Title
Essays in Trade and Spatial Economics

Author
Gorton, Nicole Emily

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays in Trade and Spatial Economics

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Nicole Emily Gorton

2022

ABSTRACT OF THE DISSERTATION

Essays in Trade and Spatial Economics

by

Nicole Emily Gorton

Doctor of Philosophy in Economics

University of California, Los Angeles, 2022

Professor Jonathan E. Vogel, Chair

This dissertation examines questions in trade and spatial economics.

The first chapter studies how changes in domestic trade costs can cause regions to decline. The agriculture-intensive states of the American Midwest (the "heartland") lost population relative to the rest of the country over the postwar period; at the same time, the price of shipping agricultural relative to manufactured goods fell considerably. I outline a simple version of a trade model and derive comparative statics of the price, production, and population effects of a decline in agricultural shipping costs to show how these two facts may be linked. I validate the model's predictions by studying how a 1963 Supreme Court ruling that sharply reduced the cost of shipping wheat versus flour affected the flour milling industry. Finally, I calibrate a multi-sector, multi-location version of the model to the U.S. in 1950 and find that observed declines in agricultural trade costs can explain nearly 10% of the postwar population decline in the heartland.

The second chapter is joint work with Pablo Fajgelbaum, Cecile Gaubert, Ed-

uardo Morales, and Edouard Schaal and studies the political economy of transportation investments. Transport networks are among the largest investments made by federal and local governments. What determines which projects are implemented? We study how people's political preferences and policymakers' preferences for redistribution or popular approval determined the implementation of California's High-Speed Rail. We combine detailed spatial data on votes for the project with a quantitative spatial model that captures its economic benefits and use our framework to estimate the weight of economic and political components in transport users' preferences. We find that votes are responsive to the expected real-income benefits of the high-speed rail but that economic benefits explain only a small fraction of the aggregate vote and the variance in welfare across space is almost entirely explained by non-economic factors.

The third and final chapter, joint with Elena Ianchovichina, uses a spatial general equilibrium framework to construct optimal transport networks and optimal expansions to existing networks in most Latin American countries, as well as within MERCOSUR and the Andean Community. The average annual welfare losses due to inefficient domestic transport networks in Latin America are around 1.6%, ranging from 2.4% in Argentina to 0.3% in El Salvador. Spatial misallocation of transnational transport networks is associated with annual welfare losses of 1.8% in MERCOSUR and 1.5% in the Andean Community. The optimal investments we identify correlate relatively well with World Bank infrastructure projects because both prioritize investments in high population areas. Optimal investments in transnational road networks benefit the least developed country in each trade bloc most.

The dissertation of Nicole Emily Gorton is approved.

Pablo D. Fajgelbaum

Ariel Burstein

Michela Giorcelli

Jonathan E. Vogel, Committee Chair

University of California, Los Angeles

2022

To Daniele, without whom this wouldn't exist, because I would have dropped out in the first year upon hearing the words "log linearize".

And to my parents, who have always supported my education: "just keep working".

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

There are many, many people without whom I would not have completed this PhD. First, my husband Daniele: I would not have made it through without his love, support, cooking, and brilliant mind. I am also very grateful to my family especially my parents and my sister, Dani whose constant encouragement and support helped me to complete this dissertation. I have been very lucky to have such a wonderful support system. Last but not least, I am grateful to our dog Mishka who has helped both me and Daniele (plus an untold number of graduate students) through stressful times.

I owe a debt of gratitude to my advisor, Jon Vogel. I met Jon at the UCLA admitted students weekend in Spring 2018. From that point on he has been, outside of my family, my most vocal and constant champion. He supported me through every step of the PhD, from the first quarter all the way to my choice of post-PhD employment. I would not have made it through the program without him, and it would be difficult to overstate the positive impact he has had on my life. He always helped me to believe in myself. Thank you.

I am also very grateful to Pablo Fajgelbaum, one of my committee members as well as my coauthor on the second chapter of this dissertation. I have learned a tremendous amount about how to do research through working with Pablo, and I am very grateful for all of the doors that Pablo opened for me. Pablo introduced me to our fantastic coauthors, who are extremely bright, creative people that I am very lucky to have spent a lot of time working with. Finally, Pablo also introduced me to Elena Ianchovichina at the World Bank, with whom I have coauthored the third

chapter of this dissertation.

Of course, two of the three chapters in this dissertation would not have existed without my fantastic coauthors. Chapter 2 is joint work with Pablo Fajgelbaum, Eduardo Morales, Cecile Gaubert, and Edouard Schaal, also known as the Superfriends. Chapter 3 is joint work with Elena Ianchovichina. In addition, Kelalani Luong-Kha provided excellent research assistance on Chapter 1, and Nan Xiang and Max Schwarz both provided excellent research assistant on Chapter 2. I am very grateful to all of them for their work on these projects.

Throughout the PhD, my research benefited greatly from the input of others, including participants in the UCLA International, Industrial Organization, and Economic History Proseminars such as John Asker, Oleg Itskhoki, and Augusto Ospital, as well as Adriana Lleras-Muney and my committee members Ariel Burstein and Michela Giorcelli. I also was lucky to be able to spend a quarter at Dartmouth College during my PhD as an International Economics PhD Fellow, and am very grateful to Treb Allen and his colleagues at Dartmouth for that experience. Financial support from the UCLA Graduate Division, as well as from the NBER/Department of Transportation Economics of Transportation in the 21st Century allowed me to focus exclusively on research for a large part of my PhD.

I was very lucky to have a wonderful cohort of classmates. Nobu Abe, Adam Baybutt, and Akria Ishide were the members of my first year comprehensive exam study group and I have them to thank for getting me through the first year. Vicky Barone, Ariadna Jou, Alvaro Boitier, Lucas Zhang, Lorenzo Cativelli, and Anthony Papac all contributed to my research through our informal reading group. And of course, I am very grateful for the other wonderful friends I made during the PhD,

including those listed here as well as Vickie Wang, Manu Navjeevan, Chang He, and Jesper Böjeryd.

Finally, I would like to thank Chiara Paz, Ed McDevitt, and everyone else in the Economics Department office for all of their help throughout the years. In particular, Chiara helped me find grading positions that would allow me to live in Stanford with my husband, and I am very grateful to her for that.

VITA

2012-2016	B.A. Economics, University of Chicago
2016-2018	Senior Research Analyst, Federal Reserve Bank of New York
2018-2020	M.A. Economics, University of California Los Angeles
2020-2021	Short Term Consultant, World Bank
2021	Dartmouth College International Economics PhD Fellow
2018-2022	Grader, Research Assistant, and Teaching Assistant, Economics Department, UCLA

CHAPTER 1

Trade Costs, Supply Chains, and the Decline of the Heartland

For the past 30 years, our population has also been growing and shifting. The result is exemplified in the vast areas of rural America emptying out of people and of promise...

— President Richard Nixon, State of the Union, 1970

1.1 Introduction

Over the past century, the spatial distribution of economic activity within many countries, especially the United States, has changed dramatically. These large shifts in where people live, where they work, and where goods are produced have welfare implications for people living in different places. Inevitably, people living in some regions end up better off than those in other regions. Why have some regions prospered while others have declined? What role have changes in trade costs within countries played in shaping these welfare gains and welfare losses across locations? To make progress on these questions, this paper examines a region over a period of significant decline: the agriculture-intensive states of the American Midwest (the "Heartland") over the postwar period.[1]

Between 1950 and 1980, the percent of the U.S. population living in Heartland states fell by 18%, from 9% to 7% (U.S. Census). By this measure, the Heartland fared worse than almost any other Census division.[2] Figure 1.1 maps the percentage change in the share of the national population living in each state between 1950 and 1980.[3] States colored in (darker) green grew (more) while states in (darker) orange shrunk (more).[4] The Heartland is outlined in black. Four Heartland states – the Dakotas plus Iowa and Nebraska – made up half of the top eight states in terms of relative population declines over the period. About one-third of all U.S. counties experienced net out-migration over the period; of those, 35% were in the Heartland, even though the Heartland includes less than 20% of all U.S. counties.

The Heartland states are sometimes referred to as the "breadbasket" because they largely specialize in producing agricultural goods, especially bulk grains like wheat (Wishart (2004)).[5] In 1950, North Dakota and South Dakota led the nation as the states with the largest share of gross output coming from the agricultural

[1] I define the Heartland as the set of states in the Census' West North Central division which corresponds to the non-Rustbelt Midwestern states. The Midwest follows the U.S. Census classification. Following Alder, Lagakos and Ohanian (2014), I define the Rustbelt as: Illinois, Indiana, Michigan, New York, Ohio, Pennsylvania, West Virginia and Wisconsin. Thus, the non-Rustbelt Midwest includes Minnesota, Iowa, Missouri, the Dakotas, Nebraska, and Kansas

[2] There are nine Census divisions. The Middle Atlantic division experienced a percentage decline in relative population of about the same magnitude (18%).

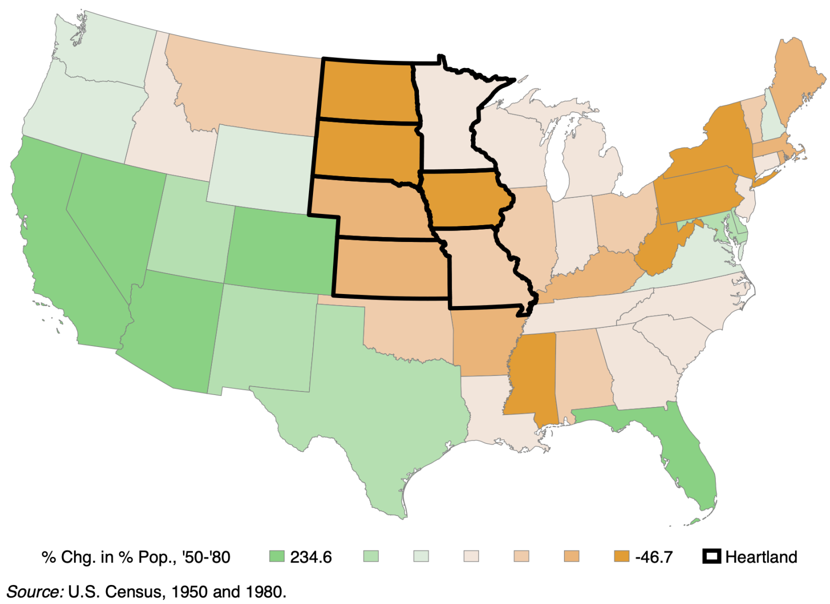
[3] Figure 1.12 plots percentage of the U.S. population living in each Census region in each year between 1900 and 2000. While there was a small decline in the Midwest's relative population preceding this period, it accelerated after the war.

[4] I focus here on state-level data because my eventual model calibration will be at the state level due to data constraints. However, the pattern is equally striking at the county level.

[5] Figure 1.13 plots the percentage of exports in agriculture for each state.

sector while Iowa and Nebraska were close behind. Although Heartland states housed only 9% of the population in 1950, together they produced nearly a quarter of the nation's agricultural output. Population declines in these rural, agriculture-intensive areas received considerable policy attention during the period. President Nixon's 1970 State of the Union remarked on the need to "stem the migration to urban centers". This rural to urban migration also motivated the passage of the Rural Development Act of 1972, an early example of a place-based policy, which provided financial support to rural areas. Its stated purpose was to "foster a balanced national development" (Rural Development Act, 1972).

Figure 1.1: The Postwar Decline of the Heartland, 1950 to 1980



Note: This figure shows the percentage change in the share of the national population living in each state between 1950 and 1980. *Source:* U.S. Census, 1950 and 1980.

This paper proposes, studies, and quantifies the importance of a novel explanation for this decline of the American Heartland: changes in the structure of domestic trade costs.[6] Freight rates for bulk (agricultural) goods fell considerably relative to those for finished (manufactured) products beginning in the 1950s. To document this fact, I digitized records from the Interstate Commerce Commission (ICC), the regulatory body for freight transport during the period. The ICC annually published the *Freight Commodity Statistics*, from 1928 to 1980, in which it reported aggregate data on tons shipped and revenue earned by major rail carriers.[7] To measure shipping costs levied by shippers on producers, I compute the revenue per ton earned by railroad companies for shipments of goods from each commodity class.[8]

In Figure 1.2, I plot annual real revenue per ton earned by the railroads in each year, separately for agricultural goods and manufactured goods.[9] I find large declines in the shipping costs of agricultural relative to manufactured goods beginning in the postwar period.[10] Revenue per ton earned in these each of these two sectors was

---

[6]This hypothesis, that changes in the structure of trade costs within the postwar U.S. had important effects on the distribution of economic activity, received some attention among economists during the period. For example, see Meyer and Morton (1975) who wrote that, "[T]he lower freight rates for transporting bulk commodities than their fabricated equivalent commonly encourages manufacturers to substitute the transport of bulk commodities for the movement of the finished goods."

[7]Total revenue is the total payments made by producers to railroad shippers.

[8]Revenue per ton is a standard measure of shipping costs used in the literature; for example, see Hyslop and Dahl (1964). However, in addition to the tariff charged by shippers, revenue per ton may also reflect changes in the underlying distance shipped. I control for this using data on bilateral trade costs in Section 1.4. Even conditional on distance, the decline in relative shipping costs is around one-third.

[9]Figure 1.14 shows the same figure using data on Class I motor carriers (trucks), starting when these data become available in 1956.

[10]This figure uses aggregate data. When I quantify the model, I use data on revenue per ton earned by each state pair for each sector to show that this relative decline in shipping costs holds up to controlling for origin-destination fixed effects.

relatively constant until around 1955 at which point it began to fall sharply in the agricultural sector. By 1967, agricultural revenue per ton was nearly one third lower than manufacturing revenue per ton. Qualitative evidence suggests that most of this decline was driven by innovations in the shipment of bulk grain products; new types of rail cars, like the unit train and the covered hopper car, allowed railroads to significantly reduce the cost of service on these products particularly over long distances.[11]

Figure 1.2: Railroad Shipping Costs Over Time

Note: This figure shows revenue per ton earned by Class I railroads in each year, separately for agricultural goods and manufactured goods. *Source:* Interstate Commerce Commission's *Freight Commodity Statistics*.

How could these changes in trade costs have caused relative population declines

[11]These new technologies are discussed more in Section 1.3. Panel (a) of figure 1.22 shows the covered hopper car, which displaced box cars over this period, as shown in panel (b) of the same figure.

in the Heartland states? To explore the channels through which these two facts may be linked, I outline a simple, Armington model of trade between two locations with three sectors. Labor is mobile across locations, and sectors are connected to each other through input-output linkages. The mechanics of this simple model are at the heart of most trade models used in the literature. The key assumption I make is that one location has a comparative advantage in the production of agricultural goods and so it relies on imports of agricultural goods from other regions relatively less than the other location in the model. I use this model to derive analytic comparative statics describing the response of prices, production, firm locations, welfare, and population to a decline in the cost of trading agricultural goods.

The model highlights a key channel through which declines in agricultural goods can cause a population decline of the heartland: the input-output structure of the economy. Agricultural goods are used by other sectors including food processing, textile mills, rubber and plastics industries, and apparel industries, as intermediate inputs (BEA, 1947). Initially, trade costs gave locations that specialized in the production of agricultural goods a comparative advantage in certain downstream sectors. But when it becomes relatively cheaper to acquire agricultural goods outside the heartland, agriculture-intensive locations' downstream comparative advantage is weakened and downstream firms move out of the agriculture-intensive location. Plus, there may be spillover effects from agriculture-intensive manufacturers to other firms through the input-ouput structure, as well as through external economies of scale in manufacturing.[12]

To show that this channel is operating in the data, I take advantage of a natural

[12]For example, as in the case of Bartelme et al. (2019).

experiment that affected flour mills in 1963. Studying this natural experiment has a few key benefits. First, it allows me to narrow in on a single industry for which I have collected location-specific data on prices for goods, plants, firms, and trade in the upstream and downstream sectors. Most of these data are not systemically available for this time period for a broader set of industries. Second, it provides a setting with a clear and significant change in the cost of shipping the upstream, agricultural good relative to the downstream, manufactured good, that was caused by an unexpected change in regulation rather than by changes in patterns of demand or supply that may also affect prices and production patterns.

Before 1963, wheat and flour, which are flour mills' agricultural input good and final manufactured good respectively, were equally costly to ship via rail as a result of railroad regulation by the Interstate Commerce Commission. Rail car innovations in the late 1950s and early 1960s reduced the railroad cost of shipping bulk products, but regulation initially prevented railroads from lowering prices accordingly. In 1963, a Supreme Court ruling allowed railroads to significantly reduce the price of wheat shipments which allowed them to compete more effectively with barges. Other railroads soon followed, and the result was that the price of shipping wheat fell by about a third while the price of shipping flour remained the same. I document this fall in the relative cost of shipping wheat versus flour using an event study design.

I then use newly digitized data and two sources of variation generated by the Court's ruling to show how this change in shipping costs affected flour prices, production, and firm locations. To estimate the causal effects of the change in trade costs on outcomes, I use a differences-in-differences strategy. The first source of variation that I use is variation across time, before and after the ruling. The second source of variation is variation across locations, as flour mills that were initially located close to wheat production were relatively less affected by the change in trade costs since they did not initially rely heavily on the railroad for wheat shipping.

I find empirical support for the model's mechanisms using this setting. I find that, after the ruling, prices of wholesale flour, consumer flour, and bread rose in the Midwest, which is close to where wheat is produced, as compared with prices in cities and states farther away from where wheat is produced. Flour milling capacity and the number of mills fell in areas close to where wheat is produced following the change in shipping rates relative to areas further away. The relative decline in the number of flour mills in wheat-producing areas was driven by an exodus of relatively less productive mills. These results are consistent with the simple model's comparative statics.

Finally, I document that, consistent with Figure 1.2, declines in the cost of shipping agricultural products over this period were not limited only to wheat and instead applied to a broad set of agricultural products. I quantify these changes in the costs of shipping agricultural goods between each pair of U.S. states over the postwar period. I then study the extent to which these observed declines in agricultural shipping costs can explain the relative population decline of the Heartland over the postwar period.

To do this, I extend the simple model to a multi-sector, multi-location model of trade between U.S. states a là Caliendo and Parro (2015). To capture the mechanism, the model includes labor mobility, trade costs that vary by sector, and input-output linkages between sectors. I include land as a fixed factor of production to correctly model the agricultural sector. I digitize data from 1950 on output and trade in order to calibrate the model. I then feed into the model *only* changes to the cost of shipping agricultural products in order to ask how changes in agricultural trade costs changed distribution of population across states over the period. I find that this channel can explain around 8% of the postwar population decline in the Heartland. Nearly half a million people would have remained in the Heartland had the structure of trade costs remained at its 1950 level.

This paper makes several contributions to the literature. First, the paper contributes to a broad literature in trade about how changes in trade costs affect different locations. Glaeser and Kohlhase (2004), Fajgelbaum and Redding (2022), Donaldson and Hornbeck (2016*b*) and Donaldson (2016) among others all use historical settings to study how changes in trade costs shape the distribution of economic activity across locations within countries. Recent work by Costinot and Donaldson (2016) studies the gains from market integration within the U.S. Gollin and Rogerson (2010) study the impact of the transport network on rural to urban migration. Relative to these papers, this paper proposes and studies a new channel – the changing structure of domestic trade costs – to explain changes in economic activity across locations.

Second, by proposing and carefully quantifying a new explanation for the decline of the American Midwest, this paper contributes to a broad literature studying historical patterns of population and production in the United States. Long and Siu (2018) and Hornbeck (2012) study how the Dust Bowl contributed to out-migration in some Midwestern counties in the 1930s. Eckert and Peters (2018) study the spatial implications of structural change within the U.S. over the past century. Caselli and Coleman II (2001) study the contribution of structural transformation in explaining the convergence of incomes across regions in the U.S. Kim (1995) explores the changing spatial distribution of manufacturing in the U.S. Alder, Lagakos and Ohanian (2014) study how competitive pressure affected the postwar, manufacturing-intensive

Rust Belt and Autor et al. (2014) study how exposure to import competition affected manufacturing-intensive locations since the 1990s. Finally, the particular population patterns I study are documented extensively, though without explanation, by Wilson (2009).

This paper also contributes to a growing body of literature, both theoretical and empirical, on supply chains and trade with input-output linkages. Caliendo and Parro (2015) and Caliendo et al. (2018) embed input-output linkages into a standard trade, Melitz (2003) style trade model. Recent theoretical work by Grossman and Helpman (2021), Antras, Fort and Tintelnot (2022), Antras and Helpman (2004) considers the effects of changes in upstream and downstream trade costs in a world connected by supply chains. This paper provides novel empirical evidence describing how downstream prices and production re-allocate in response to changes in upstream trade costs, and how this can affect population in the long run. While these mechanisms are at the heart of the models in all of these papers, there is limited evidence that they operate in the data which this paper provides. Relative to Cox (2021), I show how changes in trade costs allow supply chains can reallocate population across space.

Finally, this paper relates to work in agricultural economics on the flour milling industry over the postwar period. Kim et al. (2001), Nightingale (1967), Hyslop and Dahl (1964) and Harwood (1991) discuss the existence of the cost shock and speculate on its potential implications for the locations of flour mills. Babcock (1976) and Babcock, Cramer and Nelson (1985) use aggregate data to measure which regions will experience an increase in flour production in response to the shock. Relative to these papers, I use very granular, plant-level data on flour mills to credibly quantify the extent to which observed changes in trade costs affected prices and the locations of mills.

1.2 Simple Model and Testable Predictions

To explore the channels that link changes in the structure of trade costs with population declines, I outline a simple two-state, three-sector model. I use the model to derive comparative statics which I then test empirically.

1.2.1 Simple Model

In the initial equilibrium, there are two locations, New York, denoted as N, and Kansas, denoted as K. There are three sectors: an agricultural sector ("wheat", W), a manufacturing sector ("flour", F), and an outside manufactured good sector M. I assume that, initially, N imports more wheat from K than K imports from N, so $\pi^W_{NN} < \pi^W_{KK}$ where π^W_{in} is the share of wheat imported from i by n. This pattern will hold if, for example, K has a comparative advantage in the production of wheat as compared with N.[13] This pattern of comparative advantage is the key assumption that will drive the results.

Agents' Problem. In each location, there is a common component of utility across all agents in a location, plus an idiosyncratic component of utility associated with

[13] Consistent with this assumption, there are strong patterns of comparative advantage in the production of agricultural goods across states in the U.S., with agricultural exports making up a particularly large share of the Heartland's exports. Figure 1.13 shows the share of agricultural goods in each state's export bundle in 1949. Appendix proof 1 shows one set of parameters – that Kansas is sufficiently more productive in growing wheat than New York – that is consistent with this assumption.

each agent. Agents have quasi-linear utility over a constant elasticity of substitution (CES) aggregator of flour, plus the outside good. The outside good is homogenous and freely traded. Because agents have CES utility over flour types from each location, agents love variety; thus, they want to consume flour from every location since each location is producing its own variety.[14] The elasticity of substitution of flour across origins is σ_F. Agents in a location i choose to buy flour from each location c_{ni}^F and the outside manufactured good from each location c_{ni}^M to solve:

$$\max_{c_{ni}^F, c_{ni}^M} \left[\sum_n c_{ni}^M\right] + \ln\left(\left[\sum_n (c_{ni}^F)^{\frac{\sigma_F - 1}{\sigma_F}}\right]^{\frac{\sigma_F}{\sigma_F - 1}}\right)$$

subject to a budget constraint of $w_i = \sum_n p_{ni}^M c_{ni}^M + p_{ni}^F c_{ni}^F$. The common component of indirect utility is then $v_i = w_i + \ln\left(\frac{\sigma_F}{\sigma_F - 1}\right) - \ln\left(P_i^F\right) - \frac{\sigma_F}{\sigma_F - 1} P_i^F$.

The total indirect utility of a worker b in state i is $v_i^b = v_i + \epsilon_i^b$ where ϵ_i^b represents agent b's idiosyncratic preferences for location i. Agents choose to live in the state that gives them the largest indirect utility. Assuming that $\epsilon_n^b \sim \mathit{Gumbel}$ yields the share of agents living in state n:

$$\lambda_i = \frac{exp\left(w_i + \ln\left(\frac{\sigma_F}{\sigma_F - 1}\right) - \ln\left(P_i^F\right) - \frac{\sigma_F}{\sigma_F - 1} P_{n'}^F\right)}{\sum_n' exp\left(w_{n'} + \ln\left(\frac{\sigma_F}{\sigma_F - 1}\right) - \ln\left(P_{n'}^F\right) - \frac{\sigma_F}{\sigma_F - 1} P_{n'}^F\right)}$$

The number of people living in state i is $L_i = \lambda_i L$, where L is total population which I assume to be exogenous.

[14]This, too, is not divorced from reality. Different regions specialize in different types of flour, often related to the type of wheat that is locally grown. For example, White Lily flour is generally associated with the Southern states because it is milled there, and is important in baking biscuits.

Production. Wheat is produced using labor, $Y_{it}^W = T_{it}^W L_{it}^W$. The outside good is also produced using labor, $Y_{it}^M = T_{it}^M L_{it}^M$. Flour is produced using a CES aggregator of wheat, $Y_{it}^F = T_{it}^F \left[\sum_n \left(c_{ni}^W\right)^{\frac{\sigma_W - 1}{\sigma_W}}\right]^{\frac{\sigma_W}{\sigma_W - 1}}$ where σ_W is the elasticity of wheat across origins.

Prices. I assume that all markets are perfectly competitive and that wheat and flour are subject to sector-specific iceberg trade costs, τ_{in}^W and τ_{in}^F respectively. Thus, prices in each sector are $p_{in}^F = \tau_{in}^F p_{ii}^F$ and $p_{in}^W = \tau_{in}^W p_{ii}^W$ where p_{ii}^j is the price of producing goods from sector j in location i. Since the outside manufactured good is homogenous and freely traded, it is the numeraire good so its price is 1. Wages in each location will then be set based on productivity in this sector $w_n = T_i^M$ and thus are exogenously determined.

Extension with Heterogeneous Firms. While a causal link between declines in agricultural good trade costs and changes in the distribution of population does not require firms to be present, including firms in the model allows me to derive testable predictions governing what will happen to firm entry and average productivity of mills in different locations. The richness of my data will allow me to test these predictions. In this case, I assume monopolistic competition among flour mills, closely following Chaney (2008) and Krugman (1980). Each agent chooses c_{ni}^M and $c_{ni}^F(\omega)$ where $\omega \in \Omega$ is a flour variety to solve:

$$\max_{c_{ni}^M, c_{nit}^F(\omega)} \left[\sum_n c_{ni}^M\right] + \ln\left(\left(\sum_k \int_{\Omega_k} \left(c_{ki}^F(\omega)\right)^{\frac{\sigma_F - 1}{\sigma_F}} d\omega\right)^{\frac{\sigma_F}{\sigma_F - 1}}\right)$$

$$\text{s.t.} \quad w_i = \sum_k \int_{\Omega_k} p_{ki}^F(\omega) c_{ki}^F(\omega) d\omega + \sum_n p_{ni}^M c_{ni}^M$$

From this maximization problem, I obtain the quantity demanded of each flour variety from agents in state i:

$$c_{ki}^F(\omega) = \left(p_{ki}^F(\omega)\right)^{-\sigma_F} \left(P_i^F\right)^{\sigma_F - 1} \tag{1.1}$$

In terms of production, each mill produces its own flour variety ω, though a variety is unique conditional on a firm's productivity, which I index as φ. Entry and exit of flour mills in each state are endogenous and depend on a zero profit condition. In each market there is some endogenously given mass of potential entrants, M_i and some share of them will end up entering the market. The number of firms operating in any period is then $M_i^* = M_i \cdot (1 - G(\varphi_i^*))$ where φ_i^* is the lowest level of firm productivity for which profits are non-negative, and G(.) is the distribution of firm productivities, which I will assume to be Parteo as in Chaney (2008).

