panel correspond to estimates as the cutoff of classifying initially high- and low-manufacturing locations ranges from 50th to 90th percentiles. The right panel correspond to estimates as the treatment cutoff varies from 4 to 10 in log contract value per 1930 manufacturing employee.

spending. Second, we show that the correlation between contract spending and defense spending in later, post-war periods (1966 or 2000) at state level is quite low and insignificant, as shown in Table C.6.[12]

3.4 Quantitative Model and Structural Estimation

3.4.1 Quantitative Model

We build on the recent framework of Allen and Donaldson [2023], which endows the basic quantitative economic geography model with a dynamic structure that can accommodate multiple steady states and path dependence. We add several features: a) multiple sectors (agriculture, manufacturing, services), b) intermediate goods, c) a government sector, and d) local external economies of scale in the manufacturing sector that may affect productivity with non-constant elasticity. The first

[12]We analyze this correlation at the state level because reliable military spending data are only available at the state level in post-war years from 1960s [Nakamura and Steinsson, 2014].

three modifications help match the specific experiment in our setting or add quantitative realism, and our modelling of these elements is standard. The fourth element, external economies of scale with con-constant elasticity, is motivated by our empirical findings, as we further explain below.

Our structural estimation will show that external economies of scale that are diminishing in sector size beyond a threshold value are critical for matching the empirically observed difference in treatment effects across initial conditions. Some intuition for why this is the case can be gleaned from considering the case of a single competitive sector in partial equilibrium, with labor l the only factor of production, facing a flat labor supply curve and constant elasticity sector-level product demand with elasticity $-\sigma < -1$. The elasticity of the sector-level labor demand curve is $-\frac{\sigma}{1+(1-\sigma)\gamma(l)}$, where $\gamma(l)$ is the elasticity of productivity with respect to sectoral employment which we assume is driven by external economies. When $\gamma(l)$ is constant, then (ignoring the case of equality) either $\gamma(l) > \frac{1}{\sigma-1}$ and the labor demand curve is always downward sloping or $\gamma(l) < \frac{1}{\sigma-1}$ and the labor demand curve is always upward sloping. In the first case there is a unique equilibrium that is interior and stableand hence there is no "poverty trap." In the second scenario, there are two equilibria, a stable one at zero employment and an unstable interior equilibrium. The zero employment equilibrium could be described as a poverty trap, since a temporary shock large enough to carry employment past a threshold would result in a permanently higher level of employment and productivity. However, this model predicts that locations with with *lower* initial employment should be *less* responsive to positive demand shocks in the long run than large places, the opposite of our empirical findings. More generally, the problem with this model of a poverty trap is that it is too stylized for quantitative work, since the only two long run outcomes are an employment level of zero or infinity.

In contrast, consider an S-shaped scale function that is initially flat in the sector size, increases rapidly over some range, then flattens again as the sector grows large. Given downward sloping product demand, this type of scale function can generate a sectoral labor demand curve that crosses the labor supply curve three times in the interior, implying a stable "low" equilibrium and a stable "high" equilibrium. A large enough temporary shock can move the sector from the "low" to the "high" equilibrium, which then becomes self-sustaining; however, an economy that is initially in the high equilibrium always returns to the same steady state in response to a positive shock. This framework thus naturally implies differential responses to temporary shocks based on initial conditions.

This analysis is too simplistic in that it ignores general equilibrium considerations. As a matter of fact, the constant-elasticity model of Allen and Donaldson [2023] is capable of generating multiple stable interior steady states in general equilibrium, in both the original and our version. However, we find that the partial equilibrium *tendency* towards extreme sectoral allocations still manifests itself in general equilibrium, at least in our calibration. Thus the constant elasticity spec-

ification tends to imply that either a) no location is in a poverty trap, or b) *every* location is in a poverty trap, unless the initial equilibrium is one that is already extremely (and counterfactually) highly concentrated (a "black hole" in the terminology of Fujita et al. [2001]). This feature is in conflict with the empirical evidence from Section 3.3. For this reason, our estimation procedure strongly prefers scale economies that are sufficiently diminishing with large sector sizes.

Setup

Consider $i \in \{1, ..., N\}$ locations and discrete time $t \in \{0, 1, ...\}$. Individuals live for two periods. In the first period ("childhood"), they are born in their parents' location and consume the same as their parents. At the beginning of the second period ("adulthood"), they realize their own preferences and choose their residence as adults. In this location, they supply a unit of labor inelastically, consume, and have ζ_t children apiece.

Production

There are two tradable sectors, "agriculture" and "manufacturing," and a non-tradable sector, "services." Each location produces a differentiated variety of agricultural and manufactured goods. Agriculture and services use labor with constant returns to scale,

$$Q_{A,i,t} = T_{A,i,t} \cdot l_{A,i,t}, \tag{3.4}$$

$$Q_{S,i,t} = T_{S,i,t} \cdot l_{S,i,t}, \tag{3.5}$$

where T_A and T_S are exogenous productivities and l_A and l_S are labor inputs. Manufacturing uses labor and intermediate inputs from the manufacturing sector. Producers perceive constant returns, but external economies of scale exist in total manufacturing labor. Manufacturing output is given by

$$Q_{M,i,t} = T_{M,i,t} f\left(l_{M,i,t-1}\right) \cdot l_{M,i,t}^{\alpha} q_{M,i,t}^{1-\alpha}, \tag{3.6}$$

where T_M is exogenous productivity, $f(l)$ is the endogenous productivity depending on lagged sector scale, l_M is labor input, and q_M is a bundle of intermediate inputs.

Government

Government purchases manufacturing output from each location, then throws it away. Aggregate government consumption is

$$G_t = \sum_i g_{i,t}, \tag{3.7}$$

where $g_{i,t} = \kappa_{i,t} G_t$ are government purchases from location i. To fund these purchases, the government imposes a per-person lump-sum tax satisfying $\tau_t L_t = G_t$.

CES Demand

Preferences are given by a nested CES utility function, an upper tier across sectors and a lower tier across locations within sectors,

$$C_{i,t} = \sum_{s=\{A,S,M\}} \rho_s^{\frac{1}{\eta}} C_{s,i,t}^{\frac{\eta-1}{\eta}}, \tag{3.8}$$

$$C_{s,i,t} = \sum_{n=N} C_{s,i,n,t}^{\frac{\sigma-1}{\sigma}}, \tag{3.9}$$

where $C_{s,i,n,t}$ is the consumption in i of the sector s good produced in location n at time t. The lower tier elasticity of substitution σ is assumed to be the same for each sector, as well as for intermediate goods demand in the manufacturing sector. Let

$$\omega_{s,i,t} = \rho_s \left(\frac{P_{s,i,t}}{P_{i,t}} \right)^{(1-\eta)}, \forall s \in \{A, S, M\}, \tag{3.10}$$

where $\omega_{s,i,t}$ is the share of total expenditure $E_{i,t} = w_{i,t} L_{i,t} - \tau_t L_{i,t}$ on goods from sector s, $P_{s,i,t}$ is the sector s price index, and $P_{i,t}$ is the overall price index. Then expenditure on each sector is given by

$$E_{s,i,t} = \omega_{s,i,t} \left(w_{i,t} L_{i,t} - \tau_t L_{i,t} \right), \forall s \in \{A, S, M\} \tag{3.11}$$

where $L_{i,t}$ is the total population at location i and time t.

Goods market clearing and trade

There are three sources of demand for each good: final demand by consumers, intermediate demand, and final demand by the government. Manufactured goods trade incurs iceberg costs $\tau_{i,n,t}$.

Under perfect competition in input markets, total manufacturing wages are proportional to total sales,

$$w_{i,t}l_{M,i,t} = \alpha \cdot \left[c_{M,i,t}^{1-\sigma} \cdot \left(\sum_{n \in N} \tau_{i,n,t}^{1-\sigma} \cdot \frac{E_{M,n,t} + \frac{1-\alpha}{\alpha} w_{n,t} l_{M,n,t}}{P_{M,n,t}^{1-\sigma}} \right) + \kappa_{i,t} G_t \right], \tag{3.12}$$

where $c_{M,i,t} = \frac{w_{i,t}^{\alpha} P_{M,i,t}^{1-\alpha}}{f(l_{M,i,t}, l_{M,i,t-1}) \cdot T_{M,i,t}}$ is proportional to the unit cost in manufacturing and $P_{M,j,t} = (\sum_i (\tau_{i,j} c_{M,i,t})^{1-\sigma})^{\frac{1}{1-\sigma}}$ is the intermediate goods price index. Agricultural goods are freely traded, and their market clearing conditions are

$$w_{i,t}l_{A,i,t} = \left(\frac{w_{i,t}}{T_{A,i,t}} \right)^{1-\sigma} \cdot P_{A,t}^{\sigma-1} \cdot \sum_{n \in N} E_{A,n,t}, \forall i, t, \tag{3.13}$$

where $P_{A,t} = (\sum_i (c_{A,i,t})^{1-\sigma})^{\frac{1}{1-\sigma}}$ is the price index for agricultural goods, with unit costs $c_{A,i,t} = \frac{w_{i,t}}{T_{A,i,t}}$. Finally, market clearing in the services sector yields

$$w_{i,t}l_{S,i,t} = E_{S,i,t}, \tag{3.14}$$

with $P_{S,i} = c_{S,i} = \frac{w_{i,t}}{T_{S,i,t}}$.

Migration

Adults in i at $t-1$ have ζ_t children. At time t, children make their migration choice to maximize their adult welfare. Utility depends on location-specific amenity $H_{i,t}$, migration frictions $\xi_{i,n,t}$, and idiosyncratic preference $\epsilon_n(\ell)$,

$$U_{i,n,t}(\ell) = H_{n,t} \cdot \xi_{i,n,t} \cdot \frac{w_{n,t} - \tau_t}{P_{i,t}} \cdot \epsilon_n(\ell) \equiv U_{i,n,t} \epsilon_n(\ell). \tag{3.15}$$

We assume that $\epsilon_n(\ell)$ is distributed Fréchet with shape parameter θ, and thus the probability of choosing n is $\pi_{i,n,t} = P(U_{i,n,t} \geq \max_{j \neq n} U_{i,j,t})$. Note that $P(U_{i,n,t}(\ell) \leq u) = exp[-U_{i,n,t}^{\theta} u^{-\theta}]$. The migration flow from i to n at time t is

$$L_{i,n,t} = \zeta_t \frac{U_{i,n,t}^{\theta}}{\sum_{j \in N} U_{i,j,t}^{\theta}} L_{i,t-1}. \tag{3.16}$$

Dynamic equilibrium

Definition 3. Given initial conditions $\{L_{n,0,}, l_{M,i,0}, l_{A,i,0}\}$, an equilibrium is a sequence of wages and labor allocations such that:

1. Total sales are equal to payments to labor in each sector as given by equations (3.12) and (3.13).
2. Government budget balance is satisfied.
3. Migration flows are given by equation (3.16).
4. Local and global labor markets clear:

$$\begin{aligned} L_{i,t} &= l_{M,i,t} + l_{A,i,t} + l_{S,i,t}, \forall i, t, \\ \sum_{n \in N} L_{n,t} &= L_t = \zeta_t L_{t-1}, \forall t. \end{aligned}$$

The dynamics of the model are driven by costly migration and by the lagged effect of scale on productivity. Because the scale effects operate with a lag and there is no forward looking behavior, the within-period equilibrium conditions at time t are the same as those of a static model without scale effects, and hence we expect that there is always a unique equilibrium path from every initial allocation.[13] However, depending on the parameter values there may be multiple long-run steady states depending on the initial allocations [Allen and Donaldson, 2023].

3.4.2 Estimation using indirect inference

We calibrate the majority of the parameters of the model to standard values in the literature (Table 3.4) and use indirect inference to estimate parameters of the external economies of scale function $f(l)$. As described earlier in the section, we conjecture that matching the empirically observed difference in treatment effects across locations according to the initial size of the manufacturing sector may require $f(l)$ to diminish in sector size. We parameterize the scale function as

$$f(l) = \begin{cases} 250^{\gamma} & l_{Mit-1} \leq 250, \\ l_{Mit-1}^{\gamma} & 250 < l_{Mit-1} < l_*, \\ l_*^{\gamma} & l_{Mit-1} \geq l_*, \end{cases} \tag{3.17}$$

[13] Rigorous theoretical conditions for unique equilibria have not been established for this type of static multisector economic geography model without scale economies. However, both theoretical priors and practical experience suggest that it is reasonable to expect uniqueness.

where γ is the slope parameter and l_* is the location parameter. This function approximates an S-curve in manufacturing employment: initially flat, increasing over some intermediate range, and flat again after local manufacturing scale exceeds l_*.[14] This is a simple 2-parameter family that nests the standard constant elasticity case when the location parameter goes to infinity yet is flexible enough (as we will show) to match the empirical moments. It would certainly be ideal to estimate a more flexible functional form, but our current specification is adequate to our purpose and we do not think the data will support robust inference on additional parameters.

We estimate the slope and location parameters using classical minimum distance, sometimes referred to as "calibration with standard errors." We minimize the distance between the empirical ATTs for long term manufacturing employment, for initially low and high manufacturing locations (Table (3.2), column 2) and the corresponding moments generated by the model,

$$Loss = \left(\beta^{model} - \beta^{data}\right)\left(\beta^{model} - \beta^{data}\right)'. \tag{3.18}$$

We generate the model moments using the following procedure. First, we guess the parameters of the scale function. Next, we invert the model to obtain the exogenous local sectoral TFPs and amenities that rationalize the observed equilibrium in 1930. Then we shock the model with the empirically observed temporary local war spending shocks $\{\kappa\}_i$ for each location simultaneously, and solve forward for the steady state holding all exogenous TFPs and amenities constant. This establishes the "treatment" scenario for each location. Then, for each treated location i, we return to the observed equilibrium in 1930 and set $\kappa_i = 0$, but retain all treaments in other locations at their observed values and solve the model forward again to obtain the control observation for i, i.e. the outcome for county i in the absence of treatment in county i, all else equal. Finally, we compute the treatment effect for i as the difference between county i's manufacturing employment in the two scenarios; averaging across treated counties in each group (initially high or low manufacturing employment) gives the model-implied ATT for each group (β^{model}).

We systematically search the parameter space by repeating this procedure for different pair of parameters (scale and location) using a combination of global and local methods. Once we arrive at the pair that minimizes equation (3.18), we compute standard errors for the slope and location parameters (collected in the vector ε) using the delta method [**?**] ,

$$Z = (G'G)^{-1}G'VG(G'G)^{-1}, \tag{3.19}$$

[14]The initially flat segment plays a dual role. First, it ensures that the marginal revenue product of labor in manufacturing tends to infinity as local manufacturing employment tends to zero in each county, just as in a model with no scale economies, ensuring positive manufacturing employment in every location. Second, it avoids placing weight on the large number of counties with very small reported manufacturing employment when calibrating the model. We suspect these entries are severely contaminated with measurement error. Our results are robust to varying the cutoff for the initial flat segment.

Table 3.4: Summary of calibration

Parameter name	Parameter Notation	Value	Source
Migration elasticity	θ	3.3	
Trade elasticity	σ	8	
Manufacturing Cobb-Douglas input share	α	0.5	
EoS across sectors	η	0.33	
Sectoral share parameter	ρ_s	See notes.	Add up to 1
Population growth	ζ_t	See notes.	Population data
Slope parameter in scale function	γ	0.19	Indirect inference
Location parameter in scale function	l_*	1445	Indirect inference
Calibrated parameter in scale function	.	250	See notes

Notes: The sectoral share parameters are $\rho_{A,t} = [0.11, 0.07, 0.05, 0.03, 0.02, 0.02, 0.025, 0.019]$. Moreover, $\rho_{S,t} = [0.046, 0.056, 0.088, 0.129, 0.192, 0.254, 0.385, 0.579]$, and $\zeta_t = [1.208, 1.067, 1.156, 1.128, 1.209, 1.234, 1.224, 1.290]$. We calibrate the first location parameter of the scale function based on the average manufacturing share multiplied by the p10 of population in 1940.

where Z is the variance-covariance matrix of $\hat{\varepsilon}$, V is the variance-covariance matrix of the empirical moments (computed via block bootstrap) and $G = \partial\beta^{model}/\partial\varepsilon'$, which we compute numerically.

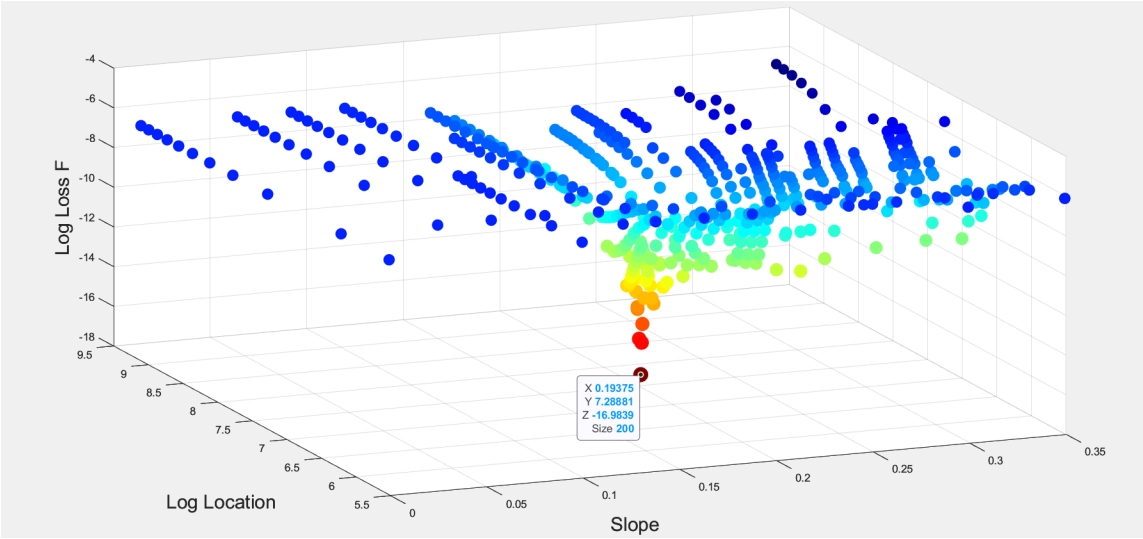
Since our estimation procedure is based on comparing steady states with identical exogenous parameters that differ only by their history, by construction any parameter combination that does not exhibit long run path dependence in response to the empirical shocks will have $\beta^{model} = 0$, which is quite far from matching the empirical moments. Thus estimation is a matter of finding a parameter vector that generates the correct empirically observed pattern of path dependence, one that primarily affects counties with initially low manufacturing employment. Speaking loosely, the size of the scale elasticity γ is almost exclusively responsible for the existence or non-existence of path dependence in the model, while the location parameter l_* primarily controls which counties exhibit path dependence in response to shocks. Both parameters influence the size of the long run ATT for those counties that exhibit path dependence, which also depends on the empirical distribution of the shocks.

Calibration and data

To invert the model we require data on $\{w_{i,t}, l_{A,i,t}, l_{M,i,t}, l_{S,i,t}, G_t, \kappa_{i,t}\}_{t\in\mathcal{T}}$ and knowledge of some parameters.[15] We calibrate the trade and migration elasticity to $\sigma = \theta = 6$, and the labor share to $\alpha = 0.5$. We set $\eta = 0.33$. Furthermore, we set the sectoral share such that aggregate labor supply adds up to aggregate labor demand in each sector, ρ_s. Then, we set population growth as $\zeta_t = \frac{\sum_i L_{it}}{\sum_i L_{it-1}}$. Finally, we set trade costs as

[15] See Appendix C.2 for details of the model inversion.

Figure 3.4: Unrestricted Loss Function



Notes: This graph plots the loss (equation (3.18)) as a function of the slope and location parameters.

$$\tau_{i,n,t} = \begin{cases} \tau_F distance_{in}^{\frac{1}{\sigma-1}} & i \neq n, \\ \tau_F & i = n, \end{cases} \tag{3.20}$$

where $\tau_F = \frac{0.3}{\sigma-1}$, and migration costs as

$$\xi_{i,n,t} = \begin{cases} distance_{in}^{-\frac{1}{\theta}} & i \neq n, \\ 1 & i = n \end{cases} \tag{3.21}$$

where the distance between i and n is defined in miles. Table 3.4 displays our calibrated parameters. Table **??** displays the variables we require for the model inversion along with their data counterparts and the sources.

The data used for simulation come from the same sources as those described in Section 3.2.1. Since we need wage information to invert the structural model, we impute missing wages for the simulation sample using a dynamic panel regression specification. The analysis is based on a harmonized county-level data set of war supply contract values, manufacturing employment, and wages for the decadal years from 1900 to 2000. Additional details are documented in Online Appendix C.1.

Table 3.5: Indirect Inference

Parameter	Estimate
Slope: γ	0.19
	(0.02)
Location: L_*	1464
	(970)

Notes: See text for details.

3.4.3 Results

Figure 3.4 plots the loss as a function of the slope and location parameters. We achieve a loss that is numerically indistinguishable from zero at the minimum, matching the empirical estimates exactly up to the fourth decimal place. We can see that both local and global identification is fairly strong in the depicted region of the parameter space, which we verify holds outside the depicted region as well.

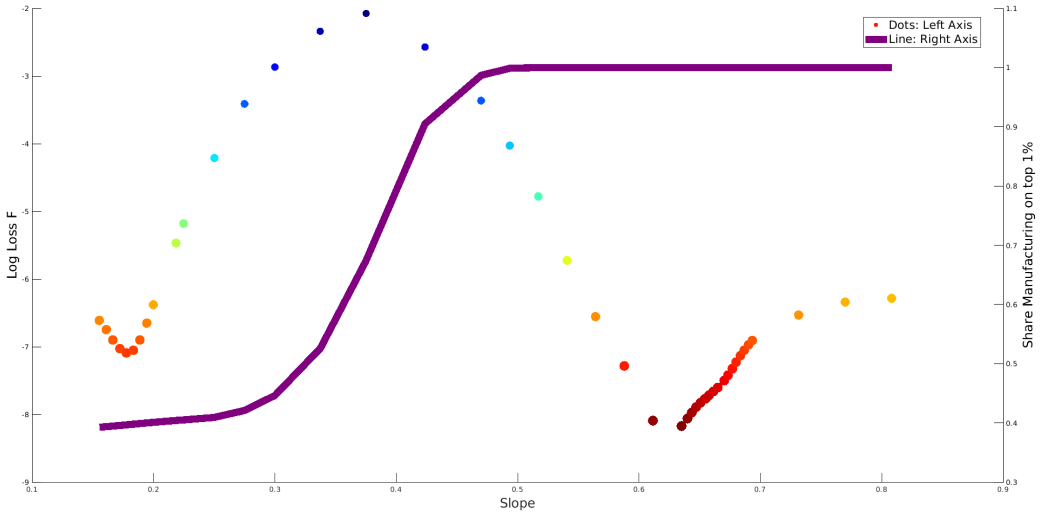
Table 3.5 shows our parameter estimates and standard errors. We estimate a slope parameter $\gamma = 0.19$ and a location parameter $L_* = 1464$. The location parameter is around the 66^{th} percentile of the county manufacturing employment distribution in 1940. Our estimates therefore imply that manufacturing productivity grows rapidly, about 140% in total, as a county moves from the 10^{th} percentile to the 66^{th} percentile of the manufacturing employment distribution, with no growth in productivity for further increases in the manufacturing labor force.

It is difficult to make a clean comparison of our estimate of γ to those in the literature. First, we are not aware of any estimates of a local scale function at this level of geographic and sectoral disaggregation. Second, estimation of "similar" functions, for example scale elasticities in manufacturing at the country level [Bartelme et al., 2019, Lashkaripour and Lugovskyy, 2023], assume a constant elasticity function which is not comparable to ours. Third, as is common in the quantitative literature, our estimation is sensitive to other details of the model calibration such as the value of the trade elasticity and our suppression of other potentially relevant agglomeration or congestion forces. Rather than focus on the specific estimated scale function, which is after all only an approximation that may be highly context-specific, we find it more illuminating to focus on how the estimated scale function fits with the rest of the model to match the observed empirical treatment effects.[16]

Due to its prominence in the literature, it is interesting to evaluate the ability of the constant elasticity model to match the empirical moments. Figure 3.5 plots the log loss function (left axis) against the scale elasticity, with the location parameter restricted to be infinite (the constant elas-

[16]With that said, our value of γ fits comfortably in the range of manufacturing scale elasticities at the country level estimated in Bartelme et al. [2019] and Lashkaripour and Lugovskyy [2023].

Figure 3.5: Restricted Loss Function



Notes: This graph plots the loss in equation (3.18) as a function of the slope parameter, assuming an infinite location parameter.

ticity case). This restricted log loss function has a local minimum at around $\gamma \approx 0.19$, which is our unrestricted estimate, and a global minimum at $\gamma \approx 0.6$. The fit, however, is very poor compared to what we achieve with a finite location parameter; indeed, the lowest loss in Figure 3.5 is about the same as the highest loss displayed in Figure 3.4. Moreover, the nature of the long run equilibrium as γ increases beyond 0.4 or so in the constant elasticity model is "black hole"-like. Figure 3.5 also plots the share of manufacturing accounted for by the top 1% of counties in the steady state, which moves from being close to the actual 1930 value of about 40% to about 100% when $\gamma > 0.4$. In fact, for larger values of γ our simulations show that 3 counties (Monroe County, MS, Gibson County, TN, and Dinwiddie County VA) account for essentially all of the nations manufacturing employment in the steady state.[17] We conclude that the constant elasticity model is unable to come close to matching the estimated heterogeneity in the long run impact of WWII spending.

In conclusion, we find strong evidence that, at least in this setting, the domain over which scale economies that are strong enough to generate path dependence is fairly limited.

[17] In contrast, at our optimal parameter estimates the top 1% share in the steady state is around 40%, as in the data.

3.5 Counterfactual Analysis

In this section, we use our estimated model to re-examine the impact of WWII spending on the long run evolution of U.S. economic geography. There is a sizable literature in history and economics on the topic, with some arguing that the WWII spending catalyzed long run manufacturing growth in the South (e.g., Wright [1986], Hooks and Bloomquist [1992]) or the Pacific West (e.g., Rhode [2003]). Others have pointed out that the manner in which spending was allocated, primarily towards counties with a large established manufacturing base, could have had the opposite effect of reinforcing existing patterns of specialization [Fishback and Cullen, 2013, Rhode et al., 2018, Fishback and Jaworski, 2016]. However, this literature has not used quantitative economic geography models to evaluate the impact, nor (at least not formally) entertained the possibility that the spending had different long run impacts on different regions, depending on initial conditions. Our exercise below does both.

Specifically, we compare the actual "treated" world given by the data from the years 1930-2000 to a counterfactual world in which the WWII spending did not occur, but all other exogenous variables evolved as they did in the data. We then simulate the model (without spending) forward to the year 2000 to generate the counterfactual equilibrium. We then compare the actual and counterfactual equilibrium outcomes in the year 2000.

Table 3.6 highlights the key finding: temporary World War II contracts had a long-term impact, inducing some counties to escape a poverty trap, but these effects were not consequential at the aggregate level. Columns 1–2 show the aggregate and average manufacturing share across counties Comparing the data (row 1) with the counterfactual (row 2) show no change in the overall share. At the individual county level, the average absolute difference in manufacturing employment (column 3) is a bit over 1%, indicating that at least some counties experienced long-term effects. Column 4 shows no major shifts in the concentration of manufacturing in rural areas. Finally, columns 5-6 depicts no effect on real income per person and ex-ante welfare per person.[18] Overall, the long-term impacts on manufacturing shares across counties were present, but not consequential for the aggregate.

We shed further light on this finding by examining the changes at the individual county level. Figure 3.6 plots the log difference in manufacturing employment between actual and counterfactual for each county, against the initial manufacturing employment in 1930. Unsurprisingly, the differences are primarily found in counties with under1500 manufacturing employees initially. Many differences are negative, implying that the long run growth of manufacturing in some smaller counties came at the expense of other small counties where manufacturing shrank. The lack of an aggregate impact implies that positive and negative effects roughly balanced out. While efficiency

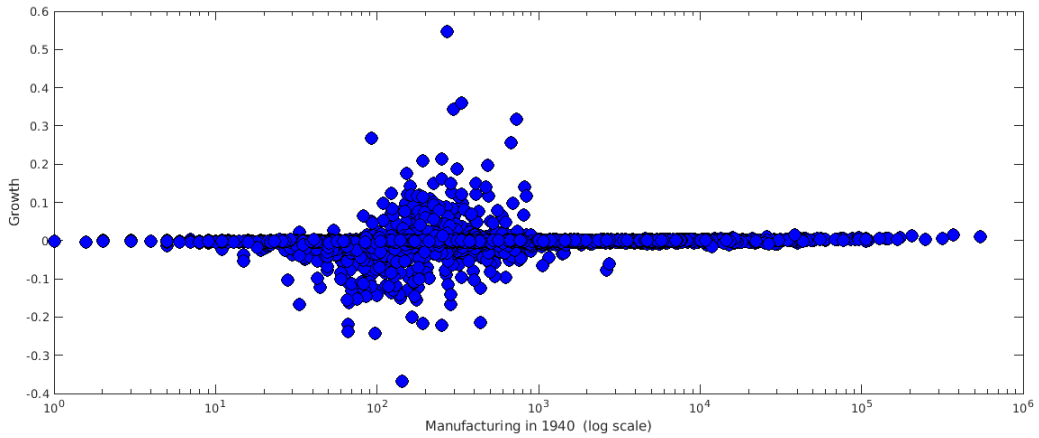
[18] Ex-ante welfare is defined as $\Phi_{i,t} = (\sum_{j \in N} U_{i,j,t}^{\theta})^{1/\theta}$.

Table 3.6: Aggregate Differences in Manufacturing Growth in 2000

	Agg. Manuf. Share	Avg. Manuf. Share (County)	Avg. Absolute Value of % Difference	Rural Share Manuf. Emp.	% Change in Avg. Real Income	% Change in Avg. Ex-ante Welfare
Data	0.1227	0.1380	.	0.1107	.	.
Counterfactual	0.1227	0.1376	0.0124	0.1103	0.00045	0.00020

Notes: This table depicts aggregate statistics. Real income is wages divided by the price index. Ex-ante welfare is given by $\Phi_{i,t} = (\sum_{j \in N} U_{i,j,t}^{\theta})^{1/\theta}$. The last two rows show the average change in real income and ex-ante welfare per person.