The threshold level of productivity above which all firms enter, and below which no firms will enter, is the productivity for which profits in a market are equal to 0. Profits made by a firm with productivity φ in location i are given by:

$$\pi_i(\varphi) = \sum_j c_{ij}^F(\varphi) \left(\underbrace{\tilde{\tau}_{ij}^F \frac{\sigma_F}{\sigma_F - 1} \frac{P_i^W}{\varphi}}_{p_{ij}^F(\varphi)} - \underbrace{\tilde{\tau}_{ij}^F \frac{P_i^W}{\varphi}}_{MC_{ij}} \right) - w_i f_e$$

where f_e is the fixed cost of entry, denominated in wages, $MC_{ij} = \tilde{\tau}_{ij}^F \frac{P_i^W}{\varphi}$ is the marginal cost of production for a firm of productivity φ in state i to produce flour for state j, and $p_{ij}^F(\varphi) = \tilde{\tau}_{ij}^F \frac{\sigma_F}{\sigma_F - 1} \frac{P_i^W}{\varphi}$ is the price charged by firm φ in state i for flour in state j. There is no fixed cost of exporting; all firms that enter can export without paying an additional fixed cost. The zero profit condition yields a closed form solution for the cut-off value of productivity. All firms above this level will enter the market:

$$\varphi_n^{*F} = \left(\frac{(\sigma_F - 1)^{1-\sigma_F} \, \sigma_F^{\sigma_j} w_n f_n \left(c_n^F\right)^{\sigma_F - 1}}{\sum_i \left(\tau_{ni}^F\right)^{1-\sigma_F} Q_i^F (P_i^F)^{\sigma_F}} \right)^{\frac{1}{\sigma_F - 1}} \tag{1.2}$$

1.2.2 Testable Predictions

I express the model in changes, where $\hat{x} = \frac{x_{post}}{x_{pre}}$ and consider a decline in the cost of shipping the agricultural good, $\hat{\tau}^W < 1$, while holding all other exogenous variable constant.

Price effects. When the price of shipping wheat falls, flour prices fall everywhere because wheat is the only input to flour production and prices are set with perfect competition. Flour prices fall by more in New York than in Kansas; formally, $\hat{p}_{NN}^F < \hat{p}_{KK}^F < 1$. I provide a proof of this result in Theorem 1. This result is driven by the initial pattern of trade, generated by Kansas' agricultural comparative advantage, plus the fact that trade within a state is costless. Because Kansas has a comparative advantage in wheat production, Kansas is importing relatively less wheat from New York in the initial equilibrium than New York is importing from Kansas. Thus, a larger share of wheat imports to New York are affected by the decline in shipping costs, so the price of importing wheat (and thus of producing flour) falls by more in New York than in Kansas.

A similar logic holds in the case of consumer flour prices: $\hat{P}_N^F < \hat{P}_K^F < 1$. Since demand is CES, agents are buying flour from all locations in the initial equilibrium. But since trade across states is costly, agents in New York are initially buying relatively more flour from New York in the initial equilibrium than agents in Kansas are buying from New York. Thus, flour becomes relatively cheaper for consumers to buy in New York than in Kansas which I show in Theorem 2. This is how the cost of living changes across locations as a result of changes in the costs of shipping agricultural goods.

Production effects. Because producer prices fall everywhere, demand for flour rises everywhere. However, demand rises by more in New York since the price of producing falls by more there; as a result, the production of flour rises more in New York than in Kansas, $\hat{Y}_N^F > \hat{Y}_K^F > 1$. The proof of this result is in Theorem 3.

Firm location effects. While my baseline model does not include firms, I turn to the version of the model with monopolistic competition to generate predictions of the effect of trade costs on firms. In this version of the model, because it becomes cheaper to produce flour everywhere and especially so in New York, the productivity threshold for firm entry falls by more in New York than it falls in Kansas. Thus, more flour milling firms enter in New York and the total number of firms in New York increases by more as well, as compared with Kansas, $\hat{M}_N^F > \hat{M}_K^F$. The proof of this result is in Theorem 5.

Productivity effects. Since the productivity entry threshold falls by more in New York, relatively less productive firms can enter the market there. Thus, the new, lower-productivity firms drag down the average productivity level in New York, so average productivity actually falls by more here than in Kansas: $\hat{\varphi}^*_N < \hat{\varphi}^*_K$, as per Theorem 4.

Decline of the heartland. Since wages are exogenous in this model, effects on welfare in this model operate through changes in the cost of living across locations. Since consumer prices of flour fall by more in New York, welfare increases by more in New York. Since relative population is linked to relative welfare, relative population thus rises there as well compared with Kansas. Since Kansas the relatively agriculture-intensive location here as is the Heartland, this effect is the "decline of the heartland". The proof of this result is in Theorem 6.

1.3 Empirical Case Study

To test whether these model predictions hold in the data, I study the flour milling industry following a sudden and sharp decline in the railroad cost of shipping wheat generated by the outcome of a Supreme Court case in 1963.

1.3.1 Historical Background

Before 1963, railroads charged identical rates for similar movements of flour and wheat (Babcock (1976), Babcock, Cramer and Nelson (1985), Harwood (1991), Held (1979), USDA (1964)). This was primarily due to regulation by the Interstate Commerce Commission (ICC), which, from its inception in 1887 until rail deregulation in 1980, controlled to a great extent how railroads could set prices. As noted by a 1963 article in the *Southwestern Miller*, a trade publication for the flour milling industry, "...the parity between rates of wheat and flour has long been in effect..." (The Southwestern Miller (1963)).

However, in the 1960s, new innovations in the shipment of bulk commodities including the covered hopper car and the unit train made it cheaper for railroads to ship bulk grains like wheat as compared with manufactured goods like flour.[15] Shipments of flour were unaffected by these innovations in part because shipments of wheat tend to be in much larger quantities than shipments of flour; bulk shipments lend themselves well to both covered hopper cars and unit trains whereas smaller shipments of more processed goods do not (USDA (1964)). In addition, hygiene requirements are much stricter among shipments of flour which is a manufactured good as compared with wheat which is a raw material. This makes flour even more difficult to ship in vast quantities. A 1965 article in the *Minneapolis Tribune* quoting a railroad executive explains: "Unit trains for grain have meant streamlined, high-speed operation– rapid turnaround and high utilization of equipment that have kept profits up in spite of lower rates... No one has come to us with any kind of a similar development for flour. Simply, more grain can be moved faster, per car, than flour".

The Southern Railroad, one of the first to develop these new technologies, proposed new, lower rates on wheat to the ICC in 1962. By law, the ICC had seven months to either approve or deny the rate change. After seven months, no decision was made. In lieu of a ruling, the reduced rates were supposed to go into effect, but a competing barge company sued. They claimed the rates could not go into effect until the ICC had made a decision and that the new rates would "irreparably

[15]For example, see figure 1.22.

injure their [the barge company's] respective economic interests". In the meantime, the ICC required the Southern to keep its rates at the initial level. In 1963, the Supreme Court ruled in *Arrow Transportation v. Southern Railway Company* that the ICC didn't have the authority to prevent proposed rates going into effect since the decision period had lapsed (United States Court of Appeals (1962)).

Once the Southern was allowed to introduce the new technology and offer lower rates, other railroads followed. Importantly, this change in railroad rates was *not* generated by changes in producer locations or characteristics, and thus is plausibly exogenous to the outcomes I will study. Many trade publications, newspapers, and economists of the period took note of changing relative trade costs and speculated on its implications for the spatial distribution of the industry. For example, the cost shock was described succinctly in the 1964 USDA Farm Index (USDA (1964)):

> With recent changes in the grain rate structure, *it now costs more to ship grain products, including flour, by rail from some locations than it does raw grain.* As a result, some millers near the city bakeries and retail outlets for flour may have new transportation advantage over those located nearer the production areas.

Local newspapers reveal how Midwest flour millers felt about the changing rates. An article published on April 30, 1963 in *The Southwestern Miller* was titled "Kansas Millers Plea Against Peril to Trade", and stated that "the Southern Railway rate cuts would place Kansas flour at a disadvantage of as much as 43c per cwt in comparison to wheat" (The Southwestern Miller (1963)). On June 27, 1965, an article entitled "Minneapolis Mills Fight for Life, Blame Transit Rates" was published in the *Minneapolis Tribune* (1965). The article included the following quote, describing how Minneapolis millers feared for their viability given the advantage that millers closer to population centers would face after this change in relative shipping costs:[16]

> A growing controversy is raging over the issue of changing transportation rates which, the [Minneapolis] millers contend, *have given Eastern flour mills an overpowering competitive advantage over the Midwest.* They say they're hurt because a disparity between transportation rates on wheat and flour makes it cheaper to ship wheat east for milling and sale in the population centers than to mill it here and ship the flour to the big Eastern markets. *The chief target of milling industry criticism are railroad freight rates which until recently were about the same for wheat and flour...*

This is shipping rate change that I exploit to study the response of prices and downstream production locations to a change in the cost of shipping agricultural inputs relative to manufactured final goods.

1.3.2 Data

Flour mill locations, sizes, flour prices. To identify the changing distribution of flour mill locations, I digitized records on the locations, sizes, and ownership of U.S. flour mills from the *Northwestern Miller*, a trade publication that published an annual directory of all U.S. flour mills that I obtained from the published, *Sosland Publishing*. Figure 1.20 shows two examples of what these directory entries look like. Given the address of each mill, I geocode mills to latitude and longitude coordinates. I then combine the coordinates with 2010 U.S. Census county boundaries to create a panel dataset of counties and years, with variable including the total number of mills, the total number of mills that belonged to a multi-unit firm, and the total

[16] Emphasis is added.

milling capacity. In addition, I use mill names and addresses to track plants over time and construct a dataset of mill entry and exit.[17]

Table 1.6 shows annual summary statistics of these data. Some key trends emerge: over the time period, the total number of mills is falling, but total capacity is rising, so the average capacity of a mill is rising considerably over the period. The share of flour produced by the states that are top producers of wheat is also falling, and the share of mills that are owned by large corporations nearly doubles over the period. One drawback is that I do not observe actual production of each mill in each year; I only observe the mill's capacity which may be a noisy indicator of demand for that mill's flour since it is costly to adjust capacity.

I measure producer prices of flour at major markets from The *Southwestern Miller*, a trade publication that included averages of flour prices from local mills for a selection of major milling markets (The Southwestern Miller (1955)). Figure 1.21 shows an example of the price listings for Kansas City. Data were published every week for many different varieties of flour. I use data from the last week in October for standard patent flour, except in the case of Portland, Oregon where the price of standard patent flour is never listed and instead I use family flour. Table 1.7 shows average prices in 1963 and 1966 for each state in my sample, which I obtain either by using the price listings from a city in that state, or from averaging over cities (as in the case of Kansas). One drawback of these data is that they are only available for a small number of cities; in total, only ten states are represented. I measure the retail prices of flour and bread in a selection of different cities from the Bureau of Labor Statistics' *Retail Prices in U.S. Cities*.

[17] I detail this matching process in 1.7.3.1.

Wheat Production and Prices. I obtain county-level data on wheat production and wheat yields from the United States Department of Agriculture (USDA) and the Census of Agriculture. 63% of wheat production in 1963 was produced in Midwestern states; 71% was produced by Midwestern states plus Montana and Idaho (USDA, 1963).

Transportation Network. I measure transport costs for agricultural products in 1950 among all pairs of counties. To using a map of the 1957 railroad network drawn by the Army Service Corps of Engineers that I have digitized.[18] To measure the cost of shipping agricultural products one ton-mile along this network, I use the reported revenue per ton-mile earned by Class I railroads in 1950 for Products of Agriculture (ICC, 1950).

Railroad Trade. I obtain data on railroad trade of wheat and flour between U.S. states and major regions from the Carload Waybill Sample Statistics. These data are a 1% sample of all terminated waybills, which are contracts between railroad companies and producers. Each observation in the data is of a commodity traded in a given year between an origin state or region and a destination state or region, and the data include the volume of goods traded as well as the revenue earned by the railroad on shipments along that route in that year. I have digitized state level data are available from 1958 through 1966. Region level data for five major regions in the U.S. are much coarser but are available for more years, 1950 through 1987,

[18] The original map is shown in Appendix Figure 1.19.

with some gaps.

1.3.3 Effects on trade costs

1.3.3.1 Empirical strategy

I first quantify the extent to which the court's ruling affected the relative cost of shipping wheat versus flour. To do this, I use railroad trade data at the route-commodity-year level to estimate the following event study:

$$\ln\left(\frac{revenue_{odct}}{ton_{odct}}\right) = \sum_{y\neq 1963} \boldsymbol{\beta_y} \cdot 1(t=y) \cdot 1(c=\text{wheat}) + \gamma_{odc} + \gamma_{odt} + \gamma_{ot} + \gamma_{dt} + \epsilon_{odct} \quad (1.3)$$

where o is an origin region, d is destination region, and $c \in$ (wheat, flour) is a commodity. Standard errors are clustered at the route level, of which there are 25. $revenue_{odct}$ measures total payments from producers to railroad shippers for goods of commodity group c shipped along route od in year t. $tons_{odct}$ measures total tons of commodity c shipped by railroad companies in year t from o to d. Each event study coefficient $\boldsymbol{\beta_y}$ measures the relative cost of shipping wheat versus flour in year y relative to 1963, which is the last untreated year.

I attribute my estimates of $\boldsymbol{\beta_y}$ to the causal effect of the Supreme Court ruling. I assume that, had the ruling in 1963 not happened, the cost of shipping wheat relative to the cost of shipping flour would have remained constant over this period. I include various fixed effects control for factors other than the court's ruling that may influence the cost of shipping wheat relative to that of shipping flour. For example, changes in the compositional quality of wheat or flour being shipped, which would affect its value, could also affect the relative revenue per ton earned by the railroads. Assuming this effect is constant across routes, commodity-time fixed effects control for any such compositional changes. Similarly, origin-destination-commodity fixed effects control for any route and commodity-specific differences in the cost of shipping that do not vary over time. I also include origin-time and destination-time fixed effects.

1.3.3.2 Results

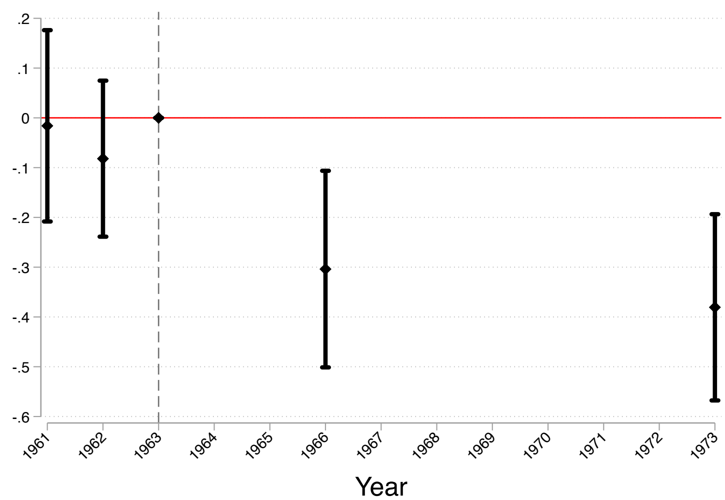
Figure 1.3 shows event study estimates of equation 1.3. While the cost of shipping wheat was not statistically different than the cost of shipping flour in 1961 and 1962, as these point estimates are small and not statistically different from zero, there is considerable wedge between the trade costs for the two commodities by 1966: the cost of shipping wheat has fallen by about 30% relative to the cost of shipping flour. This significant and sudden shift in shipping costs provides the ideal scenario to study how declines in agricultural good trade costs affect production across locations. Finally, I separately estimate equation 1.3 by commodity. Consistent with the story, the decline in relative trade costs is driven by a decline in the cost of shipping wheat, while the cost of shipping flour remained unchanged.

1.3.4 Effects on prices, production, firms

How did flour prices, flour production, and flour mill locations evolve after this change in shipping costs? Figure 1.4 shows the relationship between initial proximity to wheat and flour mill locations. Each county is colored based on its wheat market access which is the inverse distance-weighted average of wheat production in all surrounding counties.[19] Dark orange and red counties are closer to wheat-producing

[19] This measure is defined formally in Equation 1.5.

Figure 1.3: Effects of the Ruling on Shipping Costs



Note: This figure shows event study estimates of equation 1.3, in which I compare the revenue per ton earned by railroads, a measure of shipping costs, in shipping wheat versus flour each year, relative to the year of the Court's ruling.

locations; Each black dot represents a county with at least one flour mill in the indicated year; dots are sized by the number of mills in that county. Comparing 1961 (the top panel, corresponding to the last pre-shock year for which I have data) to 1975 and 1985, many (but not all) of the mills in North Dakota, Montana, and Kansas have closed. Many new mills are instead closer to population centers including New York Cities, New Orleans, Tampa, and Jacksonville. While these maps are suggestive, I estimate difference-in-differences models to quantify the effects.

1.3.4.1 Empirical strategy

To estimate the causal effects of changes in agricultural shipping costs on outcomes, I take advantage of two sources of variation. First, I take advantage of variation across time induced by the Supreme Court ruling, comparing outcomes before and after 1963. Second, I take advantage of variation across locations. Places close to where wheat is produced initially relied less on the railroad for shipping wheat than places far from where wheat is produced. For example, a flour mill in New York is far from wheat production and thus must import wheat from the Midwest while a mill in Kansas is close to wheat production and instead exports finished flour, rather than importing wheat.

Effects on prices. I first measure how prices of flour and bread evolved differently in the Midwest, where wheat is produced, versus in other places. I use a differences-in-differences specification. My estimating equation is:

$$\log\left(price_{it}\right) = \sum_{y \neq 1963}^{T} \beta_y \cdot 1(y = t) \cdot 1(i \in Midwest) + \gamma_t + \gamma_i + \epsilon_{it} \tag{1.4}$$

where i is a city and t is a year. The identifying assumption is that, in lieu of changes in trade costs, prices would have evolved in the same way in locations within the Midwest as compared with locations outside the Midwest.

Each event study coefficient β_y measures the difference in prices in the Midwest in year y relative to 1963, versus outside the Midwest. If the predictions of Section 1.2 hold, then we would expect to prices to fall by more outside the Midwest (or, in other words, rise in the Midwest relative to other locations), so $\beta_y > 0$ for $y > 1963$. There should be no difference in the evolution of prices in the Midwest versus in other locations, in which case $\beta_y = 0$ for $y > 1963$.

Effects on production and firms. When looking at production and mill locations, I have data for every county, instead of for a selection of cities as in the case of price data. This allows me to construct a continuous measure of exposure to the change in wheat shipping costs. I measure each location's "wheat market access" $WMA_{i,1959}$ before the shock:

$$WMA_{i,1959} = \sum_{n} (\tau_{ni}^{W})^{1-\sigma_W} \cdot Y_{n,1959}^{W} \tag{1.5}$$

where $Y_{n,1959}^{W}$ is production of wheat in bushels in 1959 and τ_{ni}^{W} is the iceberg cost of shipping wheat from n to i. σ_W determines the distance decay; I estimate this parameter to be 13.7 using trade data as described in Section 1.4. My main empirical specification is:

$$y_{it} = \sum_{y \neq 1961}^{T} \beta_y \cdot 1(y = t) \cdot \log(WMA_{i,1963}) + \gamma_t + \gamma_i + \epsilon_{it} \tag{1.6}$$

where i is a location, which will be either a county or a city, and t is a year.[20] The first difference is the difference between the pre-1963 and post-1963 periods while the second difference is that between counties that are relatively better or worse producers of wheat. My coefficients of interest are each β_y. I omit the last pre-treatment year for which I observe data (1961), so each estimated coefficient is relative to the 1961 level. Standard errors are clustered at the county-level, of which there are around 3,000.

Similarly to the price regressions, my identifying assumption is that, in lieu of changes in trade costs (i.e., if the Supreme Court had ruled in favor of Arrow Trans-

[20] In robustness checks, I estimate the differences-in-differences version of this specification, which is: $y_{it} = \beta \cdot 1(t > 1963) \cdot \log(WMA_{i,1963}) + \gamma_t + \gamma_i + \epsilon_{it}$.

portation Co. and the new rates had not gone into effect), flour prices, mill locations, and production would have evolved similarly across locations regardless of that location's initial specialization in wheat production. The event study nature of these regressions allows me to test, to some extent, this assumption. If this assumption holds, then we would expect there to be no difference in outcomes across locations prior to the change in trade costs after controlling for year and location fixed effects. In later sections, I will also include a series of robustness checks to control for alternative explanations of my results.

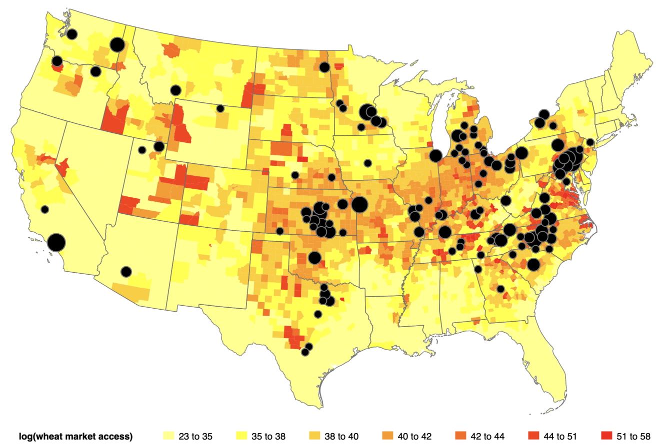
1.3.4.2 Results

Prices. Figure 1.5 plots event study estimates of equation 1.4. In panel (a), the outcome variable of interest is $log(p_{ii}^F)$, the log of the producer price of flour. I find that the producer price of flour rises in the Midwest compared to other locations by about 3% in the first year following the decline in trade costs. Three years later, the price remains about 7% higher in the Midwest relative to the rest of the country, even though there were no differential trends in flour prices between the Midwest and the rest of the country prior to 1963. In panel (b), the outcome variable of interest is $log(P_i^F)$, the log of the consumer price of flour. Here, I find a similar result as in the case of producer prices: by 1965, two years after the change in trade costs, consumer prices of flour have risen by about 6% and remain at that lower level for the decade. Panel (c) looks at bread prices and finds an effect of around 10%.[21]

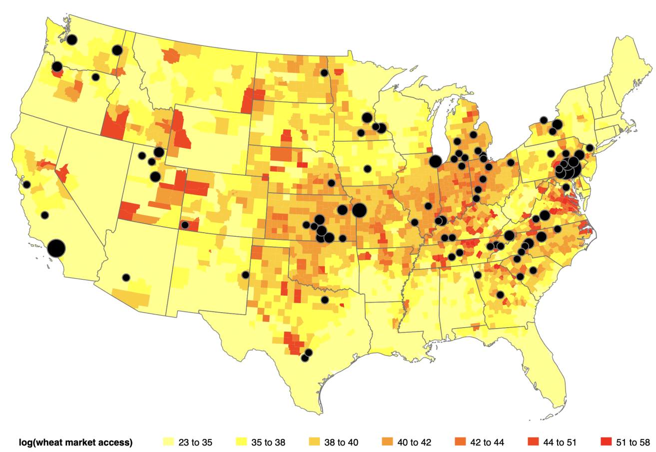
[21] Bread pricing reflects the price of flour sold to bakers, which may be priced differently than flour sold in grocery stores.

Figure 1.4: Flour Mill Locations and Capacity

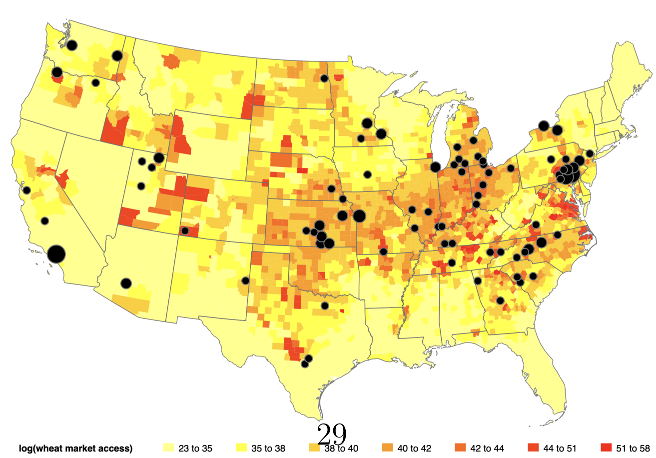
(a) 1961



(b) 1975



(c) 1985



Note: Each black dot represents the centroid of a county with at least one flour mill in the indicated year, with the size of the centroid scaled by the number of mills in that county. The background shading represent access to wheat in 1959, as measured by wheat market access in equation 1.5, with shades tending towards red representing areas closer to wheat production areas. *Source:* The Northwestern Miller, Army Map Service Map of U.S. Railroads (1957), and the Census of Agriculture (1959).

Production. Figure 1.6 plots event study estimates of equation 1.6. In Panel (a), the outcome variable is the number of mills in each county in each year. I omit the year 1961, which is the last year for which I observe data before the 1963 ruling.[22] To get a sense of the magnitudes, consider that in 1961, the average county had 0.17 flour mills. The estimated coefficient for 1975 is around -0.01. Thus, moving from a 25th percentile wheat access location (far from wheat) to a 75th percentile wheat access location (close to wheat), which is a difference of about five log points, is associated with a decline of -0.05 mills, or about 30% of the mean.

In Panel (b), I look at the total capacity of mills. Because there are many zeros, I use an inverse hyperbolic sine transformation which allows me to preserve the zeros. The pattern of changes in capacity following 1961 is almost the same as in the case of the number of mills. By 1975, flour milling capacity is about 20% lower when moving from a location with a 25th percentile wheat access location to a 75th percentile wheat access location. In Panel (c), I look at the log of the average mill size where mill size is milling capacity in a county divided by the number of mills in that county. There appears to be no differential change in mill size by location, suggesting that changes in milling capacity are driven by changes in the number of mils, not by changes in mill sizes.

In Figure 1.7, I look at outcomes related to the firms version of the model. While I do not directly observe firm productivity, I estimate how the changes in trade costs differently affected plants that were initially part of multi-unit firms versus those that were initially stand-alone plants. For example, all of General Mills' plants would be considered as part of a multi-unit firm. I separately estimate the event study for

[22]While I use OLS here, I use Poisson and Logit (with an indicator for whether there is a least one mill) models in robustness checks.

single-unit plants and multi-unit plants, or plants that are part of larger firms. I find that the effect on the number of firms is completely driven by a decline in the number of single-unit plants in initially wheat-intensive locations.

Robustness. I consider some alternative explanations for these results, and address each in turn. Tables 1.3 and 1.4 show these robustness checks for the number of mills outcome variable and the flour milling capacity outcome variable respectively. Column (1) shows the baseline difference-in-differences estimate corresponding to the event study regressions in equation 1.6 and I add additional controls in each subsequent column.

To control for state-level policies that may have changed across years, for example, any tax policies that may have incentivized firms to locate in a certain state, I include state by year fixed effects in Column (2). Another concern is that technology was increasing the returns in scale; in fact, the average size of mills grew considerably over this period (see Table 1.6). This could vary across counties depending on the initial distribution of mills: for example, a county with smaller mills initially may see a relative decline in the number of mills not due to changes in trade costs, but due to the changing returns to operating a large mill. To account for this, column (3) adds county by year linear time trends to the baseline specification.

An additional explanation of these patterns is that because mills were becoming larger, they required more labor, and thus moved closer to cities where labor may have been more easily available. While the labor share of flour milling output is small, making this story unlikely, I include dummy variables for each year interacted with a county's initial population, thus allowing the impact of population on the location of mills to vary over time in Column (4).[23]

Another possibility is that demand for U.S. flour is increasing from the rest of the world, in which case it may be becoming increasingly attractive for flour mills to locate near coasts and ports. Since the coasts are far from where wheat is grown, wheat-intensive areas locations would become relatively less attractive. To control for this possibility, I introduce dummy variables for each year interacted with the distance of that county to the coast of the U.S. an additional covariates into my main regression. This allows the impact of being close to a port to vary by year, as the level effect of a county's proximity to the port will be absorbed by the county fixed effect. These results are in column (5).

Finally, I consider some additional measures of initial proximity to wheat. I use the Global Agro-Ecological Zones data to measure the potential wheat yield in each location, given the climate and soil conditions. In this case, I simply replace the quantity of wheat production in equation 1.5 with the potential yield. I do the same using the wheat yield in each county, where yield is computed as the ratio of wheat bushels harvested to the wheat acreage planted based on the Census of Agriculture.

I also consider robustness to the functional form choice of my empirical specification. My main results are OLS estimates. In the case of the number of mills, the outcome variable is a count variable. However, the vast majority (98%) of county-year observations have either zero or one mill. Thus, my regression in this case is, in practice, nearly a linear probability model. I consider both Poisson models, a linear probability model, and a logit model. Table 1.5 shows these estimates for both the number of mills and the wheat flour capacity.

[23] In 1947, Flour and meal products' wage and salary share of total output was 4.7% (Census of Manufacturers).

1.4 Quantification

How important were these trade cost changes in shaping the distribution of the U.S. population? This section outlines and calibrates a full, quantitative model of intranational trade, which I use to asses the extent to which declining agricultural trade costs can explain the population decline of the heartland states over this period. Relative to the simple model, this full model has two key advantages. The first advantage is that it allows me to quantify the importance of changes in trade costs in shaping the distribution of population in the U.S. over the period. The second advantage is that it relaxes assumptions I made in the simple model and allows for endogenous changes in wages and rental rates.

The model set up closely follows Caliendo and Parro (2015) and Caliendo et al. (2018), and includes a few key ingredients required to capture the mechanism. First, there are input-output linkages across sectors. Second, labor is mobile across states. Third, trade costs vary by sector. Finally, to appropriately model the agricultural sector, land is an input to production.

1.4.1 Model Setup

Agents' problem. In each location n, a representative agent chooses consumption of goods C_n^j from each sector j to maximize their utility subject to a budget constraint:

$$\max u(C_n) = \Pi_{k=1}^{K} \left(C_n^k\right)^{\alpha_n^k} \text{ where } \sum_{k=1}^{K} \alpha_n^j = 1 \text{ subject to } I_n = \sum_{j=1} P_n^j C_n^j$$

where $I_n = w_n + r_n \ell_n / L_n + D_n$ is per-capita income in location n and α_n^k is the share of total income spent on products in sector k. w_n is the wage in location n, r_n is the rental rate of land in n, ℓ_n is the endowment of land in n, and D_n is the trade deficit which will be defined later.

This problem yields demand functions of $C_n^k = \frac{\alpha_n^k I_n}{P_n^k}$ and an indirect utility function: $v_n = I_n \Pi_{k=1}^K \left(\frac{\alpha_n^k}{P_n^k}\right)^{\alpha_n^k}$. v_n is the common component of utility of agents living in a location n. In addition, there is an idiosyncratic component of utility, such that indirect utility of a worker b living in location n is given by $v_n^b = v_n + \epsilon_n^b$, where $\epsilon_n^b \sim Gumbell(0, \epsilon)$, as in Eckert and Peters (2018). From this, I derive the share of agents living in each location n as $\lambda_n = \frac{exp(\epsilon \cdot v_n)}{\sum_{n'} exp(\epsilon \cdot v_{n'})}$. The population of each location is given by $L_n = \lambda_n L$. Labor and land are both freely mobile across sectors within a location and not subject to any frictions; hence, there will only be a single wage and rental rate in each location.

Intermediate goods. A continuum of intermediate goods $\omega^j \in [0, 1]$ is produced in each sector j. A firm producing good ω^j in sector j in location n combines land, labor and materials via a Cobb-Douglas production function:

$$q_n^j(\omega^j) = z_n^j(\omega^j) \left(\left(l_n^j(\omega^j)\right)^{1-\delta_n^j} \left(L_n^j(\omega^j)\right)^{\delta_n^j} \right)^{\gamma_n^j} \prod_p^K \left(m_n^{p,j}(\omega^j)\right)^{\gamma_n^{p,j}} \tag{1.7}$$

where $l_n^j(\omega^j)$ is land, $L_n^j(\omega^j)$ is labor, $m_n^{p,j}(\omega^j)$ are composite intermediate goods from sector p used in the production of intermediate good ω^j. γ_n^j is the share of value added in total output for sector j. $\gamma_n^{p,j}$ is the share of sector j total output in location n that comes from sector p such that $\sum_{p=1}^K \gamma_n^{p,j} = 1 - \gamma_n^j$. Given this production function, the unit cost of goods from sector k in location n is $p^j_n(\omega^j) = c^j_n / z^j_n(\omega^j)$ where:

$$c^j_n = \Gamma^j_n \cdot (w_n^{1-\delta^j_n} r_n^{\delta^j_n})^{\gamma^j_n} \prod_{p=1}^{K} (P^p_n)^{\gamma^{p,j}_n}$$

where $\Gamma^j_n = \gamma^j_n \left(\left((\delta^j_n)^{-\delta^j_n} \cdot (1-\delta^j_n)^{\delta^j_n - 1} \right) \right)^{-\gamma^j_n} \cdot \prod_p^K \left(\gamma^{p,j}\right)^{-\gamma^{p,j}}$. The productivity of each intermediate good producer $z^j_n(\omega^j)$ is distributed Frechet with location parameter T^j_n and shape parameter θ^j.

Composite goods. Following Caliendo and Parro (2015), composite good producers in each sector purchase intermediate goods from each firm producing varieties ω^k within that sector. They substitute across goods within a sector with elasticity of substitution σ^k to bundle goods into Q^k_n.

$$Q^k_n = \left[\int r^k_n(\omega^k)^{1-1/\sigma_k} d\omega^j \right]^{\sigma_k/(\sigma_k - 1)} \tag{1.8}$$

where $r^k_n(\omega^k)$ is the demand of intermediate goods ω^k from the lowest cost supplier across all possible origins. These composite goods are then used as materials or as final consumption. Composite good producers have the following demand for good ω^k:

$$r^k_n(\omega^k) = \left(\frac{p^k_n(\omega^k)}{P^k_n} \right)^{-\sigma_k} Q^k_n \tag{1.9}$$

where P^k_n is the unit price of the composite intermediate good:

$$P^k_n = \left[\int p^k_n(\omega^k)^{1-\sigma_k} d\omega_k \right]^{\frac{1}{1-\sigma_k}} = \left[\sum_i T^k_i \left(c^k_i \tau^k_{in} \right)^{-\theta_k} \right]^{\frac{1}{-\theta_k}} \tag{1.10}$$

where the second equality follows since the price realized for each good in each sector is the lowest price available from all locations: $p_n^k(\omega^k) = \min_i \left[\frac{c_i^k \tau_{in}^j}{z_i^k(\omega^k)}\right]$ where τ_{in}^j is the amount of a good from sector j shipped from i to n that must be shipped for one unit to arrive. Total expenditure on sector j goods in n is given by $X_n^j = P_n^j Q_n^j$.

Trade. Trade flows along a route in in sector j are defined as $X_{ni}^j = \pi_{ni}^j \cdot X_n^j$, where:

$$\pi_{in}^j = \frac{T_i^j \left(c_i^j \tau_{in}^j\right)^{-\theta^j}}{\sum_m T_m^j \left(c_m^j \tau_{mn}^j\right)^{-\theta^j}} \tag{1.11}$$

Goods market clearing. Total spending on goods from a sector j in a location n must be the total amount spent on final good consumption, plus demand for goods from that sector for all firms in that location for use as intermediate goods:

$$X_n^j = \alpha_n^j I_n + \sum_k \gamma_n^{j,k} Y_n^k \tag{1.12}$$

Land & labor market clearing. Land used in production equals land available in each location, $\sum_j l_n^j = l_n$. Labor supply equals labor demand everywhere, $L_n = \sum_j L_n^j$.

Trade imbalances. Since there are significant imbalances in trade between U.S. states, I allow for such imbalances in the model. The trade deficit for each states is total imports, minus total exports:

$$D_n = \sum_j \sum_i X_{ni}^j - \sum_j \sum_i X_{in}^j$$

1.4.2 Calibration

To calibrate the model, I express all variables in changes where $\hat{x} = x'/x$.[24] The model has $N = 48$ locations, representing each of the contiguous U.S. states, and $J = 24$ sectors. These sectors, which include 16 manufacturing sectors, 6 non-tradable sectors, and 2 raw material sectors. This set of sectors was chosen based on data availability.[25] The initial period of the model is set to 1950 and the post period is 1980.

1.4.2.1 Trade, Production, and Parameters

Trade. In the model, outcomes critically depend on trade patterns between states in the initial equilibrium. To measure these patterns, I use data on railroad trade between each pair of states for each sector from the Carload Waybill Sample Statistics in 1949. I created a crosswalk to match CWSS sectors, which use an old style of commodity classifications, to the modern commodity groupings used in my model based on the descriptions of each.[26] CWSS commodity classifications are fairly disaggre-

[24] Appendix 1.8.4 lists the set of equations that characterize the equilibrium of the model in changes.

[25] The sectors are: Food or kindred products and tobacco; Textille mill products; Apparel, leather, finished textile products; Lumber or wood products; Furniture or fixture; Pulp, paper, or allied products; Chemical or allied products; Petroleum or coal products; Rubber or plastics products; Clay, concrete, glass, or stone products; Primary metal products; Fabricated metal products; Machinery, excluding electrical; Electrical machinery, equipment, supplies; Transportation equipment; Miscellaneous products or manufacturing; Construction, wholesale and retail trade; Finance, insurance, real estate; Transportation, communications, utilities; Arts, recreation, accommodation, repair; Education, legal, health; Other services; Agriculture products; Mining products.

[26] This is trivial the case of agricultural goods and mining products, as trade is reported for these aggregate categories. However, I need to separately observe trade patterns for each manufactured good sector to calibrate the model.

gated relative to the model's sector categories: for example, CWSS sectors include "soap compounds" (mapped to the chemical products sector) and "sugar" (mapped to food and kindred products sector).

There are two main challenges with this dataset. First, many origin-destination-sector pairs don't appear in the data, suggesting zero trade flows. Some fraction of these zeros may not be "true zeros"; instead there may be small amounts of trade that are not captured in the 1% sample. In fact, some states that are listed in the Census of Manufacturers as producing goods from a given sector may appear as in the trade data as not producing any goods from that sector. Using this raw data would thus introduce considerable noise into my calibration. The second challenge is that I do not observe value shipped; instead, I observe revenue earned and tons shipped along each route. As a result, I cannot directly compute the value of trade between each location for each sector.

To address both of these issues, I measure gross output by sector and state from other sources, as I describe below. I then use the trade data to measure the share of exports from each origin state that are shipped to each destination. To address the sampling issue, I do not use the raw data. Instead, I use a Poisson model to estimate the share exported from each origin to each destination. I estimate, separately by sector:

$$\lambda_{in}^{j} = \gamma_{i}^{j} \cdot exp\left(\beta_0 + \beta_1 \cdot \log(distance_{in}) + \sum_{l} \beta_l \cdot X_{in}^{l}\right)$$

where λ_{in}^{j} is the share of sector j goods produced in i exported to n. $distance_{in}$ is the railroad distance between i and n and X_{in}^{l} is the lth covariate. I then use these estimated coefficients to compute predicted export shares. The Poisson model is consistent with the model: it will never predict that there will be zero trade flows between two locations. I use the predicted export shares in my calibration. Observing export shares, predicted from the trade data, and the total value of exports for each state and sector allows me to compute the total value shipped between each pair of locations for each sector.

In addition to distance, I use a number of different covariates to predict export flows. I use an indicator for whether that observation corresponds to a state shipping to itself, $1(i = n)$. I use an indicator for whether the destination is a coastal state, and for whether the origin and destination are both coastal states. Finally, to capture the size of the market in the destination state, I use demand for goods from that sector in the destination n.

Figure 1.17b shows the correlation between the estimated export shares and the export shares that I observe in the data. Figure 1.17a shows the relationship between that share of total imports that each state imports from itself and that state's initial specialization in agriculture, in both the model and the data. Importantly, import shares match this pattern of specialization: states that specialize in agricultural production import relatively larger shares of agricultural goods from themselves.

Gross output, value added, land shares, IO linkages. I measure input-output linkages across all sectors $\gamma^{k,j}$ as well as each manufacturing sector's share of value added in total output γ^j from the Bureau of Economic Analysis Input-Output table in 1950.[27] To measure gross output for each location and sector Y_i^j, I rely on a number of different sources. First, I digitize state-level data on value added by each

[27]These values are not disaggregated by state, so I assume that γ_n^j for manufacturing sectors and $\gamma_n^{k,j} = \gamma^{k,j}$ for all sectors.

manufacturing sector from the Census of Manufacturers (CMF, 1947).[28] When a sector has a small presence in a state, value added is not be separately reported for that sector in that state. Instead, all remaining value added is reported in a single category ("All other major industry groups"). I allocate this remaining amount to the remaining sectors that are not separately listed based on the employment share of each omitted sector in the total employment of all unlisted sectors.[29] I then use each sector's value added share of output from the BEA I-O table to convert value added, measured from the CMF, to gross output: $Y_i^j = \frac{VA_i^j}{\gamma^j}$ for every manufactured sector.. I also use the CMF to measure each state's share of manufacturing wages in manufacturing value added and assume that $\delta_n^j = \delta_n$ for every manufactured sector.

For the agricultural sector, I measure value added and gross output (total value of agricultural production) in 1949 by state from the USDA's Economic Research Service report on value added. I measure the agricultural value added share as $\gamma_n^{Ag} = VA_n^{Ag}/Y_n^{Ag}$. It is very high on average, around 75%, ranging from around 45-50% in Delaware and New Hampshire to above 95% in North Carolina and North Dakota.[30] I measure the share of wages in value added δ_n^j by dividing the total amount spent on hired labor and employee compensation by total value added for each state. I

[28] Ideally, I would measure gross output as the total value of shipments. However, the total value of shipments is not reported at the two-digit sector level for each state. There were concerns about double counting shipments between firms within the same two-digit category. While the total value of shipments is reported at the four digit industry code, using these figures would require the digitization of many hundreds of additional figures.

[29] Given this assumption the only state-sector pairs with zero gross output are those with zero value added and zero employment. Because this requires granular data on employment by sector, I rely on the 1940 U.S. Census which is available for all sectors instead of the 1950 U.S. Census which has not been fully released yet.

[30] Most of the intermediate inputs used in agricultural sector are from the agricultural sector, and include feed purchases, seed purchases, livestock purchase, plus a small amount of manufactured inputs like electricity, fertilizer, fuel, and pesticides (USDA).

measure value added, gross output, and wage payments for the mining sector from the 1954 Census of Mineral Industries. The average value added share of output is also quite high, around 75%, ranging from 52% (Nevada) to 88% (California).

I measure non-tradable gross output using a combination of two data sources. First, I measure national gross output for each non tradable sector from the BEA IO table. I then measure the share of output in each sector that is produced in each state based on each state's share of national employment in that sector, and use that to compute total output in each state for each tradable sector. I impose that $\gamma_n^j = 1, \delta_n^j = 1 \forall j \in NT$.

The data and methodology described above allow me to measure total gross output for each state and sector; however, because my model is a closed economy, I need to measure output in each state and tradable sector destined for *domestic* use. I adjust for the amount of gross output that each state and sector exports abroad by using the U.S. Department of Commerce 1966 "State Export Origin Series". It reports, for each state and manufacturing sector, total exports and total gross production in 1966. I use this to compute the share of production that is exported, and subtract this amount from gross output. Data are also provided for each state for agricultural products. For mining, there is no state-specific data, so I measure the aggregate percent of mining product shipments that are exported (5%) from U.S. Department of Commerce 1958 report on "U.S. Commodity Exports and Imports as Related to Output".[31]

Productivity dispersion. I measure the dispersion of productivity within each

[31] Both of these reports are "one-off" and exist only for the years listed here; I assume that export patterns were similar over the 1950-1966 period.

sector θ_j from the trade data and equation 1.11. I assume that trade costs take the form $\tau^j_{od} = raildist^{\kappa}_{od} \cdot X$ where X is a constant.[32] Using Poisson to handle the large quantity of zeros in the raw trade data, I estimate:

$$\pi^j_{in} = \gamma^j_i \cdot \gamma^j_n \cdot exp\left(\beta^j \cdot \log(raildist_{in}) + \epsilon^j_{in}\right) \tag{1.13}$$

where $\beta^j = \kappa \cdot \theta^j$ and $raildist_{ij}$ is the railroad distance in miles between i and j as computed using the 1957 railroad network. I measure κ, the elasticity of trade costs with respect to distance, as $\kappa = 0.169$ from Donaldson (2016). Table 1.1 shows estimates of θ^j for each sector. The median value is 8.3, which falls well within the range of conventional estimates.[33] Precisely estimated values range from 2.91 (electrical machinery) to 18.5 (miscellaneous products).[34] I assign the median value of 8.3 as the elasticities for the non-tradable sectors.

Final Consumption Shares. Following Caliendo and Parro (2015), I solve for each sector's share in final consumption in each state, α^j_n to satisfy the goods market clearing condition in equation 1.12 in the initial equilibrium. Figure 1.16a shows the median expenditure share across states for each sector.

Other Parameters. I measure total employment and the share of workers in each state and state from the 1950 U.S. Census. I measure the land area of each state

---

[32] This constant will not affect the estimation of each θ_j here, but will become important later on.

[33] For example, Eaton and Kortum (2002) also estimate a value of around 8.

[34] Estimates for textile mill products and apparel are very small and nosily measured. In my calibration, I assign the median value of θ^j to these sectors.

Table 1.1: Estimates of Trade Elasticities

Sector	$\hat{\theta}^j$
Food or kindred products and tobacco	7.48(0.45)***
Textille mill products	0.52(1.45)
Apparel, leather, finished textile products	0.97(1.55)
Lumber or wood products	9.93(0.54)***
Furniture or fixtures	6.35(0.60)***
Pulp, paper, or allied products	7.87(0.53)***
Chemical or allied products	9.30(0.64)***
Petroleum or coal products	18.58(0.62)***
Rubber or plastics products	4.30(1.09)***
Clay, concrete, glass, or stone products	19.31(0.63)***
Primary metal products	4.62(1.76)***
Fabricated metal products	11.18(0.49)***
Machinery, excluding electrical	4.38(0.55)***
Electrical machinery, equipment, supplies	2.91(1.13)***
Transportation equipment	8.73(0.73)***
Miscellaneous products or manufacturing	18.50(1.11)***
Agriculture products	13.78(0.47)***
Mining products	16.86(0.54)***
Median	8.30

Note: This table shows estimates of $\hat{\theta}^j$ from estimating equation 1.13 with Carload Waybill Sample Statistics data, and railroad distances measured from the 1957 rail network. I set $\delta = 0.169$ from Donaldson (2016), and report $\hat{\theta}^j = -\hat{\beta}^j/\eta$.

$land_n$ from the Census of Agriculture. I compute the initial rental rate on land as $r_n = \frac{\sum_j Y_n^j \gamma_n^j (1-\delta_n^j)}{\ell_n}$. Similarly, wages are $w_n = \frac{\sum_j Y_n^j \gamma_n^j \delta_n^j}{L_n}$. Finally, the elasticity of mobility across locations is critical to determining how population will change across places. I set $\epsilon = 2.38$, as estimated in the context of the historical United States by Eckert and Peters (2018).

1.4.2.2 Quantifying changes in trade costs

The key inputs to the model are bilateral changes in agricultural trade costs. To measure these, I first measure agricultural trade costs between states in the initial equilibrium. I assume that trade costs between each pair of locations are given by $\tau_{in}^{Ag} = raildist_{in}^{\kappa} \cdot X$ where $raildist_{in}$ is the railroad distance between locations i and n, and X is a constant. I solve for the constant, X, using data on revenue earned, tons shipped, railroad distance, and costs of goods at the origin, as described in Section 1.8.4.3.

Then, I estimate changes in trade costs by collecting data on revenue per ton mile earned by railroads across each pair of states for agricultural goods and manufactured goods in 1988, which is the first year of bilateral, state-to-state railroad trade data that is available after 1966. I estimate:

$$\Delta_{1949-1988} \log \left(\frac{revenue_{odc}}{tons_{odc}} \right) = \gamma_{od} + \beta_1 \cdot 1(c \in \text{Ag Good}) + \epsilon_{odc} \tag{1.14}$$

Results of estimating equation 1.14 are in Table 1.2. In column (1), all observations are weighted equally; in column (2), I weight observations by the log of the tonnage of that commodity shipped along that route in 1949. I find that overall, revenue per ton earned on agricultural products fell by roughly 33% (unweighted) to 37% (weighted).

I use these differences-in-differences point estimates to estimate the average change

Table 1.2: Changes in Agricultural Shipping Costs, 1949-1988

	Δ log(revenue per ton)	
	(1)	(2)
1(c$\in$ agricultural)	-0.329***	-0.367***
	(0.0438)	(0.0426)
N	1600	1600
Weighted	No	Yes
FE	Route	Route

Note: This table shows estimates of equation 1.14 using data from the 1949 and 1988 Carload Waybill Sample Statistics. Column (1) shows unweighted estimates; in column (2), I weight observations by the volume of trade along the route in 1949. Bulk agricultural goods include soybeans, wheat, corn, sorghum, oats, barley and rye, and rice. Standard errors clustered at the route level are reported in parentheses.

in revenue per ton earned along each location:

$$\hat{\hat{\tau}}_{in}^{Ag} = \frac{1 + 0.63 \cdot \tau_{in,1950}^{Ag}}{1 + \tau_{in,1950}^{Ag}} \tag{1.15}$$

Figure 1.15 shows $\hat{\hat{\tau}}_{in}^{Ag}$ for each pair of states. I assume that there are no changes in trade costs for shipments within states.[35]

[35] In a robustness check, I measure within-state trade costs based on the average distance between each pair of counties within the state and allow such changes to occur, based on equation 1.15. However, there is limited data to measure the extent to which trade between states occurs via rail; as these are shorter distances, goods may be more likely to be transported via trucks. Thus, this case in which I assume no changes in trade costs within states, is the most conservative choice.

1.4.3 Results

Changes in population. I use the model to compute the change in each state's share of the U.S. population given these declines in the costs of shipping agricultural goods. Figure 1.8 shows resulting the distribution of changes in population across states. To measure the extent to which the model can explain the observed population decline of the heartland, I construct the model and data implied change in the heartland's share of population between 1950 and 1980:

$$\hat{s}^v = \underbrace{\frac{\sum_n 1(n \in heartland) \cdot \hat{l}_n^v \cdot l_n}{\sum \hat{l}_n^v \cdot l_n}}_{\text{Heartland's share in 1980}} \Bigg/ \underbrace{\frac{\sum_n 1(n \in heartland) \cdot l_n}{\sum l_n}}_{\text{Heartland's share in 1950}} \tag{1.16}$$

for $v \in (model, data)$ where l_n is the population state n in 1950 as measured in the data and $\hat{l}_n^v$ is the change in state n's population in either the model or data. To compute the percentage of $\hat{s}^{data}$ that can be explained by the model, I construct $\frac{\hat{s}^{model}}{\hat{s}^{data}} \cdot 100$. I find that the observed decline in trade costs for agricultural products can explain around 8% of population declines in the heartland.

Consistent with the intuition of the simple model, the mechanism operates by reducing population in locations that initially specialized in agriculture. Figure 1.9 shows the relationship between each state's initial specialization in agriculture, as measured by the percentage of employment in agriculture in 1950, and the log of the change in relative population between 1950 and 1980.[36] Panel (a) shows the relationship in the data and panel (b) shows the relationship as generated by the model. In both, there is a negative relationship between initial agriculture specialization

[36]This is computed for state n as $\log\left(\frac{s_n^{1980}}{s_n^{1950}}\right)$ where $s_n^t = \frac{pop_n^t}{\sum_i pop_n^t}$.

and the change in relative population, with the most agriculture-intensive locations losing relatively more people.

In the data, the relationship is only marginally significant overall as some locations that are initially moderately intensive in agriculture, such as Florida, Arizona, and Nevada, saw large increases in relative population over this period. Nevertheless, there is a strong association between these two variables among the seven heartland states with the most agriculture intensive – the Dakotas, Iowa, and Nebraska – losing more relative population. Using the model, I find that declines in agricultural shipping costs generate a similar pattern in the model. While there is no strong correlation overall between the model and data generated changes in population, the correlation is strong among the heartland states (77%) as the model captures the observed relative population declines of these states.[37]

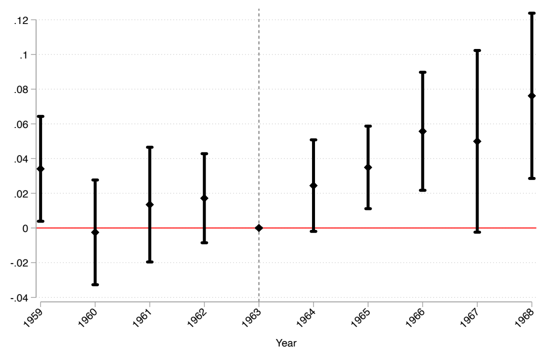
Mechanisms. Figure 1.10 shows how changes in wages and rental rates vary as a function of a location's initial specialization in agriculture. Because land is a fixed factor, and the demand for agricultural products is rising, the rental rate of land rises by more in more agriculture-intensive locations. Wages, however, move in the opposite direction, as these locations become less attractive and people move out, taking with them the consumption of non-tradable goods which are very labor intensive. Figure 1.11 shows how the composition of gross output in each state changes. In Panel (a), I plot each state's change in agriculture's share of gross output while Panel (b) plots each state's change in non-tradables' share of gross output.

[37] This is as expected. The model, which incorporates only a single mechanism occurring over a 30 year period, cannot explain population changes across all states, but can correctly replicate the population decline in the most agriculture-intensive states.

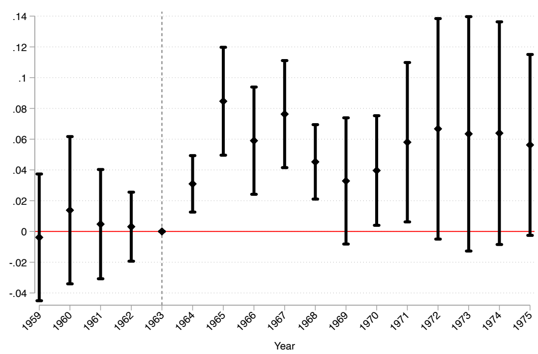
Places that were initially intensive in agriculture actually become *more* agriculture-intensive intensive, as demand for agricultural goods in these places rises. In Panel (b), we see that agriculture-intensive areas are losing non-tradable output.

Figure 1.5: Effects of Shipping Cost Changes on Flour and Bread Prices

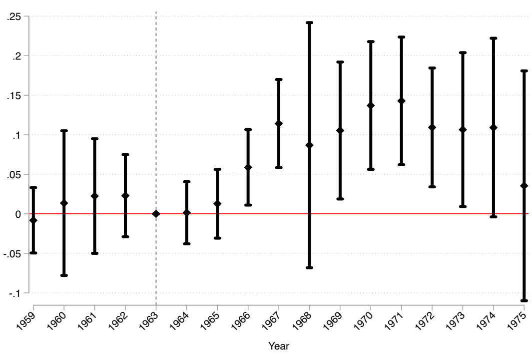
(a) Flour producer prices



(b) Flour consumer prices



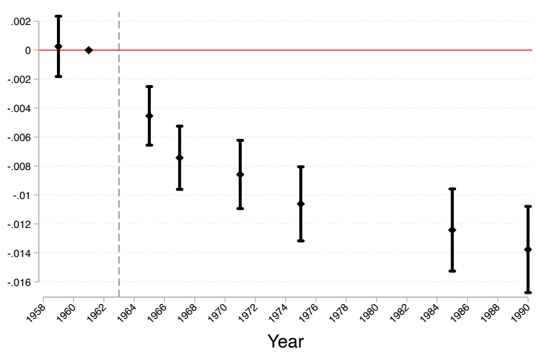
(c) Bread consumer prices



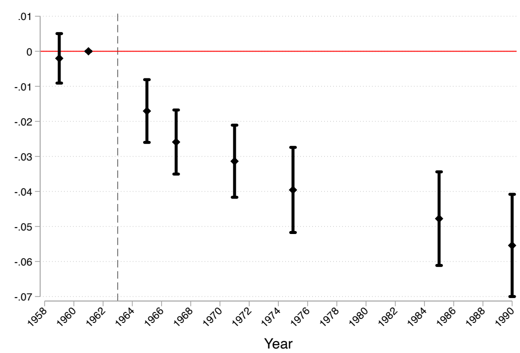
Note: These figures show event study estimates of equation 1.6. Each point estimate is the relative difference each year in prices in Midwestern cities relative to non-Midwestern cities.