Figure 3.6: Long-term Impact of WWII Manufacturing Contracts on Manufacturing Employment Across U.S. Counties

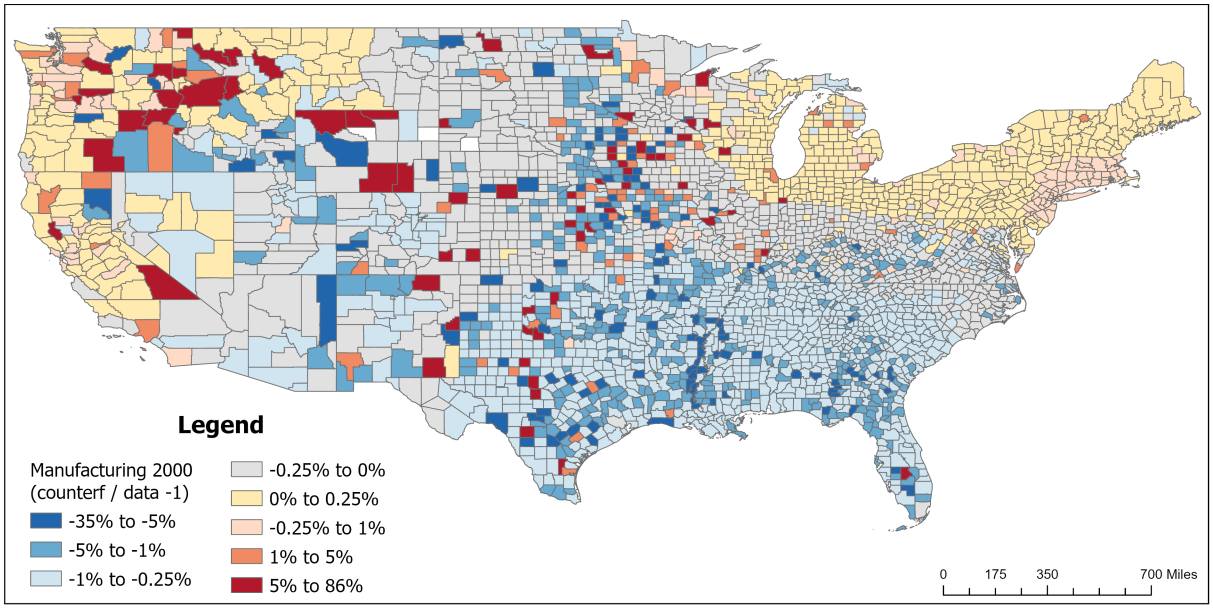


Note: This scatter compares the outcome in the treated world versus the untreated world in 2000 across counties.

considerations would tend to dictate an overall reallocation of manufacturing activity to smaller counties and a reduction in the overall manufacturing employment share (due to the EoS across sectors η being smaller than 1), the WWII spending shock did not catalyze much of a shift in this direction. This is of course consistent with the historical evidence that regional reallocation of manufacturing activity was not an objective of the wartime contract awards.

The map in Figure 3.7 highlights regional shifts in manufacturing employment. Overall, the Northeastern and Midwestern "manufacturing belt" say modestly increased manufacturing employment, consistent with the view that the WWII spending actually (modestly) reinforced their regional comparative advantage in manufacturing. Most counties in the South experienced declines in manufacturing employment relative to the counterfactual world. Regionally, the largest growth in manufacturing employment came in the Pacific Northwest, consistent with Rhode [2003], but

Figure 3.7: Long-term Impact of WWII Manufacturing Contracts on Manufacturing Employment Across U.S. Counties



Note: This map compares the outcome in the treated world versus the untreated world in 2000 across counties.

even there the regional increase in manufacturing employment was quite modest (0.5%). In all regions we see a rich pattern of impacts across counties, with some winners and some losers, even as regional outcomes are correlated on the whole. This reflects the balance between agglomeration economies at the county level, stemming from production externalities embodied in the scale function, as well as regional agglomeration economies due to the presence of trade costs and input linkages.

In sum, the counterfactual analysis reveals that while temporary World War II contracts had localized long-term impacts, prompting some counties to escape poverty traps, these effects were not significant at the aggregate level. Regional variations in manufacturing employment were observed, with modest growth in some areas and declines in others, but these shifts were not substantial enough to influence the overall economic landscape.

3.6 Conclusion

This paper examines the long-term effects of WWII spending on U.S. economic geography. Our empirical analysis reveals that while wartime spending led to short-term economic gains across counties, only those with initially low manufacturing employment experienced sustained long-term growth. Motivated by these findings, we develop a quantitative model incorporating non-

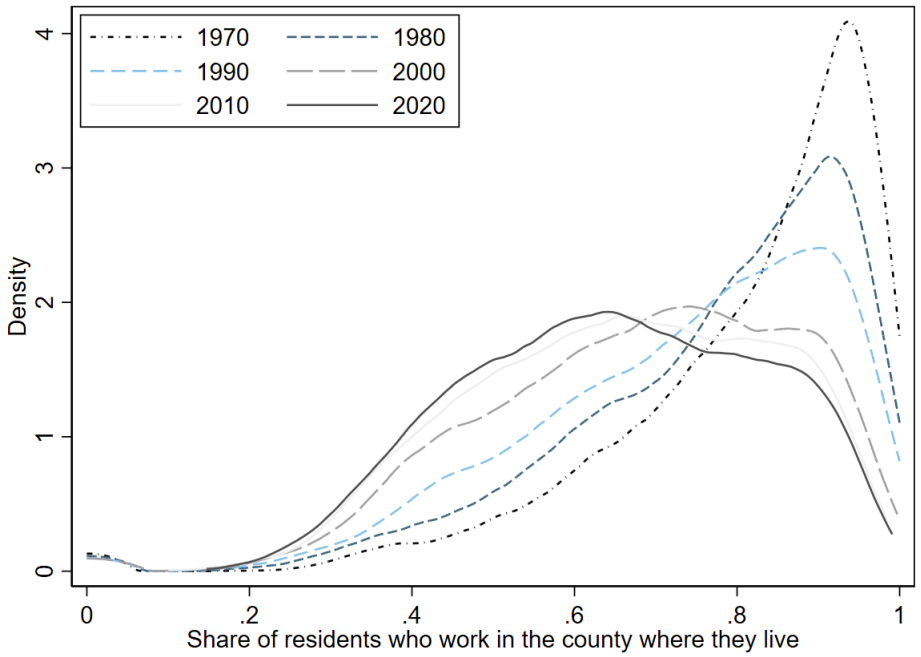
constant elasticity external economies of scale. The model successfully replicates the heterogeneous effects observed in the data, showing that coordination failures were significant only within a limited domain of manufacturing employment. Counterfactual simulations suggest that while WWII spending reshaped local economic trajectories, its aggregate impact on national manufacturing employment and welfare was minimal. These findings contribute to the broader literature on historical persistence and path dependence, emphasizing the role of initial conditions in shaping long-run economic outcomes.

APPENDIX A

Appendix to Chapter 1

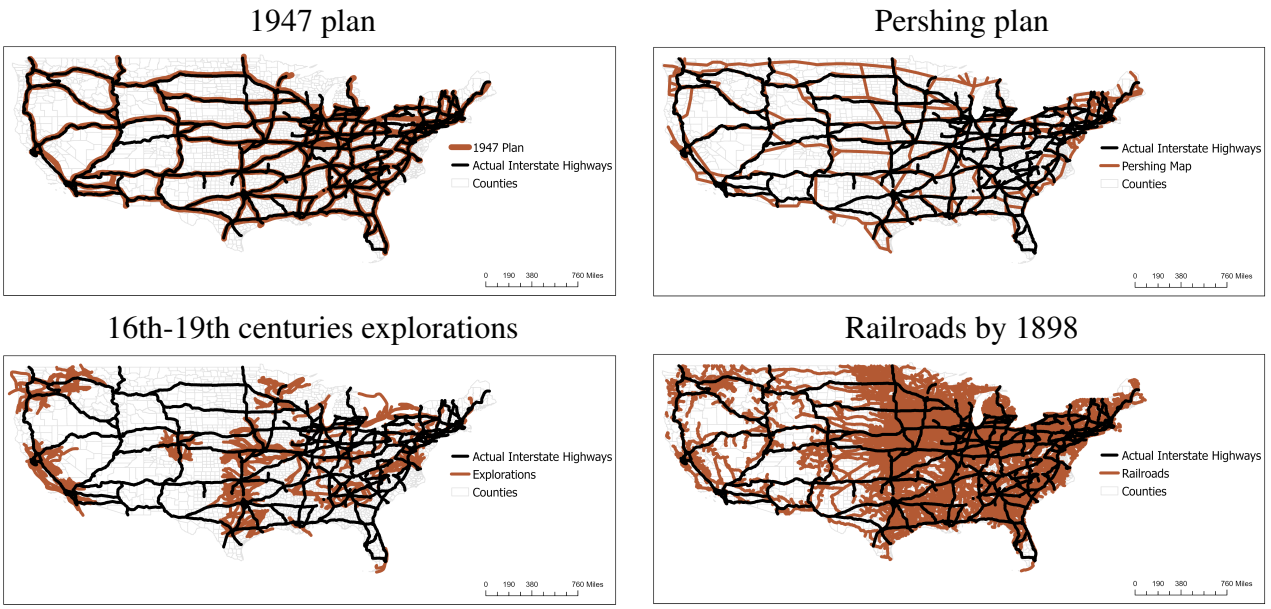
A.1 Figures and Tables

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).

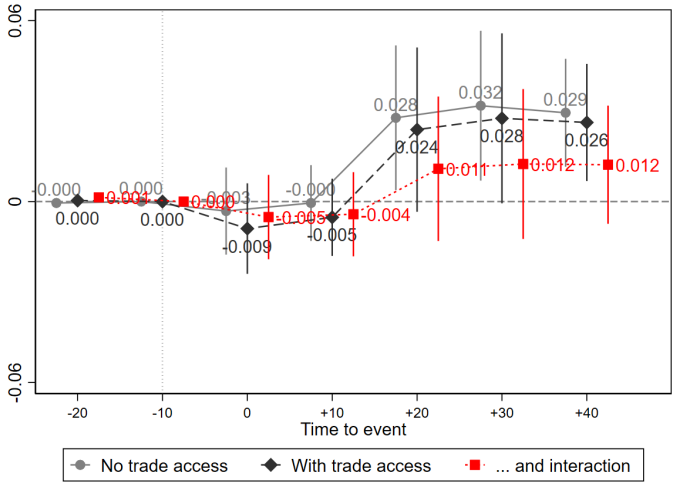
Figure A.2: Highway plans, historical routes, and actual highways



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

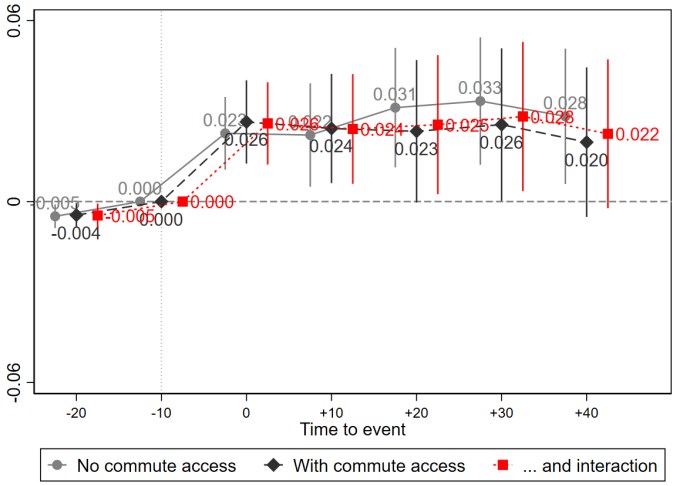
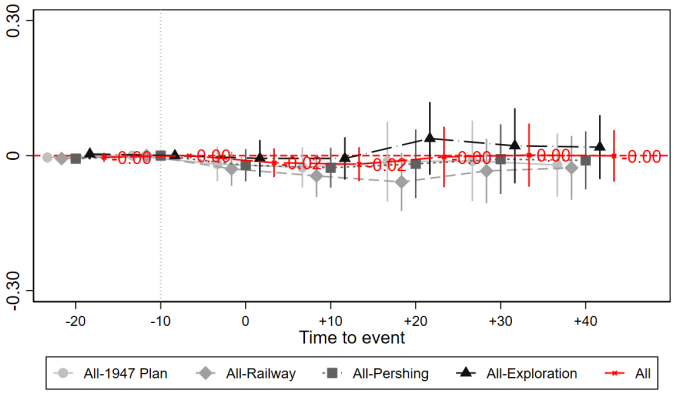


Figure A.4: 2SLS estimates of the effect of commuting and trade access

Panel A: Commute Access

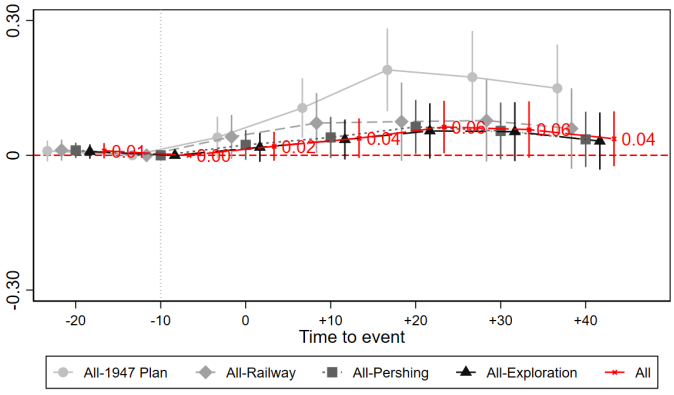
Log Employment - Log Population



Sequence of K-P F statistics is...
Excluding 1947 Plan : 31.0, ., 31.2, 30.4, 28.6, 29.1, 27.9
Excluding Railway : 31.5, ., 31.8, 26.4, 26.1, 28.1, 29.3
Excluding Pershing : 9.8, ., 9.6, 10.4, 10.0, 10.6, 10.3
Excluding Exploration : 7.0, ., 6.9, 7.6, 7.5, 8.6, 9.4
All together: 11.9, ., 11.7, 13.2, 12.9, 15.4, 17.0

Panel B: Trade Access

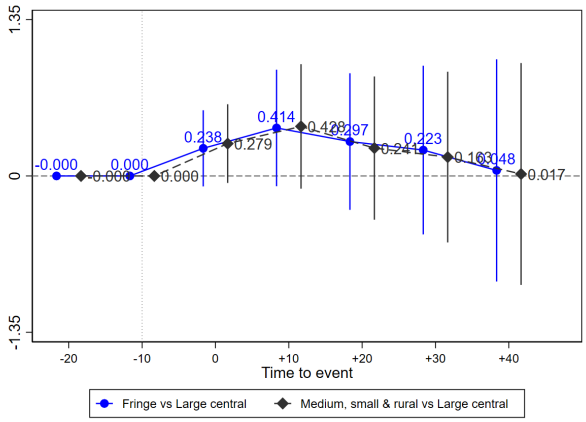
Log Employment - Log Population

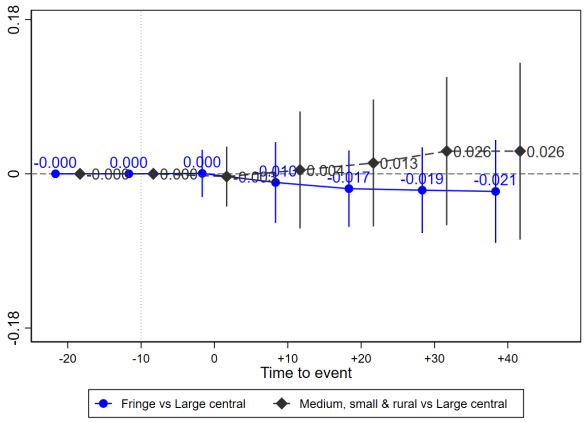


Sequence of K-P F statistics is...
Excluding 1947 Plan : 31.0, ., 31.2, 30.4, 28.6, 29.1, 27.9
Excluding Railway : 31.5, ., 31.8, 26.4, 26.1, 28.1, 29.3
Excluding Pershing : 9.8, ., 9.6, 10.4, 10.0, 10.6, 10.3
Excluding Exploration : 7.0, ., 6.9, 7.6, 7.5, 8.6, 9.4
All together: 11.9, ., 11.7, 13.2, 12.9, 15.4, 17.0

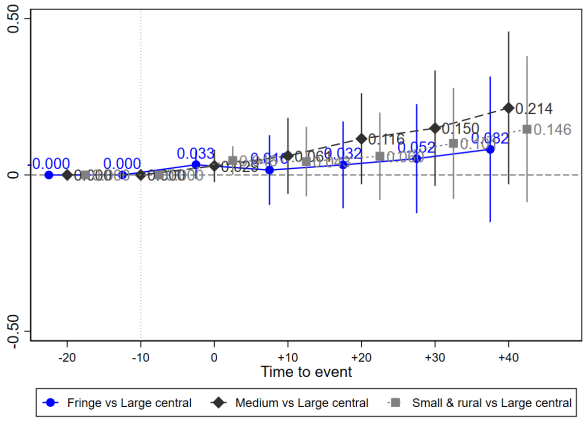
Figure A.5: OLS estimates of the effect of commute and trade access on log population by commuting zone and county NCHS classification

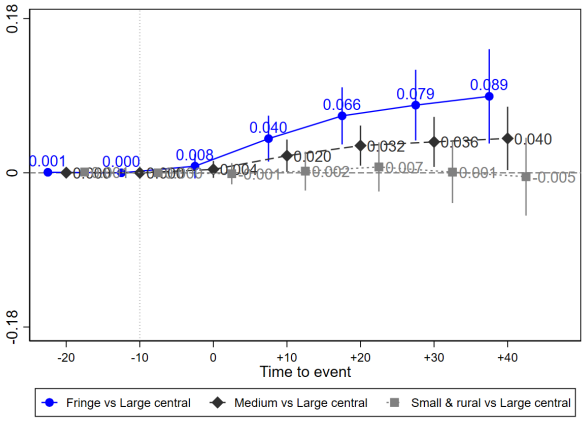
Panel A: Top 7 commuting zones



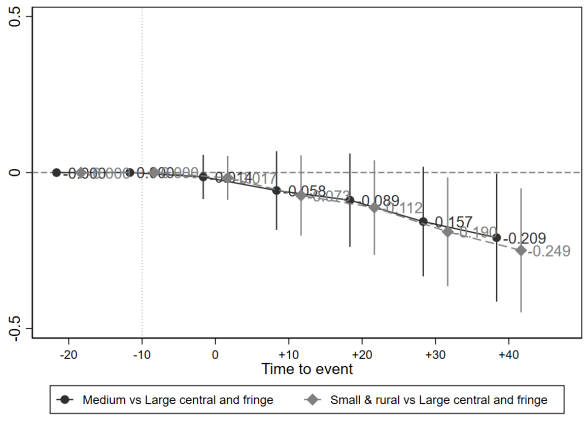


Panel B: Top 8-60 commuting zones





Panel C: Remaining commuting zones



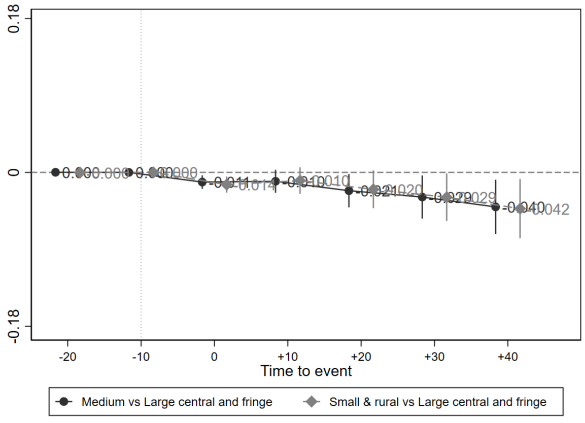
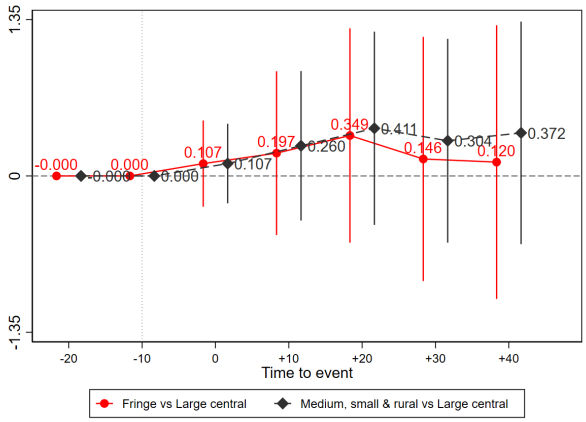
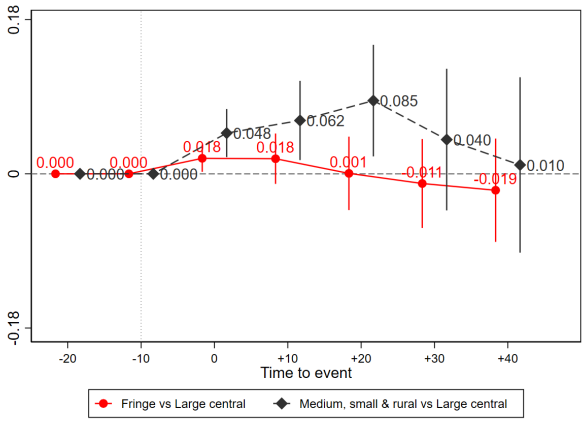


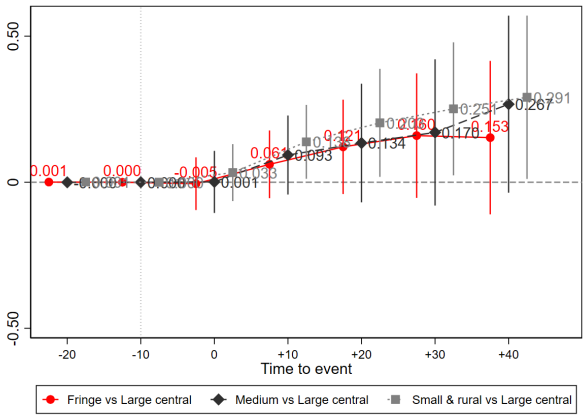
Figure A.6: OLS estimates of the effect of commute and trade access on log employment by commuting zone and county NCHS classification

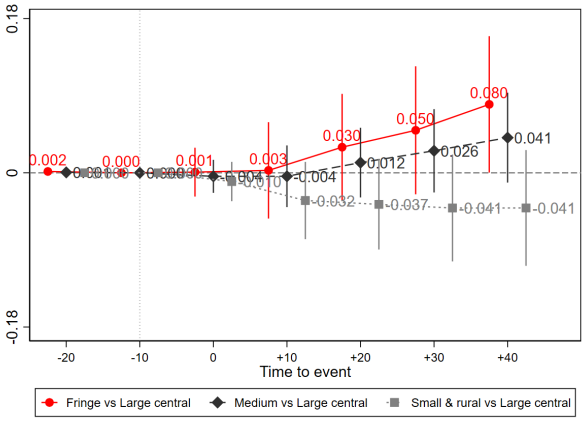
Panel A: Top 7 commuting zones



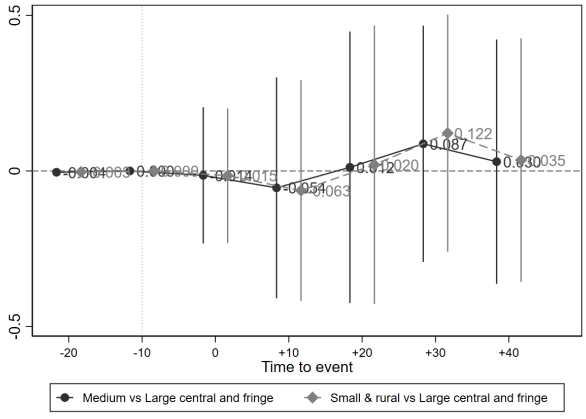


Panel B: Top 8-60 commuting zones





Panel C: Remaining commuting zones



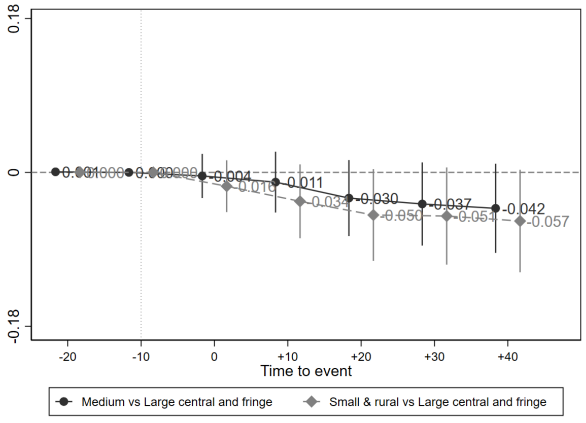
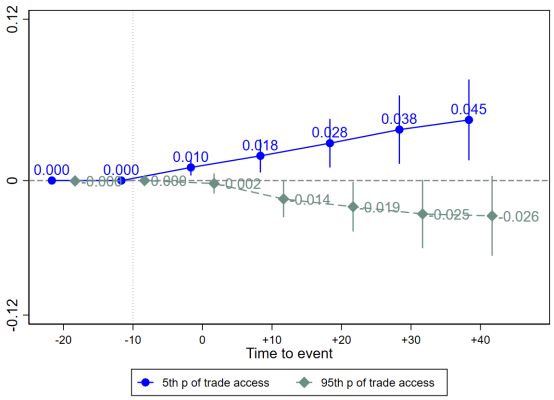


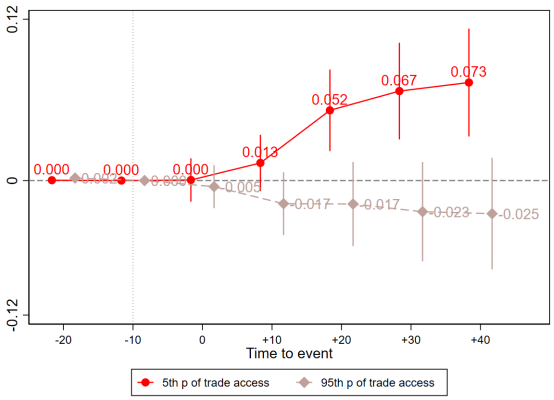
Figure A.7: OLS estimates of the effect of commuting and trade access and their interaction

Panel A: Commute Access

Log Population

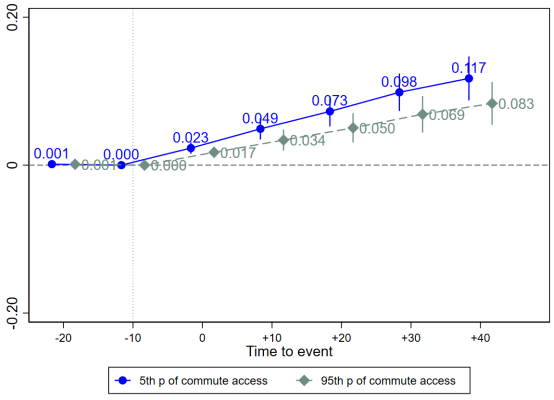


Log Employment



Panel B: Trade Access

Log Population



Log Employment

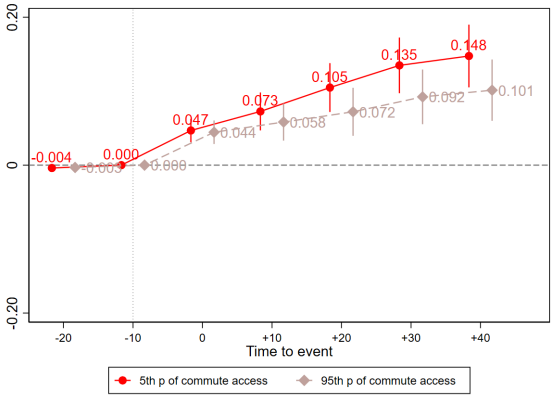
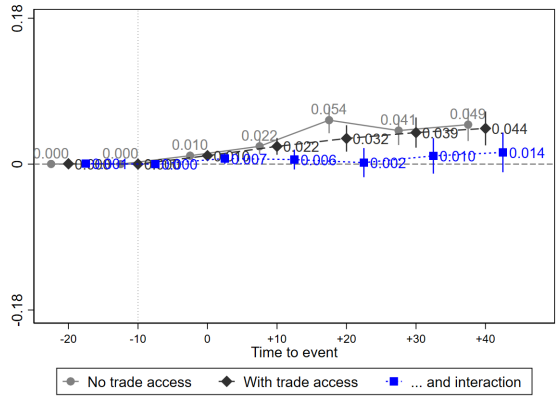


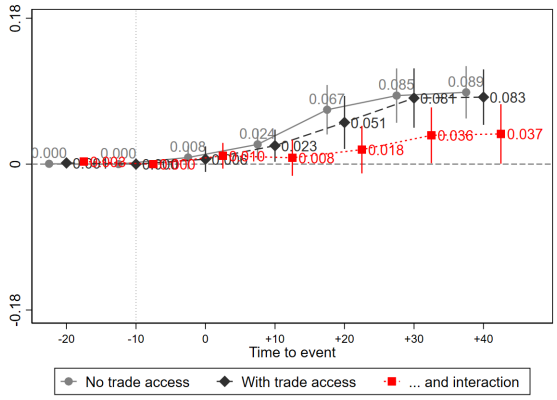
Figure A.8: OLS estimates of the effect of commuting and trade access conditioning on CZ times year-by-wave FE

Panel A: Commute Access

Log Population

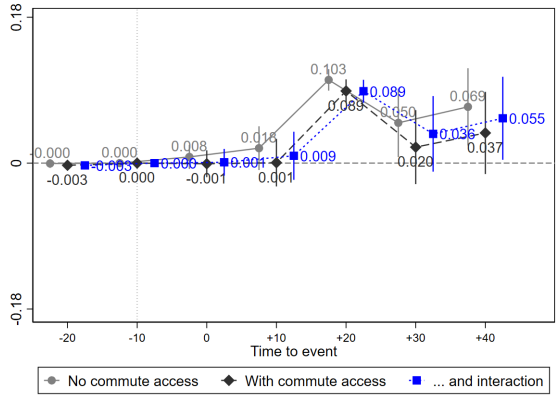


Log Employment



Panel B: Trade Access

Log Population



Log Employment

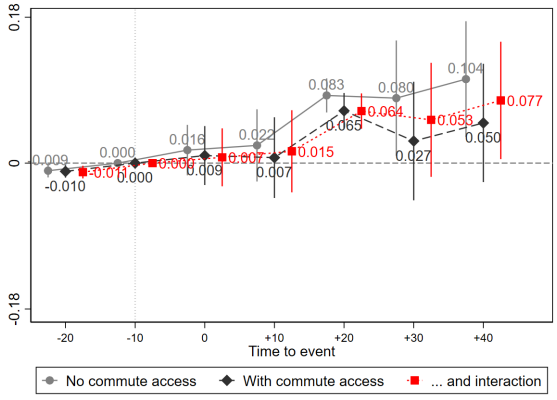
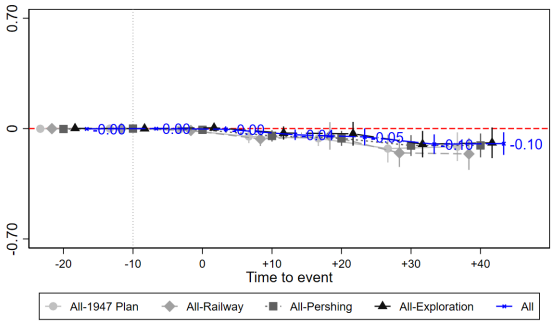


Figure A.9: 2SLS estimates of the effect of commuting and trade access conditioning on CZ times year-by-wave FE

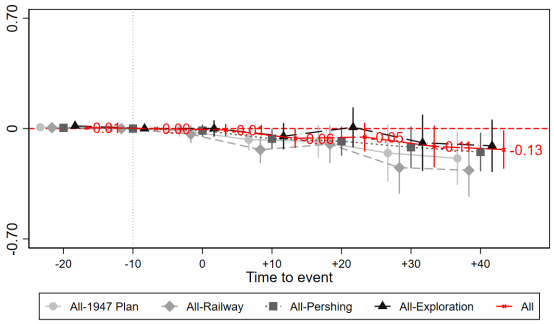
Panel A: Commute Access

Log Population



Sequence of K-P F statistics is...
Excluding 1947 Plan : 13.4, ., 13.2, 11.3, 28.6, 9.7, 9.8
Excluding Railway : 19.9, ., 18.5, 15.7, 26.1, 13.8, 14.3
Excluding Pershing : 12.1, ., 11.0, 11.1, 10.0, 9.8, 9.2
Excluding Exploration : 5.7, ., 5.5, 6.7, 7.5, 6.3, 6.6
All together: 21.3, ., 20.0, 18.9, 12.9, 16.2, 16.0