Figure 1.6: Effects of Shipping Cost Changes on Flour Mills

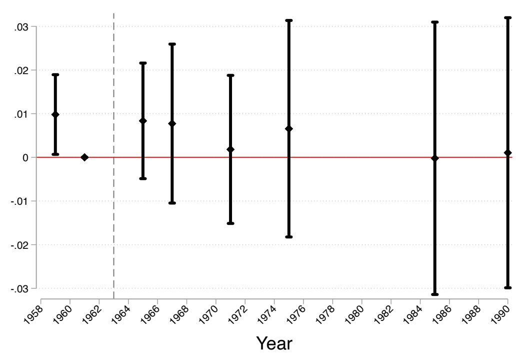
(a) Number of mills



(b) $sinh^{-1}$ (Flour milling capacity)

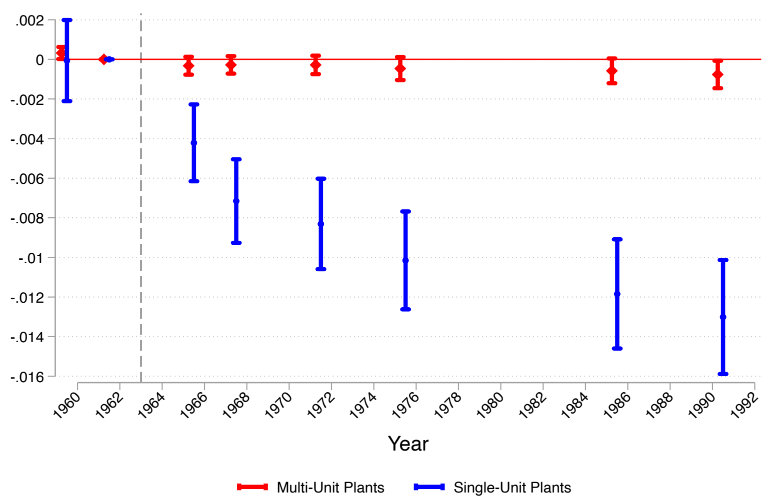


(c) log(size)



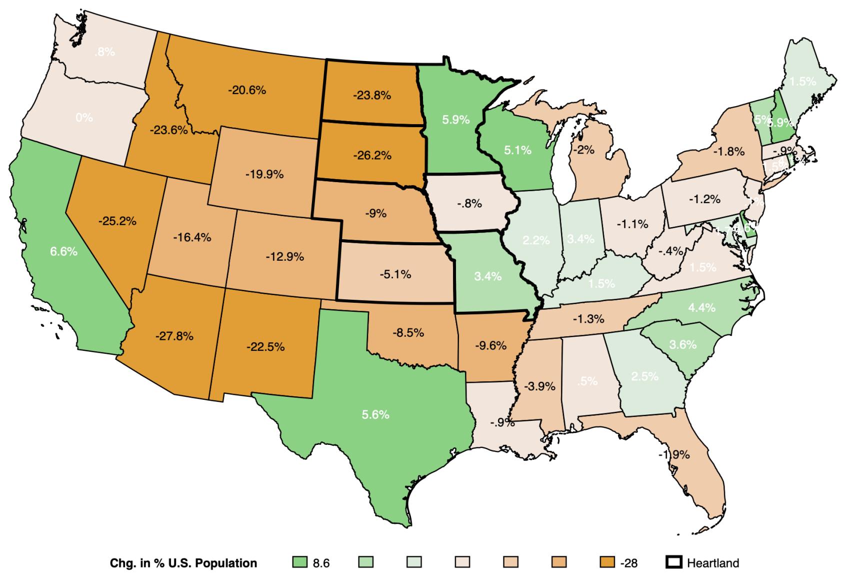
Note: These figures show event study estimates of equation 1.6. Each point estimate is the relative difference each year in the outcome variable, based on a location's initial access to wheat production.

Figure 1.7: Effects of Shipping Cost Changes on Multi- & Single-Unit Mills



Note: This figure show event study estimates of equation 1.6 where the outcome is either the number of mills that are were part of multi-unit plants (in red) and the number of mills that are part of single-unit plants (in blue) in the last year before the change in trade costs, or in their year of entry, whichever is earlier.

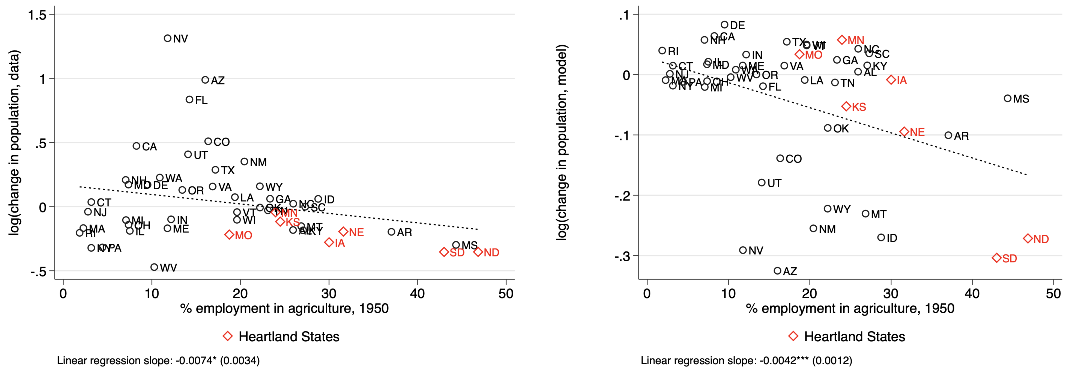
Figure 1.8: Estimated Population Changes Across States



Note: This figure shows model-predicted changes in relative population in each state, corresponding to the baseline parameterization of the model.

Figure 1.9: Population Changes and Initial Specialization in Agriculture

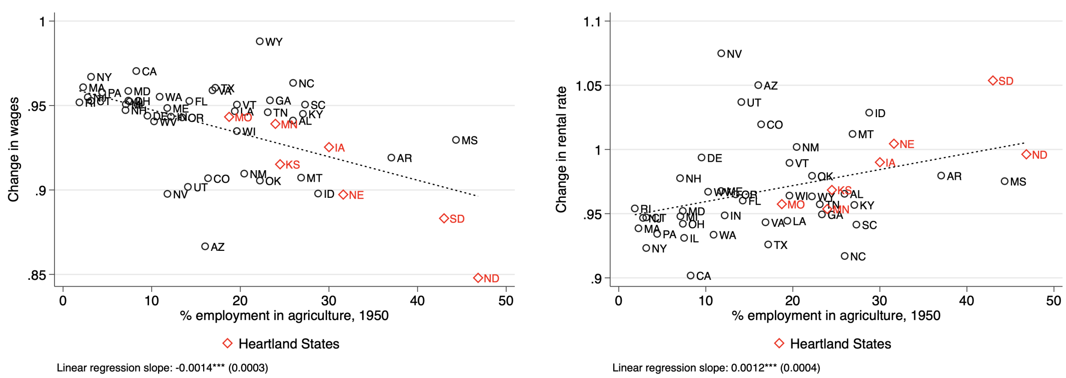
(a) Data (b) Model



Note: These figures plot, along the x-axis, the percentage of employment in agriculture in each state in 1950 and along the y-axis, the change in population as computed in the data in Panel (a) and in the model in Panel (b). The line is the best linear fit, the slope of which is reported below the figure.

Figure 1.10: Effects on Factor Prices

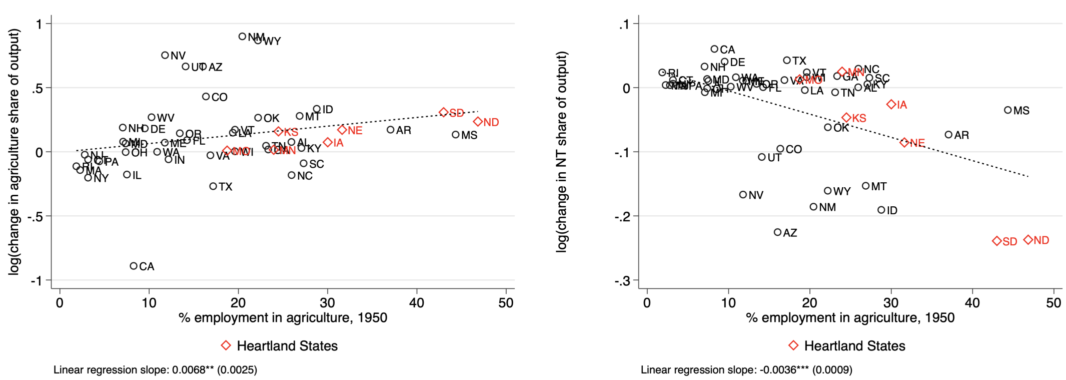
(a) Wages (b) Rental Rates



Note: These figures plot, along the x-axis, the percentage of employment in agriculture in each state in 1950 and along the y-axis, the wages in Panel (a) and in rental rates in Panel (b). The line is the best linear fit, the slope of which is reported below the figure.

Figure 1.11: Effects on Sectoral Composition of Gross Output

(a) Agriculture (b) Non-Tradables



Note: These figures plot, along the x-axis, the percentage of employment in agriculture in each state in 1950 and along the y-axis, the change in agriculture's share of gross output in Panel (a) and the change in the non-traded share of gross output in Panel (b). The line is the best linear fit, the slope of which is reported below the figure.

1.5 Conclusion

This paper studies the link between the structure of domestic trade costs across commodities and the spatial distribution of the population within countries. I use a historical setting – the American Heartland over the postwar period – to document that changes in the costs of shipping certain commodities relative to others can have substantial implications for the relative welfare of people living in different regions. I argue that rail car innovations in the U.S. over this period, which affected only bulk commodities, drove down the cost of shipping agricultural goods relative to manufactured goods and reduced the population of America's agriculture-intensive areas.

I used a simple model to explore the channels through which this could have occurred. I find that reductions in agricultural shipping costs reduce prices by more in locations farther away from where agricultural goods are produced, making other locations relatively more attractive for manufacturing production as well as for buying food. To show that this mechanism is operating in the data, I study flour mills following a sudden, significant, and exogenously generated reduction in the cost of shipping wheat versus flour. Consistent with the model's predictions, I find that this change in trade costs reduced flour and bread prices by more outside of the agricultural Heartland, and that flour milling firms entered at higher rates in locations more distant from the agricultural Heartland. Finally, I specify and calibrate a model of trade between U.S. states in 1950 and find that these changes in trade costs played a considerable role in explaining the population decline of the Heartland.

My findings suggest that how tariffs differ across commodities, and in particular how they differ between upstream and downstream goods, plays an important role in shaping the long run spatial distribution of economic activity. This was certainly the case in postwar America. As the world becomes increasingly connected via supply chains and policymakers differently apply tariffs on upstream and downstream goods, the distribution of economic activity may change substantially.

1.6 Tables and Figures

1.6.1 Tables

Table 1.3: Number of Mills Robustness Checks: Controls & Exposure Measure

	Dependent variable: Number of Flour Mills						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$1(t > 1963) \times log(WMA_i), baseline$	-0.00969***	-0.00310*	-0.00469***	-0.00999***	-0.00962***		
	(0.00112)	(0.00135)	(0.00114)	(0.00115)	(0.00110)		
$1(t > 1963) \times log(WMA_i), GAEZ$						-0.00393**	
						(0.00127)	
$1(t > 1963) \times log(WMA_i), yield$							-0.00947***
							(0.00109)
N	24856	24856	24856	24856	24848	24856	24856
County FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
State × year	✗	✓	✗	✗	✗	✗	✗
County Time Trends	✗	✗	✓	✗	✗	✗	✗
Population	✗	✗	✗	✓	✗	✗	✗
Distance to coast	✗	✗	✗	✗	✓	✗	✗

Note: This table shows robustness checks corresponding to estimating Equation 1.6 with a post indicator for all years after the trade cost shock in 1963. Column (1) shows the baseline result. Column (2) adds state by year fixed effects. Column (3) allows each county to follow its own time trend. Column (4) includes the baseline population level interacted with an indicator variable for each year. Column (5) does the same with the distance of that county to the coastline. Column (6) uses a different measure of exposure; instead of using wheat production when computing WMA_i, I use wheat yields. Column (7) uses GAEZ's wheat suitability index as the measure of wheat productivity when computing WMA_i.

Table 1.4: Milling Capacity Robustness Checks: Controls & Exposure Measure

	Dependent variable: $sinh^{-1}$ (Flour milling capacity)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$1(t > 1963) \times log(WMA_i), baseline$	-0.0352***	-0.0111^{+}	-0.0139**	-0.0351***	-0.0352***		
	(0.00519)	(0.00644)	(0.00534)	(0.00519)	(0.00517)		
$1(t > 1963) \times log(WMA_i), GAEZ$						-0.0121*	
						(0.00589)	
$1(t > 1963) \times log(WMA_i), yield$							-0.0344***
							(0.00514)
N	24856	24856	24856	24856	24848	24856	24856
County FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
State × year	✗	✓	✗	✗	✗	✗	✗
County Time Trends	✗	✗	✓	✗	✗	✗	✗
Population	✗	✗	✗	✓	✗	✗	✗
Distance to coast	✗	✗	✗	✗	✓	✗	✗

Note: This table shows robustness checks corresponding to estimating Equation 1.6 with a post indicator for all years after the trade cost shock in 1963. Column (1) shows the baseline result. Column (2) adds state by year fixed effects. Column (3) allows each county to follow its own time trend. Column (4) includes the baseline population level interacted with an indicator variable for each year. Column (5) does the same with the distance of that county to the coastline. Column (6) uses a different measure of exposure; instead of using wheat production when computing WMA_i, I use wheat yields. Column (7) uses GAEZ's wheat suitability index as the measure of wheat productivity when computing WMA_i.

Table 1.5: Robustness Checks: Functional Form

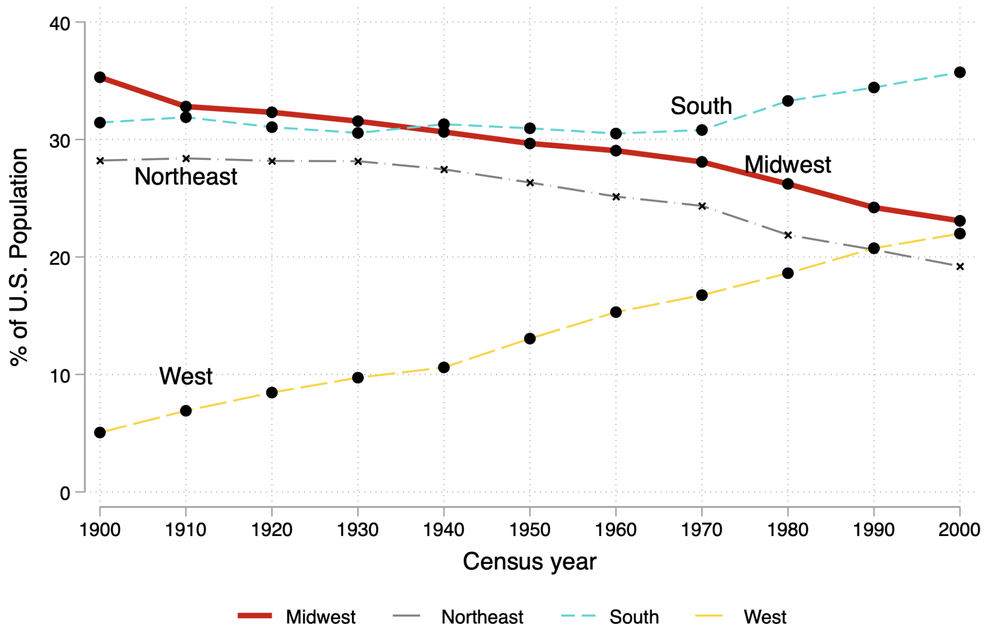
	Number of Flour Mills				Wheat Flour Capacity		
	(1) OLS	(2) PPML	(3) OLS, $1(n \geq 0)$	(4) Logit, $1(n \geq 0)$	(5) OLS, sin^{-1}	(6) OLS	(7) PPML
$1(t > 1963) \times log(WMA_i), baseline$	-0.00969***	-0.0199	-0.00578***	-0.152***	-0.0352***	-3.524	-0.0195
	(0.00112)	(0.0140)	(0.000786)	(0.0366)	(0.00519)	(2.859)	(0.0187)
N	24856	3800	24856	2800	24856	24856	3800
County FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓

Note: This table shows robustness checks corresponding to estimating Equation 1.6 with a post indicator for all years after the trade cost shock in 1963.

s

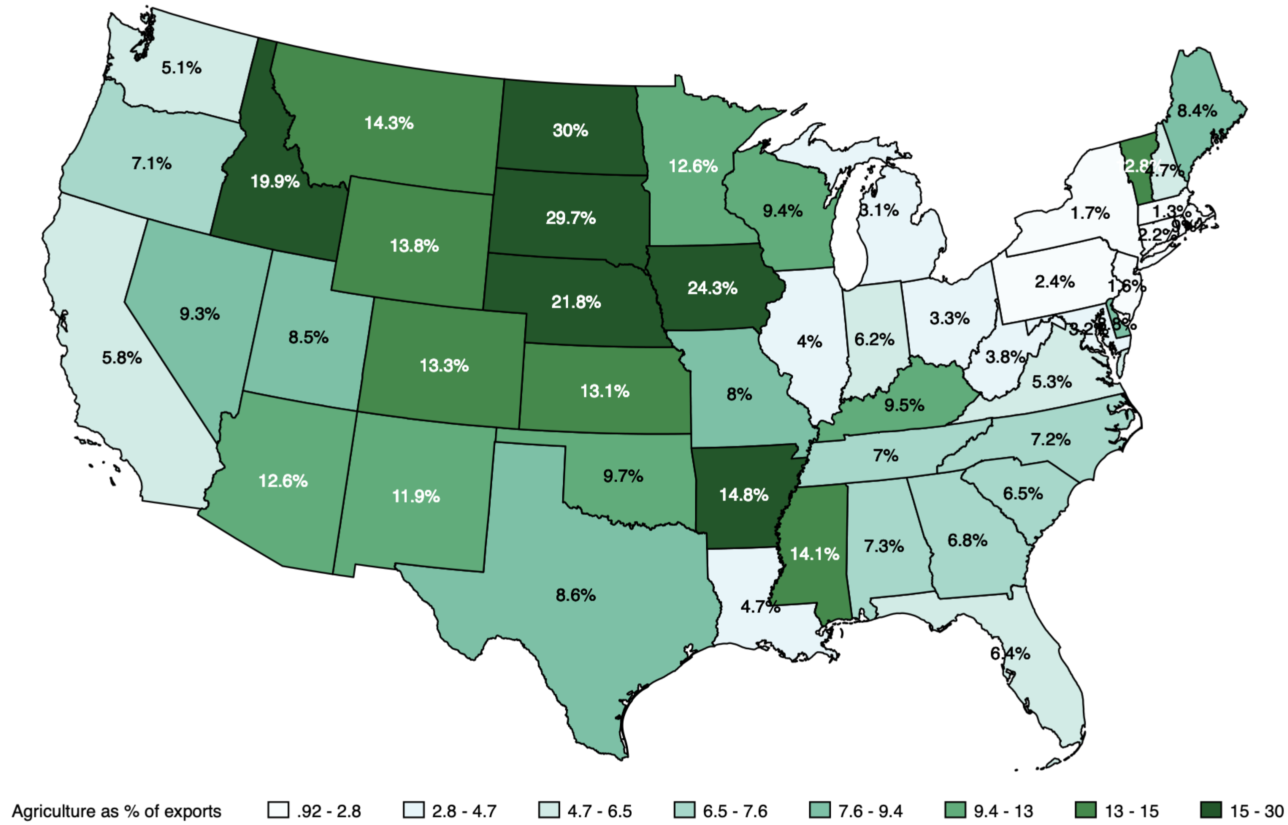
1.6.2 Figures

Figure 1.12: Distribution of U.S. Population Across Regions, Since 1900



Note: This figure shows the percentage of the U.S. population living in each region in each year. Within a year, the sum of values across the four regions is 100. Regions are defined based on U.S. Census designations. *Source*: Decennial U.S. Census.

Figure 1.13: Agricultural Goods as a Percentage of Gross Output, by State (1950)

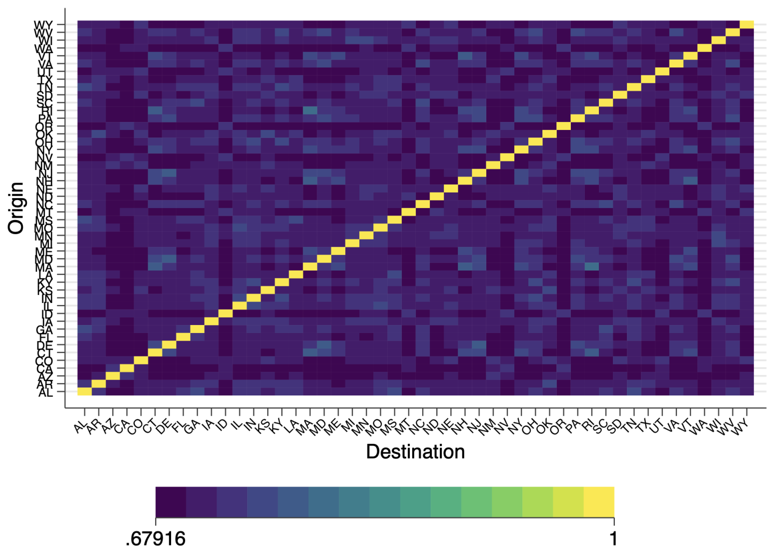


Note: This figure shows percentage of each state's gross output from the agricultural sector in the initial period.
Source: USDA (1949), Census of Manufacturers (1947), Census of Mining (1954) and author's calculations.

Figure 1.14: Annual Revenue Per Ton Earned, Motor Carriers

Note: This figure shows annual revenue per ton earned by Class I motor carriers from agricultural goods and manufactured goods. *Source:* Interstate Commerce Commission's *Freight Commodity Statistics*.

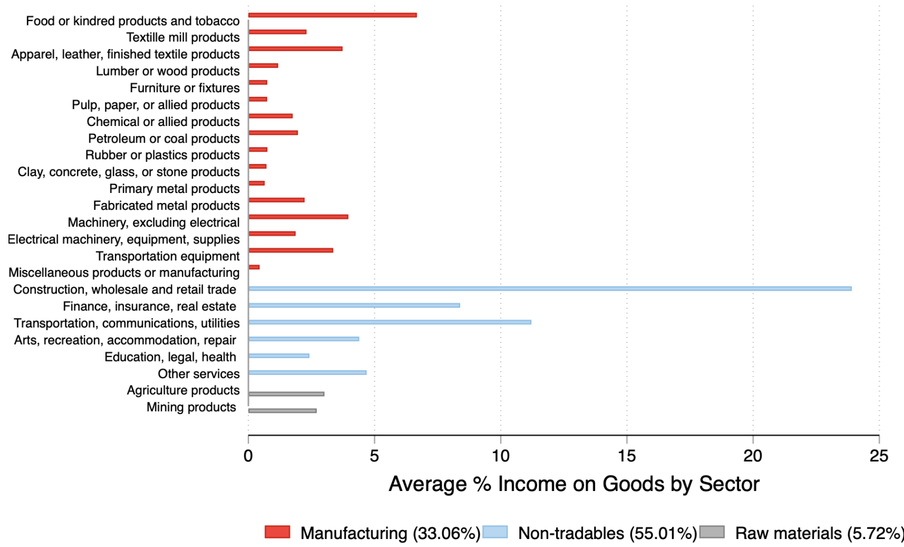
Figure 1.15: Changes in Trade Costs Across States



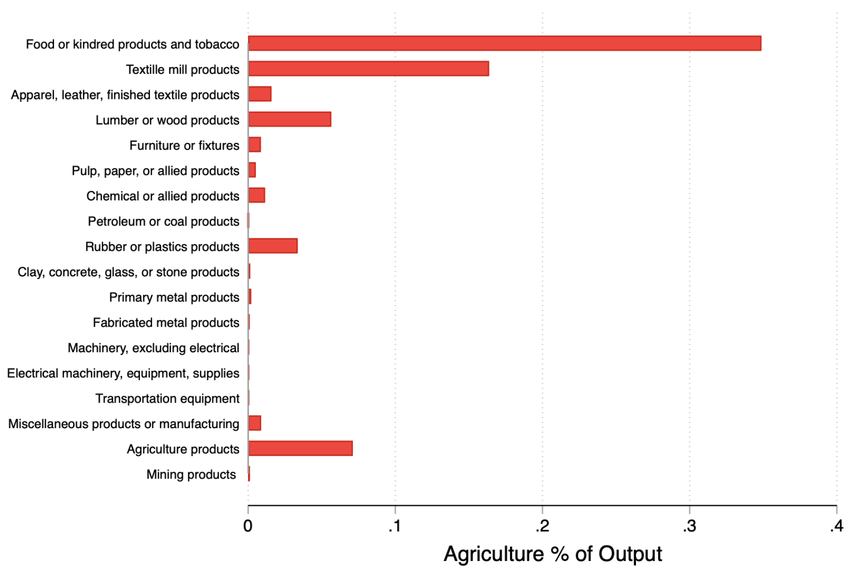
Note: This figure shows changes in trade costs between each state as fed into the model, computed from equation 1.15. *Source:* Carload Waybill Sample Statistics and author's calculations.

Figure 1.16: Key Model Inputs

(a) Sectoral Expenditure Shares



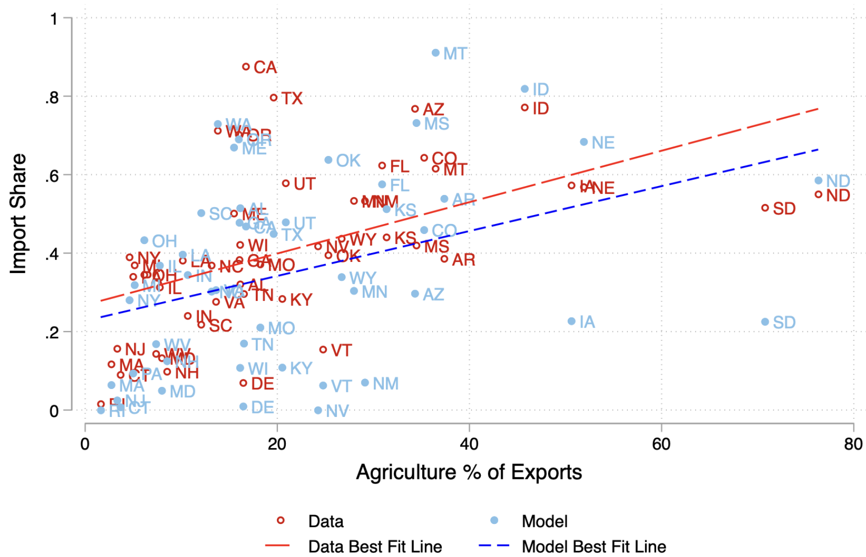
(b) Agriculture Share of Total Output by Sector



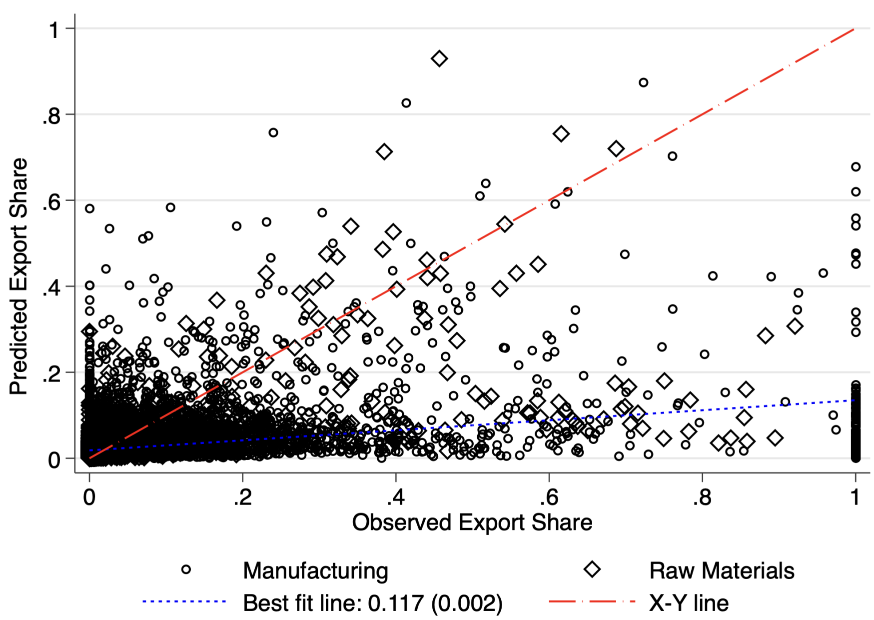
Note: Panel (a) shows the median percentage of final expenditure spent on goods from each sector across all 48 states, as computed from equation 1.12 and data as described in Section 1.4. Panel (b) shows the agricultural sector's share of final output for each sector, as computed from the 1949 BLS input-output table. *Source*: USDA, Census of Manufacturers (1947), Census of Mining (1954) and author's calculations.

Figure 1.17: Agriculture Specialization and Own-Import Shares

(a) Predicted and Observed Own-Import Shares



(b) Predicted and Observed Export Shares



Note: Panel (a) shows the correlation between each state's own import shares for agricultural goods and that state's agricultural percentage of exports in 1950. Panel (b) shows the relationship between export shares are observed in the data (on the x-axis) and the predicted export shares that I use in the calibration of the data. *Source*: Carload Waybill Sample Statistics (1949), USDA, Census of Manufacturers (1947), Census of Mining (1954) and author's calculations.

1.7 Data Appendix

1.7.1 Trade Data

I compile data on trade separately between U.S. states and regions from the Interstate Commerce Commission. The advantage to the region level data is that it more complete, while some state to state pairs are omitted if trade flows are too low. For the quantitative model, I use data on state to state trade flows. In addition to these data, I obtain aggregate data on railroad revenue earned and tons shipped in each year from the Freight Commodity Statistics, which cover Class I railroads.[38] The sampling process includes selecting all waybills, which are contracts between producers and shippers, with numbers ending in 1. This sampling procedure is supposed to result in an unbiased, representative sample of one percent of total traffic, according to the Interstate Commerce Commission.

1.7.2 Railroad Network

I obtained high resolution images of a 1957 Army Corps of Engineers Railroad Map of the United States from Stanford University Libraries, which was then hand-digitized onto a Lambert Conic projection. I use this to measure railroad distances between the centroids of each county as in Section 1.3 and each state as in Section 1.4.

[38] Class I railroads are the largest railroad companies; generally, it includes all railroad companies with annual revenues exceeding a given threshold.

Figure 1.18: Example of Carload Waybill Sample Statistics, State-to-State Trade Data

Traffic Data Averages for Specified Units of Carload Traffic with Supporting Detail, by Commodity Class,
Computed from audited Waybills Representing 1 Percent of Carload Traffic Terminated by Class I Railways during the year 1956

State-to-State Movement	Carloads	Tons (2000 lbs.)	Revenue (Dollars)	Short-line ton-miles	Short-line car-miles	Avg. tons per car	Avg. Short-line Haul per ton	Avg. Short-line Haul per car	Average revenue per 100 lbs. ¢	Average revenue per car $	Average revenue per S.L. car mile ¢	Average revenue per S.L. ton mile ¢
940 MAN AND MISC												
MIC TO IDA	1	23	1505	42400	1845	23	1843	1845	327	1505	82	3.55
MIC TO ILL	354	10568	99861	3421900	112062	30	324	317	47	282	89	2.92
MIC TO IND	217	6786	60678	1885700	66266	31	278	305	45	280	92	3.22
MIC TO IOW	26	820	10625	408800	13001	32	499	500	65	409	82	2.60
MIC TO KAN	71	1492	38238	1018600	48198	21	683	679	128	539	79	3.75
MIC TO KY	98	2548	31239	918900	34750	26	361	355	61	319	90	3.40
MIC TO LA	25	416	11488	434400	25986	17	1044	1039	138	460	44	2.64
MIC TO MAI	13	311	7115	304900	13052	24	980	1004	114	547	55	2.33
MIC TO MD	69	1931	28409	1219000	43806	28	631	635	74	412	65	2.33
MIC TO MAS	184	3339	69237	2485200	135652	18	744	737	104	376	51	2.79
MIC TO MIC	1246	46866	147146	3528800	94112	38	75	76	16	118	156	4.17
MIC TO MIN	88	1825	32460	944000	47608	21	517	541	89	369	68	3.44
MIC TO MIS	8	125	3844	114700	7235	16	918	904	154	481	53	3.35
MIC TO MO	246	5054	103370	3048500	147227	21	603	598	102	420	70	3.39
MIC TO MON	6	127	5732	188100	9297	21	1481	1550	226	955	62	3.05
MIC TO NEB	7	130	2742	98200	5315	19	755	759	105	392	52	2.79
MIC TO NEV	1	9	602	19300	2142	9	2144	2142	334	602	28	3.12
MIC TO N H	6	153	2067	118500	4603	26	775	767	68	345	45	1.74
MIC TO N J	409	8112	164679	5435300	268194	20	670	656	102	403	61	3.03
MIC TO N M	4	58	3103	83600	5688	15	1441	1422	268	776	55	3.71
MIC TO N Y	491	10307	161712	4952000	254685	21	480	519	78	329	63	3.27
MIC TO N C	31	628	13787	501900	24953	20	799	805	110	445	55	2.75
MIC TO N D	4	55	1886	43200	3140	14	785	785	171	472	60	4.37
MIC TO OHI	452	16533	112966	3443100	97465	37	208	216	34	250	116	3.28

Note: This figure shows a sample of the state to state trade data that I used to estimate changes in trade costs over time, as well as to calibrate the quantitative model in Section 1.2. *Source: Carload Waybill Sample Statistics.*

Figure 1.19: Sample of 1957 Army Corps of Engineers Railroad Map

Note: This shows the lower right hand corner of the 1957 railroad map that I have digitized and is used throughout the paper.

1.7.3 Flour Data

1.7.3.1 Mill Locations

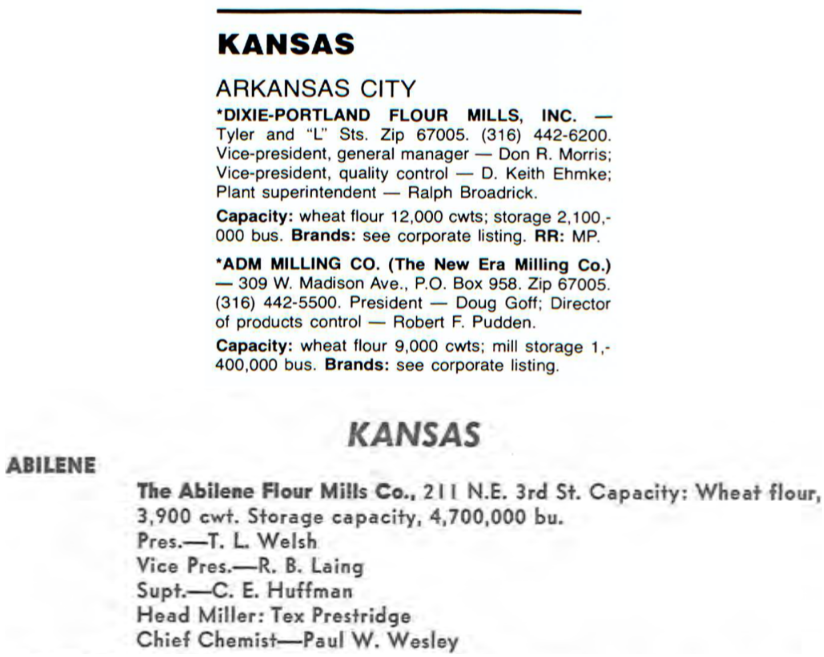
I collected data on the location, capacity, and ownership of flour mills from the *Northwestern Miller*'s annual Directory of U.S. Flour mills, which was compiled and shipped to their subscribers. Figure 1.20 gives an example of what these directories look like for two separate years (1985 and 1961 respectively). Figure 1.6 shows summary statistics describing the key variables obtained from these directories. Over this period, the number of mills falls roughly in half but total capacity of all mills rises, due to an increasing average size of the remaining mills. To construct a county-level panel of mills, and to assign a wheat market access term to each mill, I use the listed location information and the OpenCageGeocode API in Python to match plants to latitude-longitude coordinates. I then map these coordinates to modern-day county FIPS codes. I identify unique plants over time based on the

Table 1.6: Flour Mill Summary Statistics

Year	# Mills	Total Capacity	Avg. Size	Top 5 % Capacity	Top 15 % Capacity	% Big
1959	580	1072.15	1849	25.65	65.52	25.34
1961	553	1021.30	1847	23.64	60.81	25.32
1965	426	953.88	2239	22.46	60.80	25.12
1967	360	951.71	2644	21.90	60.48	28.33
1971	325	957.02	2945	20.86	59.12	28.00
1975	267	951.78	3565	19.52	59.54	34.46
1985	228	1102.62	4836	17.73	55.61	38.60
1990	207	1198.12	5788	17.49	54.09	43.00

Source: *The Northwestern Miller.* Note: Top 5 are the top 5 producers of wheat in 1963: Kansas, North Dakota, Texas, Oklahoma, and Washington which together accounted for 48% of total U.S. wheat production. Top 15 are the top 15 producers of wheat in 1963, which accounted for 75% of total wheat production. Total capacity is in 1,000 of cwt per day. Big indicates plants that were part of a multi-unit firm in their first year of entry

Figure 1.20: Milling Directory Examples



KANSAS

ARKANSAS CITY

***DIXIE-PORTLAND FLOUR MILLS, INC. —** Tyler and "L" Sts. Zip 67005. (316) 442-6200. Vice-president, general manager — Don R. Morris; Vice-president, quality control — D. Keith Ehmke; Plant superintendent — Ralph Broadrick.

Capacity: wheat flour 12,000 cwts; storage 2,100,-000 bus. **Brands:** see corporate listing. **RR:** MP.

***ADM MILLING CO. (The New Era Milling Co.)** — 309 W. Madison Ave., P.O. Box 958. Zip 67005. (316) 442-5500. President — Doug Goff; Director of products control — Robert F. Pudden.

Capacity: wheat flour 9,000 cwts; mill storage 1,-400,000 bus. **Brands:** see corporate listing.

KANSAS

ABILENE

The Abilene Flour Mills Co., 211 N.E. 3rd St. Capacity: Wheat flour, 3,900 cwt. Storage capacity, 4,700,000 bu.
Pres.—T. L. Welsh
Vice Pres.—R. B. Laing
Supt.—C. E. Huffman
Head Miller: Tex Prestridge
Chief Chemist—Paul W. Wesley

following algorithm. First, I sort the data by state, city plant name, and year. Then I check the following steps:

1. If the state, city, and name fields are the same over time and there are no duplicates of the year given the state-county-name, then this is a uniquely identified plant.

2. Suppose the state and city are the same, only one plant is listed, and there are no duplicates of the year given a state-city, but the plant name changes across years. Then I check whether there is a secondary owner listed that explains the name change. For example, if in 1971 the firm is listed as "Alabama Flour Mills" and as "Conagra" in 1975, but Conagra is listed as the secondary owner

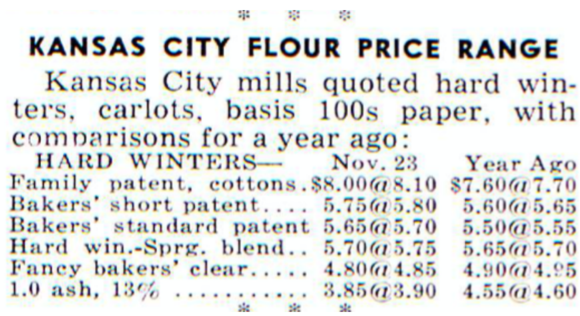
in 1971, then this is a unique plant observation (and reflects an acquisition by Conagra).

3. Suppose the state and city are the same, but many plants are listed, and names change throughout different years. If the addresses match across time, I use this to uniquely identify plants. If name changes can be connected through time with information listed on the secondary owner, then I use this to identify firms. If address information is incomplete across years, I use the firm sizes to differentiate across firms. For example, it is unlikely a plant quadrupled in size in a few years.

1.7.3.2 Prices

Producer flour prices. I collected data on flour producer prices from a few different sources. My primary source is the *Southwestern Miller*, which published weekly flour prices from a sample of flour mills in each major flour market. Because these are prices reported by mills themselves where they are produced, I consider these to be producer prices. Table 1.7 lists flour prices in 1962, the last year prior to the change, and 1966, several years after the shock for each location in the sample. I was not able to obtain price data after 1968, so that is the last year in my sample. Minneapolis flour prices are not included in later years; thus, when that series ends, I use data on flour prices from Minneapolis mills from the USDA as part of the monthly "Wheat Situation" report.

Consumer flour and bread prices. I digitized data on the annual price of wheat and flour for every city listed in the BLS's "Estimated Food Retail Prices by Cities".



KANSAS CITY FLOUR PRICE RANGE

Kansas City mills quoted hard winters, carlots, basis 100s paper, with comparisons for a year ago:

HARD WINTERS—	Nov. 23	Year Ago
Family patent, cottons	$8.00@8.10	$7.60@7.70
Bakers' short patent....	5.75@5.80	5.60@5.65
Bakers' standard patent	5.65@5.70	5.50@5.55
Hard win.-Sprg. blend..	5.70@5.75	5.65@5.70
Fancy bakers' clear.....	4.80@4.85	4.90@4.95
1.0 ash, 13%	3.85@3.90	4.55@4.60

Figure 1.21: *Southwestern Miller* Flour Price Data

While the set of cities varies slightly over time, most years include around twenty cities. Table 1.8 reports bread and flour prices for the set of cities in the sample for 1962 and 1964.

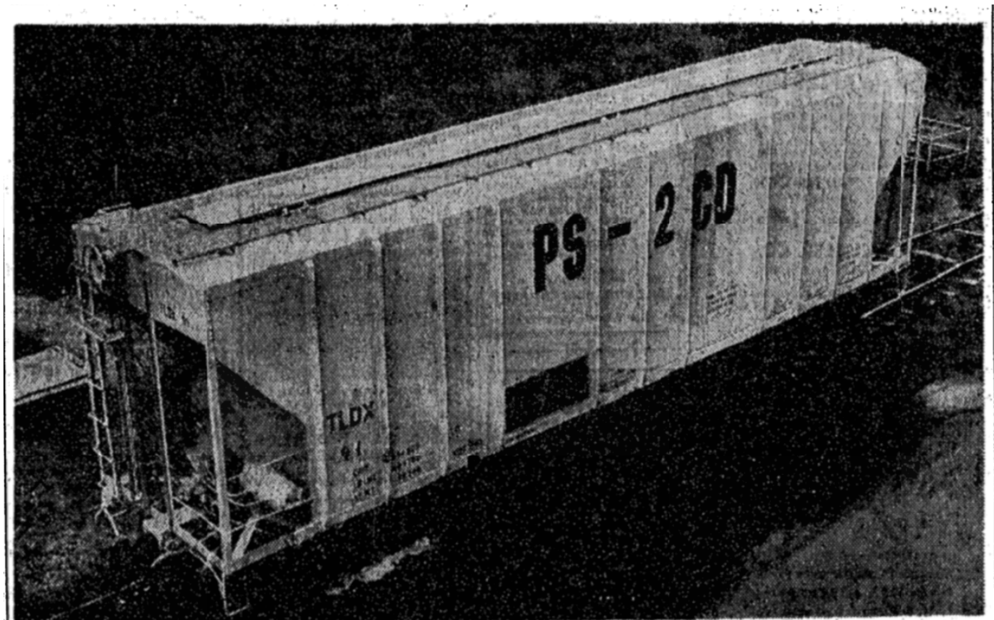
Table 1.7: Flour Producer Price Summary Statistics

State	Cities	Flour Type	Price ($) 1962	Price ($) 1966	Price Ratio
Alabama	Birmingham	Standard Patent	21.90	21.02	0.96
Illinois	Chicago	Standard Patent	20.33	20.56	1.01
Kansas	Wichita, Salina, Arkansas City, Hutchinson	Standard Patent	20.92	21.13	1.01
Minnesota	Minneapolis	Standard Patent	21.82	22.48	1.03
Missouri	Kansas City	Standard Patent	20.92	21.13	1.01
Nebraska	Omaha	Standard Patent	19.42	20.09	1.03
New York	New York	Standard Patent	23.37	22.85	0.98
Oregon	Portland	Family	27.10	28.49	1.05
Pennsylvania	Pittsburgh	Standard Patent	21.98	19.66	0.89

Note: This table shows summary statistics of flour producer prices for each location in which I observe prices. This price data is used in Section 1.3 to estimate the effects of changes in upstream trade costs on downstream prices. *Source: Southwestern Miller* and BLS.

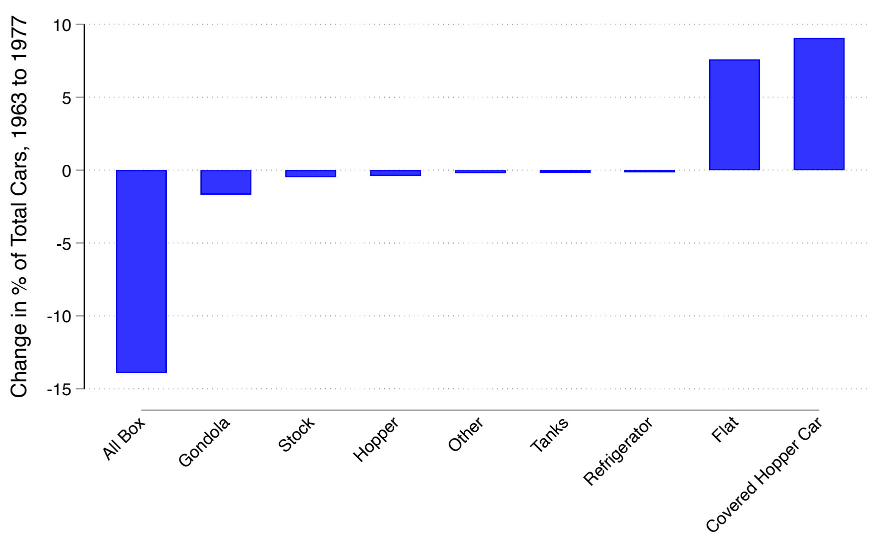
Figure 1.22: The Decline of the Box Car

(a) Covered Hopper Car



IN DEMAND: Covered hopper railroad car developed by Pullman, Inc., that is replacing boxcars being retired

(b) Changes in Rail Car Types, 1963-1977



Source: Bureau of Transport Economics and Statistics, Interstate Commerce Commission, 1979

Note: Panel (a) shows an image of a covered hopper car for transporting grain, from the New York Times (1964). Panel (b) shows the change in percentage of total cars represented by each rail car type, from the ICC (1979).

Table 1.8: Flour Consumer Price Summary Statistics

State	City	Flour Price ($) 1962	1964	Ratio	Bread Price ($) 1962	1964	Ratio
Georgia	Atlanta	1.92	1.90	0.99	0.64	0.62	0.97
Maryland	Baltimore	1.83	1.88	1.03	0.68	0.75	1.10
Massachusetts	Boston	1.83	1.87	1.02	0.70	0.69	0.99
Illinois	Chicago	1.75	1.70	0.97	0.69	0.62	0.90
Ohio	Cincinnati	1.75	1.67	0.95	0.64	0.63	0.98
Ohio	Cleveland	1.78	1.69	0.95	0.67	0.68	1.01
Michigan	Detroit	1.70	1.68	0.99	0.61	0.56	0.92
Texas	Houston	1.87	1.87	1.00	0.55	0.57	1.04
Kansas	Kansas City	1.67	1.72	1.03	0.67	0.66	0.99
California	Los Angeles	2.02	1.85	0.92	0.88	0.92	1.05
Minnesota	Minneapolis	1.84	1.82	0.99	0.62	0.59	0.95
New York	New York	1.79	1.81	1.01	0.79	0.80	1.01
Pennsylvania	Philadelphia	1.86	1.77	0.95	0.74	0.72	0.97
Pennsylvania	Pittsburg	1.81	1.73	0.96	0.71	0.68	0.96
California	San Francisco	2.11	2.04	0.97	0.84	0.88	1.05
Washington	Seattle	2.12	2.07	0.98	0.80	0.80	1.00
Missouri	St Louis	1.78	1.79	1.01	0.65	0.63	0.97
District Of Columbia	Washington	1.92	1.95	1.02	0.67	0.65	0.97

Note: This table shows summary statistics of flour and bread consumer prices for each location in which I observe prices. This price data is used in Section 1.3 to estimate the effects of changes in upstream trade costs on downstream prices. Flour prices are for five pounds of wheat flour and bread prices are for one pound of bread. All prices are in 1985 dollars. *Source:* BLS, Estimated Retail Prices of Food in cities.

1.8 Model Appendix

1.8.1 Model Setup

1.8.1.1 Flour Demand

Each agent chooses c_{ni}^F, c_{ni}^M to solve:

$$U(c_i^F, c_i^M) = c_i^M + \ln\left(c_i^F\right)$$

where:

$$c_i^F = \left[\sum_n (c_{ni}^F)^{\frac{\sigma_F - 1}{\sigma_F}}\right]^{\frac{\sigma_F}{\sigma_F - 1}} \text{ and } c_i^M = \left[\sum_n c_{ni}^M\right]$$

subject to a budget constraint of

$$w_i = \sum_n p_{ni}^M c_{ni}^M + p_{ni}^F c_{ni}^F$$

From this maximization problem we have demand functions for each variety of flour:

$$c_{ji}^F = p^M \left(p_{ji}^F\right)^{-\sigma_F} \frac{\sigma_F}{\sigma_F - 1} \left(P_i^F\right)^{\sigma_F - 1}$$

Then the total amount of spending on flour by state i is:

$$\sum_j p_{ji}^F c_{ji}^F = \sum_j p^M \left(p_{ji}^F\right)^{1-\sigma_F} \frac{\sigma_F}{\sigma_F - 1} \left(P_i^F\right)^{\sigma_F - 1} = p^M \frac{\sigma_F}{\sigma_F - 1}$$

This yields flour import shares – the share of flour imported to state i by state j – of:

$$\pi_{ji}^F = \frac{p_{ji}^F c_{ji}^F}{\sum_j p_{ji}^F c_{ji}^F} = \left(p_{ji}^F\right)^{1-\sigma_F} \left(P_i^F\right)^{\sigma_F - 1}$$

1.8.1.2 Indirect Utility

Next we want to find the indirect utility function in order to compute welfare. First sum over origins to find aggregate flour consumption:

$$c_i^F = \left[\sum_n \left(p^M \left(p_{ji}^F\right)^{-\sigma_F} \frac{\sigma_F}{\sigma_F - 1} \left(P_i^F\right)^{\sigma_F - 1}\right)^{\frac{\sigma_F - 1}{\sigma_F}}\right]^{\frac{\sigma_F}{\sigma_F - 1}} = \frac{p^M}{P_i^F} \frac{\sigma_F}{\sigma_F - 1}$$

Then to find c_i^M :

$$w_i = p^M c_i^M + \sum_n p_{ni}^F p^M \left(p_{ni}^F\right)^{-\sigma_F} \frac{\sigma_F}{\sigma_F - 1} \left(P_i^F\right)^{\sigma_F - 1}$$

$$c_i^M = \frac{w_i}{p^M} - \frac{\sigma_F}{\sigma_F - 1} \left(P_i^F\right)^{\sigma_F - 1} \sum_n \left(p_{ni}^F\right)^{1-\sigma_F} = \frac{w_i}{p^M} - \frac{\sigma_F}{\sigma_F - 1} \left(P_i^F\right)^{\sigma_F - 1} \left(P_i^F\right)^{1-\sigma_F} = \frac{w_i}{p^M} - \frac{\sigma_F}{(\sigma_F - 1)} P_i^F$$

Then, the indirect utility function for a single agents is given by (assuming income is sufficiently large):

$$V(P_i^F, I_i) = w_i + \ln\left(\frac{\sigma_F}{\sigma_F - 1}\right) - \ln\left(P_i^F\right) - \frac{\sigma_F}{\sigma_F - 1} P_i^F \tag{1.17}$$

1.8.1.3 Labor Mobility

Given the above indirect utility common across all agents in a location, and the assumption that indirect utility of a worker l in state i is $v_i^l = v_i + \epsilon_i^l$:

$$v_i^l = \left(w_i + \ln\left(\frac{\sigma_F}{\sigma_F - 1}\right) - \ln\left(P_i^F\right) - \frac{\sigma_F}{\sigma_F - 1} P_i^F \right) + \epsilon_n^l$$

Agents choose to live in the state that gives them the largest indirect utility. Assuming that $\epsilon_n^l \sim Gumbell$, the probability that an agent i chooses a state n is given by (or the share of agents living in state n):

$$\lambda_i = pr\left(v_i^l \geq \max_{n' \neq n} v_{n'}^l\right) = \frac{exp(v_i)}{\sum_{n'} exp(v_{n'})} = \frac{exp\left(w_i + \ln\left(\frac{\sigma_F}{\sigma_F - 1}\right) - \ln\left(P_i^F\right) - \frac{\sigma_F}{\sigma_F-1} P_{n'}^F\right)}{\sum_n' exp\left(w_{n'} + \ln\left(\frac{\sigma_F}{\sigma_F-1}\right) - \ln\left(P_{n'}^F\right) - \frac{\sigma_F}{\sigma_F-1} P_{n'}^F\right)}$$

Given an exogenous level of the country's population size L, the number of people living in state i is: $L_i = \lambda_i L$

1.8.1.4 Flour Millers' Problem

Flour millers in state i choose the amount of wheat to import from each state j in order to minimize costs subject to a level of production:

$$\min_{c_{ni}^W} \sum_n p_{ni}^W c_{ni}^W \text{ s.t. } Y_{it}^F = T_{it}^F \left[\sum_n \left(c_{ni}^W\right)^{\frac{\sigma_W - 1}{\sigma_W}}\right]^{\frac{\sigma_W}{\sigma_W - 1}}$$

Then, the first order conditions are:

$$[c_{ni}^W] : p_{ni}^W - \lambda_{it} T_{it}^F \left[\sum_n \left(c_{ni}^W \right)^{\frac{\sigma_W - 1}{\sigma_W}} \right]^{\frac{1}{\sigma_W - 1}} \left(c_{ni}^W \right)^{-1/\sigma_W} = 0$$

Taking the ratio of first order conditions for different origins (holding i fixed):

$$\frac{p_{ni}^W}{p_{it}^W} = \left(\frac{c_{ni}^W}{c_{ji}^W} \right)^{-1/\sigma_W}$$

Solving for c_{ni}^W:

$$\left(\frac{p_{ni}^W}{p_{ji}^W} \right)^{-\sigma_W} c_{ji}^W = c_{ni}^W$$

Substituting this into the constraint:

$$\begin{aligned} Y_i^F = T_i^F \left[\sum_n \left(c_{ni}^W \right)^{\frac{\sigma_W - 1}{\sigma_W}} \right]^{\frac{\sigma_W}{\sigma_W - 1}} &= T_i^F \left[\sum\nolimits_n \left(\frac{p_{ni}^W}{p_{ji}^W} \right)^{1-\sigma_W} \left(c_{ji}^W \right)^{\frac{\sigma_W - 1}{\sigma_W}} \right]^{\frac{\sigma_W}{\sigma_W - 1}} \\ &= T_i^F \left(p_{ji}^W \right)^{\sigma_W} c_{ji}^W \left[\sum\nolimits_n \left(p_{ni}^W \right)^{1-\sigma_W} \right]^{\frac{\sigma_W}{\sigma_W - 1}} \end{aligned}$$

Then, solving for consumption of wheat from state j in state i:

$$c_{ji}^W = \frac{Y_i^F}{T_i^F} \left(p_{ji}^W \right)^{-\sigma_W} \left(P_i^W \right)^{\sigma_W}$$

which means that total spending on wheat in state i is:

$$\sum_j p_{ji}^W c_{ji}^W = \frac{Y_i^F}{T_i^F} \left(P_i^W \right)^{\sigma_W} \sum_j \left(p_{ji}^W \right)^{1-\sigma_W} = \frac{Y_i^F}{T_i^F} P_i^W$$

Then the share of wheat imported from state j by state i, π_{ji}^W is:

$$\pi_{ji}^W = \frac{p_{ji}^W c_{ji}^W}{\frac{Y_i^F}{T_i^F} P_i^W} = \frac{\frac{Y_i^F}{T_i^F}\left(p_{ji}^W\right)^{1-\sigma_W}\left(P_i^W\right)^{\sigma_W}}{\frac{Y_i^F}{T_i^F} P_i^W} = \left(p_{ji}^W\right)^{1-\sigma_W}\left(P_i^W\right)^{\sigma_W - 1}$$

1.8.2 Proofs of Testable Predictions

1.8.2.1 Main Theorems

If $\pi_{NN}^W < \pi_{KK}^W$ and $\hat{\tau}^W < 1$, then the following theorems hold.