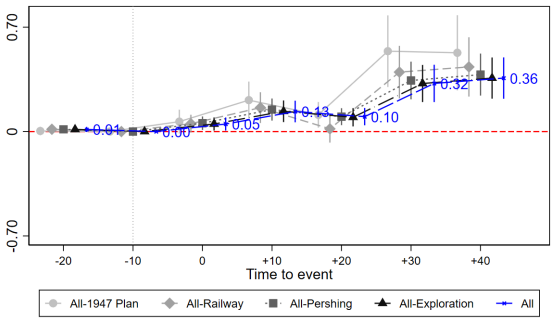
Log Employment



Sequence of K-P F statistics is...
Excluding 1947 Plan : 13.9, ., 13.8, 11.9, 27.1, 10.2, 10.3
Excluding Railway : 17.3, ., 17.0, 14.7, 25.9, 13.0, 12.9
Excluding Pershing : 10.9, ., 10.6, 10.4, 10.8, 9.3, 8.7
Excluding Exploration : 5.1, ., 5.2, 6.3, 7.8, 6.0, 6.2
All together: 19.2, ., 19.0, 17.8, 12.9, 15.5, 15.0

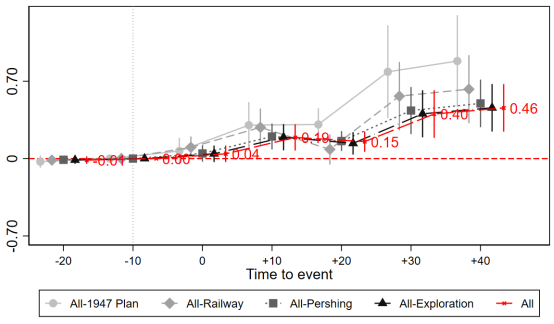
Panel B: Trade Access

Log Population



Sequence of K-P F statistics is...
Excluding 1947 Plan : 13.4, ., 13.2, 11.3, 28.6, 9.7, 9.8
Excluding Railway : 19.9, ., 18.5, 15.7, 26.1, 13.8, 14.3
Excluding Pershing : 12.1, ., 11.0, 11.1, 10.0, 9.8, 9.2
Excluding Exploration : 5.7, ., 5.5, 6.7, 7.5, 6.3, 6.6
All together: 21.3, ., 20.0, 18.9, 12.9, 16.2, 16.0

Log Employment

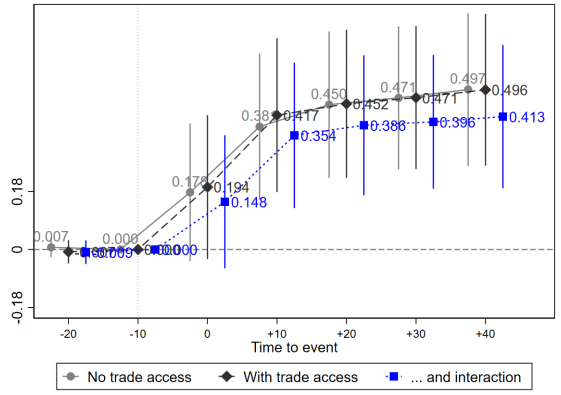


Sequence of K-P F statistics is...
Excluding 1947 Plan : 13.9, ., 13.8, 11.9, 27.1, 10.2, 10.3
Excluding Railway : 17.3, ., 17.0, 14.7, 25.9, 13.0, 12.9
Excluding Pershing : 10.9, ., 10.6, 10.4, 10.8, 9.3, 8.7
Excluding Exploration : 5.1, ., 5.2, 6.3, 7.8, 6.0, 6.2
All together: 19.2, ., 19.0, 17.8, 12.9, 15.5, 15.0

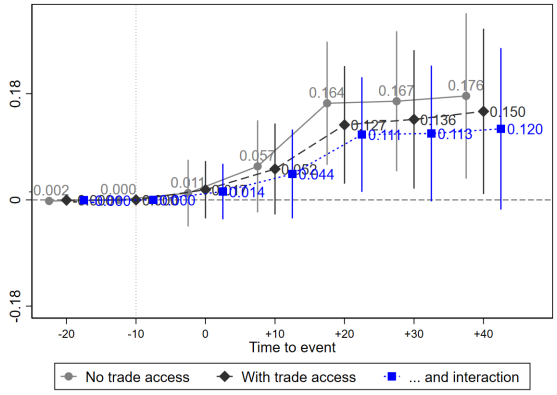
Figure A.10: OLS estimates of the effect of commuting and trade access on census-tracts' population

Panel A: Commute Access

Log Population (unweighted)

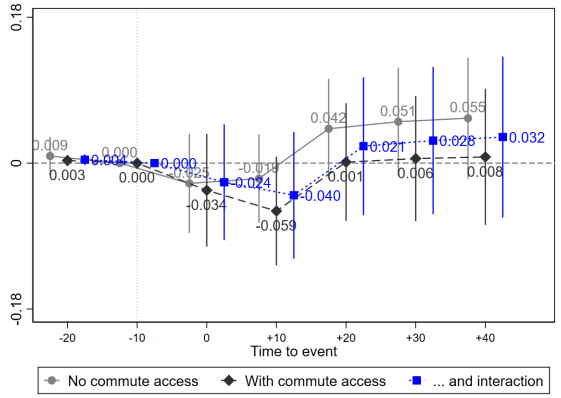


Log Population (weighted)

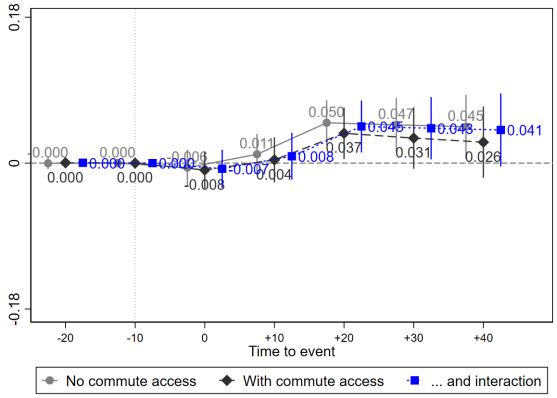


Panel B: Trade Access

Log Population (unweighted)



Log Population (weighted)



A.2 Other results

Heterogeneity across commuting zones

Facts 1 and 2 in Section 1.2.2 established that suburbs, especially in mid-sized commuting zones, emerged as key employment hubs. Section 1.3.3 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).[1] 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.[2]

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[3]. Overall, these findings support Facts 1 and 2: increased trade access has made suburbs and intermediate cities important employment hubs.

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.[4] 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

[1] 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.

[2] 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.

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

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

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.[5]

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 1.6. In contrast, the impact of trade access remains positive, even with high commute access changes.

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

[5] 2SLS results are available in Appendix, but conclusions are similar despite weak first-stage results in some subsamples.

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 1.6 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.[6] I rerun the regression using census-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.[7] 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,[8] whereas commute access effects are so in the least populous tracts.[9]

In summary, trade access decentralizes population between counties but concentrates it within

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

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

[8]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.

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

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.5 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.

A.3 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:

$$\underbrace{w_d \sum_i \pi_{ij|i} L_i}_{\text{Labor Supply}} = \underbrace{\sum_o \lambda_{do|o} \left(\bar{y}_o L_o\right)}_{\text{Labor Demand}} \tag{A.1}$$

where:

$$\lambda_{do|o} = \frac{A_d \left(c_d \varsigma_{do}\right)^{-\varepsilon}}{\Psi_o} \quad L_i = \frac{\bar{U}_i^{\eta} \bar{L}}{\sum_{i \in N} \bar{U}_i^{\eta}} \quad \bar{U}_{ni} = \frac{u_i}{P_i} \cdot \Phi_i^{1/\theta}$$

$$\bar{y}_i = \sum_j \pi_{ij|i} w_j \quad \pi_{ij|i} = \frac{\left(\frac{w_j}{d_{ij}}\right)^{\theta}}{\Phi_i} \quad \Phi_i = \sum_{l=1}^{S} \left(\frac{w_j}{d_{ij}}\right)^{\theta}$$

$$\Psi_i = \sum_j A_j \left(w_j \varsigma_{ji}\right)^{-\varepsilon} \quad P_i = \left(\Psi_i\right)^{-1/\varepsilon}$$

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

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

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). $\Psi_{RHS_i/j}$ 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 A.1, 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[(1-\pi_{AA})\,\theta + \frac{\Omega_{L_A/w_A}}{L_A}\right] + s_{a|A} \cdot \left[(1-\pi_{aA})\,\theta + \frac{\Omega_{L_a/w_A}}{L_a}\right] + s_{B|A} \cdot \left[(1-\pi_{BA})\,\theta + \frac{\Omega_{L_B/w_A}}{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} &= s_{A|A} \cdot \left[(\beta\pi_{Aa} + \pi_{AB})\,\theta + \frac{\Omega_{L_A/d}}{L_A}\right] + s_{a|A} \cdot \left[(\beta\,(\pi_{aA} - 1) + (1-\beta)\,\pi_{aB})\,\theta + \frac{\Omega_{L_a/d}}{L_a}\right] \\
&+ s_{B|A} \cdot \left[((\pi_{BA} - 1) + (1-\beta)\,\pi_{Ba})\,\theta + \frac{\Omega_{L_B/d}}{L_B}\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[(1-\pi_{Aa})\,\theta + \frac{\Omega_{L_A/w_a}}{L_A}\right] + s_{a|a} \cdot \left[(1-\pi_{aa})\,\theta + \frac{\Omega_{L_a/w_a}}{L_a}\right] + s_{Ba} \cdot \left[(1-\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} &= s_{A|a} \cdot \left[(\beta\,(\pi_{Aa} - 1) + \pi_{AB})\,\theta + \frac{\Omega_{L_A/d}}{L_A}\right] \cdot \{\kappa\} + s_{a|a} \cdot \left[(\beta\pi_{aA} + (1-\beta)\,\pi_{aB})\,\theta + \frac{\Omega_{L_a/d}}{L_a}\right] \cdot \{\kappa\} \\
&+ s_{B|a} \cdot \left[(\pi_{BA} + (1-\beta)\,(\pi_{Ba} - 1))\,\theta + \frac{\Omega_{L_B/d}}{L_B}\right] \cdot \{\kappa\} \\
\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[(1-\pi_{AB})\,\theta + \frac{\Omega_{L_A/w_B}}{L_A}\right] + s_{a|B} \cdot \left[(1-\pi_{aB})\,\theta + \frac{\Omega_{L_a/w_B}}{L_a}\right] + s_{B|B} \cdot \left[(1-\pi_{BB})\,\theta + \frac{\Omega_{L_B/w_B}}{L_B}\right]\right) \\
\Sigma_{LHS_B/d} &= s_{A|a} \cdot \left[(\beta\pi_{Aa} + (\pi_{AB} - 1))\,\theta + \frac{\Omega_{L_A/d}}{L_A}\right] + s_{a|a} \cdot \left[(\beta\pi_{aA} + (1-\beta)\,(\pi_{aB} - 1))\,\theta + \frac{\Omega_{L_a/d}}{L_a}\right] \\
&+ s_{B|a} \cdot \left[(\pi_{BA} + (1-\beta)\,\pi_{Ba})\,\theta + \frac{\Omega_{L_B/d}}{L_B}\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}
$$

$$
\begin{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)\\
&\quad+\lambda_{AB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_A}}{L_B}+\frac{\pi_{BA}w_A}{\bar{y}_B}-\left(1-\lambda_{AB|B}\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)\\
&\quad+\lambda_{AB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_a}}{L_B}+\frac{\pi_{Ba}w_a}{\bar{y}_B}+\lambda_{aB|B}\varepsilon\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)\\
&\quad+\lambda_{AB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_B}}{L_B}+\frac{\pi_{BB}w_B}{\bar{y}_B}+\lambda_{BB|B}\varepsilon\right)\\
\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)\\
&\quad+\lambda_{AB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/\varsigma}}{L_B}+\left(\lambda_{AB|B}+\lambda_{aB|B}-1\right)\epsilon\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}-\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)\\
&\quad+\lambda_{aB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_A}}{L_B}+\frac{\pi_{BA}w_A}{\bar{y}_B}-\left(1-\lambda_{AB|B}\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)\\
&\quad+\lambda_{aB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_a}}{L_B}+\frac{\pi_{Ba}w_a}{\bar{y}_B}+\lambda_{aB|B}\varepsilon\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)\\
&\quad+\lambda_{aB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_B}}{L_B}+\frac{\pi_{BB}w_B}{\bar{y}_B}+\lambda_{BB|B}\varepsilon\right)\\
\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)\\
&\quad+\lambda_{aB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/\varsigma}}{L_B}+\left(\lambda_{AB|B}+\lambda_{aB|B}-1\right)\epsilon\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}-\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)\\
&\quad+\lambda_{BB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_A}}{L_B}+\frac{\pi_{BA}w_A}{\bar{y}_B}-\left(1-\lambda_{AB|B}\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)\\
&\quad+\lambda_{BB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_a}}{L_B}+\frac{\pi_{Ba}w_a}{\bar{y}_B}+\lambda_{aB|B}\varepsilon\right)\\
\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)\\
&\quad+\lambda_{BB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/w_B}}{L_B}+\frac{\pi_{BB}w_B}{\bar{y}_B}+\lambda_{BB|B}\varepsilon\right)\\
\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)\\
&\quad+\lambda_{BB|B}L_B\bar{y}_B\left(\frac{\Omega_{L_B/\varsigma}}{L_B}+\left(\lambda_{AB|B}+\lambda_{aB|B}\right)\epsilon\right)
\end{aligned}
$$

A.4 Theorems and propositions

THEOREM A.4.1: *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:

$$\pi_{ij|i} = \frac{\left(K_j \frac{w_j}{d_{i,j}}\right)^\theta}{\sum_l \left(K_l \frac{w_l}{d_{i,l}}\right)^\theta} \tag{A.2}$$

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.4.2: *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 $c_o^{Cons} = (\beta w_o^{1-\nu} + (1-\beta) Q_o^{1-\nu})^{1/(1-\nu)}$, 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.4.3: *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:

$$c_o^{NT} = \left(\varrho^{NT} w_o^{1-\sigma} + \left(1-\varrho^{NT}\right)\tilde{R}_o^{1-\sigma}\right)^{1/(1-\sigma)} \quad \text{and} \quad P_o^{NT} = cons \cdot \left(A_d^{NT} c_o^{NT-\varepsilon}\right)^{-1/\varepsilon} \tag{A.3}$$

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

$$\underbrace{\tilde{L}_d^T w_d}_{\text{Data}} = \underbrace{\varrho_L^T \left(\frac{w_d}{c_d^T}\right)^{1-\sigma} X_d^T}_{\text{Tradable Sector}} \tag{A.4}$$

$$\underbrace{\tilde{L}_d^{NT} w_d}_{\text{Data}} = \underbrace{\varrho_L^{NT} \left(\frac{w_d}{c_d^{NT}}\right)^{1-\sigma} X_d^{NT}}_{\text{Non-Tradable Sector}} \tag{A.5}$$

where:

$$X_d^T = \sum_o \frac{A_d^T \left(c_d^T \varsigma_{do}^T\right)^{-\varepsilon}}{\Psi_o^T} \left(\underbrace{P_o^T C_o^{W,T}}_{\text{By Workers}} + \underbrace{P_o^T C_o^{H,T}}_{\text{By Landlords}} + \underbrace{\frac{\left(1-\varrho_L^T-\varrho_H^T\right)}{\varrho_L^T}\left(\frac{P_o^T}{w_o^T}\right)^{1-\sigma} w_o \tilde{L}_o^T}_{\text{By Tradable Firms}} \right) \tag{A.6}$$

$$X_d^{NT} = \underbrace{(1-b)\left(\frac{P_d^{NT}}{P_d}\right)^{1-\sigma} \alpha\bar{y}_d L_d}_{\text{By Workers}} + \underbrace{(1-b)\left(\frac{P_d^{NT}}{P_d}\right)^{1-\sigma} Q_d T_d}_{\text{By Landlords}} \tag{A.7}$$

$$P_d^T = cons \cdot \left(\sum_{o=1}^{N} A_o^T \left(c_o^T \varsigma_{od}^T\right)^{-\varepsilon}\right)^{-1/\varepsilon} \tag{A.8}$$

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:

$$c^{-(1-\sigma)/\epsilon} \frac{(1-\varrho_L^T-\varrho_H^T)}{\varrho_L^T} \left(\frac{P_o^T}{w_o^T}\right)^{1-\sigma} w_o \tilde{L}_o^T \tag{A.9}$$

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.4.4: *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 $u_{i,t}$:

$$L_{i,t} = \sum_n \zeta_t \frac{\bar{U}_{n,i,t}^{\eta}}{\sum_{i\in N} \bar{U}_{n,i,t}^{\eta}} L_{n,t-1} \tag{A.10}$$

$$\bar{U}_{n,i,t} = \frac{\xi_{n,i,t} u_{i,t}}{P_{i,t}^{\alpha} R_{i,t}^{1-\alpha}} \cdot \bar{y}_{t,|i} \tag{A.11}$$

□

A.5 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:

$$\Delta_h \log \pi_{i,j,t} = \theta_h \Delta_0 \tau_{i,j,t} + \underbrace{\phi_{i,t,h} + \phi_{j,t,h}}_{\text{FE}} + \epsilon_{i,j,t,h} \tag{A.12}$$

$$\Delta_h \tau_{i,j,t} = \alpha_h \Delta_0 \tau_{i,j,t} + \underbrace{\sigma_{i,t,h} + \sigma_{j,t,h}}_{\text{FE}} + \varepsilon_{i,j,t,h} \tag{A.13}$$

where the parameter of interest is $\tilde{\theta}_h$=θ_h/α_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: $\Delta_h \tau_{i,j,t}$.[10] 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 locations.

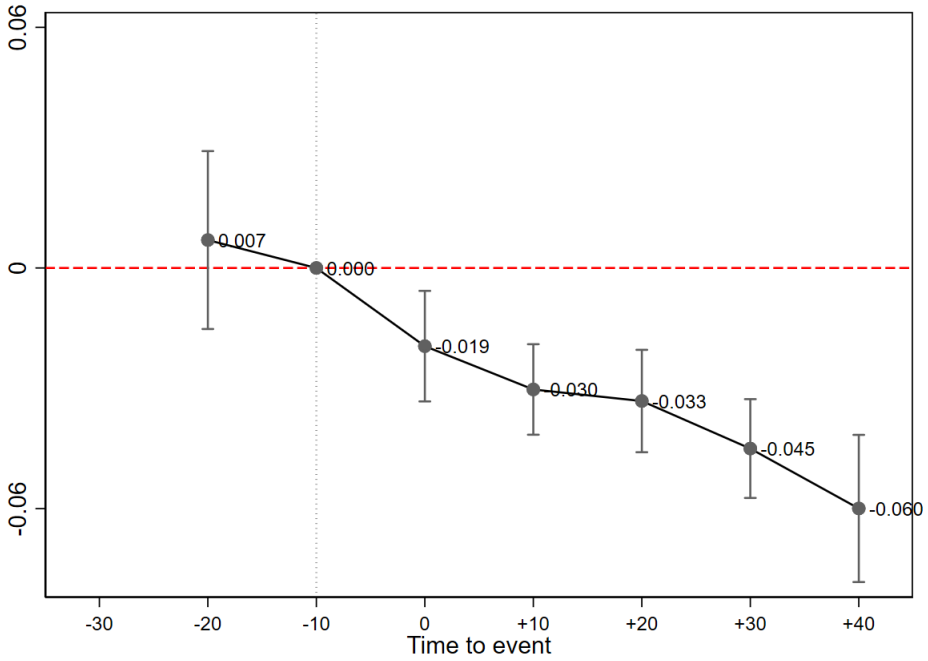
In Figure A.11, I show the results of estimating equation (A.12), which is akin to a reduced form regression. This figure reveals that when commute times rises by one minute between location i and location j, the probability of commuting to j decreases by 1.9%, approximately. However, this impact rises 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.

[10] I estimate equation (A.12) by PPML, and equation (A.13) by OLS

Figure A.11: Dynamic Gravity Equation of Commuting



APPENDIX B

Appendix to Chapter 2

B.1 Appendix Tables and Figures

B.1.1 Tables

Table B.1: The Aggregate Impact of Line 1 of the Metro and the BRT

Independent Commuting Choices								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0033	0.0040	-0.0093	0.0241	0.0233	0.0041	0.0141	0.0066
Real Income per Worker	0.0033	0.0040	0.0038	0.0058	0.0050	0.0041	0.0060	0.0036
Gender Gap	-0.0015			-0.0247		-0.0150		
Gender Gap (per Worker)	-0.0017			-0.0013		-0.0045		
# D.E. / # Married				0.0183				

Notes: I conduct a counterfactual exercise where I remove the Metro and the BRT from the city, while keeping exogenous characteristics fixed at their 2017 values. I perform this counterfactual exercise for two cases, one in which one spouse's choice is dependent on the other spouse's choice, and one in which spouses perform their commuting choices independently from each other. To perform these counterfactual exercises, I use observed data to recover unobservables under the model with interdependent choices. Then, I solve each model conditionally on this set of unobservables.

B.1.2 Figures

Figure B.1: The Line 1 of Lima's Metro

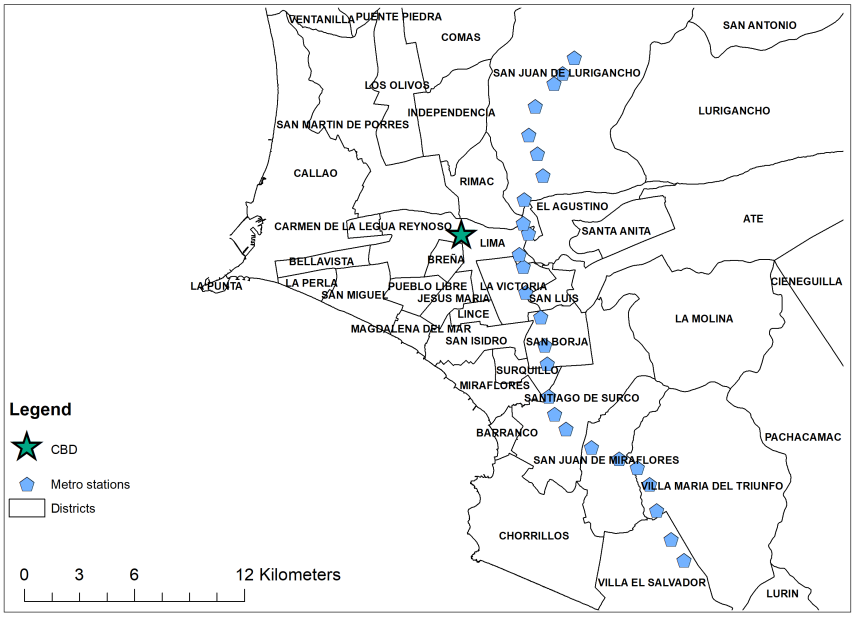


Figure B.2: The Bus Rapid Transit System

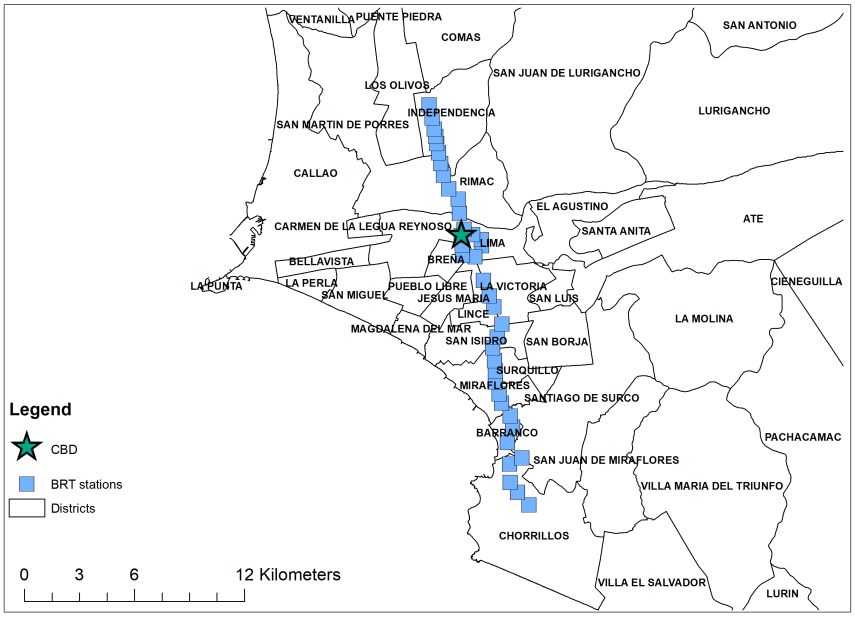
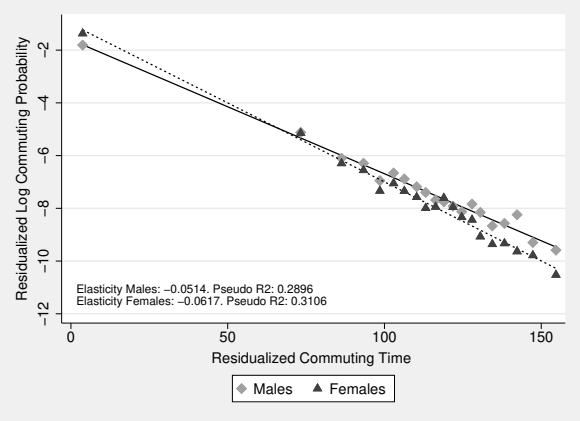
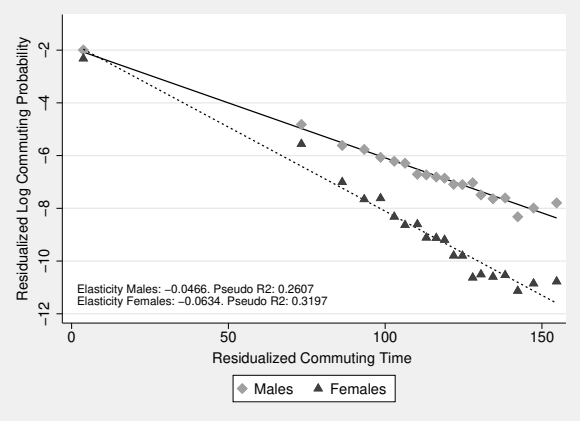


Figure B.3: Gravity in Commuting Flows among Singles and Married



Panel A: Singles



Panel B: Married

Notes: Plotting binscatters from columns (1) to (4) of Table 2.1.

Figure B.4: Constructed and Planned Lines

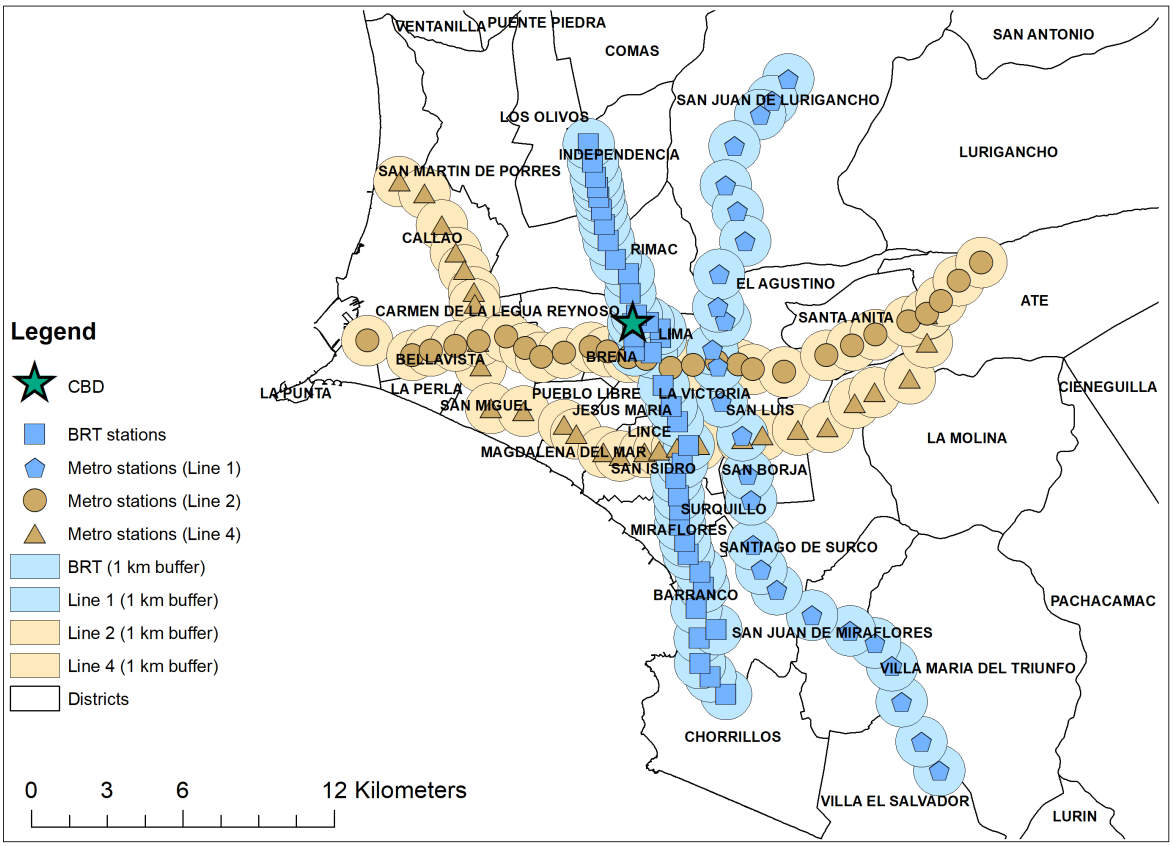
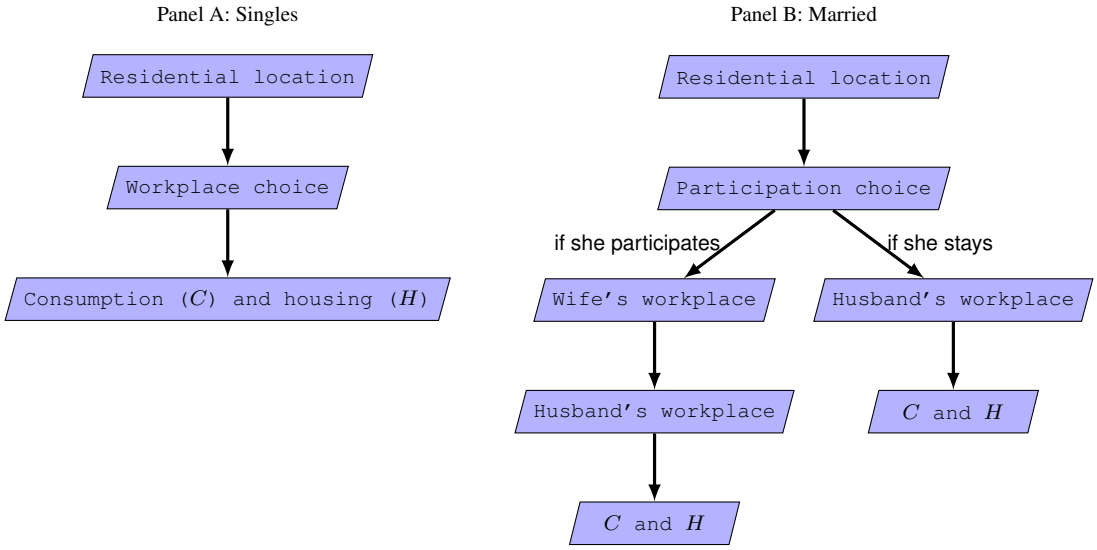


Figure B.5: Timing Assumptions



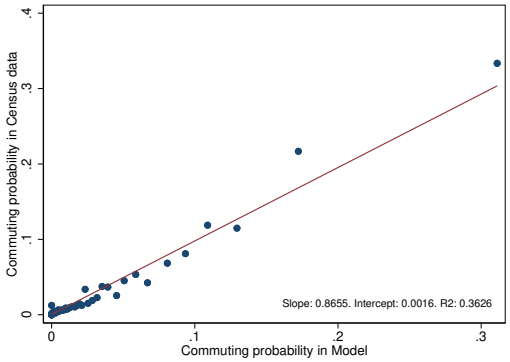
Panel A: Solving backwards, conditional on the locations that singles live and work, they choose their consumption and housing allocations. Afterwards, conditional on their residence location, they choose their workplace location. Finally, they decide where to live.