Theorem 1 (Flour producer price effect)**.** $\hat{p}_{NN}^F < \hat{p}_{KK}^F$

Proof.

$$\hat{p}_{NN}^F < \hat{p}_{KK}^F \iff \frac{\hat{P}_N^W}{\hat{T}_N^F} < \frac{\hat{P}_K^W}{\hat{T}_K^F} \iff \hat{P}_N^W < \hat{P}_K^W$$

Then note that:

$$\hat{\tau}^W < 1 \iff \left(\hat{\tau}^W\right)^{1-\sigma_W} > 1 \iff \pi_{NN}^W - \pi_{KK}^W > \left(\hat{\tau}^W\right)^{1-\sigma_W}\left(\pi_{NN}^W - \pi_{KK}^W\right)$$

since $\pi_{NN}^W - \pi_{KK}^W < 0$.

$$\iff \pi_{NN}^W - \pi_{KK}^W + \left(\hat{\tau}^W\right)^{1-\sigma_W} > \left(\hat{\tau}^W\right)^{1-\sigma_W}\pi_{NN}^W - \left(\hat{\tau}^W\right)^{1-\sigma_W}\pi_{KK}^W + \left(\hat{\tau}^W\right)^{1-\sigma_W}$$

$$\iff \pi_{NN}^W + \left(1 - \pi_{NN}^W\right)\left(\hat{\tau}^W\right)^{1-\sigma_W} > \pi_{KK}^W + \left(1 - \pi_{KK}^W\right)\left(\hat{\tau}^W\right)^{1-\sigma_W}$$

since $1 - \pi_{ii}^W = \pi_{ni}^W$:

$$\pi_{NN}^W + \pi_{KN}^W\left(\hat{\tau}^W\right)^{1-\sigma_W} > \pi_{KK}^W + \pi_{NK}^W\left(\hat{\tau}^W\right)^{1-\sigma_W}$$

$$\iff \left[\pi^W_{NN} + \pi^W_{KN}\left(\hat{\tau}^W\right)^{1-\sigma_W}\right]^{\frac{1}{1-\sigma_W}} < \left[\pi^W_{KK} + \pi^W_{NK}\left(\hat{\tau}^W\right)^{1-\sigma_W}\right]^{\frac{1}{1-\sigma_W}} \iff \hat{P}^W_N < \hat{P}^W_K$$

$\square$

Theorem 2 (Flour consumer price effect)**.** *If* $\pi^W_{NN} < \pi^W_{KK}$ *and* $\hat{\tau}^W < 1$*, then* $\hat{P}^F_N < \hat{P}^F_K$.

Proof. Define $a = 1 - \sigma_F < 0$. By proposition (2),

$$\pi^F_{KK} > \pi^F_{KN} \iff \pi^F_{KK} > 1 - \pi^F_{NN} \iff \pi^F_{KK} + \pi^F_{NN} > 1$$

Then, applying proposition (3), $\hat{P}^W_N < \hat{P}^W_K \leq 1 \iff \left(\hat{P}^W_N\right)^a > \left(\hat{P}^W_K\right)^a \geq 1 \iff \left(\hat{P}^W_N\right)^a - \left(\hat{P}^W_K\right)^a > 0$,

$$\iff \pi^F_{NN}\left(\left(\hat{P}^W_N\right)^a - \left(\hat{P}^W_K\right)^a\right) + \pi^F_{KK}\left(\left(\hat{P}^W_N\right)^a - \left(\hat{P}^W_K\right)^a\right) > \left(\hat{P}^W_N\right)^a - \left(\hat{P}^W_K\right)^a$$

$$\iff \pi^F_{NN}\left(\left(\hat{P}^W_N\right)^a - \left(\hat{P}^W_K\right)^a\right) - \left(\hat{P}^W_N\right)^a > \pi^F_{KK}\left(\left(\hat{P}^W_K\right)^a - \left(\hat{P}^W_N\right)^a\right) - \left(\hat{P}^W_K\right)^a$$

$$\iff \pi^F_{NN}\left(\hat{P}^W_N\right)^a + (1 - \pi^F_{NN})\left(\hat{P}^W_K\right)^a > \pi^F_{KK}\left(\hat{P}^W_K\right)^a + (1 - \pi^F_{KK})\left(\hat{P}^W_N\right)^a$$

$$\iff \pi^F_{NN}\left(\hat{P}^W_N\right)^a + \pi^F_{KN}\left(\hat{P}^W_K\right)^a > \pi^F_{KK}\left(\hat{P}^W_K\right)^a + \pi^F_{NK}\left(\hat{P}^W_N\right)^a$$

$$\left[\pi^F_{NN}\left(\hat{P}^W_N\right)^a + \pi^F_{KN}\left(\hat{P}^W_K\right)^a\right]^{\frac{1}{1-\sigma_F}} < \left[\pi^F_{KK}\left(\hat{P}^W_K\right)^a + \pi^F_{NK}\left(\hat{P}^W_N\right)^a\right]^{\frac{1}{1-\sigma_F}} \leq 1 \iff \hat{P}^F_N < \hat{P}^F_K$$

$\square$

Theorem 3 (Flour production effect)**.** $\hat{Y}_N > \hat{Y}^*_K$

Proof.

$$\hat{Y}^F_N > \hat{Y}^F_K \iff \lambda_{NN}\hat{c}^F_{NN} + (1 - \lambda_{NN})\hat{c}^F_{NK} > \lambda_{KK}\hat{c}^F_{KK} + (1 - \lambda_{KK})\hat{c}^F_{KN}$$

where $\lambda_{ni} = \frac{c_{ni}^F}{c_{ni}^F + c_{nn}^F}$ is the share of n's flour production shipped to i in the initial equilbirum.

$$\iff \lambda_{NN}(\hat{P}_N^W)^{-\sigma_F}\left(\hat{P}_N^F\right)^{\sigma_F-1} + (1-\lambda_{NN})(\hat{P}_N^W)^{-\sigma_F}\left(\hat{P}_K^F\right)^{\sigma_F-1}$$

$$> \lambda_{KK}(\hat{P}_K^W)^{-\sigma_F}\left(\hat{P}_K^F\right)^{\sigma_F-1} + (1-\lambda_{KK})(\hat{P}_K^W)^{-\sigma_F}\left(\hat{P}_N^F\right)^{\sigma_F-1}$$

From Proposition 3 we know that: $\hat{P}_N^W < \hat{P}_K^W$, and

$$\hat{P}_N^F < \hat{P}_K^F \iff \left[\pi_{NN}^F(\hat{P}_N^W)^{1-\sigma_F} + \pi_{KN}^F(\hat{P}_K^W)^{1-\sigma_F}\right]^{\frac{1}{1-\sigma_F}} < \left[\pi_{KK}^F(\hat{P}_K^W)^{1-\sigma_F} + \pi_{NK}^F(\hat{P}_N^W)^{1-\sigma_F}\right]^{\frac{1}{1-\sigma_F}}$$

so then it must be that $\hat{P}_N^W < \hat{P}_N^F < \hat{P}_K^F < \hat{P}_K^W$. We claim the following:

* $(\hat{P}_N^W)^{-\sigma_F}(\hat{P}_N^F)^{\sigma_F-1} > (\hat{P}_K^W)^{-\sigma_F}(\hat{P}_K^F)^{\sigma_F-1}$ which follows from the inequality, since $\frac{\hat{P}_N^W}{\hat{P}_N^F} < 1$ and $\frac{\hat{P}_K^W}{\hat{P}_K^F} > 1$
* $(\hat{P}_N^W)^{-\sigma_F}(\hat{P}_K^F)^{\sigma_F-1} > (\hat{P}_K^W)^{-\sigma_F}(\hat{P}_N^F)^{\sigma_F-1}$ which follows immediately.

Since our original inequality is a convex combination of these two conditions added together, we are done. □

Theorem 4 (Pro-competitive effect). $\hat{\varphi}_N^* < \hat{\varphi}_K^*$

Proof. First, write φ_i^* in changes:

$$\hat{\varphi}_i^* = \hat{P}_i^W\left(\hat{w}_i\hat{f}_e\right)^{\frac{1}{\sigma_F-1}}\left(\sum_j \lambda_{ij}(\hat{P}_j^F)^{\sigma_F-1}\left(\hat{\tau}_{ij}^F\right)^{1-\sigma_F}\right)^{\frac{1}{1-\sigma_F}} = \hat{P}_i^W\left(\sum_j \lambda_{ij}(\hat{P}_j^F)^{\sigma_F-1}\right)^{\frac{1}{1-\sigma_F}}$$

Then, proceed:

$$\hat{\varphi}_N^* < \hat{\varphi}_K^* \iff \left((\hat{P}_N^W)^{1-\sigma_F}\lambda_{NN}(\hat{P}_N^F)^{\sigma_F-1} + (\hat{P}_N^W)^{1-\sigma_F}\lambda_{NK}(\hat{P}_K^F)^{\sigma_F-1}\right)^{\frac{1}{1-\sigma_F}}$$
$$< \left((\hat{P}_K^W)^{1-\sigma_F}\lambda_{KK}(\hat{P}_K^F)^{\sigma_F-1} + (\hat{P}_K^W)^{1-\sigma_F}\lambda_{KN}(\hat{P}_N^F)^{\sigma_F-1}\right)^{\frac{1}{1-\sigma_F}} \tag{1.18}$$
$$\iff (\hat{P}_N^W)^{1-\sigma_F}\lambda_{NN}(\hat{P}_N^F)^{\sigma_F-1} + (\hat{P}_N^W)^{1-\sigma_F}(1-\lambda_{NN})(\hat{P}_K^F)^{\sigma_F-1}$$
$$> (\hat{P}_K^W)^{1-\sigma_F}\lambda_{KK}(\hat{P}_K^F)^{\sigma_F-1} + (\hat{P}_K^W)^{1-\sigma_F}(1-\lambda_{KK})(\hat{P}_N^F)^{\sigma_F-1} \tag{1.19}$$

We know that $\hat{P}_N^W < \hat{P}_K^W$, and

$$\hat{P}_N^F < \hat{P}_K^F \iff \left[\pi_{NN}^F(\hat{P}_N^W)^{1-\sigma_F} + \pi_{KN}^F(\hat{P}_K^W)^{1-\sigma_F}\right]^{\frac{1}{1-\sigma_F}} < \left[\pi_{KK}^F(\hat{P}_K^W)^{1-\sigma_F} + \pi_{NK}^F(\hat{P}_N^W)^{1-\sigma_F}\right]^{\frac{1}{1-\sigma_F}}$$

so then it must be that $\hat{P}_N^W < \hat{P}_N^F < \hat{P}_K^F < \hat{P}_K^W$. We claim the following:

* $(\hat{P}_N^W)^{1-\sigma_F}(\hat{P}_N^F)^{\sigma_F-1} > (\hat{P}_K^W)^{1-\sigma_F}(\hat{P}_K^F)^{\sigma_F-1}$ which follows from the inequality, since $\frac{\hat{P}_N^W}{\hat{P}_N^F} < 1$ and $\frac{\hat{P}_K^W}{\hat{P}_K^F} > 1$
* $(\hat{P}_N^W)^{1-\sigma_F}(\hat{P}_K^F)^{\sigma_F-1} > (\hat{P}_K^W)^{1-\sigma_F}(\hat{P}_N^F)^{\sigma_F-1}$ which follows immediately.

Since the last expression is a convex combination of these two conditions added together, we are done.

$\square$

Theorem 5 (Flour mill location effect)**.** $\hat{M}_N^* > \hat{M}_K^*$

Proof. $\hat{M}_N^* > \hat{M}_K^* \iff \hat{M}_N\frac{(1-G(\varphi'_N))}{(1-G(\varphi'_N))} > \hat{M}_K\frac{(1-G(\varphi'_K))}{(1-G(\varphi'_K))} \iff \frac{(1-G(\varphi_N\hat{\varphi}_N))}{(1-G(\varphi_N))} > \frac{(1-G(\varphi_K\hat{\varphi}_K))}{(1-G(\varphi_K))}$. If $G(.)$ is pareto, then $1-G(\varphi) = A^{\theta_i}\varphi^{\theta_i}$. Then we need to show that:

$$\iff \frac{A^{\theta_N}\varphi_N^{\theta_N}\hat{\varphi}_N^{-\theta_N}}{A^{\theta_N}\varphi_N^{-\theta_N}} > \frac{A^{\theta_K}\varphi_K^{-\theta_K}\hat{\varphi}_K^{-\theta_K}}{A^{\theta_K}\varphi_K^{-\theta_K}} \iff \hat{\varphi}_N^{-\theta_N} > \hat{\varphi}_K^{-\theta_K} \iff \hat{\varphi}_N < \hat{\varphi}_K$$

which we show in Theorem 4. □

Theorem 6 (Decline of the Heartland)**.** *$\Delta v_N > \Delta v_K$ and $\hat{L}_N > \hat{L}_K$.*

Proof, Welfare.

$$\Delta v_N > \Delta v_K \iff v_N' - v_N > v_K' - v_K \iff -\ln\left(\hat{P}_N^F\right) - \frac{\sigma_F}{\sigma_F - 1}\Delta P_N^F > -\ln\left(\hat{P}_K^F\right) - \frac{\sigma_F}{\sigma_F - 1}\Delta P_K^F$$

From Theorem 2 we know that $\hat{P}_N^F < \hat{P}_K^F \iff -\ln\left(\hat{P}_N^F\right) > -\ln\left(\hat{P}_K^F\right)$. Then $\pi_{NN}^W < \pi_{KK}^W \iff P_N^W > P_K^W \iff P_N^F > P_K^F$, $P_N^F > P_K^F \iff -P_N^F < -P_K^F$. By proposition 2, $\hat{P}_N^F < \hat{P}_K^F \iff \hat{P}_N^F - 1 < \hat{P}_K^F - 1 < 0$. Multiplying these together (since both are negative) yields,

$$-P_N^F(\hat{P}_N^F - 1) > -P_K^F(\hat{P}_K^F - 1)$$

Multiplying through by the positive constant $\frac{\sigma_F}{\sigma_F - 1}$ and adding to our original condition yields the result. □

Proof, Population.

$$\begin{aligned} & \hat{L}_N > \hat{L}_K \iff \\ \frac{exp(v_N' - v_N)}{\lambda_N exp(v_N' - v_N) + (1-\lambda_N) exp(v_K' - v_K)} > & \frac{exp(v_K' - v_K)}{\lambda_N exp(v_N' - v_N) + (1-\lambda_N) exp(v_K' - v_K)} \iff \\ & exp(v_N' - v_N) > exp(v_K' - v_K) \end{aligned}$$

which is true by the above welfare result. $\square$

1.8.2.2 Auxiliary Propositions

Proposition 1 (Primitives imply condition)**.** *Suppose that* $\frac{T_K^M}{T_N^M} < \frac{T_K^W}{T_N^W}$*. Then,* $\pi_{NN}^W < \pi_{KK}^W$.

Proof.

$$\pi_{NN}^W < \pi_{KK}^W \iff \frac{\left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}}{\left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1} + \left(\tau^W\right)^{1-\sigma_W}\left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1}} <$$

$$\frac{\left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1}}{\left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1} + \left(\tau^W\right)^{1-\sigma_W}\left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}}$$

$$\iff \left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}\left(\left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1} + \left(\tau^W\right)^{1-\sigma_W}\left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}\right)$$

$$< \left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1}\left(\left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1} + \left(\tau^W\right)^{1-\sigma_W}\left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1}\right)$$

$$\iff \left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}\left(\left(\tau^W\right)^{1-\sigma_W}\left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}\right) < \left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1}\left(\left(\tau^W\right)^{1-\sigma_W}\left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1}\right)$$

$$\iff \left(w_N T_{\mathrm{N}}^W\right)^{2(\sigma_W - 1)} < \left(w_K T_{\mathrm{K}}^W\right)^{2(\sigma_W - 1)} \iff \left(w_N T_{\mathrm{N}}^W\right)^{\sigma_W - 1}$$

$$< \left(w_K T_{\mathrm{K}}^W\right)^{\sigma_W - 1} \iff w_N T_{\mathrm{N}}^W < w_K T_{\mathrm{K}}^W$$

Then, since $w_i = \frac{1}{T_i^M}$,

$$\iff \frac{1}{T_N^M} T_{\mathrm{N}}^W < \frac{1}{T_K^M} T_{\mathrm{K}}^W \iff \frac{T_K^M}{T_N^M} < \frac{T_{\mathrm{K}}^W}{T_{\mathrm{N}}^W}$$

which is true by assumption. $\square$

Proposition 2. $\pi_{KK}^F > \pi_{KN}^F$

Proof. Denote $v_i = w_i^{\omega}(P_i^W)^{1-\omega}/T_i^F$.

$$\pi_{KK}^F > \pi_{KN}^F \iff \frac{v_K^{1-\sigma_F}}{v_K^{1-\sigma_F} + (\tau^F)^{1-\sigma_F} v_N^{1-\sigma_F}} > \frac{(\tau^F)^{1-\sigma_F} v_K^{1-\sigma_F}}{(\tau^F)^{1-\sigma_F} v_K^{1-\sigma_F} + \tau^F v_N^{1-\sigma_F}}$$

$$\iff \frac{1}{v_K^{1-\sigma_F} + (\tau^F)^{1-\sigma_F} v_N^{1-\sigma_F}} > \frac{(\tau^F)^{1-\sigma_F}}{(\tau^F)^{1-\sigma_F} v_K^{1-\sigma_F} + v_N^{1-\sigma_F}}$$

$$\iff (\tau^F)^{1-\sigma_F} v_K^{1-\sigma_F} + v_N^{1-\sigma_F} > \left(v_K^{1-\sigma_F} + (\tau^F)^{1-\sigma_F} v_N^{1-\sigma_F}\right)(\tau^F)^{1-\sigma_F}$$

$$\iff 1 > (\tau^F)^{1-\sigma_F} \iff 1 < \tau^F$$

which is true since trade is costly. □

Proposition 3. *If* $\pi_{NN}^W < \pi_{KK}^W$ *and* $\hat{\tau}^W < 1$, *then* $\hat{P}_N^W < \hat{P}_K^W$.

Proof.

$$\hat{\tau}^W < 1 \iff \left(\hat{\tau}^W\right)^{1-\sigma_W} > 1 \iff \pi_{NN}^W - \pi_{KK}^W > \left(\hat{\tau}^W\right)^{1-\sigma_W} \left(\pi_{NN}^W - \pi_{KK}^W\right)$$

since $\pi_{NN}^W - \pi_{KK}^W < 0$.

$$\iff \pi_{NN}^W - \pi_{KK}^W + \left(\hat{\tau}^W\right)^{1-\sigma_W} > \left(\hat{\tau}^W\right)^{1-\sigma_W} \pi_{NN}^W - \left(\hat{\tau}^W\right)^{1-\sigma_W} \pi_{KK}^W + \left(\hat{\tau}^W\right)^{1-\sigma_W}$$

$$\iff \pi_{NN}^W + (1 - \pi_{NN}^W)\left(\hat{\tau}^W\right)^{1-\sigma_W} > \pi_{KK}^W + (1 - \pi_{KK}^W)\left(\hat{\tau}^W\right)^{1-\sigma_W}$$

since $1 - \pi_{ii}^W = \pi_{ni}^W$:

$$\pi_{NN}^W + \pi_{KN}^W \left(\hat{\tau}^W\right)^{1-\sigma_W} > \pi_{KK}^W + \pi_{NK}^W \left(\hat{\tau}^W\right)^{1-\sigma_W}$$

$$\iff \left[\pi_{NN}^W + \pi_{KN}^W \left(\hat{\tau}^W\right)^{1-\sigma_W}\right]^{\frac{1}{1-\sigma_W}} < \left[\pi_{KK}^W + \pi_{NK}^W \left(\hat{\tau}^W\right)^{1-\sigma_W}\right]^{\frac{1}{1-\sigma_W}} \iff \hat{P}_N^W < \hat{P}_K^W$$

$\square$

1.8.3 Model in Changes: System of Equations

Given parameter estimates for σ_F and σ_W, changes in total population $\hat{L}$, changes in productivities $\hat{T}_i^F$, $\hat{T}_i^W$, and $\hat{T}_i^M$, changes in trade costs $\hat{\tau}_{ji}^F$, $\hat{\tau}_{ji}^W$, measures of the initial import shares for wheat and flour π_{ji}^W, π_{ji}^F, the initial export shares for wheat and flour Π_{ji}^W, Π_{ji}^F, population shares, λ_i and the initial price of flour, changes in allocations $\{\Delta v_i, \hat{\lambda}_i, \hat{L}_i, \hat{c}_{ni}^F, \hat{c}_{ni}^W, \hat{Y}_i^F, \hat{Y}_i^W\}$ and in prices $\{\hat{w}_i, \hat{p}_{ni}^F, \hat{p}_{ni}^W, \hat{P}_i^F, \hat{P}_i^W\}$ are given by:

$$\hat{w}_i = \hat{T}_i^M \tag{1.20}$$

$$\hat{p}_{ni}^W = \frac{\hat{w}_n}{\hat{T}_n^W}\hat{\tau}_{ni}^W \tag{1.21}$$

$$\hat{p}_{ij}^F = \frac{\hat{P}_i^W}{\hat{T}_i^F}\hat{\tau}_{ni}^F \tag{1.22}$$

$$\hat{P}_i^F = \left(\sum_n \pi_{ni}^F (\hat{p}_{ni}^F)^{1-\sigma_F}\right)^{\frac{1}{1-\sigma_F}} \tag{1.23}$$

$$\hat{P}_i^W = \left(\sum_n \pi_{ni}^W (\hat{p}_{ni}^W)^{1-\sigma_W}\right)^{\frac{1}{1-\sigma_W}} \tag{1.24}$$

$$\hat{L}_i = \hat{L}\hat{\lambda}_i \tag{1.25}$$

$$\hat{\lambda}_i = \frac{exp(\Delta v_i)}{\sum_n \lambda_n exp(\Delta v_n)} \tag{1.26}$$

$$\Delta v_i = \Delta w_i - \ln\left(\hat{P}_i^F\right) - \frac{\sigma_F}{\sigma_F - 1}\Delta P_i^F \tag{1.27}$$

$$\hat{Y}_i^F = \sum_j \hat{\tau}_{ij}^F \Pi_{ij}^F \hat{c}_{ij}^F \tag{1.28}$$

$$\hat{Y}_i^W = \sum_j \hat{\tau}_{ij}^W \Pi_{ij}^W \hat{c}_{ij}^W \tag{1.29}$$

$$\hat{c}_{ni}^W = \frac{\hat{Y}_i^F}{\hat{T}_i}\left(\hat{p}_{ji}^W\right)^{-\sigma_W}\left(\hat{P}_i^W\right)^{\sigma_W} \tag{1.30}$$

$$\hat{c}_{ji}^F = \left(\hat{p}_{ji}^F\right)^{-\sigma_F}\left(\hat{P}_i^F\right)^{\sigma_F - 1} \tag{1.31}$$

1.8.4 Quantitative Model

The baseline model is outlined in Section 1.4. Below I define an equilibrium in changes, where $\hat{x} = x'/x$.

1.8.4.1 System of Equations in Changes

Given parameters $\{\epsilon, \theta_j, \gamma_n^j, \gamma_n^{j,p}, \alpha_n^j, \delta_n^j\}$, data on $\{L_n, l_n, P_n, \lambda_n, w_n, r_n, D_n, \pi_{in}^j\}$, and changes in the exogenous variables $\{\hat{L}, \hat{\tau}_{ni}^k\}$, an equilibrium is a set of changes in allocations

$\{\hat{v}_n, \hat{I}_n, \hat{c}_n^k, \hat{\pi}_{ni}^k, \hat{L}_n, X_n^{k\prime}, \hat{Y}_n^k, \hat{D}_n\}$ and prices $\{\hat{P}_n^k, \hat{w}_n, \hat{r}_n\}$ for a total of $6N + 4KN + KN^2$ unknowns that are defined by the following set of $6N+4KN+KN^2$ equations:

$$N \qquad \hat{v}_n = \hat{I}_n \cdot \Pi_{k=1}^{K} \left(\hat{P}_n^k\right)^{-\alpha_n^k} \tag{1.32}$$

$$N \qquad I_n' = \hat{w}_n w_n + \frac{r_n \hat{r}_n l_n + D_n'}{L_n \hat{L}_n} \tag{1.33}$$

$$N \qquad \hat{L}_n = \hat{L} \frac{\hat{v}_n^{\epsilon}}{\sum_n \lambda_n \hat{v}_n^{\epsilon}} \tag{1.34}$$

$$N^2 K \qquad \hat{\pi}_{in}^k = \frac{(\hat{\tau}_{in}^k \hat{c}_i^k)^{-\theta_k}}{\sum_i \pi_{in}^k (\hat{\tau}_{in}^k \hat{c}_i^k)^{-\theta_k}} \tag{1.35}$$

$$NK \qquad \hat{c}_n^j = \hat{w}_n^{\gamma^j} \hat{r}_n^{\delta_j} \prod_p^K \left(\hat{P}_n^p\right)^{\gamma^{p,j}} \tag{1.36}$$

$$NK \qquad X_n^{k\prime} = \alpha^k L_n \hat{L}_n I_n' + \sum_{j=1}^K \gamma_n^{k,j} \sum_{i=1}^N X_i^{j\prime} \pi_{in}^j \hat{\pi}_{in}^j \tag{1.37}$$

$$NK \qquad Y_n^k \hat{Y}_n^k = \sum_i X_i^{k\prime} \hat{\pi}_{ni}^k \pi_{ni}^{k\prime} \tag{1.38}$$

$$N \qquad D_n' = \sum_k \sum_i \pi_{in}^k \hat{\pi}_{in}^k X_n^{k\prime} - \sum_k \sum_i \hat{\pi}_{ni}^k X_i^{k\prime} \pi_{ni}^k \tag{1.39}$$

$$N \qquad \hat{w}_n w_n \hat{L}_n L_n = \sum_j \omega_n^j \gamma_n^j Y_n^{j\prime} \tag{1.40}$$

$$N \qquad \hat{r}_n r_n l_n = \sum_j (1 - \omega_n^j) \gamma_n^j Y_n^{j\prime} \tag{1.41}$$

$$NK \qquad \hat{P}_n^k = \left[\sum_i \pi_{in}^k \left(\hat{c}_i^k \hat{\tau}_{in}^k\right)^{-\theta_k}\right]^{\frac{1}{-\theta_k}} \tag{1.42}$$

1.8.4.2 Solving the Model

I solve the model using Python's optimization routines and the following algorithm:

1. Guess values for wages $\hat{w}_n$, rents $\hat{r}_n$, population $\hat{L}_n$, prices, $\hat{P}_n^j$, and spending X_n^j.

2. Given the guess for prices, construct $\hat{c}_n^j$ from equation 1.36.

3. Construct a model-implied value for prices $\tilde{\hat{P}}_n^j$ from equation 1.42.

4. Construct $\hat{\pi}_{in}^k$ from equation 1.35.

5. Construct $Y_n' = Y_n \hat{Y}_n$ from equation 1.38.

6. Construct trade imbalances D_n' from equation 1.39 using the guess for spending.

7. Using population, wage, and rental rate guesses, plus trade imbalance from above, construct per-capita income in the post period I_n' from equation 1.33.

8. Compute changes in welfare from equation 1.32, and model-implied changes in population $\tilde{\hat{L}}_n$ from equation 1.34.

9. Compute a model-implied value of spending $\tilde{X}_n^{k'}$ from 1.37.

10. Form objective function based on:

$$\begin{aligned}
0 &= \hat{P}_n^j - \tilde{\hat{P}}_n^j \\
0 &= \hat{L}_n - \tilde{\hat{L}}_n \\
0 &= \tilde{X}_n^{k'} - X_n^{k'} \\
0 &= \hat{r}_n - \frac{\sum_j (1-\omega_n^j)\gamma_n^j Y_n^{j'}}{r_n l_n} \\
0 &= \hat{w}_n - \frac{\sum_j \omega_n^j \gamma_n^j Y_n^{j'}}{w_n L_n \hat{L}_n}
\end{aligned}$$

where the last two lines come from equations 1.41 and 1.40 respectively. The optimization procedure then iterates through steps 1 through 10 until all values of 10 are zero.

1.8.4.3 Quantifying Trade Costs

To feed bilateral changes in trade costs into the model, I want to measure initial trade costs along each pair of locations. I assume that trade costs between two locations are the product of the railroad distance of the route and a constant X, which coverts distance to revenue: $\tau_{od} = km_{od}^{\kappa} \cdot X$. If I observed revenue per ton earned along every route, then I could use that to measure initial trade costs along each pair of states. However, I do not observe this in the data for every route since there is no trade in agricultural products along some routes. Aggregating over every route where $revenue_{od} > 0$, I can compute total RTM:

$$RTM = \frac{\sum_{od} revenue_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}}$$

Then, since $revenue_{od} = c_o \tau_{od} Ton_{od}$, I can re-write this as:

$$RTM = \frac{\sum_{od} c_o \tau_{od} Ton_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}} = \frac{\sum_{od} c_o km_{od}^{\kappa} \cdot X \cdot Ton_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}} = X \frac{\sum_{od} c_o km_{od}^{\kappa} \cdot Ton_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}}$$

Setting these two equal to each other and solving for X:

$$\frac{\sum_{od} revenue_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}} = X \frac{\sum_{od} c_o \cdot km_{od}^{\kappa} \cdot Ton_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}} \implies X = \frac{\frac{\sum_{od} revenue_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}}}{\frac{\sum_{od} c_o \cdot km_{od}^{\kappa} \cdot Ton_{od}}{\sum_{o'd'} tons_{o'd'} \cdot km_{o'd'}}}$$

I measure c_o, which is the cost of producing agricultural goods at location o as the dollar value received for a unit of wheat in state o, measured from the USDA in 1950. Given this estimate of X, I solve for initial trade costs across all locations.

CHAPTER 2

Political Preferences and the Spatial Distribution of Infrastructure: Evidence from California's High-Speed Rail

(with Cecile Gaubert, Pablo Fajgelbaum, Eduardo Morales, and Edouard Schaal)

PRELIMINARY: Click link to updated version

.

2.1 Introduction

Transport networks are among the largest investments made by federal and local governments. What determines the projects that are implemented? The sheer size of transport infrastructure projects often makes them a focus of public debate, with people and politicians taking strong stances on their suitability. These views are partially driven by distributional considerations, as transportation networks shape the spatial organization of economic activity and, with it, the distribution of income across regions or demographic groups. However, as with other forms of government spending, people's preferences over whether a transport project is a good idea depend on more than the usual economic variables that will likely be impacted (such as access to jobs or house prices, in this case). Instead, they also encompass political considerations such as preferences for public goods, environmental concerns, or loyalty to the political party championing a project. When deciding whether and how to invest in a large public good such as transport networks, policymakers are likely to take these extra considerations of constituencies into account alongside motivations such as their own preferences for redistribution.

In this paper, we study how the preferences of transport users and policymakers determine the benefits and the design of a large infrastructure project. Specifically, we ask: how important are the political components in people's preferences for infrastructure, vis-a-vis the more often studied real-income effects? In general, this question is hard to answer because preferences are not observed. To make progress, we choose as our setting the California's High-Speed Rail (CHSR), an electric high-speed rail system designed to connect urban centers in California. The CHSR, arguably the second-largest transport project in US history after the Inter-State Highway System, presents an ideal setting both because of its importance and because we observe how Californians voted for or against it: as part of the 2008 general election, Californians were asked whether they approved of the state issuing bonds worth about 10 billion USD to finance part of the project. The proposition passed with 52.6% in favor and 93% of presidential election voters casting a vote. Across approximately 8,000 census tracts in California, we observe voting for or against financing the HSR, as well as for other ballots in the general election.

To quantify the importance of economic factors in shaping preferences, we first estimate the rate at which the expected real-income effects of the HSR translate into voter approval. Estimating this rate is essential in assigning a real income equivalent value to proxies for political ideology, such as party affiliation. A central challenge in this endeavor is that the expected economic returns from the HSR are not observed. For this, we develop and estimate a quantitative spatial framework extended to included specific features of the HSR.

The model of economic gains from the HSR builds upon urban models in the style of Ahlfeldt et al. (2015) and Monte, Redding and Rossi-Hansberg (2018). In these models, transport infrastructure impacts the commuting decisions that shape the demand and supply of floor space (for housing and production) and the production of urban spillovers through efficiency or amenities. We build upon these models to include specific features of the HSR. First, while some planned segments of the HSR facilitate commuting, its connections can be used for long-distance non-regular travel such as business trips or leisure travel. We show that these travel purposes, rather than regular commuting, are the most likely beneficiaries of the HSR, and incorporate them into the decisions of households and firms. Second, as the HSR competes with other travel modes (car, air, and standard rail), we incorporate a transport-mode decision that accounts for trade-offs between time savings and money by transport mode.[1]

Our estimation deals with a number of potential concerns, such as the possibly spurious relationship between the economic benefits of the HSR and political preferences, and the potentially erroneous predictions of voters about the economic returns to the HSR. To deal with the former concern, we first introduce a host of controls including observed votes for other political issues that absorb variation in

[1]This version of the paper outlines the most general model with all modes of transport. In our current estimation of the model, we pool all modes together.

values. We then construct an instrument for welfare gains from the HSR using the welfare gains implied by potential, but not implemented, HSR routes and randomly allocated stations. For the latter, we use an instrumental variables approach that deals with endogeneity of expectational errors and we implement the analysis under different assumptions on the sophistication of voters in terms of the economic model that they use to forecast incomes. We consider the robustness of our results to a variety of modeling assumptions.

We find that voters are quite responsive to the expected economic impact of the HSR. Had voters ignored the real-income effects of the HSR, the 2.6% difference in its favor in the actual vote would have fallen by about 1/3. Still, this economic component explains only a very small fraction of the variance of votes across space, with the observed presidential vote absorbing a large fraction of this variance. Overall, these results paint a picture where differences across space in people's attitudes towards the HSR were shaped more strongly by non-economic considerations, but where the real income factor is still taken into consideration to a degree that can influence the aggregate approval of the project.

Our paper contributes to several different literatures. First, we build off of existing work that studies the real income effects of infrastructure, including Redding and Turner (2015), Tsivanidis (2019), Donaldson (2018), Faber (2014), Severen (2019), Bernard, Moxnes and Saito (2019), and Donaldson and Hornbeck (2016*b*) to estimate the real income effects of the high speed rail. The model we develop to estimate the economic effects of the HSR builds off of existing papers that develop quantitative spatial equilibrium models including Ahlfeldt et al. (2015), Monte, Redding and Rossi-Hansberg (2018), and Dingel and Tintelnot (2020). Other projects have also studied the political economy of transportation projects, including work by

Brueckner and Selod (2006) and Glaeser and Ponzetto (2018). Finally, Van Patten and Méndez (2022), Holian and Kahn (2015), and Kahn and Matsusaka (1997) use voting on data to study how individuals' preferences shape economic outcomes.

2.2 Background on California High Speed Rail

In 1996, the California state legislature established the California High Speed Rail Authority (CHSRA) to explore the creation of a high speed rail network that would travel at least 200 miles per hour to connect northern California, including San Francisco and Sacramento, with southern California, including Los Angeles and San Diego (U.C. Hastings, 2008). A significant portion of funding for the rail line would come from state-issued bonds. By law, issuing such bonds requires voter approval that must be obtained through a successful ballot measure. In August 2008, the California legislature voted to approve that Proposition 1a would appear on the ballot in the November 2008 general election, asking California voters to approve the issuance of nearly $10 billion in bonds to fund the HSR.

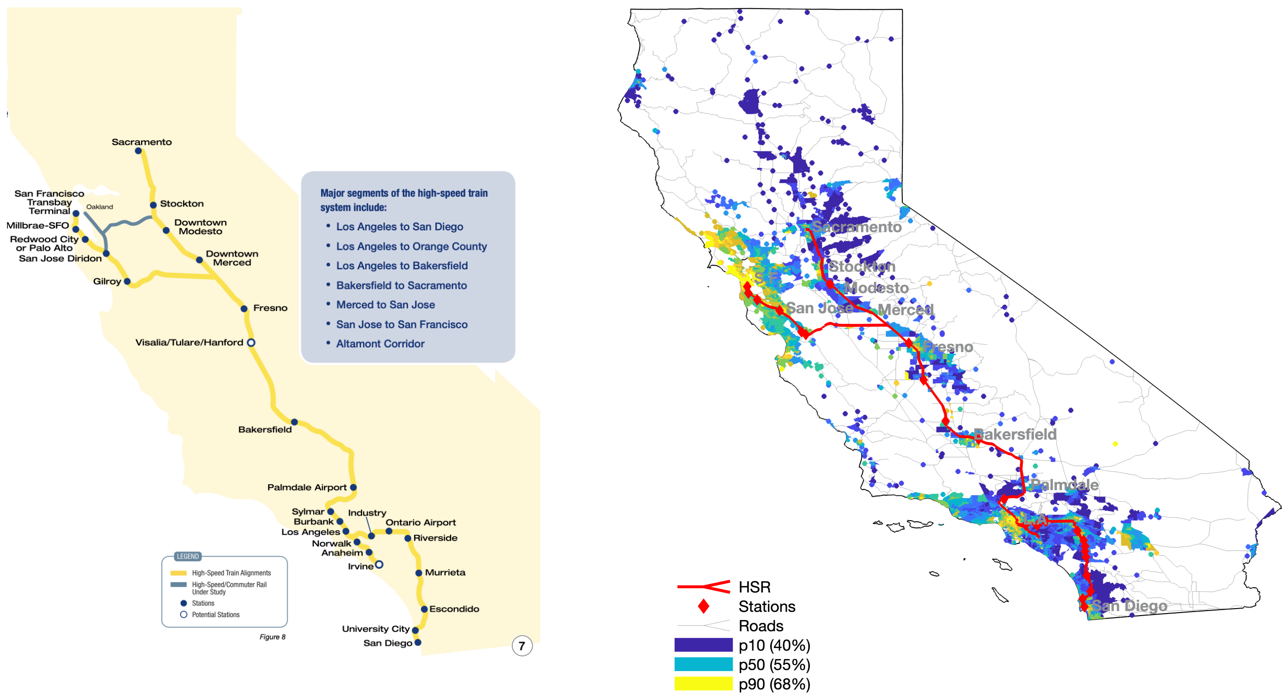
According to the legislature's vote to approve Proposition 1a, the HSR would be required to meet several criteria. First, it would connect San Francisco to Los Angeles and Anaheim, and would also include Sacramento, the San Francisco Bay Area, the Central Valley, Los Angeles, the Inland Empire, Orange County, and San Diego. Second, it would travel at at least 200 miles per hour, making the trip from San Francisco to Los Angeles Union Station in at most two hours and 40 minutes. Third, there would be no more than 24 stations across the entire network. Finally, the anticipated completion date as of 2008 was 2020. The network planned as of November 2008 is shown in Figure 2.1a. The first phase of building would focus on the LA-SF corridor, while Phase 2 would extend the network north to Sacramento and south to San Diego. Proposition 1A was one of twelve measures on the ballot in the November 2008 general election. The ballot text read:

> To provide Californians a safe, convenient, affordable, and reliable alternative to driving and high gas prices; to provide good-paying jobs and improve California's economy while reducing air pollution, global warming greenhouse gases, and our dependence on foreign oil, shall $9.95 billion in bonds be issued to establish a clean, efficient high-speed train service linking Southern California, the Sacramento/San Joaquin Valley, and the San Francisco Bay Area, with at least 90 percent of bond funds spent for specific projects, with federal and private matching funds required, and all bond funds subject to independent audits?

Figure 2.1b hows support for the HSR, as measured by votes on Proposition 1a, varied across Census tracts in California. Each point is the population centroid of a tract; in denser areas where tracts are smaller in size, the entire tract is colored. Bright yellow areas are those that were more supportive; (dark) blue areas are those that were (less) supportive. San Francisco and LA are more supportive of the HSR compared with Fresno and Sacramento. Proposition 1a was ultimately approved with a vote in favor of 52.6%. Participation was high: 94% of voters casting a vote in the presidential race also cast a vote on Proposition 1a.

Figure 2.1: HSR Route and Proposition 1a Votes

(a) Planned HSR Route as of November 2008 (b) % Yes on Prop 1a



Note: Panel (a) shows the CHSR network as proposed in the November 2008 CHSR Authority business plan. Panel (b) shows the percentage of voters in each tract voting in favor of the HSR.

Planned routes. The route that was ultimately selected was not the only option considered: the CHSRA identified several potential main routes for the train. The three main routes are shown in Figure 2.2 below. One route, in navy blue, travels near the Amtrak line along the California coast, while another, in bright blue, follows Interstate 5 and a third, in green, tracks State Route 99 through the major cities of the Central Valley like Fresno and Madera. This map also highlights different possibilities within each option: for example, the green route along SR-99 shows a few possible routes from Palmdale to Bakersfield and to LA.

The route that was ultimately selected at the time of the 2008 vote runs through the Central Valley along SR-99 and is shown in in detail in Figure 2.1a as reproduced from the CHSR's 2008 Business plan. While funding constraints and other issues have forced officials to reconsider what proportion of this line will be built and some small changes have been made to the set of stations (i.e., with the introduction of a station in Madera), the route itself has not been materially altered. A variety of factors affected the CHSRA's ultimate choice of the SR-99 route in the Central Valley. One factor was that the coastal route is longer in terms of distance than the other two routes. The topology of the coastal area also makes it more difficult for trains to reach top speed and is more costly in terms of building; a high speed train traveling along this route would not be able to reach the top speeds that a train traveling in land would be able to reach.

Funding sources. The proposed funding sources for the HSR included funds from bond issuances, private investors, the federal government, and other local sources (LA Times, 2007). Specifically, the nearly $10 billion in Proposition 1A bonds were expected to be coupled with $2-3 billion in local funding, $12-16 billion in federal

Figure 2.2: Potential HSR Routes (1996)

Note: This figure shows the possible HSR routes as outlined by the 1996 CHSR Commission, and reproduced in the CHSR Environmental Impact Report (2005).

funding, and \$6.7-7.5 billion in public-private partnership funding, which was supposed to be sufficient to build a complete San Francisco to Anaheim system for a total of \$33.6 billion in 2008 dollars (ENO).

Current status. Although construction on the HSR began in 2005, progress has been slower than anticipated. The estimated cost has risen from the 2008 estimate of \$40 billion to more than \$100 billion. Today, construction on 180 miles of the Central Valley segment has begun, and is expected to be completed between 2023 and 2025 (Fresno Bee, 2022).

2.3 Framework

This section gives an overview of the theoretical framework. A complete discussion of equations and derivations is provided in Appendix Section 2.7. We start by presenting voters' preferences and a model of the voting decision driven by political and economic considerations. We then zoom in on these economic considerations by developing a quantitative model of the impact of the CHSR on the distribution of real incomes across census tracts in California.

2.3.1 Utility and Voting

Preferences. Consider a resident ω of a location i. Her utility $u_\omega(s)$ captures both her economic and political preferences, which depend on whether an infrastructure project, in our case the California high-speed rail, is planned to be built $(s = B)$ or not $(s = NB)$:

$$u_\omega(s) = \mathbb{E}\left[\ln W(i,s) \mid \mathcal{I}_i\right] + \ln a(i,s) + \varepsilon^u_\omega(s). \tag{2.1}$$

The first component, $\mathbb{E}\left[\ln W(i,s) \mid \mathcal{I}_i\right]$ measures the average expected real income of residents of i when the project has status s. It captures economic forces through which an infrastructure project can impact real income, such as better market access or higher house prices, and will be discussed in more detail in section 2.3.2. The expectation on $\ln W(i,s)$ is taken over shocks that may affect the economy, including shocks over fundamental economic characteristics of different locations i and uncertainty about the transportation project itself, conditional on the information $\mathcal{I}_i$ of residents of i. The second component, $\ln a(i,s)$, captures the political component of preferences on average across residents of i. Finally, the shock $\varepsilon_{\omega}^{u}(s)$ captures idiosyncratic (mean-zero) variation in utility across residents of tract i stemming from either economic or political considerations.

Given (2.1), we define the average *utility impact* of the transport project across residents of location i as $\Delta U(i) \equiv \mathbb{E}_{\omega}[u_{\omega}(B) - u_{\omega}(NB)]$ and get:

$$\Delta U(i) = \mathbb{E}\left[\ln \hat{W}(i) \mid \mathcal{I}_i\right] + \ln \hat{a}(i) . \tag{2.2}$$

In this expression, $\hat{W}(i) \equiv \frac{W(i,B)}{W(i,NB)}$ measures real income differences between a world with and without the planned infrastructure project, while $\hat{a}(i)$ measures the net political preferences for the project. Both components affect which locations win or lose from the project ($\Delta U(i) \gtrless 0$). The existing work on distributional impacts of infrastructure measures $\hat{W}(i)$. One of the central goals in this paper is to measure the relative importance of $\hat{W}(i)$ and $\hat{a}(i)$ in shaping this distributional impact of a transport project across locations i.

In practice, neither the real income or the political components are directly observed. To make progress, we impose more structure on the problem. First, we assume that residents form rational expectations so that:

$$\mathbb{E}\left[\ln \hat{W}(i) \mid \mathcal{I}_i\right] = \ln \hat{W}(i) - \epsilon_W(i) , \tag{2.3}$$

where $\ln \hat{W}(i)$ is the value of the real-income gains brought about by the HSR under the actual realization of shocks and $\epsilon_W(i)$ is a mean-zero expectational error. In the next section, we proxy for $\hat{W}(i)$ using a range of economic models with varying degrees of sophistication in the forces that they incorporate.

Second, the political component $\hat{a}(i)$ is not directly observed but can be proxied for with location-specific variables $X_k(i)$ that bear a relationship to people's political ideology, such as people's party affiliation (we discuss these proxies later on). Formally, we assume that:

$$\ln \hat{a}(i) = \sum_{k=1}^{K} \tilde{\beta}_k X_k(i) + \epsilon_a(i), \tag{2.4}$$

where $X_0(i) = 1$ so that $\tilde{\beta}_0$ captures a common component of ideology across space, and where $\epsilon_a(i)$ are additional unobserved determinants of political ideology. Given our goals, a central question is how to measure the rate $\tilde{\beta}_k$ at which a given proxy (people's concerns about the environment, say) can be meaningfully compared with the real-income impacts of a given project. Unlike in the case of $\ln \hat{W}(i)$, no family of models exists to translate proxies of values into a welfare metric that is comparable to real income. To overcome this issue, we turn next to the voting model. We rely on the simple insight, which we formalize below, that voting outcomes directly reveal this trade-off between political values and economic gains.

Voting. Faced with a choice between B and NB, households vote in favor of B if and only if $u_\omega(B) > u_\omega(NB)$. Aggregating over individuals, given utility (2.1), our setup corresponds to a probabilistic voting model (Dixit and Londregan, 1996), with the fraction of positive votes in location i defined as $v(i) = \Pr[u_\omega(B) > u_\omega(NB)]$. We assume that these idiosyncratic shocks are Type-I extreme-value distributed across residents, with shape parameter θ_V:

$$\Pr\left(\varepsilon_\omega^u(s) < x\right) = e^{-e^{-\theta_V x}}. \tag{2.5}$$

As a result, the fraction of voters in i that support building the rail takes the standard logit form:

$$v\left(i\right)=\frac{e^{\theta_V\left(\mathbb{E}\left[\ln\hat{W}(i)|\mathcal{I}_i\right]+\ln\hat{a}(i)\right)}}{1+e^{\theta_V\left(\mathbb{E}\left[\ln\hat{W}(i)|\mathcal{I}_i\right]+\ln\hat{a}(i)\right)}}. \tag{2.6}$$

Combining this expression with (2.3) and (2.4), and re-writing the left-hand side as the log odds-ratio, we obtain the equation that we will bring to the data:

$$\ln\left(\frac{v\left(i\right)}{1-v\left(i\right)}\right)=\theta_V\ln\hat{W}\left(i\right)+\sum_{k=1}^{K}\beta_k X_k\left(i\right)+\epsilon\left(i\right), \tag{2.7}$$

where $\beta_k\equiv\theta_V\tilde{\beta}_k$ and where $\epsilon\left(i\right)\equiv-\epsilon_W\left(i\right)+\epsilon_a\left(i\right)$ includes the expectational error and the unobserved components of political ideology.[2]

Equation 2.7 shows the trade-off between real income and political preferences in determining votes. The parameter θ_V is the rate at which real income translates into votes. Having identified this parameter, we can recover the $\tilde{\beta}_k$'s in (2.4) as $\tilde{\beta}_k=\beta_k/\theta_V$ thus revealing how a given proxy for political preferences translates into real-income equivalent terms. We can then construct $\ln\hat{a}\left(i\right)$ and compute how much it drives the heterogeneity in $\Delta U\left(i\right)$. The elasticity of votes to real income θ_V does not matter per-se for the average welfare of a tract $\Delta U\left(i\right)$, but estimating this parameter is a crucial step to translate ideology proxies into a common real-income metric.

[2]While notation has not been introduced for it, this error term also captures that voters may use an economic model that diverges from the range of models that use to proxy for $\hat{W}\left(i\right)$. This source of error, as well as the sources of error that we have previously explicitly specified, bring about identification concerns that we tackle in the empirical section through various instrumental variables.

2.3.2 Real Income Gains from the HSR

2.3.2.1 Model Ingredients

Delayed benefits. The real income impacts of the HSR are twofold: (i) Until the HSR is built and operational, households expect to pay a tax t to fund the HSR; (ii) once the HSR is operational, households still pay the tax but also benefit from the corresponding real income gains. The overall real income effect of the HSR is the annualized flow combining these two streams.

At the time of voting, in 2008, the HSR business plan stated a specific year in which the HSR would be operational. To account for the fact that voters may have doubted that timing, we assume that they expected the HSR to be implemented no sooner than T years after the vote, and that they attributed a yearly completion probability equal to p afterwards, conditional on the HSR not being built until then. We consider a range of scenarios with varying degrees of optimism over p and T. Assuming a yearly discount rate r, these assumptions imply that the annualized real-income effects of the HSR, if approved, are a weighted average of $\ln(1-t)$ (the annual real-income change from paying a yearly tax t to finance to the HSR) and $\ln \hat{W}^Y(i)$ (the yearly real-income change, including taxes, associated with the HSR being operational for residents of location i). Specifically:

$$\ln \hat{W}(i) = \left(1 - (1+r)^{-T} \frac{p}{r+p}\right) \ln(1-t) + (1+r)^{-T} \frac{p}{r+p} \ln \hat{V}(i) \, . \qquad (2.8)$$

Economic benefits from HSR. We next turn to evaluating the real-income change associated with the HSR being operational, $\hat{V}(i)$ for residents of tract i. We follow a large body of research that uses quantitative spatial models in the style of Ahlfeldt et al. (2015) to estimate the distributional impacts of infrastructure improvement. We tailor a model from this class to the specifics of the high-speed rail. One may worry that voters rely on simpler heuristics rather than on the predictions of models that may include fairly rich spatial interactions. Therefore, as we describe next, we consider two polar cases of such models with different levels of complexity: a baseline that only includes the direct economic impacts of the HSR, on time savings and cost of travel; and a full model with several equilibrium feedback loops. In all versions of the model, we assume that there is a fixed number $N_R(i)$ of residents in location i. The development of the HSR impacts their commuting and travel choices, but not their residential choice.

The *raison d'être* of the HSR is that it allows passengers to save time on their long-distance travel. At the heart of the model, therefore, is the time saved by workers on daily commutes (for workers who have especially long commutes) and on less frequent long-distance leisure trips. The HSR also facilitates long distance business trips, making firms more productive as they connect with suppliers and customers (Bernard, Moxnes and Saito, 2019). We augment a commuting model a la Ahlfeldt et al. (2015) to incorporate these infrequent long-distance trips.

The "baseline model" only captures direct effects of the HSR. Specifically, in each census tract, a continuum of residents spends income on tradeable goods, housing, services, commuting, and leisure trips. Each resident receives idiosyncratic preference shocks for commuting or traveling for leisure to different destinations. Each of these activities entails a payoff (a salary when commuting or a utility flow when traveling for leisure) in exchange for a time cost and a monetary cost of travel. Based on these, residents make discrete choices of destinations for commuting and for leisure trips. For the latter, they also choose the number of trips taken. When deciding where to travel, people incorporate a choice of transport mode (such as car, public transit, air, or walking/biking), with residents of different locations potentially varying in their preferences for different travel modes. Building the HSR affects choices and welfare by lowering travel times and changing the price of travel on some routes. However, wages, the cost of housing, and costs of services remain constant in this baseline. As shown below, the real income gains stemming from these forces amount to the expected time-savings over likely destinations for travel, adjusted by the travelers' valuation of time and willingness to substitute across destinations.

The "full model" captures, in addition, equilibrium impacts on wages, land rents, cost of services, amenities, and productivity. First, to the previous setup, this more complex model adds endogenous wages. Tradeable goods are produced with labor, land, and business trips while housing uses land and local services use labor. Firm productivity is impacted by agglomeration spillovers that depend on the (endogenous) spatial distribution of worker density. So, the HSR impacts the productivity of firms and equilibrium wages through its impact on the time and monetary cost of business trips and on the employment distribution. Second, spillovers from endogenous labor density also impact the amenity enjoyed by residents and leisure travelers to a given destination. Third, because producers compete with residents to use land, land rents respond to the development of the HSR. Land rents capitalize the local productivity enhancement and the increased demand for productive space. A share of the residents of each tract are homeowners, and thus benefit from this capitalization. Together, changes in land rents and wages capture that the HSR may lead to local economic development.

2.3.2.2 Key Equations

We show here the key equations governing the real income impact of the HSR, formalizing the discussion above. A full description of the underlying model is discussed in the Appendix section 2.7. The equations below apply to the full model, and nest the baseline model under some parameter restrictions.

The annual real-income change for residents of tract i due to the HSR is:

$$\hat{V}\left(i\right) = \left(\frac{\hat{W}_C\left(i\right)}{\hat{P}_L\left(i\right)^{\mu_L(i)}}\right)\left(\frac{\hat{B}\left(i\right)}{\hat{P}_S\left(i\right)^{\mu_L(i)}\,\hat{r}\left(i\right)^{\mu_H(i)}}\right), \tag{2.9}$$

The first term in parenthesis includes in the numerator the change in expected income net of commuting costs and in the denominator the change in the price index for leisure trips ($\hat{P}_L\left(i\right)$), adjusted by the spending share of tract i's residents on leisure travel ($\mu_L\left(i\right)$). The second term in parenthesis includes in the numerator the change in residential amenities stemming from urban spillovers ($\hat{B}\left(i\right)$) and in the denominator the change in the cost of housing ($\hat{r}\left(i\right)$) and in the cost of services ($\hat{P}_S\left(i\right)$), adjusted by their respective spending share in tract i ($\mu_H\left(i\right)$ for housing and $\mu_S\left(i\right)$ for services). In the baseline model, this second term in parenthesis equals one by assumption.

The change in income net of commuting cost $\hat{W}_C\left(i\right)$ captures the expected time savings by commuters based on their typical commuting patterns, adjusted by the value of time and augmented by the flexibility to substitute across employment destinations:

$$\hat{W}_C(i) \equiv \left(\sum_{j\in\mathcal{J}} \sum_{m\in\mathcal{M}} \lambda_C^{pre}(i,j,m) \left(\frac{\hat{I}(i,j,m)}{\hat{\tau}(i,j,m)^{\rho_C}} \right)^{\theta_C} \right)^{\frac{1}{\theta_C}}. \tag{2.10}$$

In this expression, $\hat{I}(i,j,m)$ is the change in disposable income conditional on commuting from one's residence i to one's workplace j using transport mode m. This term depends on the changes in the monetary cost of commuting from i to j as well as (in the full model only) on the changes in wage at destination and in land rent income at origin. Second, $\hat{\tau}(i,j,m)$ is the change in travel time from i to j using m, converted into a dollar-equivalent value by the elasticity ρ_C. Third, θ_C captures the extent to which residents substitute across commuting destinations and travel modes when the relative appeal of destination or modes changes.[3] Finally, the weights $\lambda_C^{pre}(i,j,m)$ on these changes are the fraction of tract-i residents that commute for work to j through mode m before the HSR is built.