Panel B: Solving backwards, married households start by deciding how much they consume in the final good and in residence floorspace, conditional on the workplace location of both spouses and the residential location. Afterwards, conditional on the wife's labor status and the wife's workplace location, households decide the workplace of the husband. Then, they decide the workplace of the wife. If the wife does not participate, households only decide the workplace location of the husband. Next, conditional on the household's residence location, households compare the utility they would receive if the wife works against the utility they would receive if the wife stays at home performing household production. Based on this comparison, households decide the labor status of the wife. Finally, households choose their residence location.

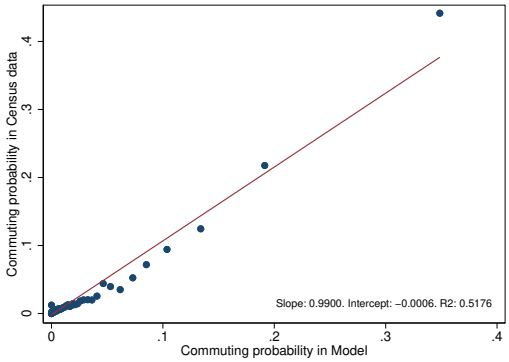
Figure B.6: Commuting and Over-identification

Panel A: Females in Dual-Earner Households

Independent Commute

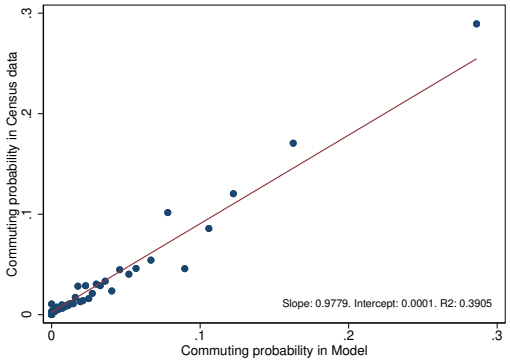


Interdependent Commute

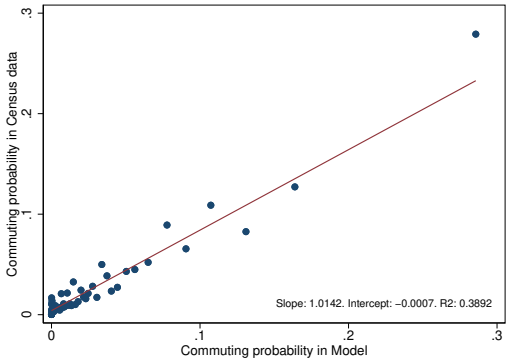


Panel B: Males in Dual-Earner Households

Independent Commute

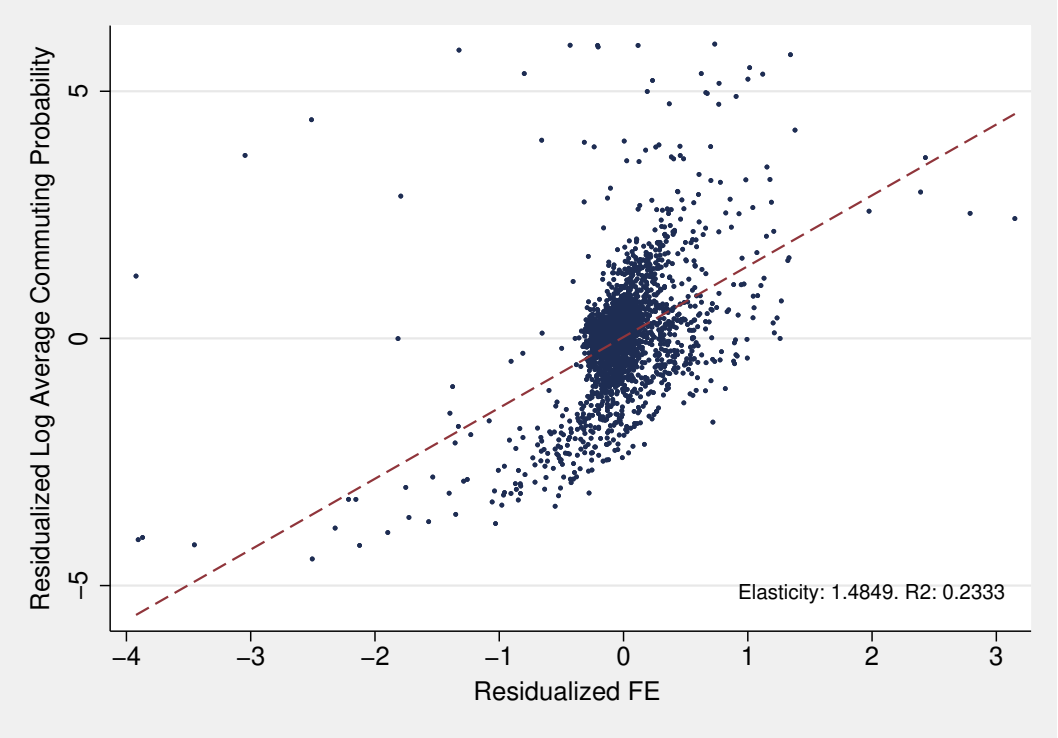


Interdependent Commute



See Section 2.5.5.

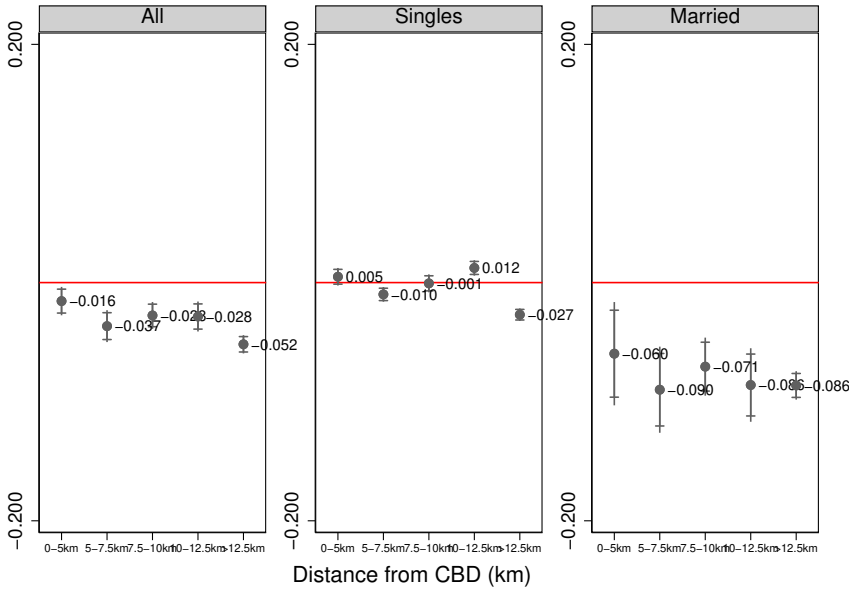
Figure B.7: Interdependence in Commuting Choices within Dual-earner Households



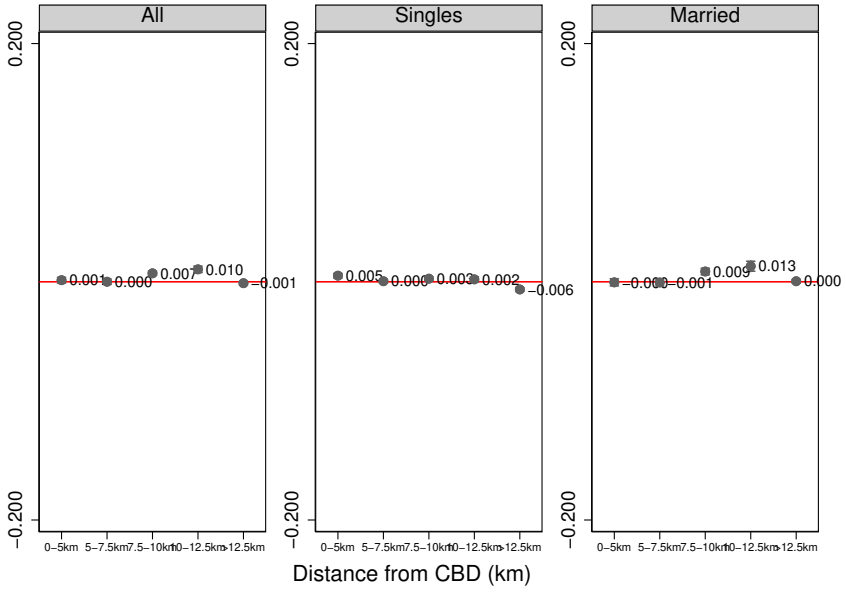
Notes: Using data generated by the model with interdependent commute choices, I estimate $\log \pi_{j|ij',g} = \beta \text{time}_{ij} + \text{FE}_{i,j',g} + \text{FE}_{j,j',g} + \varepsilon_{ijj',g}$ for males in dual-earner households. Then, I recover the estimates of $\text{FE}_{j,j',g}$. I also control for the time it takes to go from workplace in j to workplace in j'. Finally, I estimate the coefficient ψ from $\log \pi_{j|j',g} = \psi_g \widehat{\text{FE}}_{j,j'} + \text{FE}_j + \text{FE}_{j'}$. This figure depicts a scatter plot between residualized commuting probabilities and the residuals of $\text{FE}_{j,j'}$. I only depict this plot for husbands because this formulation is only valid for them, as per equations 2.14 and 2.15. This result is a by-product of the timing of idiosyncratic shocks.

Figure B.8: Impact with Independent Commute Choices by Distance to the CBD

Panel A: Gap in Aggregate Real Earnings



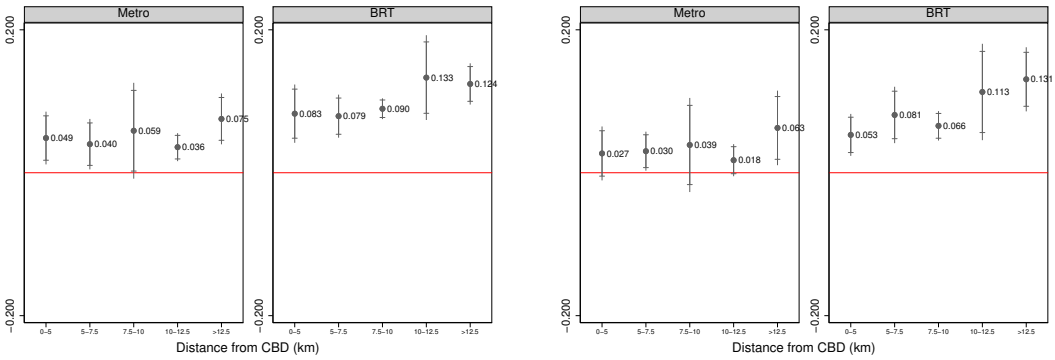
Panel B: Gap in Real Earnings per Worker



See Section 2.6.2.

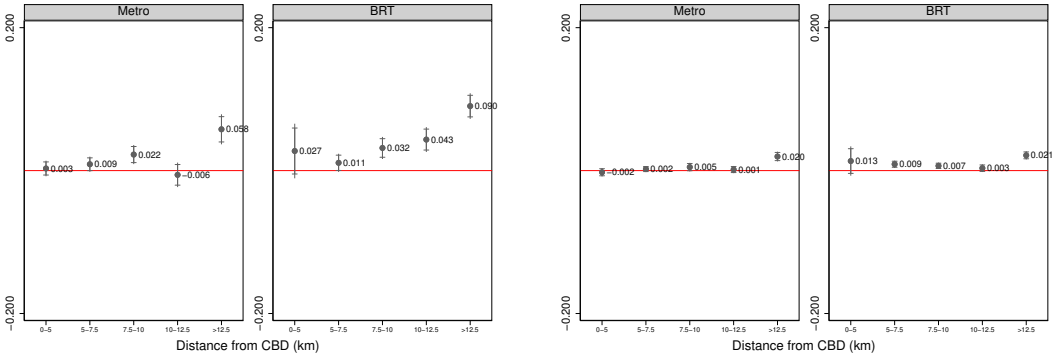
Figure B.9: Reduced Form Impact of Transit Infrastructure by Distance to CBD

Panel A: Dual-Earners to Male Breadwinners



Interdependent Independent

Panel B: Female Labor Earnings relative to Male Labor Earnings



Interdependent Independent

Notes: I use data generated by the model with interdependent and independent commuting to estimate reduced-form responses. For Panel A, I interact treatment dummies in equation 2.3 with a complete set of dummies indicating the distance to the CBD. Then, I plot β^{Metro} and β^{BRT} by distance to CBD. For Panel B, I interact treatment dummies in equation 2.4 with distance dummies. I also add the interaction of distance dummies and the female dummy. Then, I plot β_F^{Metro} and β_F^{BRT} by distance to CBD.

B.2 Fact 1 Extended

Heterogeneity. Having a kid at home can potentially induce higher sensibility to commute times. Every minute outside home is costly since one has to take care of the kids. To explore these ideas, as in Section 2.3.2, using the 2017 Population Census, I classify households into different groups. The first two consist of (i) *single females* and (ii) *single males*. Then I consider (iii) *male breadwinner* households, that is, households where the wife stays at home while the husband goes to work. For the sake of completeness, I also consider (iv) *female breadwinner* households. Finally, I consider (v) men in *dual-earner* households, and (vi) women in dual-earner households. I further divide these groups into two: if there is at least one kid below 8 years old living in the same household, and if there is not. This gives a total of twelve groups. For each one of these groups I reconstruct commuting flows $\pi_{j|i,k}$, and run the following reduced form regression:

$$\log \pi_{j|i,k} = \beta_k time_{ij} + FE_{i,k} + FE_{j,k} + \varepsilon_{ij} \tag{B.1}$$

where $\pi_{j|i,k}$ is the share of members of civil status-by-gender-by presence of kid k that live in origin i that commute to j. $time_{ij}$ is the time that takes to commute from i to j. Moreover, $FE_{i,k}$ and $FE_{j,k}$ are fixed effects by origin i and destination j. Finally, β_k is the *commuting elasticity*. It measures how the probability of commuting to j would change in relative terms after increasing the commuting time by one minute.

Results are shown in Table B.2. Panel A depicts the commuting elasticities when there are no kids at home, whereas Panel B do so when there is at least one kid at home. There are two takeaways from these results. First, women in breadwinner and dual-earner households are much more sensitive to commute times than men. This is true whether there is a kid at home or not. However, single women are much more sensitive to commute times than men only when there is a kid at home. Second, having a kid at home increases sensibility to commute times to all women, and only marginally does so for men. Thus, this exercise further validates one observation of this paper: given current social norms, it is more costly for women than for men to keep the balance between family and work responsibilities, specially if married.

Robustness. In Section 2.3.2, I estimated commuting elasticities from a gravity equation connecting commuting flows from origins i to destinations j among married and single households. Given census data, I can control for the average characteristic of individuals using the route $\{i, j\}$. This would ameliorate the concern that, for example, single and married households can differ in several characteristics, such as age and education.

Nonetheless, in principle for my analysis, I am agnostic about the source of these differences between types of households. I am taking the marriage market as given. My target population are

current married and singles even if they are different along some dimension or characteristic. Any difference between the two groups is summarized by how sensitive they are to commute times. This is the objective of my analysis: to quantify the impact of new infrastructure conditional on singles' and married's observed sensitivity to commute times.

Table B.2: Commuting Elasticity by Presence of Kids

	Single HH		Breadwinner HH		Dual-Earner HH	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No kids	Males	Females	Males	Females	Males	Females
Travel Time	-0.0547	-0.0588	-0.0449	-0.0628	-0.0470	-0.0626
	(0.0037)***	(0.0037)***	(0.0033)***	(0.0041)***	(0.0032)***	(0.0038)***
Gap (Female/Male-1)		7%		40%		40%
Panel B: At least one kid	Males	Females	Males	Females	Males	Females
Travel Time	-0.0568	-0.0727	-0.0475	-0.0701	-0.0479	-0.0717
	(0.0048)***	(0.0045)***	(0.0032)***	(0.0044)***	(0.0033)***	(0.0045)***
Gap (Female/Male-1)		28%		48%		50%
Gap (Kid/No Kid - 1)	4%	24%	6%	12%	2%	15%
Origin FE	X	X	X	X	X	X
Destination FE	X	X	X	X	X	X
N	2500	2500	2500	2500	2500	2500

Notes: Observations are at the municipality-by-municipality level. The independent variable in all regressions is the median travel time between origin i and destination j is t, with origin and destination fixed effects included. Column (1) of the regression analysis uses the share of single male workers who reside in origin municipality i and work in destination municipality j as dependent variable. Column (2) uses the corresponding share for single female workers. Columns (3) and (4) focus on households where only the male or female breadwinner works, respectively. In columns (5) and (6), the shares are defined for dual-earner households, specifically for males and females separately. The difference between coefficients in column (1) and (2), between column (3) and (4), and between column (5) and (6) is statistically significant at the 1 percent level, except for single men and women in Panel A.

Admittedly, if differences in the sensitivity to commute times between men and women persist after introducing controls, this would be informative about one of the main mechanism explored in this paper. That is, given current social norms, it is more costly for women to keep the balance between family and work responsibilities, specially if they are married. So, to further validate this mechanism, I re-estimate equation 2.1 but controlling for the average characteristic at the $\{i, j\}$ pair. In particular, I control for age, education, religion, language, recent moving history, and dwelling characteristics such as the number of rooms and whether sewage is available.

Results are shown in Table B.3. The main takeaway from these results is that, even when accounting for origin-destination controls, women are much more sensitive to commute times than men, specially when they are married. In comparison to my baseline results, women's sensitivity relative to men's sensitivity is somewhat smaller, but still sizeable. For instance, women in breadwinner households are now 34% more sensitive than men. However, according to my baseline results, relative sensitivity was 36%.

Table B.3: Commuting Elasticity and Origin-Destination Controls

	Single HH		Breadwinner HH		Dual-Earner HH	
	Males	Females	Males	Females	Males	Females
	(1)	(2)	(3)	(4)	(5)	(6)
Travel Time	-0.0459 (0.0027)***	-0.0515 (0.0030)***	-0.0375 (0.0022)***	-0.0503 (0.0034)***	-0.0389 (0.0024)***	-0.0513 (0.0029)***
Origin-Destination Controls	X	X	X	X	X	X
Origin FE	X	X	X	X	X	X
Destination FE	X	X	X	X	X	X
Gap (Female/Male-1)		12%		34%		32%
N	2500	2500	2500	2500	2500	2500

Notes: Observations are at the municipality-by-municipality level. The independent variable in all regressions is the median travel time between origin i and destination j is t, with origin and destination fixed effects included. I also control for the average characteristic at the origin-destination level. In particular, I control for age, education, religion, language, recent moving history, and dwelling characteristics such as the number of rooms and whether sewage is available. Column (1) of the regression analysis uses the share of single male workers who reside in origin municipality i and work in destination municipality j as dependent variable. Column (2) uses the corresponding share for single female workers. Columns (3) and (4) focus on households where only the male or female breadwinner works, respectively. In columns (5) and (6), the shares are defined for dual-earner households, specifically for males and females separately. The difference between coefficients in column (1) and (2), between column (3) and (4), and between column (5) and (6) is statistically significant at the 1 percent level.

B.3 Fact 2 Extended

One could still argue that the procedure used for Fact 2 in the main text is capturing anything that varies at the destination-pair level, but that is unrelated to the joint decision that partners make regarding their commuting. I now provide suggestive evidence of the cross-spouse income effect. Using firm-survey data, I compute average wages by gender across locations. I drop the destination-pair fixed effect in equation 2.2 and instead add the average wage at destination j (where g commutes to) and the average wage at destination j' (where the spouse commutes to). Moreover, I interact wages at one's destination with the total income the couple would obtain at a destination pair j, j'. I show the results of this specification in Table B.4. These results show that the greater the household income, the lesser the sensitivity to further increases in one's own wage. Moreover, there is some heterogeneity around the average own-gender wage elasticity.[1] In the case of husbands in dual-earner households, this elasticity goes between 1.856% and 3.054% (column 1), whereas for wives this elasticity goes between 1.178% and 1.914% (column 2).

[1] I obtain average wages for each gender at destinations from the 2015-2017 waves of the national firm-level survey data. I have demeaned the variable for log total income, hence the coefficient associated to one's own gender's wages yields the average elasticity.

Table B.4: Commuting and Wages

	Husband in D.E. Households	Wife in D.E. Households
	(1)	(2)
Wage at one's destination	2.5647	1.6135
	(0.3202)***	(0.1699)***
Wage at one's destination ×	-0.4466	-0.2744
HH income at destination pair	(0.0588)***	(0.0297)***
Origin-spouses' destination FE	X	X
Min Elasticity	1.856	1.178
St.Dev	(0.2288)	(0.1242)
Max Elasticity	3.054	1.914
St.Dev	(0.3839)	(0.2019)
N	90477	94773

B.4 Fact 3 Extended

In this section, I extend the set of reduced-form results. As it will become clear in the theoretical section of the paper, the new transit infrastructure can impact households through many channels. The main channel is the improvement in commuting times leading to greater access to jobs. However, the Metro and BRT can also affect local exogenous amenities, for instance by bringing more crime or noise, and the price of the land (Brinkman and Lin, 2022). In this section, I explore these issues in the context of reduced-form specifications.

B.4.1 Planned stations

I perform a reduced-form analysis by exploiting the planned network of Metro lines. Specifically, I leverage Line 2 and Line 4 of the Metro.[2] I compare locations closer to Line 1 of the Metro or the BRT (treatment locations) against locations closer to Line 2 and Line 4 (control locations). The basic idea is that we can account for location selection effects by comparing treated locations to locations closer to planned-but-not-constructed stations. Thus, these should constitute a better control group than other locations.

To illustrate how I define the treatment and control groups, Figure B.10 depicts the map of the constructed and planned stations, in addition to buffers around each station. In practice, I define the treatment group as all locations that are within 1.5 kilometers of any Line 1 or BRT station. Then, I define the control group as all locations that are within 3.2 kilometers from any Line 2 or Line 4 station. To avoid contamination from the treated group, I add the restriction that all locations in the control group should also be at least 2 kilometers away from all Line 1 and BRT stations. I use a bigger radius to define the control group because some locations are lost due to contamination

[2] I leave out Line 3 because it will run over the same locations as the BRT.

which hurts statistical power. However, results are robust to alternative definitions.

[Figure B.10 here]

B.4.2 Empirical strategy

I estimate time-event studies by comparing outcomes in locations close to the Metro and BRT stations to outcomes in locations close to yet-to-be-constructed stations. I utilize two data sources: the 1993, 2007 and 2017 Population Censuses, and the 2007-2017 waves of the National Household Survey.

The 1993, 2007 and 2017 Population Censuses. First, due to data limitations I start by performing a time event study with a subset of the data.[3] Then, after showing that pre-trends are parallel, I move to a differences-in-differences specification employing only the 2007 and 2017 data, which allows me to fully leverage the cross-sectional variation of the data.

I start by running the following time event study:

$$\log y_{lt} \quad = \quad \sum_{k} \beta_k \left[Treat_l \cdot Year_t^k \right] + FE_l + FE_t + \epsilon_{it}$$

where y_{lt} is an outcome constructed from Census data, $Treat_l$ is a treatment dummy taking the value of one if the location l is closest to either the Metro or the BRT station. I interact these variables with year dummies and so β_{2017} measures the effect of the new transit infrastructure on locations that are closer to the stations, relative to locations that are closer to yet-to-be-constructed stations. $\beta_{1993} = 0$ would be indicative of parallel pre-trends. Finally, I introduce a set of location fixed effects FE_l, in addition to year fixed effects FE_t.

Next, I use the whole sample of locations in 2007 and 2017 to fully leverage the cross-sectional variation and estimate a differences-in-differences model. First, I break the treatment dummy into two: $Treat_l^{Metro}$, which takes the value of one if location l's closest station is a Metro station, and $Treat_l^{BRT}$ if the closest station is a BRT station.[4] Second, I argue that being closer to a station

[3] To perform the time-event study I need a panel of locations identified in 1993, 2007, and 2017. The problem is that the match between the 1993 data and the 2007 data is imperfect, partially, because many locations that existed in 2007 did not exist in 1993. Out of 605 locations in either the treatment or control group, I can correctly match 125. Matched locations tend to be at a lower altitude, and are closer to the CBD than unmatched locations, as expected. However, there are no statistically significant differences in the size of the block, in the slope, nor in the overall treatment status. I argue that while this subset of the data is non-random and might introduce a selection effect, it is still reassuring to find parallel pre-trends. Also, out of the 605 locations, I can match 600 in years 2007 and 2017.

[4] To control for differential trends among a set of covariates, I include a set X_{lt} of block-level controls, i.e. the physical size of the block (i.e. the area), the elevation of the block, the median slope, a continuous measure of distance to the Central Business District, and a continuous measure of distance to the closest station (either constructed or planned) each interacted with year dummies FE_t.

does not guarantee greater access to relevant destinations. Gains in access should be higher in the fringe of the city since these areas were more isolated at the beginning of the sample. Conversely, gains in access in locations closer to the CBD may be less pronounced since workers were already close to valuable destinations. Also, impacts on amenities and other outcomes could potentially be heterogeneous across locations. Thus, I allow for the treatment effects to vary by the distance of households to the CBD. I construct five dummies indicating if the household is located within 0-5, 5-7.5, 7.5-10, 10-12.5 kilometers, and more than 12.5 kilometers.

The 2007-2017 waves of the National Household Survey. I use 2007-2017 rounds of the Peruvian National Household Survey, georeferenced at the block level, to precisely locate households and classify them into treatment and control groups. Then, I run the following time-event study at the household level i and block level l at year t:

$$\begin{aligned} \log y_{il(i)t} &= \sum_k \beta_k \left[Treat_{l(i)}^{Metro} \cdot Time_t^k \right] + \sum_k \beta_k \left[Treat_{l(i)}^{BRT} \cdot Time_t^k \right] \\ &+ \gamma' X_{il(i)t} + FE_{l(i)} + FE_t + \epsilon_{it} \end{aligned} \tag{B.2}$$

where y_{it} is a household-level outcome such as rents and income.[5] $Treat_{l(i)}^{Metro}$ and $Treat_{l(i)}^{BRT}$ are treatment dummies taking the value of one if the location l is closest to either the Metro or the BRT station. $Time_t$ are a set of time dummies indicating if the data comes from (i) any year preceding 2009 i.e. 2007 or 2008, (ii) the year just before the introduction of the new transit infrastructure i.e 2009, or (iii) any year after the introduction of the new transit infrastructure, i.e. 2010 or after. I set the base category to be the year 2009 since the Metro and the BRT started operating in 2010. X_{1it} is a set of household-level controls comprising age, years of education, civil status, mother's language, and moving history. Finally, X_{2lt} is a set of block-level controls, i.e. the physical size of the block (i.e. the area), the elevation of the block, the median slope, a continuous measure of distance to the Central Business District, and a continuous measure of distance to the closest station (either constructed or planned) each interacted with year dummies FE_t.

The impact of being close to a station can hide substantial heterogeneity. In fact, gains in access are not uniformly distributed across stations. I allow for the treatment effects to vary by the distance of households to the CBD. I construct five dummies indicating if the household is located within 0-5, 5-7.5, 7.5-10, 10-12.5 kilometers, and more than 12.5 kilometers. As we will see, pre-trends do not seem to be an issue. Hence, I simplify equation B.2 by replacing the time dummies with a dummy variable taking the value of one if the data comes from any year after 2009 and zero otherwise.

[5]Unfortunately, the size of the sample does not allow me to have enough statistical power to explore the heterogeneity in the impact across different types of households.

B.4.3 Results

Additional Fact: Rents and household expenditures tended to increase in the fringe, especially due to the BRT. Table B.5 shows the results of estimating equation B.2 using survey data. Column (1) depicts the results on a variable constructed by the Bureau of Statistics which considers paid rents plus hypothetical rents. I show that on average there were no statistically detectable pre-trends, but also no average impacts of either the Metro or the BRT.[6]

Moving to measures of household income and expenditure (columns 2 and 3 of Table B.5), I find no detectable pre-trends of either the Metro or the BRT. Furthermore, I find that, on average, locations closer to the Metro increased their expenditures by 10.6%. I find a similar coefficient when looking at the effect on income, but it is not statistically different from zero. I do not find any average impact of the BRT. In column 4 of Table B.5, I complement previous outcomes with dummies indicating whether household i is poor.[7] Once more, I do not find any pre-trends. Furthermore, I find that the Metro decreased poverty in nearby locations by 4.96 percentage points.

Next, I explore the heterogeneity of the effect across locations. Regarding rents, Panel A of Figure B.11 reveals that the estimated treatment effects are greater the farther away you move from the CBD. This is true for both the Metro and especially for the BRT.[8] Panel B of Figure B.11 shows the heterogenous impact on expenditures. Regarding the Metro, point estimates tend to be quite homogeneous across locations. For the BRT, I find that expenditures increased especially in locations farther away from the CBD.[9]

[Table B.5 here]

[Figure B.11 here]

Additional Fact: On average, people moved out from the stations. This effect is reinforced in the fringe for the BRT, and in the CBD for the Metro. The basic intuition is that location choices can summarize to some extent the impact of the new infrastructure on households. As it will become clear in the theoretical section of the paper, new transit infrastructure can impact households through many channels. The main channel is the improvement in commuting times leading to greater access to jobs. However, the metro can also affect local amenities, for instance by bringing crowds, crime, and noise, and thus can also affect the price of the land. All of these

[6] I obtain similar estimates if I use only paid rents instead.

[7] A household is defined as poor if it is unable to consume 2318 kilocalories per day as well as spend on basic services such as clothing, rent, health, transportation, and education, among others (INEI, 2000).

[8] See Panel A of Figure B.12 for impacts on actual paid rents.

[9] See Figure B.12 for impacts on income and poverty rates.

will impact the location chosen by households. If we observe population decreasing in treated locations relative to control locations, it would be indicative that the new infrastructure is harming welfare in some way in these locations (e.g. by worsening amenities).

Results for the time-event study utilizing Census data are reported in Table B.6. Column (1) of Panel A shows the impact on overall population living in each block. We can see that relative to locations near yet-to-be-constructed stations, population in locations closer to stations shrunk by 12%. Importantly, I do not find evidence for differential pre-trends. In Table B.7, I show the corresponding differences-in-difference specification using the whole sample of 2007 and 2017 locations. Results are similar albeit coefficients are somewhat smaller. While the negative impact may seem surprising given the improvement in commuting times, it is not if we consider that rents were increasing especially in the fringe, and that amenities could be deteriorating.

Next, Figure B.13 shows the heterogeneity of the impact across locations. On the one hand, estimates reveal that the Metro decreased the population located in the CBD. On the other, the BRT had a negative impact on locations in the fringe. This could be because the BRT and the Metro affected access to opportunities, rents, and amenities in different magnitudes across locations. For instance, given that rents increased mainly in the fringe but not in the CBD, and that the Metro had a negative impact on population only in the CBD, this would imply that the Metro deteriorated amenities in the CBD.[10]

[Table B.6 here]

[Figure B.13 here]

Additional Fact: Males are better off in intermediate locations (between the CBD and the fringe) but married women are better off in the fringe. In Panel A of Table B.6, I exhibit the impact on other outcomes. Column (2) shows the impact on the ratio of singles to married, while column (3) shows the impact on the ratio of female to males. Estimates indicate that singles were more benefited than married households: in locations closer to stations relative to locations near yet-to-be-constructed stations, the population of singles relative to married increased by 8.1%. However, on average, the ratio of males and female stayed constant. Finally, column (4) shows the impact on the number of dual earner households relative to male breadwinner households. On

[10]In the main text, I am concerned about the impact of improving commute times since this interacts with the new mechanism I am proposing: interdependence in commute choices. However, one may argue that the impact of the new transit infrastructure on the geography of amenities may have induced reallocation of households in such a way that resulted in changes on gender earnings gap. While the Metro and the BRT may have affected the distribution of exogenous amenities, in Section B.6, I show that it did not have a major impact on the gender earnings gap. So, the analysis in the main text which abstracts away from changes in amenities is without loss of generality.

average I do not find any impact. Also note that, importantly, pre-trends are always statistically indistinguishable from zero.[11]

Panel B of Figure B.13 reveals that locations around 7.5-10 km from the CBD experienced a decrease of singles relative to married for both the Metro and the BRT. However, both at the CBD and at the fringe, some increases were detectable for the BRT. Panel A of Figure B.14 shows the impact on the ratio of females to males. I find that women were more prone to leave locations around 7.5-10 km from the CBD. This implies that males were better off in intermediate locations. Finally, in Panel B of Figure B.14 the ratio of dual-earners to male breadwinner households increased in the fringe. So, while single households are equally or better off in the fringe (and worse off between 7.5-10 km from the CBD), dual-earner households are better off compared to male breadwinner households, implying that opportunities to women—and in particular to married women—expanded in the fringe.

[Figure B.14 here]

In conclusion, first, these results reveal that the new transit infrastructure fostered households to move out from areas closer to the stations. However, this impact was heterogeneous across household groups. For instance, results suggest that males' welfare expanded in intermediate locations, whereas females' welfare did so in the fringe, especially married women's welfare. Second, as it will become clear with the model, this rich heterogeneity comes from a combination of facts related to the geography of improvements in commuting times, the initial distribution of households and jobs across space, potential changes in amenities, and other general equilibrium impacts such as on rents and on wages. The model allows me to tease out the impacts that come exclusively from having better commute times (and the resulting GE impacts). Third, comparing stations to yet-to-be-constructed stations seems to take care of some endogeneity concerns since I do not find strong evidence for pre-trends. Not a single pre-trend coefficient is significant. Fourth, these results point towards the need of a model that takes the geography of Lima very seriously since impacts vary across locations.[12]

Finally, another word of caution when interpreting these impacts is that the introduction of the new transit infrastructure may have induced households to relocate to other parts of the city, as the Census data seems to suggest. In other words, the impacts I estimate are conflating partial and general equilibrium effects. It is hard to extract definitive conclusions from this analysis since the introduction of these transit investments have potentially affected the whole city, blurring the lines between what control and treatment groups are. To better understand these issues and tease out

[11] In Table B.7, I show the corresponding differences-in-difference specification using the whole sample of 2007 and 2017 locations.

[12] One could argue that Lima cannot be fully described by a monocentric model were only one CBD exists. Hence, the geographical dimension in the model should be flexible enough to account for this heterogeneity.

Table B.5: Reduced Form Impact of New Transit Infrastructure on Rents and Income

	A+H Rents	HH Income	HH Exp.	Poverty
	(1)	(2)	(3)	(4)
$Treat_l^M \times Before2009_t$	-0.0010	0.0311	0.0687	-0.0152
	(0.0698)	(0.0781)	(0.0468)	(0.0325)
$Treat_l^M \times After2009_t$	-0.0216	0.0809	0.1059	-0.0496
	(0.0671)	(0.0651)	(0.0464)**	(0.0275)*
$Treat_l^{BRT} \times Before2009_t$	-0.0229	0.0376	-0.0651	0.0448
	(0.0830)	(0.0937)	(0.0516)	(0.0307)
$Treat_l^{BRT} \times After2009_t$	-0.0414	0.1052	-0.0183	-0.0380
	(0.0686)	(0.0900)	(0.0486)	(0.0249)
Block FE	X	X	X	X
Controls	X	X	X	X
N	24246	24246	24246	24246
Est. Method	PPML	PPML	PPML	OLS

the different forces that are interacting in this context, in the main text I start building intuition by developing a partial equilibrium model of commuting in married households. Then, in the quantitative section of the paper I set up a general equilibrium model of city structure based on Ahlfeldt, et al. (2015) and Tsivanidis (2021) but with the addition of interdependent commuting choices within couples.

Table B.6: Time-Event Study of New Transit Infrastructure on Census Outcomes

Panel A:	Total Population	Singles to Married	Females to Males
	(1)	(2)	(3)
$Treat \times (Year = 1993)$	0.0938	0.0396	0.0234
	(0.1070)	(0.0444)	(0.0256)
$Treat \times (Year = 2017)$	-0.1242	0.0812	-0.0033
	(0.0354)***	(0.0276)***	(0.0063)
Panel B:	**Single to Married (Males)**	**Single to Married (Females)**	**Dual-Earners to Male Breadwinner**
	(1)	(2)	(3)
$Treat \times (Year = 1993)$	0.0287	0.0461	-0.0073
	(0.0502)	(0.0442)	(0.0491)
$Treat \times (Year = 2017)$	0.0894	0.0764	-0.0516
	(0.0298)***	(0.0288)***	(0.0315)
Block FE	X	X	X
Year FE	X	X	X
N	375	375	375

Figure B.10: Constructed and Planned Lines

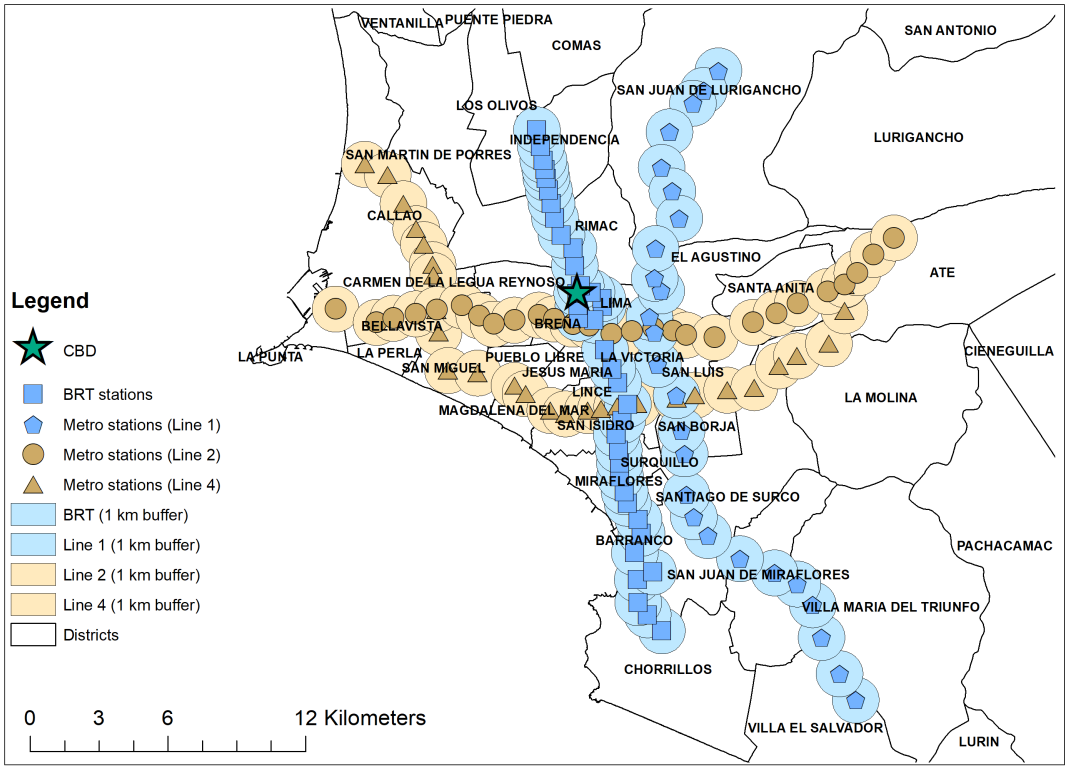


Table B.7: Differences-in-Differences of New Transit Infrastructure on Mobility

Panel A:	Total Population	Singles to Married	Females to Males
	(1)	(2)	(3)
$Treat \times (Year = 2017)$	-0.0964 (0.0212)***	0.0477 (0.0137)***	-0.0041 (0.0030)
Panel B:	Single to Married (Males)	Single to Married (Females)	Dual-Earners to Male Breadwinner
	(1)	(2)	(3)
$Treat \times (Year = 2017)$	0.0534 (0.0147)***	0.0418 (0.0142)***	-0.0070 (0.0135)
Block FE	X	X	X
Year FE	X	X	X
N	1190	1190	1190

Figure B.11: The Reduced Form Impact of Transit Infrastructure on Rents and Expenditure by Distance to the CBD

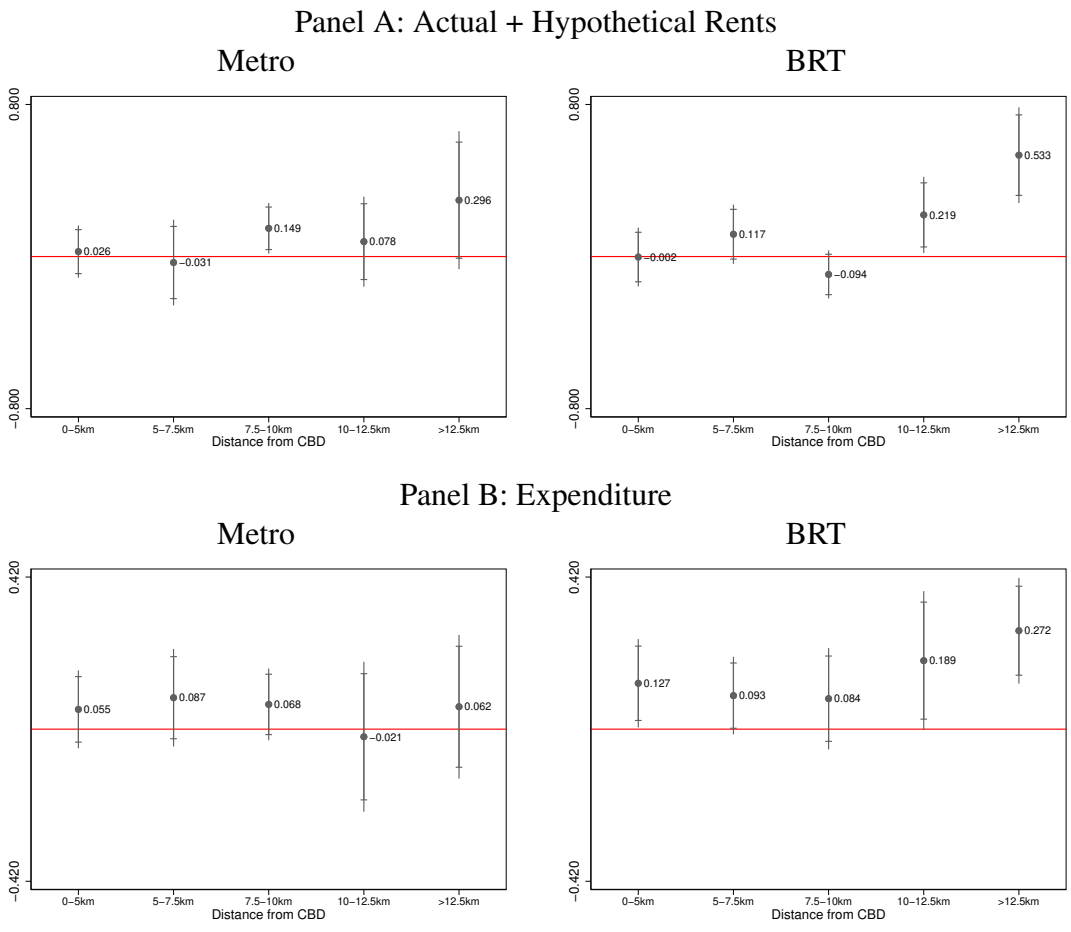
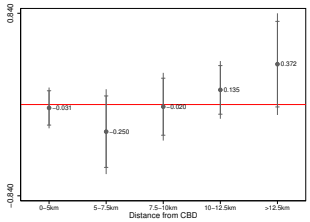


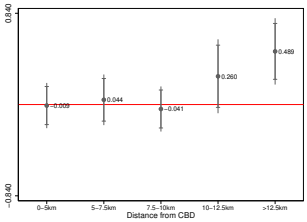
Figure B.12: The Reduced Form Impact of Transit Infrastructure on Rents, Income, and Extreme Poverty

Panel A: Actual Rents

Metro

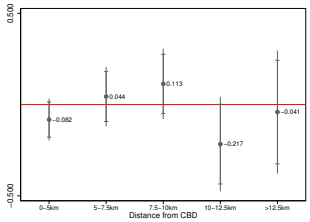


BRT

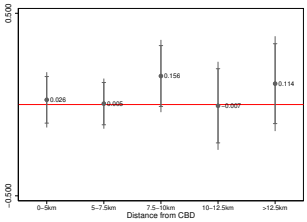


Panel B: Income

Metro

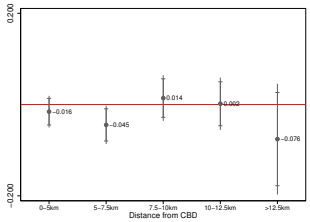


BRT

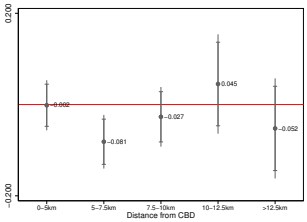


Panel C: Poverty

Metro

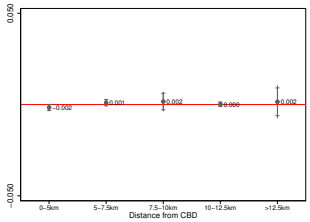


BRT



Panel D: Extreme Poverty

Metro



BRT

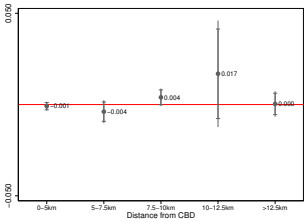


Figure B.13: The Reduced Form Impact of Transit Infrastructure on Census Outcomes by Distance to CBD, Part 1

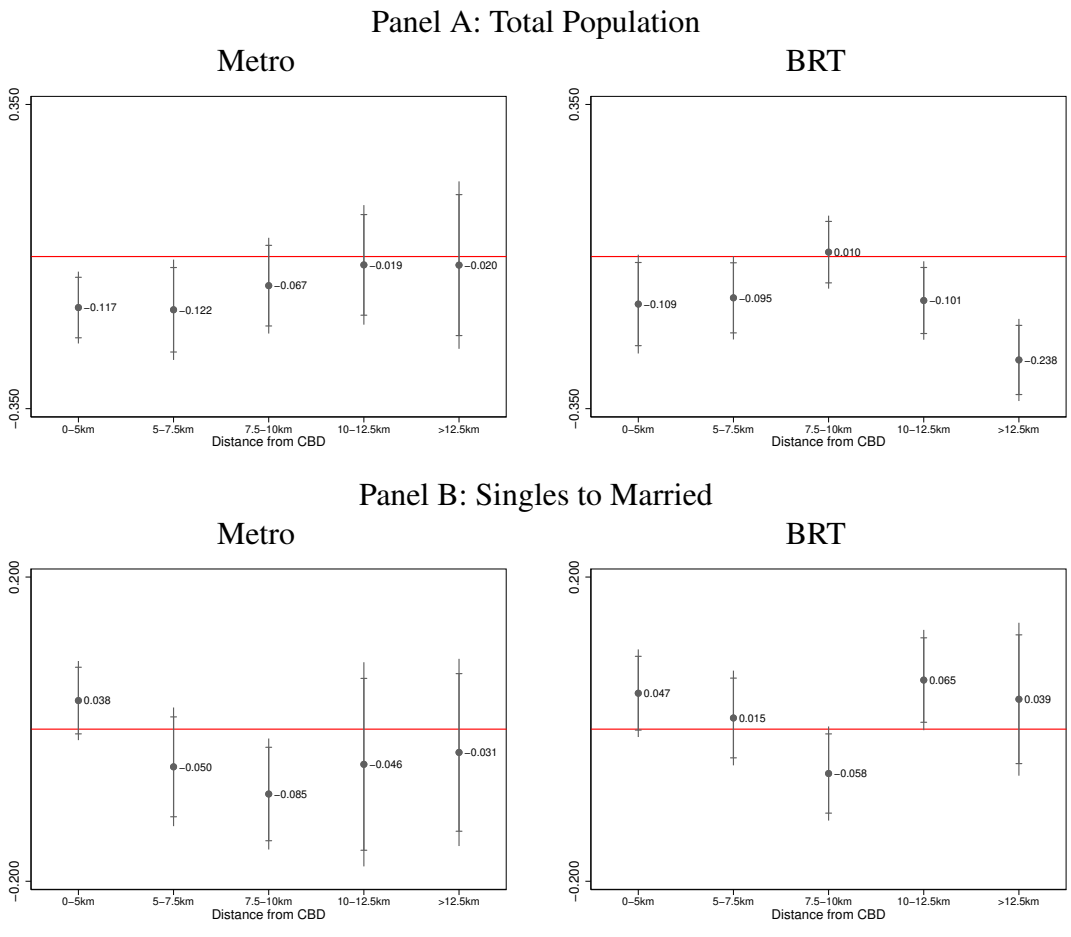
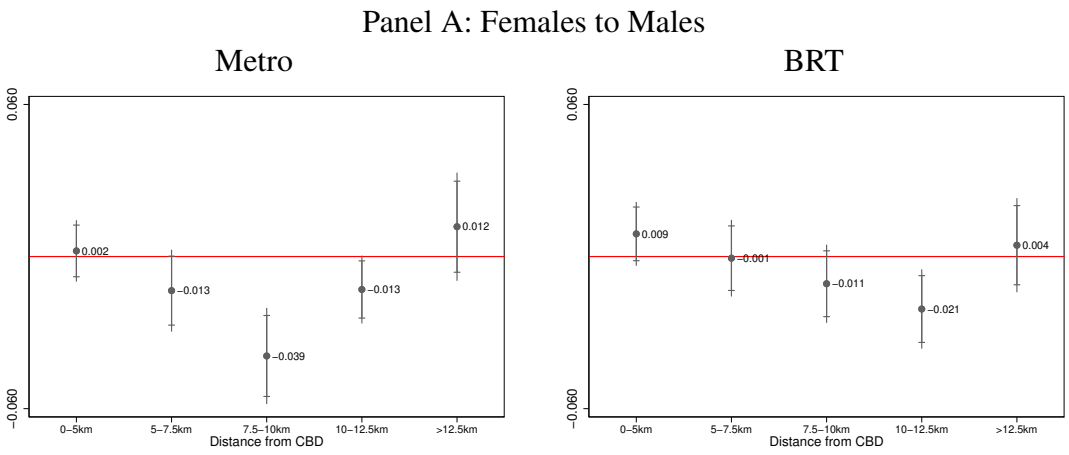
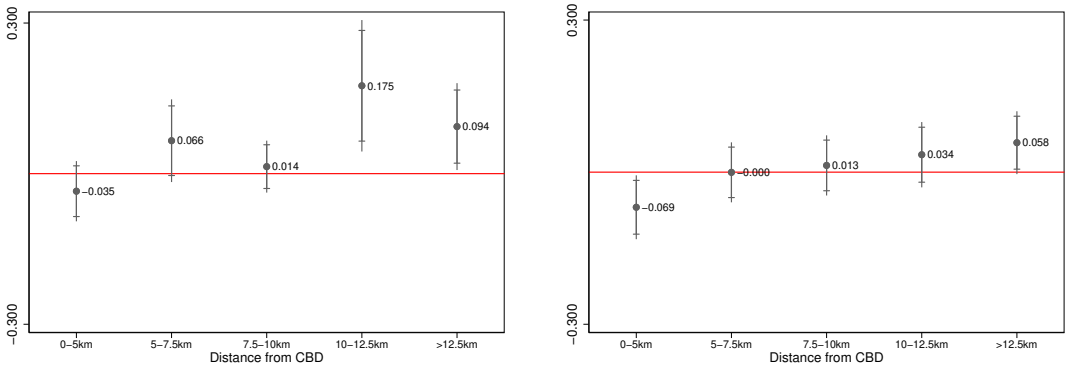


Figure B.14: The Reduced Form Impact of Transit Infrastructure on Census Outcomes by Distance to CBD, Part 2

Panel A: Females to Males



Panel B: Dual-Earners to Male Breadwinner



B.5 Mechanisms

In this section I turn off mechanisms from the full model in order to assess their quantitative relevance. Results are shown in Table B.8. For comparison, the first column shows the aggregate impact on real income and the gender gap using the complete model.

The second column removes externalities, and assume that productivity and amenity terms observed in 2017 remain fixed. This column reveals that the impact on real income and real income per worker rises to 1.39% and 1.02%, respectively. In other words, when endogenous externalities are allowed, real earnings grow slower. This is driven mainly by changes in the price index. In the full model, the price index decreased by 1.78%, whereas with no externalities it decreased by 3.02%. This is despite average productivity across the seven sectors grew by 0.53% in the full model. The explanation is that agglomeration externalities make the labor demand more elastic. Then, shifts in the labor supply have milder effects on wages, which prevents the price index to fall. In fact, males and female wages at destinations decreased by 1.61% and 2.28% in the full model, whereas they decreased by 2.52% and 3.17% in the model without externalities. Finally, the impact on the aggregate gender gap remains around of -2.00%, although the impact on the gender gap in earnings per worker becomes -0.84% rather than -1.05% as it was in the full model.

The third column keeps the distribution of households across locations fixed at its 2017 value. In other words, it assumes that households cannot reallocate in response to the removal of the transit infrastructure. Compared to column (1), the impacts on real income and on real income per capita are 42% and 79% lower, which suggest that mobility of households is an important mechanism. Furthermore, while the impact on the aggregate gender gap remains quite constant, the impact on the gender gap in terms of average earnings becomes -0.78% instead of -1.05%. This is driven by singles for which the gap increases by 0.83% rather than 0.48% as in the full model. Thus, households' mobility allows single women to better adjust to improvements in commute times.

Column (4) assumes that married couples cannot endogenously change their labor force participation. This reduces the impact on aggregate real income by $1 - 0.0096/0.0107 = 10\%$ which speaks to the importance of this margin in the aggregate. Moreover, preventing households to change their labor participation choice has some distributional consequences through the general equilibrium. In particular, when this margin is active, increased supply in female labor tends to decrease females wages relative to male wages. This is reflected on the gender gap in earnings per worker. When married couples have the possibility of changing their labor supply choice, the overall gender gap decreases by 1.05% as a consequence of the new transit infrastructure. Preventing this choice also prevents female wages to decrease even further, leading to a steeper shrinking of the gender gap in earnings per worker, as it decreases by 1.59%.

Finally, Column (5) turns off the three mechanisms above. Doing so decreases the growth

Table B.8: The Aggregate Impact of the Line 1 of the Metro and the BRT - Mechanisms

	Full	No Externalities	No Mobility	No LFP	Neither of the three
	(1)	(2)	(3)	(4)	(5)
Real Income	0.0107	0.0139	0.0062	0.0096	0.0075
Real Income per Worker	0.0070	0.0102	0.0015	0.0099	0.0075
Gender Gap	-0.0215	-0.0200	-0.0233	-0.0112	-0.0114
Gender Gap (per Worker)	-0.0105	-0.0084	-0.0078	-0.0159	-0.0149
Gender Gap, Singles	0.0048	0.0089	0.0083	-0.0018	0.0058
Gender Gap (per Worker), Singles	0.0056	0.0104	0.0098	-0.0022	0.0068
Gender Gap, Married	-0.0413	-0.0401	-0.0457	-0.0200	-0.0227
Gender Gap (per Worker), Married	-0.0254	-0.0237	-0.0230	-0.0285	-0.0303
# DE/# Couples	0.0174	0.0174	0.0223	0	0

Notes: To assess the significance of individual mechanisms, I disable them from the full model and present their impacts in separate columns. The first column reports the overall effect on real income and the gender gap using the complete model. The second column eliminates externalities and assumes that productivity and amenity terms remain constant at their 2017 levels. The third column fixes the distribution of households across locations at the 2017 level. The fourth column assumes that married couples cannot change their labor force participation in response to economic changes.

in aggregate real income from 1.07% to 0.75%. So, between the three they explain 30% of the growth. However, earnings per worker remains almost unchanged. While the mobility margin is an important mechanism through which improved commute times increase real income per worker, the LFP and externalities tended to have the opposite effect. Turning to the gender gap, without the three mechanisms, the reduction in the aggregate earnings gap becomes 1.14% instead of 2.15%. However, the reduction in gender gap in earnings per worker increases from 1.05% to 1.49%. This is mainly driven by the LFP margin. Increased female supply decreases women's average earning since the labor demand is downward sloping.

B.6 Understanding the reduced-form results in light of the model

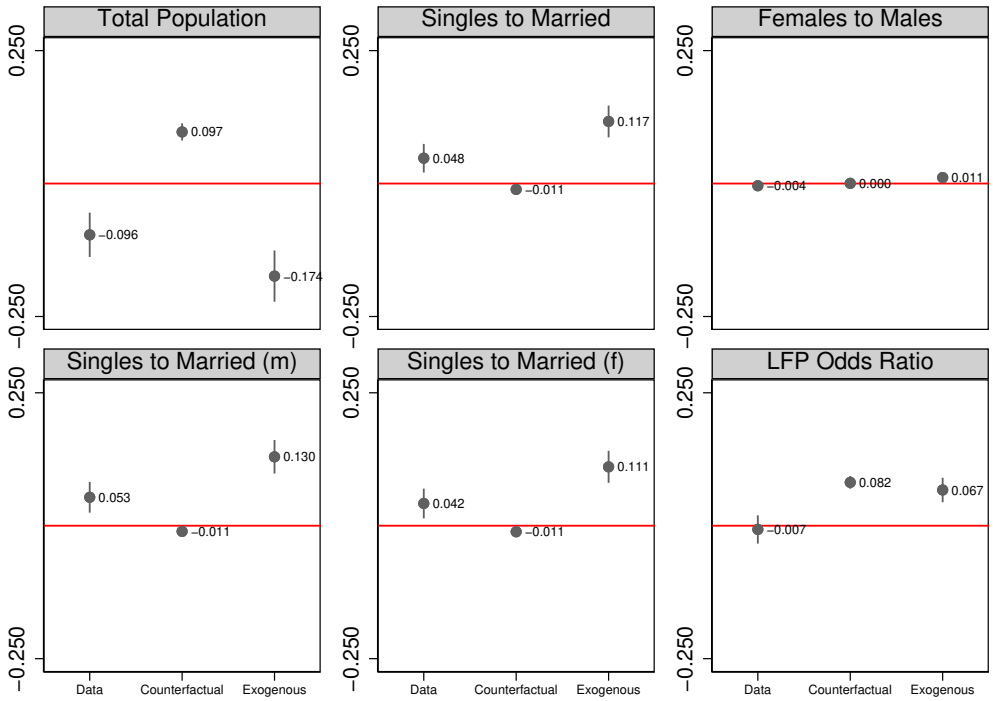
In the reduced-form section of the paper, I explained that, while location choices might be indicative of overall welfare, they conflate changes in market access and amenities. In this section, I re-estimate the differences-in-differences framework using location outcomes comparing locations closer to stations against yet-to-be-constructed stations before and after the commute times were improved. These capture the effect of the new transit infrastructure that goes through better commuting times, but not through exogenous characteristics changing. Thus, I also estimate the differences-in-differences specification using backed-out exogenous amenities from 2007 and 2017 as dependent variable.

Results are shown in Figure B.15. These results show that amenities worsened in locations near

to station. This explains apparent divergence between the counterfactual exercise and the data in terms of total population. While better commuting times fostered households to move closer to stations, amenities pushed households out. For completeness, I also show this exercise by civil status and gender and a similar pattern emerges: exogenous amenities for singles and married changed in non-negligible ways. In Figure B.16, I estimate the differences-in-differences specification but allowing heterogeneous impacts by distance to the CBD. These results reveal that amenities worsened especially in areas farther away from the CBD. However, when comparing singles to married, females to males, and dual-earners to male breadwinner households, counterfactuals and data moved in similar ways.

Finally, I solve the model keeping 2017 commute times constant but with 2007 amenities. The objective of this exercise is to evaluate whether differential changes in amenities in locations close to stations had an impact on the gender gap in earnings per worker. In Figure B.17, I show the results using as outcome the gender gap in earnings per worker in a one difference specification.[13] This figure reveals that earnings per worker remain virtually unchanged for singles and married.

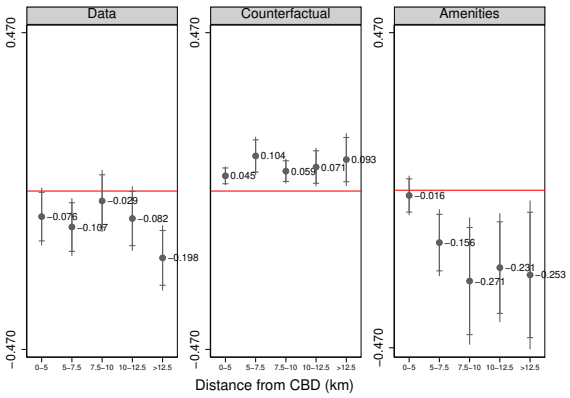
Figure B.15: The Reduced-form Impact of Transit Infrastructure, Data and Model



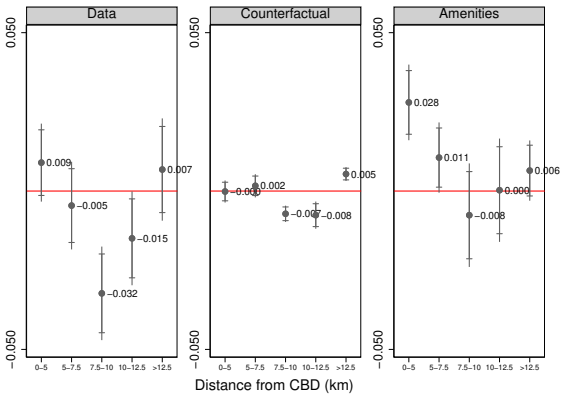
[13] I compare outcomes with 2017 amenities against outcomes with 2007 amenities.

Figure B.16: Reduced-form Impact by Distance to the CBD, Data and Model

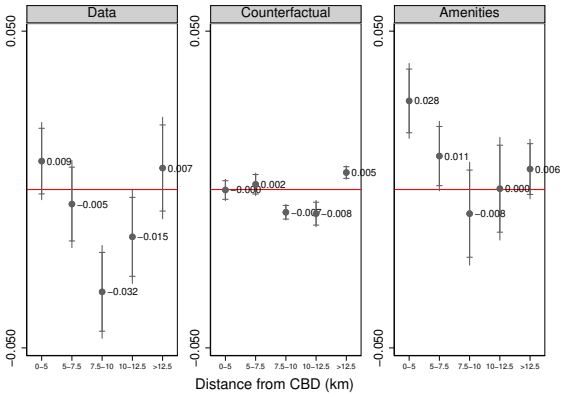
Panel A: Total Population



Panel B: Singles to Married



Panel C: Females to Males



Panel D: Dual-Earners to Male Breadwinner

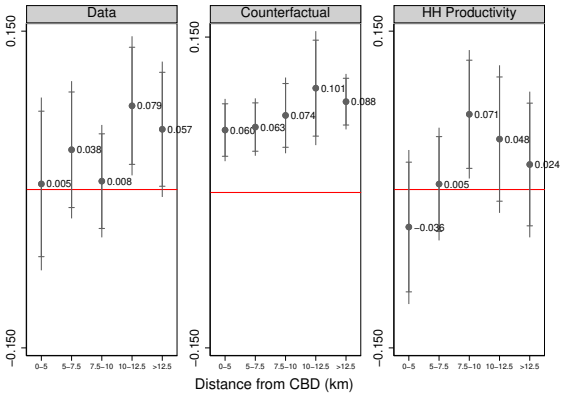
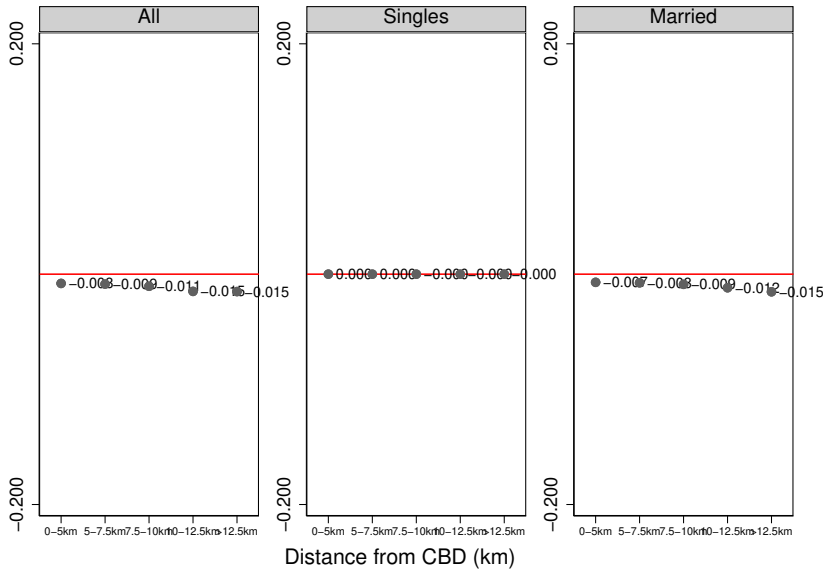


Figure B.17: Reduced-form Impact with Independent Commute Choices by Distance to the CBD

Gap in Real Earnings per Worker



B.7 Policy: Completing the Metro's Network

In this section, I use the model to evaluate what would happen if the Lines 2 and 4 of the Metro are completed (see Online Supplement B.10). To do so, I assume that speeds in these lines are similar to those of Line 1. I then keep 2017 exogenous characteristics constant and incorporate Lines 2 and 4 to the city. I compare this counterfactual scenario against (a) a scenario where the Line 1 of the Metro and the BRT are active, and (b) a scenario where neither Metro Lines or the BRT are active. The first comparison informs us about the incremental impact of completing the network, whereas the second comparison informs us about the overall impact of the Metro and the BRT.

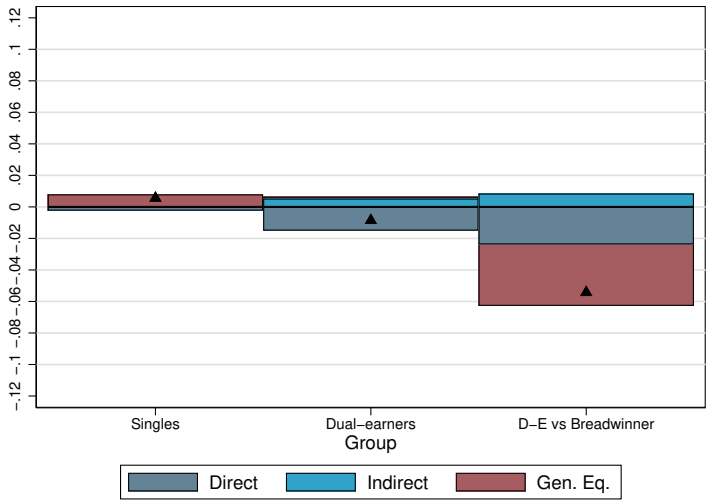
Results are shown in Table B.9. Panel A reveals that on average real income would grow an additional 0.95% compared to the situation where only the Line 1 and the BRT are active. Also, income per worker would increase by 0.56%. Moreover, the gap would decrease by 1.87%. The gap in average earnings would shrink by 0.56%. The effect on the gap is driven by married couples. For these households, the gender earnings gap would decrease by 4.05%. In terms of income per worker, the gap would decrease by 2.26%. Overall, impacts are comparable in magnitude and sign to those of the Line 1 of the Metro and the BRT. The main difference is that in this case, the gap among singles increases in a non-negligible way, i.e. by 1.12%, or 1.31% if looking at the gap per worker.

Panel B exhibits the overall impact of the Metro and the BRT. The complete Metro network plus the BRT increases real income by 2.03%, and average earnings per worker by 1.26%. Also,

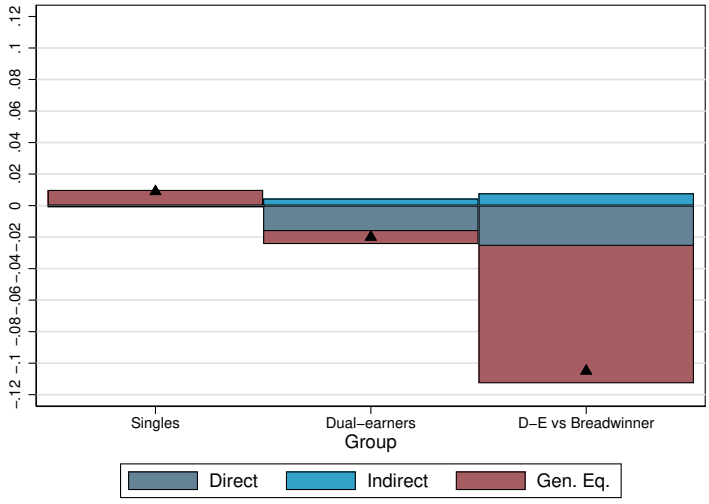
the gap decreases by 3.98% on aggregate. The gap in income per worker decreases by 1.60%. Unsurprisingly, this is due to a decrease in the gap among married. In these households, the gap in aggregate earnings decreases by 8.01%, whereas the gap in earnings per worker decreases by 4.75%.[14]

Figure B.18: Decomposition of the Aggregate Impacts

Panel A: Incremental Impact of Metro Lines 2 and 4



Panel B: Overall Impact of Metro Lines (1, 2, 4) and the BRT



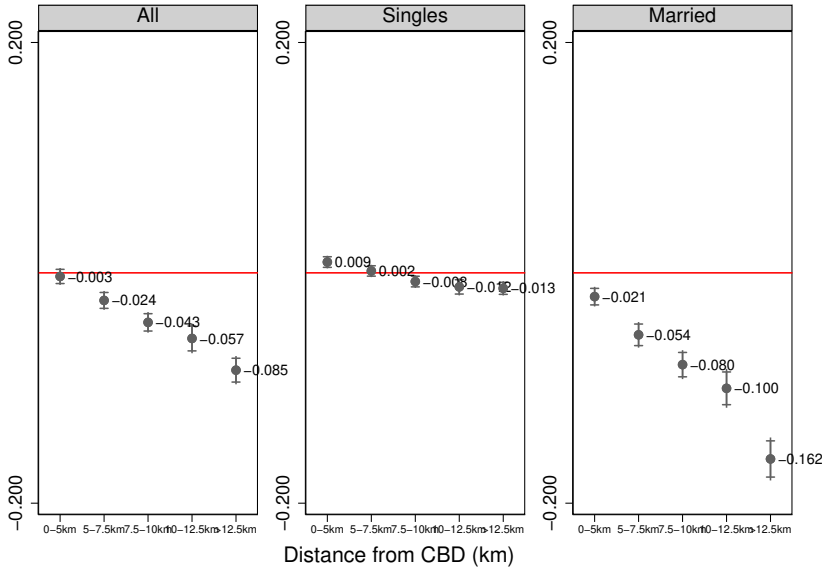
[14]Figure B.18 shows the decomposition of these aggregate impacts. Compared to the analysis of the impact of the Line 1 and the BRT, I observe similar patterns. Finally, in Figure B.19 I show average impacts across locations. Once again, I find that impacts on the gap are concentrated on areas closer to remote stations.

Table B.9: The Aggregate Impact of Completing the Metro Network

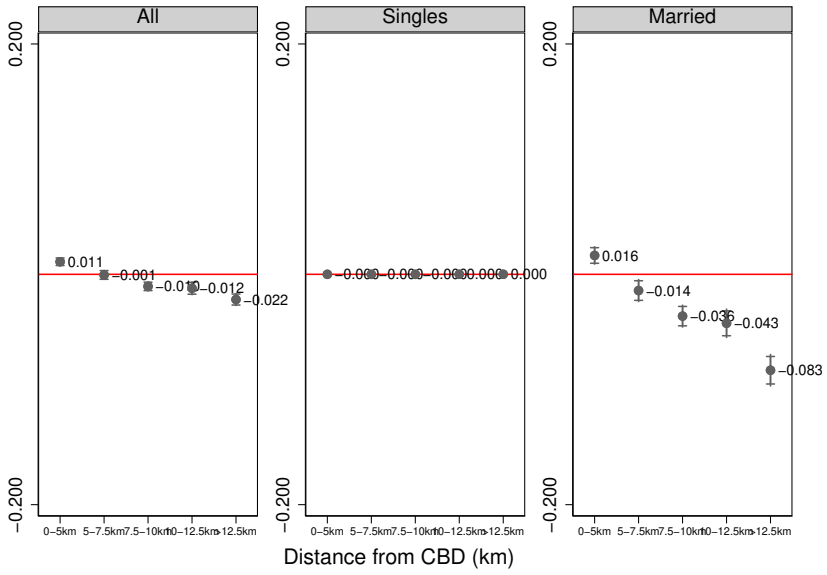
Panel A: Incremental Impact of Metro Lines 2 and 4								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0039	-0.0011	-0.0175	0.0295	0.0386	0.0063	0.0185	0.0095
Real Income per Worker	0.0039	-0.0011	0.0072	0.0112	0.0201	0.0063	0.0090	0.0056
Gender Gap	0.0112			-0.0405		-0.0187		
Gender Gap (per Worker)	0.0131			-0.0226		-0.0056		
# D.E. / # Married				0.0181				
Panel B: Overall Impact of Metro Lines (1,2, 4) and the BRT								
	Singles		Married			All		
	Males	Females	Males in BW HH	Males in D.E. HH	Females in D.E. HH	Males	Females	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Real Income	0.0099	0.0028	-0.0333	0.0596	0.0797	0.0133	0.0402	0.0203
Real Income per Worker	0.0099	0.0028	0.0139	0.0230	0.0424	0.0133	0.0213	0.0126
Gender Gap	0.0161			-0.0801		-0.0398		
Gender Gap (per Worker)	0.0188			-0.0475		-0.0160		
# D.E. / # Married				0.0358				

Figure B.19: Incremental Impact of Metro Lines 2 and 4 by Distance to the CBD

Panel A: Gap in Real Aggregate Income



Panel B: Gap in Real Earnings per Worker



B.8 Online Supplement: Solution Algorithm

Corner solutions and observed prices

Given that I have agglomeration forces in the model, it is possible to have corner allocations for some locations. Without these externalities it is only possible if the exogenous component of productivity or of amenities is equal to zero. With externalities, even if the exogenous component of productivity or of amenities is positive, I can have corner solutions as the endogenous part, which enters multiplicatively, can be equal to zero when either employment or residence are equal to zero.

This implies that the groups $\mathcal{I}_{FR}$, $\mathcal{I}_F$, and $\mathcal{I}_R$ are endogenously determined:

$$\begin{aligned}
\mathcal{I}_{FR} &= \{i : u_{ik} > 0, A_{is} > 0 \text{ for some } k \in mf, m, f \text{ and for some } s\} \\
\mathcal{I}_R &= \{i : u_{ik} > 0, A_{is} = 0 \text{ for some } k \in mf, m, f \text{ and for all } s\} \\
\mathcal{I}_F = &\quad \{i : u_{ik} = 0, A_{is} > 0 \text{ for all } k \in mf, m, f \text{ and for some } s\}
\end{aligned}$$

Following Ahlfeldt et al. (2015), we can define the observed price of floor space (r_i). We can summarize the relationship between the observed price of floor space (r_i), the price of commercial floor space (r_{F_i}), the price of residential floor space (r_{R_i}), and land use as (ϑ_i):

$$\begin{aligned}
r_i = &\begin{cases} \zeta_{F_i} r_{F_i} & \zeta_{F_i} = 1, \quad i \in \mathcal{I}_F = \{i : u_{ik} = 0 \text{ for all k}, A_{is} > 0 \text{ for some s}\} \\ \zeta_{F_i} r_{F_i} & \zeta_{F_i} = 1 - \tau_i, \quad i \in \mathcal{I}_{FR} = \{i : u_{ik} > 0 \text{ for some k}, A_{is} > 0 \text{ for some s}\} \end{cases} \\
r_i = &\begin{cases} \zeta_{R_i} r_{R_i} & \zeta_{R_i} = 1, \quad i \in \mathcal{I}_R = \{i : u_{ik} > 0 \text{ for some k}, A_{is} = 0 \text{ for all s}\} \\ \zeta_{R_i} r_{R_i} & \zeta_{R_i} = 1, \quad i \in \mathcal{I}_{FR} = \{i : u_{ik} > 0 \text{ for some k}, A_{is} > 0 \text{ for some s}\} \end{cases}
\end{aligned} \tag{B.3}$$

where ζ_{F_i} and ζ_{R_i} relate observed floor prices to commercial and residential floor prices. $\mathcal{I}_F$ is the set of locations specialized in commercial activity ($1-\vartheta_i = 1$), $\mathcal{I}_R$ is the set of locations specialized in residential activity ($\vartheta_i = 1$), and $\mathcal{I}_{FR}$ is the set of locations with both commercial and residential activity ($\vartheta_i \in (0, 1)$). Note that according to equation B.3, the relationship between he observed price of floor space (r_i), the price of commercial floor space (r_{F_i}), the price of residential floor space (r_{R_i}), and land use as (ϑ_i), are a function only of the exogenous locational characteristics given by the vector defined by A_{is}, u_{ig}, and τ_i.

Algorithm

The system of equations defined above can be solved using the following algorithm.

1. Guess vector $w_g^0, r^0, \vartheta^0, u_{mf}^0, u_m^0, u_f^0, A_s^0, \mu^0$.

   - Vector of prices should have normalization, e.g. $w_{1m} = 1$, so divide the vector of wages and rents by w_{1m}.

2. Given a vector $w_g^t, r^t, \vartheta^t, u_{mf}^t, u_m^t, u_f^t, A_s^t, \mu^t$:

   (a) Given our guess for productivity and amenities, we can determine the group indicator variables:

$$\begin{aligned}
\mathcal{I}_{FR} &= \{i : u_{ik} > 0, A_{is} > 0 \text{ for some } k \in mf, m, f \text{ and for some } s\} \\
\mathcal{I}_{R} &= \{i : u_{ik} > 0, A_{is} = 0 \text{ for some } k \in mf, m, f \text{ and for all } s\} \\
\mathcal{I}_{F} &= \{i : u_{ik} = 0, A_{is} > 0 \text{ for all } k \in mf, m, f \text{ and for some } s\}
\end{aligned}$$

   (b) Construct rents for this iteration:

$$r_{R_i} = \begin{cases} r_i \times (1 - \tau_i) & \text{if} i \in \mathcal{I}_{FR} \\ r_i & \text{if} i \in \mathcal{I}_{R} \\ . & \text{if} i \in \mathcal{I}_{F} \end{cases} \quad r_{F_i} = \begin{cases} r_i & \text{if} i \in \mathcal{I}_{FR} \\ . & \text{if} i \in \mathcal{I}_{R} \\ r_i & \text{if} i \in \mathcal{I}_{F} \end{cases}$$

   (c) Compute total rents:

$$E = \sum_i r_{R_i} H_{R_i} + r_{F_i} H_{F_i}$$

   (d) Compute residence market access for single households:

$$\Psi_{R_i^g} = \sum_{l'} \left(\frac{d_{il'}^g}{w_{l'g}} \right)^{-\theta^g}, \forall g \in \{m, f\}$$

   (e) Compute residence market access for couples:

      i. In male breadwinner households:

$$\Phi_{R_{i,\ell=0}}^h = \sum_{l'} \left(\frac{d_{il}^h}{w_{lm}} \right)^{-\theta^h}$$

ii. In dual-earner households:

$$\Phi^{h}_{R_{i,\ell=1,j'}} = \sum_{l}\left(\frac{d^{h}_{il}}{w_{lm}+w_{j'f}}\right)^{-\theta^{h}}$$

$$\Phi^{w}_{R_{i,\ell=1}} = \sum_{l}\left(\frac{\left(\Phi^{h}_{R_{i,\ell=1,l}}\right)^{1/\theta^{h}}}{d^{w}_{il}}\right)^{\theta^{w}}$$

iii. On average:

$$\Psi_{R^{mf}_{i}} = \left[\left(\Phi_{R^{w}_{i,\ell=1}}\right)^{\frac{\nu}{\theta^{w}}} + \left(\xi^{mf}_{i\ell=0}\right)^{\nu}\left(\Phi_{R^{h}_{i,\ell=0}}\right)^{\frac{\nu}{\theta^{h}}}\right]^{1/\nu}$$

(f) Compute supply of residents to each location for each gender-by-household type:

$$N_{R^{k,k}_{i}} = \pi^{k,k}_{i}N^{k,k}, k\in\{m,f\}$$

$$N_{R^{mf}_{i}} = \underbrace{(1-\mu_{i})\,\pi^{mf}_{i}N^{mf}}_{N_{R^{mf}_{i,0}}} + \underbrace{(\mu_{i})\,\pi^{mf}_{i}N^{mf}}_{N_{R^{mf}_{i,1}}}$$

where

$$\pi^{g}_{i} = \frac{\left(\frac{r^{1-\beta^{g}}_{R_{i}}}{u^{g}_{i}\Psi^{1/\theta g}_{R^{g}_{i}}}\right)^{-\eta^{g}}}{\sum_{l}\left(\frac{r^{1-\beta^{g}}_{R_{l}}}{u^{g}_{l}\Psi^{1/\theta g}_{R^{g}_{l}}}\right)^{-\eta^{g}}}$$

and

$$\pi^{mf}_{i} = \frac{\left(\frac{r^{1-\beta^{mf}}_{R_{i}}}{u^{mf}_{i}\Psi^{1/\nu}_{R^{mf}_{i}}}\right)^{-\eta^{mf}}}{\sum_{l}\left(\frac{r^{1-\beta^{mf}}_{R_{i}}}{u^{mf}_{l}\Psi^{1/\nu}_{R^{mf}_{l}}}\right)^{-\eta^{mf}}}$$

(g) Compute commuting probabilities:

i. For singles:

$$\pi^{g}_{j|i} = \frac{\left(\frac{d^{g}_{ij}}{w_{jg}}\right)^{-\theta^{,g}}}{\Phi_{R^{g}_{i}}}$$

ii. For males in male-breadwinner households:

$$\pi^{h}_{j|i,\ell=0} = \frac{\left(\frac{d^{h}_{ij}}{w_{jm}}\right)^{-\theta^{h}}}{\Phi^{h}_{R_{i,\ell=0}}}$$

iii. For females in dual-earner households:

$$\pi^{w}_{j'|i,\ell=1} = \frac{\left(\frac{d^{w}_{ij'}}{\left(\Phi^{h}_{R_{i,\ell=1,j'}}\right)^{1/\theta^{h}}}\right)^{-\theta^{w}}}{\Phi^{w}_{R_{i,\ell=1}}}$$

iv. For males in dual-earner households:

$$\pi^{h}_{j|i,\ell=1} = \sum_{l} \pi^{h}_{j|i,\ell=1,l} \cdot \pi^{w}_{l|i,\ell=1}$$

where

$$\pi^{h}_{j|i,\ell=1,j'} = \frac{\left(\frac{d^{h}_{ij}}{w_{jm}+w_{j'f}}\right)^{-\theta^{h}}}{\Phi^{h}_{R_{i,\ell=1,j'}}}$$

(h) Compute labor supply:

$$\begin{aligned} L_{F_{jm}} &= \sum_{i}\left\{\pi^{m,m}_{j|i} N_{R^{m,m}_{i}}\right\} + \sum_{i}\left\{\pi^{h}_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\} + \sum_{i}\left\{\pi^{h}_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} \\ L_{F_{jf}} &= \sum_{i}\left\{\pi^{f,f}_{j|i} N_{R^{f,f}_{i}}\right\} + \sum_{i}\left\{\pi^{w}_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} \end{aligned}$$

(i) Compute total income:

$$\begin{aligned} y^{mf}_{i,\ell=1} &= \underbrace{\sum_j \left(w_{j,m}\pi^h_{j|i,\ell=1}\right) N_{R^{mf}_{i,1}}}_{y^h_{i,\ell=1}} + \underbrace{\sum_j \left(w_{j,f}\pi^w_{j|i,\ell=1}\right) N_{R^{mf}_{i,1}}}_{y^w_{i,\ell=1}} \\ y^{mf}_{i,\ell=0} &= \sum_j \left(w_{j,m}\pi^h_{j|i,\ell=0}\right) N_{R^{mf}_{i,0}} \\ y^g_i &= \sum_j \left(w_{j,g}\pi^g_{j|i}\right) N_{R^g_i} \end{aligned}$$

(j) Compute price index:

$$\begin{aligned} p_{js} &= W^{\alpha_s}_{js} r^{1-\alpha_s}_{Fj} / A_{js} \\ P &= \left(\sum \left(W^{\alpha_s}_{js} r^{1-\alpha_s}_{Fj} / A_{js}\right)^{1-\sigma_D}\right)^{\frac{1}{1-\sigma_D}} \\ W_{js} &= \left(\sum_g \alpha^{\sigma}_{sg} w^{1-\sigma}_{jg}\right)^{\frac{1}{1-\sigma}} \end{aligned}$$

(k) Update guesses using demand side:

i. Update labor force participation:

$$\mu^{t+1}_{i,mf} = \frac{\left[\left(\Phi^w_{R_{i,\ell=1}}\right)^{-1/\theta^w}\right]^{-\nu}}{\Psi_{R^{mf}_i}}$$

ii. Update endogenous productivities:

$$A^{t+1}_{js} = \bar{A}_{js}\left(L_{F_{js}}\right)^{\phi_{Fs}}$$

iii. Plugging labor supply into labor demand to update new wages:

$$\tilde{w}_{jg} = \left(\frac{\sum_s \left(\frac{1}{\alpha_{sg} W_{js}}\right)^{-\sigma} M_{js}}{L_{F_{jg}}}\right)^{1/\sigma}$$

where:

$$\begin{aligned} M_{js} &= \alpha_s \frac{X_{js}}{W_{js}} \\ X_{js} &= \frac{p_{js}^{1-\sigma_D}}{P^{1-\sigma_D}} X \\ X = PC &= \sum_i PC_i + \sum_i \left(H_{R_i} r_{R_i} + H_{F_i} r_{F_i}\right) \\ &= \beta^m y_i^{m,m} + \beta^f y_i^{f,f} + \beta^{mf} y_{i,\ell=1}^{mf} + \beta^{mf} y_{i,\ell=0}^{mf} + E \end{aligned}$$

iv. Update observed rents. To do so, note that aggregate expenditure on housing in location i by households is:

$$\begin{aligned} E_{Ri} &= r_{R_i} H_{R_i} \\ &= \left(1-\beta^m\right) y_i^{m,m} + \left(1-\beta^f\right) y_i^{f,f} + \left(1-\beta^{mf}\right) y_{i,\ell=1}^{mf} + \left(1-\beta^{mf}\right) y_{i,\ell=0}^{mf} \end{aligned}$$

Similarly, note that aggregate spending on floorspace by firms is:

$$H_{F_i} = \textstyle\sum_s \left(1-\alpha_s\right) \frac{X_{js}}{r_{F_j}}$$

Therefore, our new guess for observed rents would be as follows:

$$\begin{aligned} \text{if i}\in \mathcal{I}_R, \text{then} r_i^{t+1} &= E_{Ri}/H_i \\ \text{if i}\in \mathcal{I}_F, \text{then} r_i^{t+1} &= X_i/H_i \\ \text{if i}\in \mathcal{I}_{FR}, \text{then} r_i^{t+1} &= \frac{\frac{E_{Ri}}{(1-\tau_i)} + X_i}{H_i} \end{aligned}$$

v. Update land allocation:

$$\begin{aligned} \text{if i}\in \mathcal{I}_R, \text{then} \vartheta_i &= 1 \\ \text{if i}\in \mathcal{I}_F, \text{then} \vartheta_i &= 0 \\ \text{if i}\in \mathcal{I}_{FR}, \text{then} \vartheta_i &= \frac{\frac{E_{Ri}}{(1-\tau_i)}}{\frac{E_{Ri}}{(1-\tau_i)} + X_i} \end{aligned} \tag{B.4}$$

(l) Normalize new vector of prices.

(m) If $|| \left(w_g^{t+1}, r^{t+1}, \vartheta^{t+1}, u_{mf}^{t+1}, u_m^{t+1}, u_f^{t+1}, A_s^{t+1}, \mu^{t+1}\right)$ –

$\left(w_g^t, r^t, \vartheta^t, u_{mf}^t, u_m^t, u_f^t, A_s^t, \mu^t\right) || \leq tolerance$, then stop. Otherwise, set:

$$\begin{aligned}\left(w_g^{t+1}, r^{t+1}, \vartheta^{t+1}, u_{mf}^{t+1}, u_m^{t+1}, u_f^{t+1}, A_s^{t+1}, \mu^{t+1}\right) &= \\ \rho\left(w_g^{t+1}, r^{t+1}, \vartheta^{t+1}, u_{mf}^{t+1}, u_m^{t+1}, u_f^{t+1}, A_s^{t+1}, \mu^{t+1}\right) & \\ +(1-\rho)\left(w_g^t, r^t, \vartheta^t, u_{mf}^t, u_m^t, u_f^t, A_s^t, \mu^t\right) &\end{aligned}$$

where $\rho \in (0,1)$.

(n) Return to step 2.a

B.9 Proofs

B.9.1 Proposition 1

PROPOSITION: *Given data on residence by gender and household type, $N_{R_{i,\ell}^k}$, employment by gender $L_{F_{jg}}$ and commute costs d_{ij}^k, and parameters θ^k, there exists a vector of wages w_{jg} up to scale that rationalizes the observed data as a model equilibrium. Additionally, if one does not observe $L_{F_{jg}}$ but rather employment by industry $L_{F_{js}}$ and parameters α_{sg} and σ, there exists a vector of wages w_{jg} up to scale that rationalizes the observed data as a model equilibrium.*

Proof. Recall that labor supply is:

$$\begin{aligned}L_{F_{jm}} &= \sum_i \left\{\pi_{j|i}^{m,m} N_{R_i^{m,m}}\right\} + \sum_i \left\{\pi_{j|i,\ell=0}^h N_{R_{i,0}^{mf}}\right\} + \sum_i \left\{\pi_{j|i,\ell=1}^h N_{R_{i,1}^{mf}}\right\} \\ L_{F_{jf}} &= \sum_i \left\{\pi_{j|i}^{f,f} N_{R_i^{f,f}}\right\} + \sum_i \left\{\pi_{j|i,\ell=1}^w N_{R_{i,1}^{mf}}\right\}\end{aligned}$$

Or

$$\begin{aligned}L_{F_{jm}} &= w_{jm}^{\theta^m} \sum_i \left[\frac{N_{R_i^m}\left(d_{ij}^m\right)^{-\theta^m}}{\sum_{l'} w_{jm}^{\theta^m}\left(d_{il'}^m\right)^{-\theta^m}}\right] + w_{jm}^{\theta^h} \sum_i \left[\frac{N_{R_{i,0}^{mf}}\left(d_{ij}^h\right)^{-\theta^h}}{\sum_l w_{lm}^{\theta^h}\left(d_{il}^h\right)^{-\theta^h}}\right] \\ &+ \sum_i \left[\frac{N_{R_{i,1}^{mf}}\left(d_{ij}^h\right)^{-\theta^h}}{\sum_{l'}\left(\sum_n \left(\frac{d_{in}^h}{w_{nm}+w_{l'f}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h}\left(d_{il'}^w\right)^{-\theta^w}} \times\right. \\ &\times \left.\sum_l \frac{\left(d_{il}^w\right)^{-\theta^w}\left(\sum_n \left(\frac{d_{in}^h}{w_{nm}+w_{lf}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h-1}}{\left(w_{jm}+w_{lf}\right)^{-\theta^h}}\right]\end{aligned}$$

and

$$
\begin{aligned}
L_{F_{jf}} &= w_{jf}^{\theta^f} \sum_i \left[\frac{N_{R_i^f} \left(d_{ij}^f \right)^{-\theta^f}}{\sum_{l'} w_{jf}^{\theta^f} \left(d_{il'}^f \right)^{-\theta^f}} \right] \\
&+ \sum_i \left[\frac{N_{R_{i,1}^{mf}} \left(d_{ij}^w \right)^{-\theta^w}}{\sum_{l'} \left(\sum_n \left(\frac{d_{in}^h}{w_{nm}+w_{l'f}} \right)^{-\theta^h} \right)^{\theta^w/\theta^h} (d_{il'}^w)^{-\theta^w}} \left(\sum_n \left(\frac{d_{in}^h}{w_{nm}+w_{jf}} \right)^{-\theta^h} \right)^{\theta^w/\theta^h} \right]
\end{aligned}
$$

Inuitively, if one could observe residence supply by gender and household type, $N_{R_{i,k}}^{\ell}$, and employment by gender, $L_{F_{jg}}$, then one could solve the system defined above.

Additionally, if one does not observe $L_{F_{jg}}$ but rather $L_{F_{js}}$ and knows parameters α_{sg} and σ, from FOCs of the firm, one gets:

$$
L_{F_{jg}} = \sum_s \frac{\left(w_{jg}/\alpha_{sg} \right)^{-\sigma}}{\sum_{h \in m,f} \left(w_{jh}/\alpha_{sh} \right)^{-\sigma}} L_{F_{js}}
$$

We can plug this expression to the equations above, and solve the system. In what is next, I follow Mas-Collel, Whinston, and Green (1995) and Allen, and Arkolakis (2015) to prove existence. While existence is assured, uniqueness is not unless additional conditions are imposed. Let $\mathcal{Z}_{jm}$ and $\mathcal{Z}_{jf}$ be defined as:

$$
\begin{aligned}
\mathcal{Z}_{jm} &= \sum_s \frac{\left(w_{jm}/\alpha_{sm} \right)^{-\sigma}}{\sum_{g \in m,f} \left(w_{jg}/\alpha_{sg} \right)^{-\sigma}} L_{F_{js}} - \sum_i \left\{ \pi_{j|i}^m N_{R_i^m} \right\} + \sum_i \left\{ \pi_{j|i,\ell=0}^h N_{R_{i,0}^{mf}} \right\} \\
&+ \sum_i \left\{ \pi_{j|i,\ell=1}^h N_{R_{i,1}^{mf}} \right\} \\
\mathcal{Z}_{jf} &= \sum_s \frac{\left(w_{jf}/\alpha_{sf} \right)^{-\sigma}}{\sum_{g \in m,f} \left(w_{jg}/\alpha_{sg} \right)^{-\sigma}} L_{F_{js}} - \sum_i \left\{ \pi_{j|i}^f N_{R_i^f} \right\} + \sum_i \left\{ \pi_{j|i,\ell=1}^w N_{R_{i,1}^{mf}} \right\}
\end{aligned}
$$

Part I: Existence. $\mathcal{Z}_z = [\mathcal{Z}_{jm}, \mathcal{Z}_{jf}]$ can be decomposed into $\mathcal{Z}_z(w) = \sum_s \mathcal{L}_{sj}(w) - \sum_k \mathcal{N}_{kj}(w)$ where $\mathcal{L}_{sj}(w) \geq 0$ and $\mathcal{N}_{kj}(w) \geq 0$ are, respectively, homogeneous of degree ζ_s and $\varkappa_k$ with $\zeta_s = \varkappa_k = 0$. Since I am stacking $\mathcal{Z}_{jm}$ and $\mathcal{Z}_{jf}$, then the index z corresponds to each location. So, there is a mapping between the index z in vector $\mathcal{Z}_z$, and the index of a location j in $\mathcal{Z}_{jm}$ and $\mathcal{Z}_{jf}$:

$$
j = \begin{cases} z & z \leq J \\ z - J & z \geq J + 1 \end{cases}
$$

First, notice that

$$\mathcal{L}_{sz}\left(w_m, w_f\right) = \begin{cases} \frac{\left(w_{jm}/\alpha_{sm}\right)^{-\sigma}}{\sum_{g\in m,f}\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}} L_{F_{js}} & z \leq J \\ \frac{\left(w_{jf}/\alpha_{sf}\right)^{-\sigma}}{\sum_{g\in m,f}\left(w_{jg}/\alpha_{sg}\right)^{-\sigma}} L_{F_{js}} & z \geq J+1 \end{cases}$$

is homogeneous of degree zero. Second,

$$\mathcal{N}_{kz}\left(w_m, w_f\right) = \begin{cases} \sum_i \left\{\pi^m_{j|i} N_{R^m_i}\right\} & z \leq J, k = s \\ \sum_i \left\{\pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\} & z \leq J, k = mf0 \\ \sum_i \left\{\pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} & z \leq J, k = mf1 \\ \sum_i \left\{\pi^f_{j|i} N_{R^f_i}\right\} & z \geq J+1, k = s \\ \sum_i \left\{\pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} & z \geq J+1, k = mf1 \end{cases}$$

When $k = s$, it is easy to see that either $\sum_i \left\{\pi^m_{j|i} N_{R^m_i}\right\}$ and $\sum_i \left\{\pi^f_{j|i} N_{R^f_i}\right\}$ are homogeneous of degree zero. Similarly, when $k = mf0, \sum_i \left\{\pi^h_{j|i,\ell=0} N_{R^{mf}_{i,0}}\right\}$ is homogeneous of degree zero. Finally, by inspection,

$$\sum_i \left\{\pi^w_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} = \sum_i \left[\frac{N_{R^{mf}_{i,1}} \left(d^w_{ij}\right)^{-\theta^w}}{\sum_{l'} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{l'f}}\right)^{-\theta^{mf,m}}\right)^{\theta^w/\theta^h} \left(d^w_{il'}\right)^{-\theta^w}} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{jf}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h}\right]$$

and

$$\sum_i \left\{\pi^h_{j|i,\ell=1} N_{R^{mf}_{i,1}}\right\} =$$

$$\sum_i \left[\frac{N_{R^{mf}_{i,1}} \left(d^h_{ij}\right)^{-\theta^h}}{\sum_{l'} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{l'f}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h} \left(d^{mf,f}_{il'}\right)^{-w}} \sum_l \frac{\left(d^w_{il}\right)^{-\theta^w} \left(\sum_n \left(\frac{d^h_{in}}{w_{nm}+w_{lf}}\right)^{-\theta^h}\right)^{\theta^w/\theta^h-1}}{k^{-\theta^h}\left(w_{jm}+w_{lf}\right)^{-\theta^h}}\right]$$

are homogeneous of degree zero. Thus, we can normalize the vector w. Since Walras' Law is satisfied, and the excess demand function is continuous, existence is assured by Brouwer's fixed point theorem.

□

Next I discuss why the uniqueness proof fails.

Part II: Uniqueness. If $\mathcal{Z}_z$ satisfies gross substitution, that is $\frac{\partial \mathcal{Z}_z}{\partial w_z} < 0$ and $\frac{\partial \mathcal{Z}_z}{\partial w_{-z}} > 0$, where

$w_z = \begin{cases} w_{jm} & z \leq J \\ w_{jf} & z \geq J+1 \end{cases}$, then uniqueness is assured. We need to compute the derivative of labor supply and labor demand to wages. Starting with labor demand, we get that:

$$\begin{aligned} \frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{1m}} &= \sum_s \frac{\partial}{\partial \log w_{1m}} \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g\in m,f}(w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1s}} \\ &= -\sigma \sum_s \left(1 - \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g\in m,f}(w_{1g}/\alpha_{sg})^{-\sigma}}\right) L_{F_{1sm}} < 0 \end{aligned}$$

And

$$\begin{aligned} \frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{1f}} &= \sum_s \frac{\partial}{\partial \log w_1} \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g\in m,f}(w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1s}} \\ &= \sigma \sum_s \frac{(w_{1f}/\alpha_{sf})^{-\sigma}}{\sum_{g\in m,f}(w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1sf}} > 0 \end{aligned}$$

And,

$$\frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{jm}} = \frac{\partial \sum_s \mathcal{L}_{s1m}(w)}{\partial \log w_{jf}} = 0$$

Now, focusing on labor supply of single households, we get that:

$$\begin{aligned} \frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{1f}} &= \sum_i \frac{\pi^f_{1|i} N_{R^f_i}}{\sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}} \left(1 - \pi^f_{1|i}\right) \theta^f > 0 \\ \frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{jf}} &= -\sum_i \frac{\pi^f_{1|i} N_{R^f_i}}{\sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}} \pi^f_{j|i} \theta^f < 0 \\ \frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{1m}} &= \frac{\partial \log \sum_i \left\{\pi^f_{1|i} N_{R^f_i}\right\}}{\partial \log w_{jm}} = 0 \end{aligned}$$

For males in dual-earner households we get a similar expression. For women in dual-earner households we get that:

$$\frac{\partial \log \sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}}{\partial \log w_{1f}} = \sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}\right\}} \left(1 - \pi^w_{1|i,\ell=1}\right) \sum_l \pi^h_{l|i,\ell=1,1} \frac{w_{1f}}{(w_{lm} + w_{1f})} \theta^w \geq 0$$

$$\frac{\partial \log \sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}}{\partial \log w_{1m}} = \sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \left\{ \pi^h_{1|i,\ell=1,1} \frac{w_{1m}}{(w_{1m}+w_{1f})} - \pi^h_{1|i,\ell=1} \tilde{\omega}_{1m} \right\} \theta^w$$

with

$$\tilde{\omega}_{1m} = \sum_l \frac{\pi^w_{l|i,\ell=1} \pi^h_{1|i,\ell=1,l}}{\pi^h_{1|i,\ell=1}} \frac{w_{1m}}{(w_{1m}+w_{lf})}$$

$$\frac{\partial \log \sum_i \left\{ \pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}}{\partial \log w_{jf}} = -\sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \pi^w_{j|i,\ell=1} \sum_l \pi^h_{l|i,\ell=1,j} \frac{w_{jf}}{(w_{lm}+w_{jf})} \theta^w < 0$$

$$\frac{\partial \log \sum_i \left\{ \pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}}{\partial \log w_{jm}} = \sum_i \frac{\pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^{mf,f}_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \left\{ \pi^{mf,m}_{j|i,\ell=1,1} \frac{w_{jm}}{(w_{jm}+w_{1f})} - \pi^{mf,m}_{j|i,\ell=1} \tilde{\omega}_{jm} \right\} \theta^{mf,f}$$

with

$$\tilde{\omega}_{jm} = \sum_l \frac{\pi^{mf,f}_{l|i,\ell=1} \cdot \pi^{mf,m}_{j|i,\ell=1,l}}{\pi^{mf,m}_{j|i,\ell=1}} \frac{w_{jm}}{(w_{jm}+w_{lf})}$$

I stop here because it should be already clear where the gross-substitution property fails to hold. In particular, it can be shown that $\frac{\partial \log \mathcal{Z}_z}{\partial \log w_{-z}} > 0$ may not always hold. For instance, taking the derivative of excess demand for women in location 1 with respect to wages of men in location 1, that is $\frac{\partial \log \mathcal{Z}_z}{\partial \log w_{-z}}$, with $z = J+1, -z = 1$ we get:

$$\begin{aligned}
\frac{\partial \log \mathcal{Z}_z}{\partial \log w_{-z}} &= \underbrace{\sigma \sum_s \frac{(w_{1m}/\alpha_{sm})^{-\sigma}}{\sum_{g \in m,f} (w_{1g}/\alpha_{sg})^{-\sigma}} L_{F_{1sm}}}_{>0} - \underbrace{\frac{\text{single women}_j}{\text{all women}_j} \left(\frac{\partial \log \sum_i \left\{ \pi^f_{1|i} N_{R^f_i} \right\}}{\partial \log w_{1m}} \right)}_{=0} \\
&- \underbrace{\frac{\text{married women}_j}{\text{all women}_j} \left(\sum_i \frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \left\{ \pi^h_{1|i,\ell=1,1} \frac{w_{1m}}{(w_{1m}+w_{1f})} - \pi^h_{1|i,\ell=1} \tilde{\omega}_{1m} \right\} \theta^w \right)}_{?}
\end{aligned}$$

The supply of women to location 1 can increase even if male wages increase when

$$\frac{\pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}}}{\sum_i \left\{ \pi^w_{1|i,\ell=1} N_{R^{mf}_{i,1}} \right\}} \left(\pi^h_{1|i,\ell=1,1} \frac{w_{1m}}{(w_{1m}+w_{1f})} - \pi^h_{1|i,\ell=1} \tilde{\omega}_{1m} \right)$$

is negative enough such that the positive effect coming from other origins and from the downward-sloping demand get both offset.

B.9.2 Proposition 2

PROPOSITION: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$, rents r_{F_j},and commute costs d^k_{ij}, and parameters θ^k, α_s, σ_D there exists a vector of sales X_{js} and productivities A_{js} that rationalizes the observed data as an equilibrium of the model. Additionally, we get the same result if instead of employment by gender we observe employment by industry $L_{F_{js}}$ and parameters σ, α_{sg}.*

Proof. From proposition 1, we get wages. Afterwards, we obtain the total wage bill:

$$W_{js}N_{js} = w_{jm}\left[L_{F_{jm}}\right] + w_{jf}\left[L_{F_{jf}}\right]$$

Or if one observes employment by industry and parameters α_{sg} and σ:

$$W_{js}N_{js} = w_{jm}\left[\frac{(w_{jm}/\alpha_{sm})^{-\sigma}}{\sum_h (w_{jh}/\alpha_{sh})^{-\sigma}} L_{F_{js}}\right] + w_{jf}\left[\frac{(w_{jf}/\alpha_{sf})^{-\sigma}}{\sum_h (w_{jh}/\alpha_{sh})^{-\sigma}} L_{F_{js}}\right]$$

where first to second line comes from FOC: $\frac{\tilde{L}_{F_{jhs}}}{\tilde{L}_{F_{jgs}}} = \left(\frac{\alpha_{sh}/w_{jh}}{\alpha_{sg}/w_{jg}}\right)^{\sigma}$ and by summing over h and rearranging, I get $\frac{\tilde{L}_{F_{jgs}}}{\tilde{L}_{F_{js}}} = \frac{(w_{jg}/\alpha_{sg})^{-\sigma}}{\sum_h \left(w_{jh}/\alpha_{sh}\right)^{-\sigma}}$.

Then, we can gets sales X_{js} from the firm's FOCs because: $\alpha_s X_{js} = W_{js}N_{js}$.

Once we obtain sales, we can recover productivity:

$$X_{js} \propto \left(\frac{W_{js}^{\alpha_s} r_{F_j}^{1-\alpha_s}}{A_{js}}\right)^{1-\sigma_D} \rightarrow A_{js} \propto \left(\frac{W_{js}^{\alpha_s} r_{F_j}^{1-\alpha_s}}{X_{js}^{\frac{1}{1-\sigma_D}}}\right)$$

□

B.9.3 Proposition 3

PROPOSITION: *Given data on residence by gender and household type, $N_{R^k_{i,\ell}}$, employment by gender $L_{F_{jg}}$, available floorspace H_i, floorspace allocation ϑ_i, rents r_{R_i} and r_{F_j} and commute costs d^k_{ij}, and parameters θ^k, η^k, ν, σ^D, α_s, β^k, there exists a vector of amenities u^k_i, productivities A_{js}, household productivity $\xi_{i\ell=0}$, sales X_{js}, floorspace wedge τ_i, and total rents E that rationalizes the observed data as an equilibrium of the model. Additionally, we get the same result if instead of*

employment by gender we observe employment by industry $L_{F_{js}}$ and parameters σ, α_{sg}.

Proof. From proposition 1, we get wages. Then, we can compute residential market access:

- For single households: $\Phi_{R_i^g} = \sum_{l'} \left(\frac{d_{il'}^g}{w_{l'g}} \right)^{-\theta^g}, \forall g \in \{m, f\}$
- For couples in male breadwinner households: $\Phi^h_{R_{i,\ell=0}} = \sum_{l'} \left(\frac{d_{il}^h}{w_{lm}} \right)^{-\theta^h}$
- In dual-earner households: $\Phi^h_{R_{i,\ell=1,j'}} = \sum_l \left(\frac{d_{il}^h}{w_{lm}+w_{j'f}} \right)^{-\theta^h} \Phi^w_{R_{i,\ell=1}} = \sum_l \left(\frac{d_{il}^w}{\left(\Phi^h_{R_{i,\ell=1,l}} \right)^{1/\theta^h}} \right)^{-\theta^w}$
- On average:

$$\Phi_{R_i^{mf}} = \left\{ \left[\frac{1}{\gamma^w} \left(\Phi_{R^w_{i,\ell=1}} \right)^{-1/\theta^w} \right]^{-\nu} + \left[(\xi_{i\ell=0})^{-1} \left(\Phi_{R^h_{i,\ell=0}} \right)^{-1/\theta^h} \right]^{-\nu} \right\}$$

where $\gamma^w = \Gamma\left(\frac{\theta^w - 1}{\theta^w}\right) (T^w)^{1/\theta^w} = 1$

Using data on labor force participation choices, we can use the following equation to recover household productivities:

$$\frac{\mu_i^{mf}}{1-\mu_i^{mf}} = \left(\frac{\left(\Phi^w_{R_{i,\ell=1}} \right)^{-1/\theta^w}}{(\xi_{i\ell=0})^{-1} \left(\Phi_{R^h_{i,\ell=0}} \right)^{-1/\theta^h}} \right)^{-\nu}$$

i,e,

$$\xi_{i\ell=0} = \frac{\left(\Phi^w_{R_{i,\ell=1}} \right)^{1/\theta^w}}{\left(\Phi_{R^h_{i,\ell=0}} \right)^{1/\theta^h}} \left(\frac{\mu_i^{mf}}{1-\mu_i^{mf}} \right)^{\frac{1}{-\nu}}$$

Then using the residential supply conditions, from which we can recover implied amenities

once that we know wages and rents:

$$N_i^g \equiv \pi_i^g N^g = \sum_l \left(\frac{r_{R_l}^{1-\beta^g}}{u_l^g \Phi_{R_l^g}^{1/\theta^g}} \right)^{-\eta^g} N^g \left(u_i^g \Phi_{R_i^g}^{1/\theta^g} r_{R_i}^{\beta-1} \right)^{\eta^g}$$

$$N_i^{mf} \equiv \pi_i^{mf} N^{mf} = \sum_l \left(\frac{r_{R_i}^{1-\beta^{mf}}}{u_l^{mf} \Phi_{R_l^{mf}}^{1/\nu}} \right)^{-\eta^{mf}} N^{mf} \left(u_i^{mf} \Phi_{R_i^{mf}}^{1/\nu} r_{R_i}^{\beta-1} \right)^{\eta^{mf}}$$

Then we need to solve for unobservables on the housing side. First, we need to get back sales using proposition 2. Then, it is needed to introduce a new pair of location characteristics since the floorspace market clearing condition $E_{Ri} = r_{R_i} H_{R_i}$ will not necessarily hold at the values for data and wages. Therefore, we need to introduce an additional unobservable that can be interpreted as quality of housing $\tilde{H}_{R_i} = H_{R_i} q_{R_i}$ where H_{R_i} are the physical units of floorspace. From the housing market clearing condition we get that $E_{Ri} = r_{R_i} \tilde{H}_{R_i} = r_{R_i} H_{R_i} q_{R_i} \rightarrow q_{R_i} = \frac{E_{Ri}}{r_{R_i} H_{R_i}}$. Similar residuals can be defined for commercial floorspace: $q_{F_i} = \frac{X_i}{r_{F_i} H_{Fi}}$.

Finally, we just need to solve for land use wedge, which can be identified from:

$$(1 - \tau_i) = \frac{r_{R_i}}{r_{F_i}}$$

for locations with mixed land use. For locations with single land use, these wedges cannot be identified but are rationalized by zero productivities for all sectors or zero amenities for all worker groups. □

B.10 Online Supplement: Data Appendix

In this section, I explain the choices I made when processing the datasets available in this setting.

B.10.1 Population and Household Census, 1993, 2007 and 2017

Household Classification. The primary source of population data is the Population and Household Census of 1993, 2007 and 2017. These were conducted by the National Institute of Statistics and Informatics (INEI). Crucially, these data sets contain the population in each block by gender and civil-status. Moreover, the household's roster is observed. Hence, I can observe who lives with whom.

I define a couple as a male and a female adult living in the same two-members-household. In 2017 around 77% of households had two or fewer adult members. In 2007, 75% of households had two or fewer adult members. So, these account for the majority of households. I also include the restriction that at least one of the two should work. For singles, I consider households with one adult member. This lone member should work.

For simplicity, I treat individuals living with two or more additional adults as singles (i.e. as if they were roommates and therefore not sharing their income). I make this simplification because if more than two adults live together, then it is hard to infer who is the couple of whom, and I require this knowledge to compute conditional commute probabilities. However, in some cases, it could potentially be inferred. For example, when the household head lives with his or her partner. Or if only two of the $N > 2$ members report to be married. One can also make some assumptions about the average age gap between the spouses. To avoid further complications, I make the choice of simply treating these individuals as singles. If anything, this choice makes singles to look a bit more alike to couples since I am potentially misidentifying some couples in multi-members households as singles.[15] By considering singles and couples in this simple way I am already improving on what the literature has done. In any case, results are similar if I drop those that report to be married among multi-members households. Finally, my main results consider as adults those between 25 and 65 years old, but results are robust to alternative definitions.

This classification leads to the following distribution of households: (i) single females (31.9% of households in 2017, 29% of households in 2007), (ii) single males (35.9% and 36.1%), (iii) male breadwinner households, that is, households where the wife stays at home while the husband goes to work (13.3% and 17.3%), (iv) female breadwinner households (1.3% and 1.5%), and (v) dual-earner households (17.6% and 16.2%). I calculate the number of households for each of these five groups for each block.

For consistency, I use the same classification across databases.

[15] Less than 50% of individuals in these households report to be married.

Commuting Flows. In the 2017 Census, employed individuals were asked to provide the municipality in which they work. Using this information, along with data on the locations where these individuals reside, I calculate commute probabilities as follows:

$$\pi^g_{j|i} = \frac{N^g_{j|i}}{\sum_k N^g_{k|i}}$$

where $N^g_{j|i}$ is the number of workers of group g that live in municipality i and work in municipality j. For conditional commute probabilities in dual-earner households,

$$\pi^g_{j|il} = \frac{N^g_{j|il}}{\sum_k N^g_{k|il}}$$

where $N^g_{j|il}$ is the number of dual-earner workers of gender g working in j, living i, and whose partners work in l.

B.10.2 Economic Census, 2008

The 2008 Economic Census of Peru was conducted by the National Institute of Statistics and Informatics (INEI). This database covers firms of all sizes and all industries, except agriculture and financial services, even those that had not filed their taxes. It focuses exclusively on the firms' operations in 2007. Information and consistency checks were collected by census officers visiting business establishments. To my convenience, the location of the businesses were recorded at a very detailed level, including the specific census block they were located in. I aggregate sectors into seven sectors so that it is easier to match to the Firms' Administrative Data of 2015. In particular, I consider textiles, other manufacture, services, business services, wholesale, retail trade, and transportation services. I aggregate employment counts at the block level for each of the seven industries described above. This way, I get the distribution of economic activity per industry across locations for the first period of the analysis. I scale-up employment by the ratio of employment in the Economic Census to total employment in Lima according to the Population Census in 2007.

B.10.3 Firms' Administrative Data, 2015

The Ministry of Production used their raw data and classified more than 1 million firms by industry, geographic location, sales, and number of workers. The database is available at this link. The information relevant to my analysis is the one concerning the industrial classification, the geographic location and the number of workers. Firms were classified into seven categories: 0-5 workers, 6-

10 workers, 11-20 workers, 21-50 workers, 51-100 workers, 101-200 workers, and more than 200 workers. To compute the number of workers per firm I simply impute the lower bound of each interval.[16] Then, I aggregate these counts at the block level for each of the following broad industries: textiles, other manufacture, services, business services, wholesale, retail trade, and transportation services. This way, I get the distribution of economic activity per industry across locations for the second period of the analysis. I scale-up employment by the ratio of employment in the Economic Census to total employment in Lima according to the Population Census in 2017.

B.10.4 Road Network Data and Commute Times

Road network data. I use road network data from Open Street Map. Open Street Map (OSM) is a free, editable map of the world that is created and maintained by a community of volunteer contributors. The project was started in 2004 with the goal of creating a free and open alternative to proprietary mapping services, such as Google Maps. The data in OSM is collected from a variety of sources, including GPS tracks, aerial imagery, and manual surveying. The data is then added to the OSM database and is available for anyone to use. The quality of OpenStreetMap data can vary depending on the region and the level of community engagement. In general, OSM data tends to be more accurate and up-to-date in areas with a high level of community engagement, such as densely populated urban areas and regions with a strong OSM community. In these areas, the map is regularly updated with new data and errors are quickly corrected. Importantly, the OSM data contains information on road type or classification, which I use to impute road speeds.

Commute times. I use the data on road type to input speeds during rush hour according to the following classification: motorway (50 km/h), primary (30 km/h), residential (15 km/h), secondary (20 km/h), and tertiary (10 km/h). This classification is based on a report for the Ministry of Transport and Communications done by the Japan International Cooperation Agency in 2013 (JICA, 2013). Additionally, I manually input speeds in certain roads, the BRT, and the Metro following the same report (see Table B.10). Then, I run ArcGis's Network Analyst tool to compute the optimal route from origins to destination to generate an origin-destination matrix at the block level. I utilize this tool to compute commute times in absence of the Metro and the BRT and when the two are functional. Sometimes, I aggregate these times to the municipality level by taking the median value across zones within municipalities. I show that these times are well correlated with times reported in commuting surveys (see Figure B.20).

The matrix of commute times used in the analysis is an approximation of actual commute times. Ideally, the routes of public buses and road speeds would be used to calculate how long it takes

[16] Results are similar if I instead use the midpoint between the two bounds.

to travel between locations using public transportation. However, data on bus routes is not readily available in this context. Despite this limitation, the commute elasticities, which link commute times to costs, are still estimated using the approximation of commute times and probabilities. These elasticities already take into account any potential differences between the estimated and actual commute times. For example, if actual commute times are twice as long as the estimated times, the estimated elasticities would be half as large. The only concern would be if the error in the approximation of commute times is correlated with a blocks' location or some characteristic, which could lead to problems.

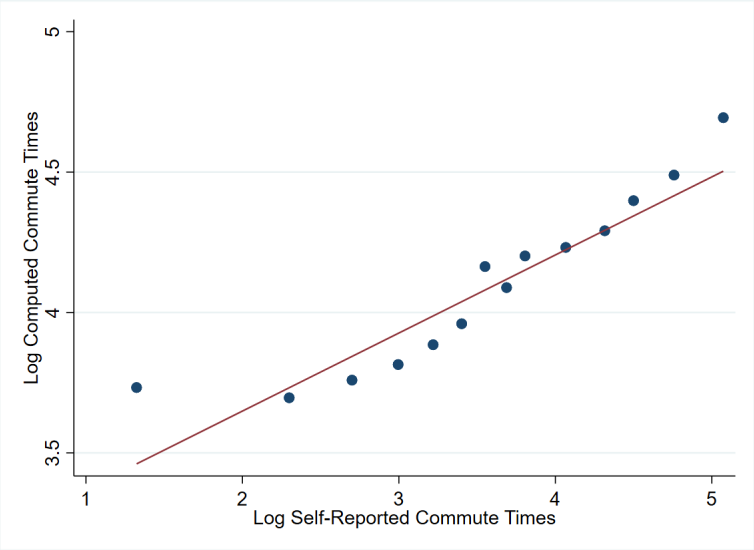
Table B.10: Manually imputed speeds

Road	Speed
Avenida Brasil	15 km/h
Avenida Javier Prado Oeste	15 km/h
Avenida La Marina	15 km/h
Avenida Faustino Sanchez Carrión	15 km/h
Avenida Almirante Miguel Grau	15 km/h
Avenida Abancay	15 km/h
Avenida Arequipa	15 km/h
Avenida de Tomás Marsano	15 km/h
Avenida República de Panamá	15 km/h
Avenida Arica	15 km/h
Avenida Aviación	15 km/h
Avenida República de Venezuela	15 km/h
Avenida Bolognesi	15 km/h
Avenida Prolongación Paseo de la República	15 km/h
Avenida Defensores del Morro	15 km/h
Avenida Universitaria	15 km/h
Metro Lines	45 km/h
Bus Rapid Transit System	35 km/h

B.10.5 National Household Survey, 2007-2017

The National Institute of Statistics and Informatics (INEI) in Peru conducts the Peruvian National Household Survey (ENAHO) annually. The sample size of the survey varies each year, with a range of 26,000 to 40,000 households. The survey provides a wealth of information on individual income and demographics, including details on primary and secondary activities such as wages and hours worked. One of the survey's key benefits is that households are geocoded at the cluster level, with each cluster containing about 140 households. This level of granularity allows for accurate assignment of households to census blocks. Additionally, the survey collects valuable information

Figure B.20: Imputed versus reported commuting times

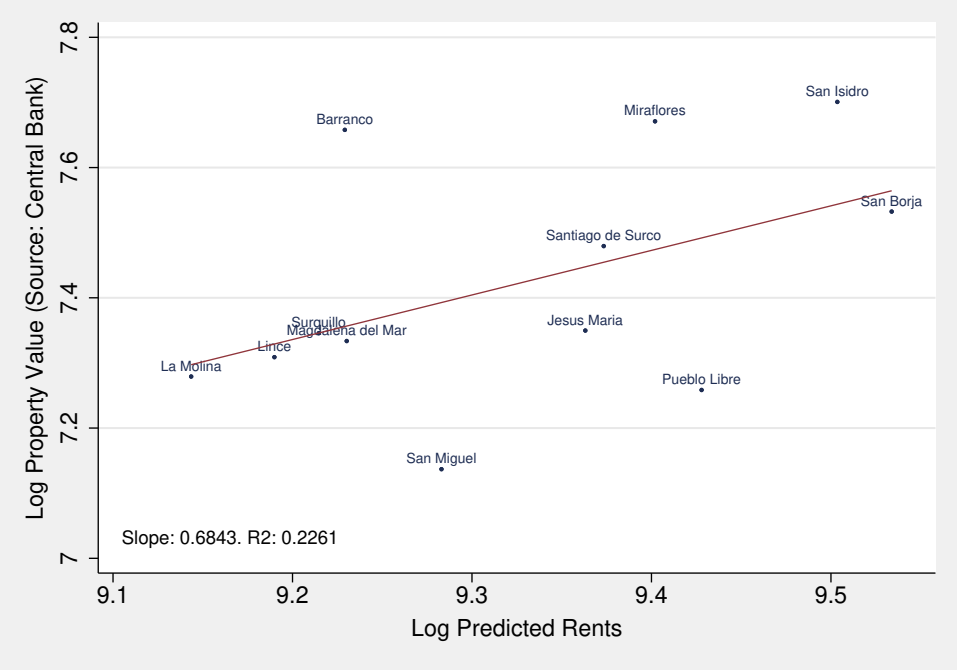


Self-reported commute times, at the municipality level, come from the 2010-2018 waves of the Commuting Surveys. Imputed commute times come from the Open Street Road Network data in par with Imputed speeds. Imputed commute times were computed at the zone level, and then aggregated to the municipality level using median values. Scatter was generated using the *binscatter* command in Stata. The correlation between the two measures of commute times is 53%.

from household heads, such as the estimated rental value of their home, total household income, and expenditures.

I use this data to predict how property values vary with centrality, and other block and dwelling characteristics.[17] Then, I use the estimated coefficients to project property values into census' blocks. I assign data on years 2007-2010 to the first period considered in the analysis, and data on years 2014-2017 to the second period. In Figure B.21, I compare the log of the average value of predicted rents by municipality to data on the log of property values compiled by the Central Bank of Peru for a subset of municipalities. These averages are computed over years 2014-2017. Figure B.21 shows that the two measures are well correlated. Finally, since geocoded data on rents by firms is not available in this setting, I assume that there are no differential rents due to land regulations, and so, the returns to floorspace use in residential and commercial activities are the same, conditional on a block's location.

Figure B.21: Property Values versus Predicted Rents by Municipality



B.10.6 Land Use Data, 2020

The Metropolitan Planning Institute collects actual land use data, but it is only available for the year 2020 and not for previous years. As a result, when required, I am assuming that land use in 2007 is

[17] I achieve and R^2 of around 60% across the two periods considered in the analysis.

the same as it was in 2017.[18] Despite this assumption, it does not have any impact on the estimation of any parameters in the model. This is because this data is only used to recover unobservables on the housing side, which is modeled in a straightforward manner following Tsivanidis (2021). These unobsevables are not an input in the estimation of any of the parameters. Additionally, the counterfactual analysis is not affected as I am using 2017 data as a baseline when removing the Metro and BRT from Lima. Therefore, the lack of land use data for 2007 does not affect this exercise. The lack of this data would be of significance if I were to perform a counterfactual analysis using the year 2007 as baseline.

[18] An alternative option would be to use data on floor space zoning, which is available for earlier years.

APPENDIX C

Appendix to Chapter 3

C.1 Data Used in Estimation and Simulation Samples

The data for WWII supply contracts come from volumes originally issued by the Civilian Production Administration (formerly the War Production Board), which record all defense contracts greater than $50,000 issued between June 1940 and September 1945. Our measure of government defense spending is the aggregate monetary value of war supply contracts in each county during that period.

The empirical analysis is based on county-level data for the years 1900 to 2000. Data on county characteristics are the same as those in Kline and Moretti [2014] and come from a variety of sources including the Census of Population and Housing and Agricultural Census. In particular, the data for the years 1900 to 1930 comes from the Historical, Demographic, Economic, and Social Data: The United States, 1790-2000, ICPSR 2896 (specifically Parts 20, 22, 24, 26, 29, and 85, corresponding to the 1900, 1910, 1920, and 1930 Censuses, as well as the 1930 Census Parts I and IV Families, and the 1910 Census of Agriculture). However, variables for manufacturing and agricultural employment were constructed from individual-level data sourced from a 1% sample extract of the 1900, 1910, 1920, and 1930 Censuses, available through IPUMS.

Variables for the period from 1940 to 1970 were developed using two primary data sources: the County and City Data Book [United States] Consolidated File: County Data, 1947-1977 (ICPSR 7736) and the Historical, Demographic, Economic, and Social Data: The United States, 1790-2000 (ICPSR 2896, specifically Parts 70, 72, 74, and 76, corresponding to the County Data Books for 1947, 1952, 1962, and 1972). The definitions of the variables were mostly consistent across these two data sources. In case of discrepancies, we use the definitions from ICPSR 2896. The data from 1980 to 2000 was derived from census extracts provided by the National Historical Geographic Information System.

The quality of some key variables is imperfect, particularly at the beginning of our sample period, where significant measurement errors are likely. In the early years, there is no direct data

on workers' wages at the county level. As in Kline and Moretti [2014], we proxy the average wage in manufacturing by dividing the total annual wage bill in manufacturing by the estimated number of workers in the industry. Additionally, in some counties, the wage bill data is incomplete or missing. For agriculture, there is no available county-level wage bill data, making it difficult to compute an average agricultural wage.

Since we need wage information to invert the structural model, we impute missing wages for the simulation sample. Specifically, we impute missing wage values for county i in decade t using a dynamic panel regression specification as follows:

$$w_{it} = \alpha w_{it-1} + \beta X_{it} + \gamma t + \theta_i,$$

where we control for lagged wage w_{it-1}, county characteristics X_{it} (which include population, median house value, median rent, ratio of white population, agricultural employment share, urban population share, when possible), county fixed effects θ_i, and a linear time trend.[1]

To avoid issues with changes in county definitions over time due to splits and merges, we use the crosswalk provided by Eckert et al. [2020a] to obtain a sample of harmonized counties between 1900 and 2000. We also exclude counties in Alaska or Hawaii as well as underpopulated counties with population fewer than 1,000 in any decade in the sample period or manufacturing employment fewer than 10 in 1930 from the empirical analysis.

[1] If lagged wage is missing, we drop the first term from the regression equation.

C.2 Model Inversion

We infer TFP in manufacturing, agriculture, and services ($T_{M,i,t}, T_{A,i,t}, T_{S,i,t}$), and amenities ($H_{i,t}$) using observed data on population, sectoral employment, and wages. Given the data $\{w_{i,t}, l_{M,i,t}, l_{A,i,t}, l_{S,i,t}, L_{i,t}, G_t, \kappa_{i,t}\}_{t\in\mathcal{T}}$, where $\mathcal{T} = \{0, 1, ..., T\}$, and parameters $\{\sigma, \eta, \epsilon_s, \theta, \alpha, \tau_{i,n,t}, \xi_{i,n,t}, \zeta_t\}_{t\in\mathcal{T}}$, we invert the model and solve for $\{T_{M,i,t}, T_{A,i,t}, T_{S,i,t}, H_{i,t}\}_{t\in\mathcal{T}}$ for each location, up to a normalization for each variable.

1. The first step is to solve for $c_{M,i,t}, c_{A,i,t}$, and $c_{S,it}$ up to a normalization. Given the data, $c_{M,i,t}, c_{A,i,t}, c_{S,it}$ solves the implicit equations:

$$\begin{aligned} w_{i,t}l_{M,i,t} &= \alpha \cdot \left[c_{M,i,t}^{1-\sigma} \cdot \left(\sum_{n\in N} \tau_{i,n,t}^{1-\sigma} \cdot \frac{E_{M,n,t} + \frac{1-\alpha}{\alpha} w_{n,t} l_{M,n,t}}{\sum_j c_{M,j,t}^{1-\sigma} \tau_{j,n,t}^{1-\sigma}} \right) + \kappa_{i,t} G_t \right], \forall i, t \\ w_{i,t}l_{A,i,t} &= (c_{A,i,t})^{1-\sigma} \cdot P_{A,t}^{\sigma-1} \sum_{n\in N} E_{A,n,t}, \forall i, t \\ w_{i,t}l_{S,i,t} &= E_{S,i,t}, \forall i, t \end{aligned}$$

   where

$$\begin{aligned} E_{s,i,t} &= \omega_{s,i,t} \left(w_{i,t} L_{i,t} - \tau_t L_{i,t} \right) \\ \omega_{s,i,t} &= \rho_s \left(\frac{P_{s,i,t}}{P_{i,t}} \right)^{(1-\eta)}, \forall s \in \{A, S, M\} \\ \omega_{A,i,t} &+ \omega_{S,i,t} + \omega_{M,i,t} = 1 \end{aligned}$$

   and ρ_s are chosen such that market clearing is assured in all sectors for any unit cost vector $\{c_{M,i,t}, c_{A,i,t}, c_{S,it}\}$. Note that this equation is homogeneous of degree zero in $\{c_{M,i,t}, c_{A,i,t}, c_{S,it}\}$, and that all other variables are known from the data. There is a unique solution (up to scale) of this equation, $\frac{c_{M,i,t}}{c_{M,1,t}}, \frac{c_{S,i,t}}{c_{S,1,t}}, \frac{c_{A,i,t}}{c_{A,1,t}}$.[2]

[2]Our main data source are the Decennial Censuses. Since most of the contract expenditures happened between 1940 and 1950, and we do not observe any data point in those years, we assume $G_t = 0$ when backing out sectoral TFP.

2. The second step is to recover the price index $P_{M,n,t}^{1-\sigma} = \sum_j \frac{c_{M,j,t}^{1-\sigma}}{c_{M,1,t}^{1-\sigma}} \tau_{j,n,t}^{1-\sigma}$ and compute:

$$\begin{aligned}
\frac{T_{M,i,t} \cdot f(l_{M,i,t-1})}{T_{M,1,t} \cdot f(l_{M,1,t-1})} &= \frac{c_{M,1,t}}{c_{M,i,t}} \cdot \frac{w_{i,t}^{\alpha} \left(\sum_j \frac{c_{M,j,t}^{1-\sigma}}{c_{M,1,t}^{1-\sigma}} \tau_{j,i,t}^{1-\sigma} \right)^{\frac{1-\alpha}{1-\sigma}}}{w_{1,t}^{\alpha} \left(\sum_j \frac{c_{M,j,t}^{1-\sigma}}{c_{M,1,t}^{1-\sigma}} \tau_{j,1,t}^{1-\sigma} \right)^{\frac{1-\alpha}{1-\sigma}}}, \forall i, t \\
\frac{T_{S,i,t}}{T_{S,1,t}} &= \frac{w_{i,t}/c_{S,i,t}}{w_{1,t}/c_{S,1,t}}, \forall i, t \\
\frac{T_{A,i,t}}{T_{A,1,t}} &= \frac{w_{i,t}/c_{A,i,t}}{w_{1,t}/c_{A,1,t}}, \forall i, t
\end{aligned}$$

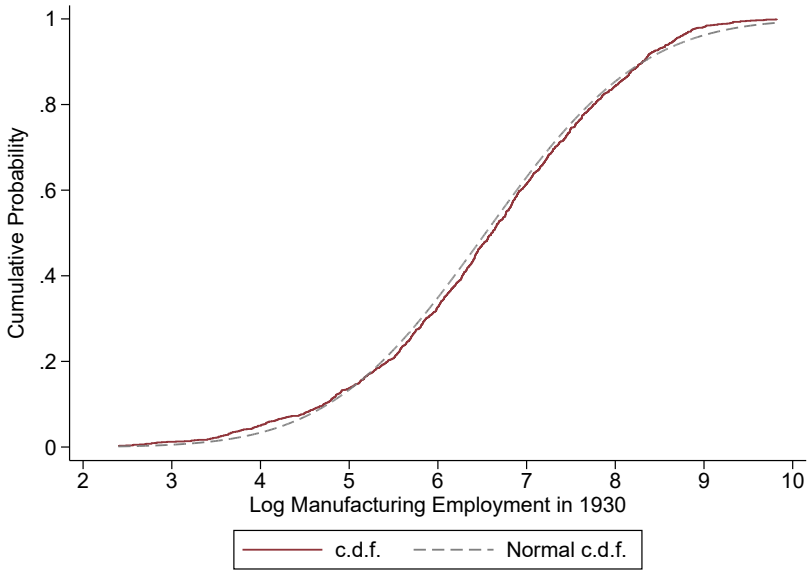
3. The final step is to recover exogenous amenities. Armed with previous solutions, we compute price index $P_{i,t}^{(1-\eta)} = \rho_s \left(P_{s,i,t} \right)^{(1-\eta)}$ and use the migration equation:

$$\begin{aligned}
L_{n,t} &= \zeta_t \sum_i \frac{U_{i,n,t}^{\theta}}{\sum_{j \in N} U_{i,j,t}^{\theta}} L_{i,t-1}, \forall n, t \\
U_{i,n,t} &= H_n \cdot \xi_{i,n,t} \cdot \frac{w_{n,t} - \tau_t}{P_{i,t}}
\end{aligned}$$

which can be solved uniquely (up to scale) for $\frac{H_{n,t}}{H_{1,t}}$.

C.3 Additional Tables and Figures

Figure C.1: Cumulative Probability Distribution of Initial Manufacturing Employment



Notes: This figure shows the cumulative probability distribution of manufacturing employment in 1930 in the sample.

Table C.1: Summary Statistics of Pre-WWII Regional Characteristics in 1930

	Treated Counties		Control Counties	
	Mean	Std. Dev.	Mean	Std. Dev.
Log population	10.428	0.596	10.031	0.520
Log manufacturing employment	7.105	1.286	6.138	1.316
Log employment	9.445	0.622	9.015	0.539
Agricultural employment share	0.359	0.180	0.486	0.192
Manufacturing employment share	0.138	0.106	0.094	0.095
Log average manufacturing wage	2.738	0.296	2.631	0.362
% Urban	0.359	0.203	0.216	0.195
Log average farm value	6.815	0.652	6.602	0.619
Log median housing values	10.683	0.382	10.462	0.391
Log median contract rents	10.045	0.414	9.834	0.407
% White	0.880	0.166	0.852	0.203
% Illiterate	0.047	0.048	0.057	0.059
% households owning radio	0.310	0.166	0.247	0.173
% Urbanization Rate	0.033	0.018	0.025	0.018
% Whites foreign born	0.051	0.056	0.043	0.053
Max elevation (meters)	1,675.159	2,084.700	1,761.862	2,292.476
Elevation range (max-min)	1,177.056	1,964.030	1,186.685	2,006.357
Area (squared miles)	780.856	846.258	769.884	691.974

Notes: Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms. To better balance treated and control samples, we drop counties whose pre-WWII characteristics yield extreme propensity scores of treatment higher than 99% or lower than 1%. All monetary values are in constant 2000 USD.

Table C.2: Contract Allocation and Initial Manufacturing Conditions

	(1) Treat	(2) ContractValue	(3) Treat	(4) ContractValue
Auto and Air Share	0.009	0.080		
	(0.004)	(0.024)		
Auto and Air Share Big			0.127	0.896
			(0.023)	(0.185)
Log Manuf. Emp. 1930	0.015	-0.017	0.025	0.074
	(0.058)	(0.565)	(0.060)	(0.596)
Controls	Yes	Yes	Yes	Yes
Obs.	1,777	1,777	1,780	1,780
Adj. R^2	0.377	0.530	0.382	0.532

Notes: This table shows the relationship between wartime contract allocation and initial manufacturing conditions in 1930 across counties. The dependent variables are treatment indicators and log contract value per capita. Key independent variables are log manufacturing employment levels, predicted log contract value, and employment shares in sectors highly related to the war effort including automobile factories, ship and boat building, wagon and carriage factories, and air transportation. All regressions adjust for the control variables described in the text. Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms.

Table C.3: Robustness Check Controlling for Industry Composition

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat*LowM	0.026	0.036	0.025	0.083	0.027	−0.004
	(0.008)	(0.010)	(0.008)	(0.020)	(0.009)	(0.010)
Treat*HighM	0.009	0.012	0.019	0.060	0.004	−0.022
	(0.008)	(0.009)	(0.009)	(0.015)	(0.008)	(0.009)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows robustness checks controlling for employment shares in sectors highly related to the war effort including automobile factories, ship and boat building, wagon and carriage factories, and air transportation. All regressions adjust for the control variables described in the text. Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms.

Table C.4: Reduced-Form Estimation Based on Employment Shares in War-Related Sectors

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
AA Share Big*LowM	0.059	0.067	0.068	0.104	0.052	0.025
	(0.015)	(0.015)	(0.017)	(0.034)	(0.015)	(0.017)
AA Share Big*HighM	0.003	0.006	0.023	0.032	−0.009	−0.011
	(0.006)	(0.008)	(0.009)	(0.014)	(0.006)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table shows reduced-form effects based on employment shares in sectors highly related to the war effort including automobile factories, ship and boat building, wagon and carriage factories, and air transportation. *AA Share Big* is an indicator variable for employment shares in war-related sector above the median level if counties with positive employment shares in those sectors. Treated counties correspond to those that received WWII war supply contract value per capita exceeds the 50th percentile in logarithmic terms. Standard errors are clustered at the state level.

Table C.5: Long-Term Impact on Growth Rate of Outcomes by Contract Types

	(1) Pop 1930-2000	(2) ManufEmp 1930-2000	(3) Pop 1930-1960	(4) ManufEmp 1930-1960	(5) Pop 1960-2000	(6) ManufEmp 1960-2000
Treat*LowM	0.020	0.052	0.041	0.037	0.075	0.058
	(0.006)	(0.019)	(0.013)	(0.008)	(0.019)	(0.013)
Treat*HighM	0.006	−0.000	0.001	0.012	0.002	0.005
	(0.005)	(0.007)	(0.006)	(0.006)	(0.008)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Treated counties correspond to those that received WWII war supply contract value per capita exceeds the median in logarithmic terms. Treatment status is classified into three categories of contract spending as follows: 1) "footwear, textile, chemicals, miscellaneous" (columns 1–2); 2) "metal" (columns 3–4); and 3) "machinery, electronics, transportation" (columns 5–6). Robust standard errors are reported in parenthesis.

Table C.6: Relationship Between Post-WWII Military Spending and War Supply Contracts

	(1) MilitaryExp 1966	(2) MilitaryExp 2000	(3) MilitaryExp 1966	(4) MilitaryExp 2000	(5) MilitaryExp 1966	(6) MilitaryExp 2000
Contract value	0.002	0.027				
	(0.023)	(0.021)				
Treat			0.096	0.438		
			(0.413)	(0.380)		
Treat (median)					-0.093	0.271
					(0.404)	(0.393)
Obs.	50	50	50	50	50	50

Notes: This table shows coefficients of regressing post-WWII military spending in 1966 and 2000 at the state level on war time contract spending or treatment status. Treatment status is defined based on whether the log contract value is above 8 (columns 3 and 4) or the median value (columns 5 and 6). Robust standard errors are reported in parenthesis.

BIBLIOGRAPHY

D. Acemoglu, D. H. Autor, and D. Lyle. Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. Journal of political Economy, 112(3):497–551, 2004.

G. M. Ahlfeldt and D. P. McMillen. New estimates of the elasticity of substitution between land and capital. 2014.

G. M. Ahlfeldt and D. P. McMillen. Tall buildings and land values: Height and construction cost elasticities in chicago, 1870–2010. Review of Economics and Statistics, 100(5):861–875, 2018.

G. M. Ahlfeldt, S. J. Redding, D. M. Sturm, and N. Wolf. The economics of density: Evidence from the berlin wall. Econometrica, 83(6):2127–2189, 2015.

T. Allen and C. Arkolakis. Trade and the topography of the spatial economy. The Quarterly Journal of Economics, 129(3):1085–1140, 2014.

T. Allen and C. Arkolakis. Economic activity across space: A supply and demand approach. Journal of Economic Perspectives, 37(2), 2023.

T. Allen and D. Donaldson. Persistence and path dependence in the spatial economy. Technical report, National Bureau of Economic Research, 2023.

T. Allen, C. Arkolakis, and X. Li. Optimal city structure. Yale University, mimeograph, 2015.

M. Almagro and T. Domínguez-Iino. Location sorting and endogenous amenities: Evidence from amsterdam. Technical report, 2024.

J. Atack. Historical geographic information systems (gis) database of u.s. railroads. Technical report.

C. Azariadis and J. Stachurski. Poverty traps. Handbook of economic growth, 1:295–384, 2005.

D. Bartelme, Z. Li, and D. Velasquez. Initial conditions, size, and the biggest push in us history. 2024.

D. G. Bartelme, A. Costinot, D. Donaldson, and A. Rodriguez-Clare. The textbook case for industrial policy: Theory meets data. Technical report, National Bureau of Economic Research, 2019.

N. Baum-Snow. Did highways cause suburbanization? The quarterly journal of economics, 122 (2):775–805, 2007.

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

N. Baum-Snow and D. Hartley. Accounting for central neighborhood change, 1980–2010. Journal of urban economics, 117:103228, 2020.

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

P. Bayer, F. Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. Journal of political economy, 115(4):588–638, 2007.

K. Behrens, F. Mayneris, and T. N. Biéda. How Densely Do Manufacturing Establishments Occupy Land? 2022.

D. A. Black, N. Kolesnikova, and L. J. Taylor. Why do so few women work in new york (and so many in minneapolis)? labor supply of married women across us cities. Journal of Urban Economics, 79:59–71, 2014.

H. Bleakley and J. Lin. Portage and path dependence. The Quarterly Journal of Economics, 127 (2):587–644, 2012.

A. S. Blinder. Wage discrimination: reduced form and structural estimates. Journal of Human Resources, pages 436–455, 1973.

R. Blundell, P.-A. Chiappori, T. Magnac, and C. Meghir. Collective labour supply: Heterogeneity and non-participation. The Review of Economic Studies, 74(2):417–445, 2007.

C. E. Boehm, A. A. Levchenko, and N. Pandalai-Nayar. The long and short (run) of trade elasticities. American Economic Review, 113(4):861–905, 2023.

L. P. Boustan. Was postwar suburbanization "white flight"? evidence from the black migration. The Quarterly Journal of Economics, 125(1):417–443, 2010.

A. Brinatti. Third-country effects of us immigration policy. Available at SSRN 4892498, 2024.

J. Brinkman and J. Lin. Freeway revolts! the quality of life effects of highways. Review of Economics and Statistics, 106(5):1268–1284, 2024.

J. K. Brueckner. Urban sprawl: Diagnosis and remedies. International regional science review, 23 (2):160–171, 2000.

S. Cai, L. Caliendo, F. Parro, and W. Xiang. Mechanics of spatial growth. Technical report, 2022.

L. Caliendo, M. Dvorkin, and F. Parro. Trade and labor market dynamics: General equilibrium analysis of the china trade shock. Econometrica, 87(3):741–835, 2019.

B. Callaway and P. H. Sant'Anna. Difference-in-differences with multiple time periods. Journal of econometrics, 225(2):200–230, 2021.

B. Callaway, A. Goodman-Bacon, and P. H. Sant'Anna. Difference-in-differences with a continuous treatment. Technical report, 2024.

F. Carta and M. De Philippis. You've come a long way, baby. husbands' commuting time and family labour supply. Regional Science and Urban Economics, 69:25–37, 2018.

P.-A. Chiappori, J. I. Giménez-Nadal, J. A. Molina, and J. Velilla. Household labor supply: Collective evidence in developed countries. Handbook of Labor, Human Resources and Population Economics, pages 1–28, 2022.

A. Ciccone and R. E. Hall. Productivity and the density of economic activity. The American Economic Review, pages 54–70, 1996.

P.-P. Combes, G. Duranton, and L. Gobillon. Spatial wage disparities: Sorting matters! Journal of Urban Economics, 63(2):723–742, 2008.

P.-P. Combes, G. Duranton, and L. Gobillon. The production function for housing: Evidence from france. Journal of Political Economy, 129(10):2766–2816, 2021.

J. B. Cullen and S. D. Levitt. Crime, urban flight, and the consequences for cities, 1999.

D. R. Davis and D. E. Weinstein. Bones, bombs, and break points: The geography of economic activity. American Economic Review, 92(5):1269–1289, 2002.

J. De la Roca, A. Parkhomenko, and D. Velásquez-Cabrera. Skill allocation and urban amenities in the developing world. Technical report, 2023.

R. Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. American Economic Review, 106(3):479–524, 2016.

D. Donaldson. Railroads of the raj: Estimating the impact of transportation infrastructure. American Economic Review, 108(4-5):899–934, 2018.

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

A. Dube, D. Girardi, O. Jorda, and A. M. Taylor. A local projections approach to difference-in-differences event studies. Technical report.

G. Duranton and D. Puga. Micro-foundations of urban agglomeration economies. In Handbook of regional and urban economics, volume 4, pages 2063–2117. 2004.

G. Duranton and M. A. Turner. Urban growth and transportation. Review of Economic Studies, 79(4):1407–1440, 2012.

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

F. Eckert, T. C. Fort, P. K. Schott, N. J. Yang, et al. County business patterns database. 2020a.

F. Eckert, A. Gvirtz, J. Liang, and M. Peters. A method to construct geographical crosswalks with an application to us counties since 1790. Technical report, 2020b.

D. Epple, B. Gordon, and H. Sieg. A new approach to estimating the production function for housing. American Economic Review, 100(3):905–924, 2010.

R. Ewing, T. Schmid, R. Killingsworth, A. Zlot, and S. Raudenbush. Relationship between urban sprawl and physical activity, obesity, and morbidity. American journal of health promotion, 18 (1):47–57, 2003.

B. Faber. Trade integration, market size, and industrialization: evidence from china's national trunk highway system. Review of Economic Studies, 81(3):1046–1070, 2014.

J. Fan. Internal geography, labor mobility, and the distributional impacts of trade. American Economic Journal: Macroeconomics, 11(3):252–288, 2019.

R. C. Feenstra, P. Luck, M. Obstfeld, and K. N. Russ. In search of the armington elasticity. Review of Economics and Statistics, 100(1):135–150, 2018.

E. Field and K. Vyborny. Women's mobility and labor supply: Experimental evidence from pakistan. Asian Development Bank Economics Working Paper Series, 2022.

P. Fishback and J. A. Cullen. Second World War spending and local economic activity in US counties, 1939–58. The Economic History Review, 66(4):975–992, 2013.

P. V. Fishback and T. Jaworski. World War II and US economic performance. In Economic History of Warfare and State Formation, pages 221–241. Springer, 2016.

D. Frye. Transportation networks and the geographic concentration of employment. Review of Economics and Statistics, pages 1–34, 2024.

M. Fujita, P. R. Krugman, and A. Venables. The spatial economy: Cities, regions, and international trade. MIT press, 2001.

E. L. Glaeser and M. E. Kahn. Sprawl and urban growth. In Handbook of regional and urban economics, volume 4, pages 2481–2527. 2004.

E. L. Glaeser and D. C. Maré. Cities and skills. Journal of labor economics, 19(2):316–342, 2001.

J. Greenwood, N. Guner, and R. Marto. The great transition: Kuznets facts for family-economists. Technical report, National Bureau of Economic Research, 2021.

Y. Gu, N. Guo, J. Wu, and B. Zou. Home location choices and the gender commute gap. Journal of Human Resources, pages 1020–11263R2, 2021.

S. Heblich, S. J. Redding, and D. M. Sturm. The making of the modern metropolis: evidence from london. The Quarterly Journal of Economics, 135(4):2059–2133, 2020.

I. Heiland and W. Kohler. Heterogeneous workers, trade, and migration. European Economic Review, 144:104077, 2022.

G. Hooks and L. E. Bloomquist. The legacy of world war ii for regional growth and decline: The cumulative effects of wartime investments on us manufacturing, 1947–1972. Social Forces, 71 (2):303–337, 1992.

ILO. World employment and social outlook: Trends for women 2017. International Labour Office – Geneva: ILO, 2017.

G. W. Imbens and J. M. Wooldridge. Recent developments in the econometrics of program evaluation. Journal of Economic Literature, 47(1):5–86, 2009.

T. Jaworski. World war II and the industrialization of the American South. The Journal of Economic History, 77(4):1048–1082, 2017.

JICA. Data collection survey on urban transport for lima and callao metropolitan area, final report. Technical report, 2013. https://openjicareport.jica.go.jp/pdf/12087516_01.pdf. Accessed January 2020.

M. Johnson and M. P. Keane. A dynamic equilibrium model of the us wage structure, 1968–1996. Journal of Labor Economics, 31(1):1–49, 2013.

R. Juhász. Temporary protection and technology adoption: Evidence from the napoleonic blockade. American Economic Review, 108(11):3339–3376, 2018.

M. Kawabata and Y. Abe. Intra-metropolitan spatial patterns of female labor force participation and commute times in tokyo. Regional Science and Urban Economics, 68:291–303, 2018.

G. Khanna, C. Medina, A. Nyshadham, D. Ramos, J. Tamayo, and A. Tiew. Spatial mobility, economic opportunity, and crime. 2021.

P. Kline. Oaxaca-Blinder as a reweighting estimator. American Economic Review, 101(3):532–537, 2011.

P. Kline and E. Moretti. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. The Quarterly journal of economics, 129(1):275–331, 2014.

P. Krugman. Increasing returns and economic geography. Journal of political economy, 99(3): 483–499, 1991.

E. Kwon. Why do improvements in transportation infrastructure reduce the gender gap in south korea? Technical report, 2022.

A. Lashkaripour and V. Lugovskyy. Profits, scale economies, and the gains from trade and industrial policy. American Economic Review, 113(10):2759–2808, 2023.

T. Le Barbanchon, R. Rathelot, and A. Roulet. Gender differences in job search: Trading off commute against wage. The Quarterly Journal of Economics, 136(1):381–426, 2021.

S. Lee and J. Lin. Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. The Review of Economic Studies, 85(1):663–694, 2018.

M. Leonardi and E. Moretti. The agglomeration of urban amenities: Evidence from milan restaurants. American Economic Review: Insights, 5(2):141–157, 2023.

Z. Li and D. Koustas. The long-run effects of government spending on structural change: Evidence from Second World War defense contracts. Economics Letters, 178:66–69, 2019.

S. Liu and Y. Su. The geography of jobs and the gender wage gap. The Review of Economics and Statistics, pages 1–27, 2020.

R. E. Lucas and E. Rossi-Hansberg. On the internal structure of cities. Econometrica, 70(4): 1445–1476, 2002.

A. Manning. The real thin theory: monopsony in modern labour markets. Labour economics, 10 (2):105–131, 2003.

E. G. Mendoza. Real business cycles in a small open economy. The American Economic Review, pages 797–818, 1991.

G. Michaels. The effect of trade on the demand for skill: Evidence from the interstate highway system. The Review of Economics and Statistics, 90(4):683–701, 2008.

F. Monte, S. J. Redding, and E. Rossi-Hansberg. Commuting, migration, and local employment elasticities. American Economic Review, 108(12):3855–3890, 2018.

A. Moreno-Maldonado. Mums and the city: Household labour supply and location choice. In Mums and the City: Household Labour Supply and Location Choice: Moreno-Maldonado, Ana. [Sl]: SSRN, 2023.

I. Mulalic, J. N. Van Ommeren, and N. Pilegaard. Wages and commuting: Quasi-natural experiments' evidence from firms that relocate. The Economic Journal, 124(579):1086–1105, 2014.

K. M. Murphy, A. Shleifer, and R. W. Vishny. Industrialization and the big push. Journal of political economy, 97(5):1003–1026, 1989.

E. Nakamura and J. Steinsson. Fiscal stimulus in a monetary union: Evidence from US regions. The American Economic Review, 104(3):753–792, 2014.

R. Oaxaca. Male-female wage differentials in urban labor markets. International Economic Review, pages 693–709, 1973.

OECD. Oecd family database. https://www.oecd.org/els/family/database.htm, 2023.

D. Oviedo, L. Scholl, M. Innao, and L. Pedraza. Do bus rapid transit systems improve accessibility to job opportunities for the poor? the case of lima, peru. Sustainability, 11(10):2795, 2019.

R. Owens III, E. Rossi-Hansberg, and P.-D. Sarte. Rethinking detroit. American Economic Journal: Economic Policy, 12(2):258–305, 2020.

H. S. Pellegrina and S. Sotelo. Migration, specialization, and trade: Evidence from brazil's march to the west. Technical report, 2024.

J. E. Rauch. Productivity gains from geographic concentration of human capital: evidence from the cities. Journal of urban economics, 34(3):380–400, 1993.

S. J. Reber. Court-ordered desegregation: Successes and failures integrating american schools since brown versus board of education. Journal of Human resources, 40(3):559–590, 2005.

S. J. Redding and E. Rossi-Hansberg. Quantitative spatial economics. Annual Review of Economics, 9(1):21–58, 2017.

P. Rhode. After the war boom: Reconversion on the us pacific coast, 1943-49, 2003.

P. W. Rhode, J. M. Snyder Jr, and K. Strumpf. The arsenal of democracy: Production and politics during WWII. Journal of Public Economics, 166:145–161, 2018.

J. D. L. Roca and D. Puga. Learning by working in big cities. The Review of Economic Studies, 84(1):106–142, 2017.

P. N. Rosenstein-Rodan. Problems of industrialisation of eastern and south-eastern europe. The economic journal, 53(210-211):202–211, 1943.

S. S. Rosenthal and W. C. Strange. Female entrepreneurship, agglomeration, and a new spatial mismatch. Review of Economics and Statistics, 94(3):764–788, 2012.

C. Severen. Commuting, labor, and housing market effects of mass transportation: Welfare and identification. Review of Economics and Statistics, pages 1–99, 2021.

S. Sotelo. Domestic trade frictions and agriculture. Journal of Political Economy, 128(7):2690–2738, 2020.

A. C. Stockman and L. L. Tesar. Tastes and technology in a two-country model of the business cycle: explaining international comovements. The American Economic Review, 85(1):168–185, 1995.

L. Sun and S. Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of econometrics, 225(2):175–199, 2021.

T. Tombe and X. Zhu. Trade, migration, and productivity: A quantitative analysis of china. American Economic Review, 109(5):1843–1872, 2019.

N. Tsivanidis. Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio. University of California, Berkeley, 2023.

A. Valentinyi and B. Herrendorf. Measuring factor income shares at the sectoral level. Review of Economic Dynamics, 11(4):820–835, 2008.

D. Velásquez. Transit infrastructure, couples' commuting choices, and gender inequality. Technical report, Working paper, 2024.

P. Warnes. Transport infrastructure improvements and spatial sorting: Evidence from buenos aires. 2021.

L. Weiwu. Unequal access: Racial segregation and the distributional impacts of interstate highways in cities. Technical report, Working paper, 2023.

L. Weiwu. Opportunity in motion: Equilibrium effects of a place-based policy on economic mobility. Technical report, Working paper, 2024.

G. Wright. Old South, new South: Revolutions in the southern economy since the Civil War. Basic Books, 1986.

R. D. Zárate. Spatial misallocation, informality, and transit improvements. 2021.