The change in the leisure price index adopts a similar functional form:

$$\hat{P}_L(i) \equiv \left(\sum_{j\in\mathcal{J}} \sum_{m\in\mathcal{M}} \lambda_L^{pre}(i,j,m) \left(\frac{\hat{p}_L(i,j,m)\,\hat{\tau}(i,j,m)^{\rho_L}}{\hat{B}(j)} \right)^{-\mu_L(i)\theta_L} \right)^{-\frac{1}{\mu_L(i)\theta_L}}, \tag{2.11}$$

where now $\lambda_L^{pre}(i,j,m)$ is the fraction of leisure travelers from i choosing destination j using mode m before the HSR is built, $\hat{p}_L(i,j,m)$ is the corresponding change in cost of the trip including the monetary cost of travel and the cost of services at destination), $\hat{B}(j)$ are changes in amenities in the leisure destination j, and (ρ_L, θ_L)

[3]Formally, as detailed in the appendix, this parameter is the inverse of the dispersion in idiosyncratic preference draws across residents of i about where to commute and how to commute there.

capture respectively the value of time when traveling for leisure and the substitution across destinations for leisure.[4]

Our baseline model with only direct effects assumes away endogenous changes in amenities or land rents in (2.9) (i.e., $\hat{B}(i) = \hat{r}(i) = 1$), assumes that changes in disposable income in (2.10) stem only from the tax to finance the HSR (i.e., $\hat{I}(i,j,m) \approx 1-t$),[5] and assumes away endogenous changes in amenities in (2.11) (i.e., $\hat{B}(j) = 1$). As a result, in our baseline model where only direct effects are included, we do not need any additional equations beyond (2.10) and (2.11). In contrast, the full model with indirect effects generates additional changes in amenities $\hat{B}(i)$, land rents $\hat{r}(i)$, and disposable income $\hat{I}(i,j,m)$. The latter is a function of changes in the entire wage distribution, with changes in wages themselves a function of agglomeration spillovers and the greater ease of sending workers on business travel. The full system of equations describing these forces is presented in Appendix 2.7.7.

---

[4] When comparing (2.10) and (2.11), one can notice two asymmetries: in (2.10), the monetary cost enters as a negative additive shifter to disposable income while in (2.11) it enters multiplicatively; and the latter is shaped by the intensity of leisure spending $\mu_L(i)$. As detailed in the appendix, these differences arise due to the non-homothetic nature of spending on commuting : assuming a fix number of commuting days in the year, travelers spend a fix amount of money in commuting and the remaining income is divided into consumption, housing, and leisure trips. In contrast, leisure travelers decide how many trips to make to their preferred destination, with homothetic preferences (over leisure trips, consumption, and housing) with spending shares possibly varying across tracts.

[5] The change in disposable income is defined as the change in after-tax income net of commuting costs: $\hat{I}(i,j,m) = (1+\chi^{pre}(i,j,m))(1-t)\,\hat{y}(i,j,m) - \chi^{pre}(i,j,m)\,\hat{p}_C(i,j_C,m_C)$, where $\hat{y}(i,j,m)$ is the change in pre-tax income, $\chi^{pre}(i,j,m)$ is the share of commuting costs in disposable income for someone traveling from i to j through mode m before the HSR is operational, and $\hat{p}_C(i,j_C,m_C)$ is the change in the monetary cost of this commuting route. Because the model with direct effects assumes no changes in wages and land rents, in that case disposable income changes only with taxes, meaning that $\hat{y}(i,j,m) = 1-t$.

2.4 Economic Impacts of the High-Speed Rail

In this section, we describe the procedure we follow to estimate the parameters of the model described in Section 2.3 and present the resulting estimates. In Section 2.4.1, we describe our sources of data, and present some summary statistics on the key variables entering our analysis. In Section 2.4.2, we describe the estimation of the parameters of the economic model described in Section 2.3.2.

2.4.1 Data

Voting data. We obtain data on votes and voter registration from the 2006 and 2008 California general elections by precinct, as well as a crosswalk file linking voting precincts (of which there are around 20k) to Census tracts (of which there are around 8k) from the University of California at Berkeley's Statewide Database. We use the provided crosswalk to construct a tract-level dataset of votes.

Travel times. We measure travel times between each pair of Census tracts as the fastest route through the transport network of a given mode between the centroid of the origin tract and the centroid of the destination tract for each mode. To approximate the population center of each tract, the centroid of each tract is defined as the geographic centroid of the most populous Census block within that tract.

To measure car travel times, we construct the road network based on the 2010 primary and secondary road network in California using a shapefile obtained from the U.S. Census, and calibrate speeds along rural/urban and primary/secondary road segments to match observed travel times from Google Maps. Appendix 2.9.2 provides more details on this calibration approach.

We allow air travel by assuming that agents will drive to the airport closest to their Census tract of origin, fly to the airport closest to their tract of destination, and then drive to their destination tract. We allow air routes operating within California, based on the Bureau of Transportation's 2008 airline ticket dataset and flight times from Google Maps. Thus, total air travel time is the sum of driving time at either end of the journey, plus flight time. We measure flight times on each route from Google Maps. In addition, we assume that it takes 45 minutes to move from the road network to the air network which reflects the time cost associated with moving through the airport.

We measure walking speeds based on road network distances, assuming people walk at a speed of 5 km per hour. For public transit travel, we assume that travelers can either take the bus, which is the road network where speeds are calibrated to match bus travel times (as detailed in Appendix 2.9.2), or they can drive a car to the nearest rail network. We impose that travelers will always take the bus along a route where the closest rail station to the origin tract is the same rail station that is closest to the destination tract, and that they will take bus if the bus option is faster than the car plus rail option. Appendix 2.9.2 describes how we use timetables from each passenger rail network, which includes both intercity networks like Amtrak as well as intracity networks like Caltrain, within California to construct the existing pre-HSR rail network.

To estimate travel times when the HSR is available, we obtain a shapefile of the planned route and stations of the HSR from the University of California at Davis at the time of the vote in November of 2008. The HSR map as of 2008 includes 26

Figure 2.3: Planned HSR Travel Speeds (2008)

Note: This figure shows planned travel speeds along each segment of the planned HSR route, as of the November 2008 CHSR Business Plan. We use these reported speeds to construct travel times between each pair of Census tracts when the HSR is built.

stations, two of which (Irvine and Tulare) are marked as potential stations and we exclude them in our baseline analysis. This total of 24 stations is consistent with the description of the network in the original HSR bill that was passed by the California legislature earlier in 2008.[6]

We measure planned travel speeds along each segment of the route from the

[6]Assembly Bill No. 3034 which officially allowed Proposition 1A to appear on the ballot, was passed in August 2008 and stated that "[t]he total number of stations to be served by high-speed trains for all of the corridors described in subdivision (b) of Section 2704.04 shall not exceed 24". It also included a number of other stipulations about travel times between stations.

November 2008 Business Plan as shown in Figure 2.3. Combining travel speeds on each segment with the distance of each segment allows us to compute travel times between all pairs of stations. The travel times that we estimate with this approach closely match those reported between major stations in the November 2008 CHSR Business Plan. We calibrate the time it takes to transfer between the road network and the HSR network based on transfers taken within the existing public transit system, as described in Appendix 2.9.

Travel costs. We measure the cost of traveling between each origin and destination pair on each mode. For car travel, we assume that the cost is a constant function of time. To convert minutes to dollars, we assume that the cost of gas is $3.52 per gallon in 2008[7], that the fuel economy of the average car is 21 miles per gallon[8], and that the average travel speed is 50 miles per hour. Combining these, we find that the cost of traveling via car for one minute is $0.17 in 2019 dollars.

For air, we assume that the cost is the sum of the cost of driving to and from the airport and the cost of the plane ticket. We measure the cost of traveling along each air route in California from the Bureau of Transportation Statistics. The average cost is $150, and since there is very little variation in price across routes, we assign this cost to every air route.

For rail, we assume that the cost is the sum of the cost of driving to and from the nearest rail station, plus the cost of traveling between the origin and destination railroad stations which we assume to be a function of time traveled on rail. To

[7]This figure is obtained from the Los Angeles Almanac at `https://www.laalmanac.com/energy/en12.php`.

[8]This figure is obtained from fueleconomy.gov.

parameterize this relationship between time and fare on rail, we estimate:

$$fare_i = \alpha + \beta \cdot time_i + \epsilon_i \tag{2.12}$$

where i is a route. We estimate $\hat{\alpha} = 2.9$ and $\hat{\beta} = 0.17$ using data on ticket fares ($fare_i$) and travel times ($time_i$) from the Capitol Corridor, a passenger train in Northern California operated by Amtrak. The fit of this functional form is quite good, with an R^2 of 94%. We assume that the cost of traveling via bus varies at the origin county level and measure the cost of a single adult bus ticket for each county in California from the American Public Transportation Association, supplemented with data from individual county webpages when data from APTA is missing.[9] We assume that walking is free.

Finally, we measure the cost of traveling per minute HSR using the projected estimate of a ticket at 77% of airfare, for a total of \$115, between the Los Angeles and San Francisco stations (2 hrs 38 minutes).[10] Assuming the fixed cost of traveling by HSR is the same as that of traveling by rail (\$2.92), we obtain a cost per minute of HSR travel equal to \$0.70 which is about four times as expensive as other rail lines.

Commuting flows. The American Community Survey reports tract-to-tract data on commuting flows as a part of the Census Transportation Planning Products.[11]

---

[9] This procedure is described in more detail in Appendix Section 2.9.2.

[10] This estimate of the ticket price was one proposal in HSR planning documents: https://hsr.ca.gov/wp-content/uploads/docs/about/business_plans/BPlan_2008_SRC_RiderRevenue.pdf.

[11] The U.S. Census' Longitudinal Employer-Household Dynamics (LEHD) also reports commuting flows between each origin and each destination but these tabulations are based on administrative data linking employees' home locations (origin) with their employer's location (destination). We do

The question asked of respondents to measure workplace location is: "At what location did this person work LAST WEEK?". The question asked to measure mode of travel is: "How did this person usually get to work LAST WEEK?". We exclude workers who work from home, since we cannot correctly measure the wage at their workplace without any information on where their firm is located.[12]

Leisure and business trips. To measure the demand for non-commuting travel across locations, we use the California Household Travel Survey which was administered between 2010 and 2012. 18,008 households were asked to record all trips taken over the eight week survey period that were more than 50 miles long; 68,193 trips are included in the dataset. The data include information on the origin census tract, the destination census tract, the residence census tract, the number of people on the trip, the travel mode, and the purpose of the trip. We classify each trip based on its stated purpose into either a leisure trip or a business trip; these trips together account for 84% of all trips taken in data. The remaining trips include combined business and please trips, medical trips, school-related activities, and trips for which the purpose is not stated.

We classify leisure trips as those taken for entertainment, vacation, shopping, or to visit friends and family. The top leisure destinations are Disneyland, Yosemite,

not observe the frequency with which these trips are taken. As such, these origin-destination flows may not reflect commute flows that are taken regularly, which is why we use ACS data to measure commute flows separately by each mode of transportation. The differences between commuting flows measured in the ACS and the LEHD are discussed at length here: https://www2.census.gov/ces/wp/2014/CES-WP-14-38.pdf. Our own analysis suggests, consistent with this, that the LEHD may over-represent very long commutes "flows" are based on a workers' firm location and residence location even if the worker does not travel to that location on a daily basis. Thus, we use the ACS to measure commuting flows.

[12] During this time period, workers working at home made up less than 5% of statewide workers.

Mission Beach (San Diego), Downtown San Francisco, and Downtown San Diego. Business trips are those taken for meetings, conventions, and seminars. The top business destinations are the State Capitol in Downtown Sacramento, Downtown Los Angeles, Downtown San Francisco, and Downtown San Diego.

Wage data. Our primary sources of data on wages are the 2008 and 2019 samples of the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics published by the U.S. Census. The LEHD reports the number of workers by origin and destination Census tract pair who have monthly earnings below $1,250, between $1,250 and $3,333, and above $3,333. To construct an average wage for each tract, we first measure average earnings within each of these three bins in California using the individual level American Community Survey samples in 2008 and 2019. We use these bins and our estimates of average earnings within each bin to compute average earnings among workers within tract pair.[13]

Other covariates. We use data on covariates from a number of different sources. We measure the share of homes in each tract that are owner-occupied, population, average income, racial composition, occupational composition, the share of residents with a college degree, and other demographic covariates at the tract level from the 2006-2010 and 2015-2019 American Community Survey five-year estimates. We measure population from the 2010 U.S. Census. We measure county-level unemployment rates from the Bureau of Labor Statistics. We measure average house prices and the share of floor space used for residential purposes in each tract from Zillow's ZTRAX

[13]This procedure is described in detail in Appendix Section 2.9.1.

data. We measure the share of income spent on leisure travel, and the share of leisure travel spending on transport from the Bureau of Labor Statistics' Consumer Expenditure Survey.

Sample. There are 8,057 census tracts in California. We eliminate tracts from our sample if they have no residents, have no workers, are on islands off the coast of the state, or are missing wage data. Our final sample comprises 7,866 census tracts which together house 98.5% of the statewide population.

2.4.2 Estimation of Parameters of Economic Model

In Section 2.4.2.1, we describe the estimation of the parameter vector (θ_C, ρ_C), and the commuting shares $\lambda_C^{pre}(i, j, m, NB)$ for every pair of census tracts (i, j) and every mode of transport m before the HSR is built (so $s = NB$) which enter 2.10. In Section 2.4.2.2, we describe the estimation of the parameter vector $(\theta_B, \rho_B, \mu_B)$, and the traveling shares $\lambda_B^{pre}(i, j, m, NB)$ for every pair of census tracts (i, j) and every mode of transport m. In Section 2.4.2.3, we describe the estimation of the parameter vector (θ_L, ρ_L) and $\mu_L(i)$ for every census tract i, and the traveling shares $\lambda_L^{pre}(i, j, m, NB)$ for every pair of census tracts (i, j) and every mode of transport m. In this version, we specify all equations in the case with different travel modes, but our estimation pools all modes together.[14]

[14] In future versions, we will estimate the parameters for the multi-mode case.

2.4.2.1 Commuting Equation

We base the estimation of the parameter vector (θ_C, ρ_C) on the following model-implied relationship

$$\lambda_C^{pre}(i,j,m,s) = \frac{\left(\frac{I(i,j,m)}{d_C(i,j,m)}\right)^{\theta^C}}{\sum_{j'}\sum_{m'\in M_{ij}}\left(\frac{I(i,j',m')}{d_C(i,j',m')}\right)^{\theta^C}} \tag{2.13}$$

where, as a reminder, $\lambda_C^{pre}(i,j,m)$ denotes the share of residents of location i that commute to location j using mode of transport m in a setting prior to the construction of the high-speed rail, $I(i,j,m)$ denotes the net income that residents of location i would have if they were to commute to location j through mode of transport m, and $d_C(i,j,m)$ is the monetary-equivalent cost of the time spent commuting from i to j through mode m. We assume that

$$I(i,j,m) = y(i,j) - p_C(i,j,m), \tag{2.14}$$

where $y(i,j)$ denotes the income of residents of location i who work in location j and $p_C(i,j,m)$ is the monetary cost of commuting from i to j though mode m. We also assume that

$$d_C(i,j,m,s) = D_C(i,m)\,\tau(i,j,m,s)^{\rho^C} \tag{2.15}$$

where $D_C(i,m)$ is the monetary-equivalent value of an amenity term that residents of census tract i experience when commuting through model of transport m, and $\tau(i,j,m)$ denotes the commute time between locations $i$and j when using mode of transport m under state s.[15]

To measure $y(i,j)$, we equate it to labor income, and assume that it is the product of an origin-specific component $w(i)$, which accounts for the possiblity that workers that reside in different locations have different human capital, and a destination-specific component $w(j)$, which accounts for productivity differences across work-places. Specifically, we assume that $y(i,j) = w(i)w(j)$, and we estimate these origin- and destination-specific components using the following estimating equation

$$w^{data}(i,j) = \exp\left(ln(w(i)) + ln(w(j)) + \varepsilon(i,j)\right) \tag{2.16}$$

where $w^{data}(i,j)$ denotes the observed average wage of workers who reside in i and work in j , and $\varepsilon(i,j)$ accounts for all other factors affecting observed average wages that cannot be accounted by an origin-specific and a destination-specific fixed effect.[16] We assume that $\varepsilon(i,j)$ does not impact workers' commuting decisions and is mean-independent of the origin- and destination-specific components; e.g., it captures measurement error in wages as well as wage shocks unexpected to workers when making their commuting decisions. For every pair of census tracts i and j and every mode of transport m, we measure the commuting shares $\lambda_C^{pre}(i, j_C, m, s)$, commuting costs $p_C(i,j,m)$ and travel time $\tau(i,j,m)$ as indicated in Section 2.4.1. Importantly, as our information on commuting shares come from a finite sample of residents in each origin census tract i, we allow for the possibility that the observed commuting shares for every pair of census tracts i and j and every mode of transport m,

[15]In Appendix 2.8 we show results using an alternative assumption where $d_C(i,j,m,s) = D_C(i,m)\, exp\left(\tau(i,j,m,s)\right)^{\rho^C}$.

[16]We take this approach instead of setting $y(i,j) = w^{data}(i,j)$ because we do not observe $w^{data}(i,j)$ for every (i,j) pair.

$\lambda_C^{obs}(i, j_C, m)$, differ from the true commuting share, which we assume is determined according to equations 2.13 to 2.16, in a term $e_C^s(i, j, m, s)$ that captures sampling error; i.e.,

$$\lambda_C^{obs}(i, j_C, m, s) = \lambda_C^{pre}(i, j, m, s) + e_C^s(i, j, m, s) \tag{2.17}$$

Given equations 2.13 to 2.17, we rewrite our estimating equation as

$$\lambda_C^{obs}(i, j_C, m, s) = \exp(\psi_C(i, m) + \theta^C \ln((w(i)w(j)) - p_C(i, j_C, m)) - \theta^C \rho^C \ln \tau(i, j, m, s)) + e_C^s(i, j, m, s) \tag{2.18}$$

where $\psi_C(i, m)$ is an origin- and mode-specific effect that accounts both for the denominator in equation (2.13) and for the term $D_C(i, m)$ entering equation (2.15).

When estimating the parameters in this equation, we consider that any of the 7,866 census tracts in CA is a potential origin or destination of a commuting trip, and we account for three possible transportation modes m: car only, public transit, and walking. Given these choices, we use information on $\lambda_C^{obs}(i, j_C, m)$ for every possible triplet (i, j_C, m) (including those for which $\lambda_C^{obs}(i, j_C, m) = 0$), treat the terms $\psi_C(i, m)$ as origin- and transport mode-specific fixed effects, and compute Poisson Pseudo Maximum Likelihood (PPML) estimator of these fixed effects and structural parameters θ_C and ρ_C.

In this version of the estimation, we pool all modes together.[17] We sum across all modes within an origin-destination pair, and measure travel times along each route as the fastest time on the transport network. We report in columns (1) to (4) in Table 2.1 the corresponding estimates both when using data from the years 2008 (i.e., when the vote on Proposition 1a took place) and in columns (5) to (8) from 2019 (i.e., when the HSR was expected to be built when the vote on Proposition 1a took place). As $e_C^s(i, j, m, s)$ is assumed to capture sampling noise, assuming we observe data from a random sample

[17]Estimation of the full, more general model is in progress.

Table 2.1: Commuting Equation Estimates

	2008				2019			
	(1) PPML	(2) PPML	(3) FS	(4) IV	(5) PPML	(6) PPML	(7) FS	(8) IV
log(pre-HSR travel time)	-3.671***	-3.704***			-3.741***	-3.773***		
	(0.00452)	(0.00443)			(0.00454)	(0.00446)		
log(workplace earnings, LEHD, 2008)	2.634***			2.749***				
	(0.00884)			(0.118)				
log(earnings, ACS, 2006-2010)			0.347***					
			(0.00870)					
log(workplace earnings, LEHD, 2019)					3.091***			2.669***
					(0.00973)			(0.108)
log(earnings, ACS, 2012-2016)							0.317***	
							(0.00624)	
ρ_C	1.39			1.35	1.21			1.41
θ_A	2.63			2.75	3.09			2.67
N	61873956	61873956	7866	7866	61873956	61873956	7866	7865
R2			0.227	0.186			0.291	0.234
F-stat				1600				2600
Destination FE	No	Yes			No	Yes		

Source: LEHD, 2008 and 2019; ACS, 2006-2010; ACS, 2012-2016.

Note: This table shows PPML estimates of the commuting equation in 2.18. Table 2.6 shows estimates under a different functional form of the disutility of travel time in 2.15.

of workers residing in each census tract implies this unobserved term has mean zero for each triplet (i, j, m), and converges to zero as the sample size for each origin tract i goes to infinity. Therefore, our PPML estimator yields estimates of all parameters entering in equation 2.18 that converge to their true values as the sample size in each census tract grows arbitrarily large.[18]

We estimate $\hat{\rho_C}$, which measures the value of time in commuting, to be around 1.2-

[18] Assuming that $e^s_C(i, j, m, s)$ accounts exclusively for sampling noise implies that our PPML estimates converge to their true values even if we use a small set of origin and destination pairs of census tracts (i, j) and modes of transport m, as long as the sample size in those origin census tracts we use for estimation grows arbitrairy large. However, in practice, we compute our estimates using a very large number of possible origin and destination census tracts. This implies that our estimates of the parameters entering (2.18) will also converge to their true values as long as the error term $e^s_C(i, j, m, s)$ is mean zero and exhibits weak correlation patterns across destination tracts j for each pair of origin tracts and modes of transport (i, m). Consequently, our estimation procedure also yields valid estimates if one interprets the error term in equation (2.18) not as sampling error but as some unobserved determinant of commuting flows between census tracts $i$and j through mode of transport m.

1.4. We estimate $\hat{\theta}_C$, which measures the elasticity of commuting to any given destination with respect to wages in that destination, to be around 2.7 in both years of data. These estimates are consistent with the literature. Severen (2019), for example, also computes a PPML estimator of θ_C using tract-level data for Los Angeles, and obtains an estimate of 2.2. Monte, Redding and Rossi-Hansberg (2018) use county-to-county commuting data covering all of the US, and obtain estimates of $\theta_C = 3.3$ and $\rho_C = 1.34$.

Using the estimates of (θ_C, ρ_C) reported in Column (4) of Table 2.1, the estimates of $\psi_C(i)$ (since we are pooling all modes together here), and the expression for $\lambda_C^{pre}(i, j, NB)$ in equation (2.18) after setting the sampling noise to zero, we generate model-predicted values of the share of commuters between every pair of census tracts (i, j). As these predicted shares are functions of consistent estimators of the parameters of interest, they are themselves consistent.

2.4.2.2 Business Travelers

Our modelling of business trips implies the following expression for the total number of business trips that originate in a census tract j, destination in a census tract j_B, and are done using a mode of transit m under state s:

$$R_B(j, j_B, m, s) = \psi_0(j) \left(\frac{q_B(j, j_B) A(j_B, s)}{d_B(j, j_B, m, s)} \right)^{\mu_B \theta^B} p_B(j, j_B, m)^{-\mu_B \theta^B - 1} \tag{2.19}$$

where $\psi_0(j)$ is an origin effect, $q_B(j, j_B)$ is a shifter (expressed in monetary units) of the number of trips done from tract j to tract j_B, $d_B(j, j_B, m, s)$ is the monetary-equivalent cost of the time spent traveling for business from j to j_B through mode m, and $p_B(j, j_B, m)$ is the direct monetary cost of traveling for business from j toj_B though mode m. $A(j_B, s)$

is the amenity value of the destination tract j_B under state s. We assume

$$d_B(j, j_B, m, s) = D_B(j, m)\,\tau(j, j_B, m, s)^{\rho^B} \tag{2.20}$$

where $D_B(j,m)$ is an unobserved shifter affecting the number of business trips that originate from location j through mode of transport m, and, as a reminder, $\tau(j, j_B, m, s)$ denotes the travel time between locations j and j_B when using mode of transport m under state s. Due to limitations in the size of the sample we use to estimate the parameters entering equation (2.19), we do not treat the shifters $q_B(j, j_B)$ and $D_B(j, m)$ as origin-destination and origin-mode specific fixed effects, respectively. Instead, we write these shifters as a function of observable characteristics; specifically, we assume

$$\mu_B\theta_B(\ln(q_B(j, j_B)\,A(j_B, s)) - \ln(D_B(j, m))) = \gamma_B(m)X_B(j, j_B, s) \tag{2.21}$$

where $\gamma_B(m)$ is a mode-specific parameter, and $X(j, j_B, s)$ is a vector of observed characteristics that, in our empirical specification, includes three covariates: (a) a measure of the similarity in industry composition between tracts j and j_B, (b) the share of workers in management roles in the destination tract j_B, and (c) the amenity value of the destination $A(j_B, s)$ as defined in equation (2.51).[19] For every pair of census tracts j and j_B, a state s, and every mode of transport m, we measure the number of business trips $R_B(j, j_B, m, s)$, travel time $\tau(j, j_B, m, s)$, and commuting costs $p_B(j, j_B, m)$ as indicated in Section 2.4.1. Our measure of $R_B(j, j_B, m, s)$ for every triplet (j, j_B, m) in each state s, which we denote as $R_B^{obs}(j, j_B, m, s)$, is based on a small random sample of residents of Califonia and, consequently, we account for the possibility that the number of trips reported in our sample

[19] We measure the similarity in industry composition between any two tracts j and j_B as the Euclidean distance between the vectors that capture the employment shares by sector in each of the two tracts: $odsim(j, j_B) = \left(\sum_k \left(emp_k^j - emp_k^{j_B}\right)^2\right)^{1/2}$, where emp_k^j is the employment share in census tract j in sector k.

for any pair of tracts and mode of transport differs from the corresponding true number of trips; i.e.,

$$R_B^{obs}\left(j, j_B, m, s\right) = R_B\left(j, j_B, m, s\right) + e_B^s(j, j_B, m, s) \tag{2.22}$$

where $e_B^s(j, j_B, m, s)$ denotes the potential sampling error in our measure of $R_B\left(j, j_B, m, s\right)$.

Combining equations (2.19) to (2.22), we rewrite our estimating equation as:

$$\begin{aligned} R_B^{obs}\left(j, j_B, m, s\right) = \exp(\psi_B\left(j\right) + \gamma_B(m) X_B(j, j_B, s) - \\ \mu_B \theta^B \rho_B \ln \tau(j, j_B, m) - (\mu_B \theta^B + 1) p_B\left(j, j_B, m\right)) + e_B^s(j, j_B, m) \end{aligned}$$

where $\psi_B\left(j\right) \equiv \ln \psi_0(j)$ is treated as an origin fixed effect in our estimation, and accounts for the term $\psi_0\left(j\right)$ in equation (2.19). Importantly, as $R_B^{obs}\left(j, j_B, m, s\right)$ is obtained from a survey that restricts the set of business trips collected in the sample to those involving two census tracts j and j_B that are at least 50 miles away from each other, when estimating the parameters in equation (2.23), we restrict our estimation sample to all (including those with observed business trips equal to zero) possible pairs of tracts (j, j_B) such that their bilateral distance is at least 50 miles. Concerning the modes of transport, we consider three feasible modes of transport for business trips: car only, public transit and airplane.

As the expression in equation (2.23) illustrates, the parameters μ_B and θ^B are not separately identified from this estimating equation alone. We thus calibrate μ_B, the share of a firm's revenue spent on its employees' business travel, using external data sources. Specifically, we set $\mu_B = 0.1$ using information from industry reports.[20] Given this calibrated value of μ_B, we use a PPML estimator to compute the estimates of the remaining parameters entering equation (2.23), treating the origin-specific terms $\psi_B\left(j\right)$ for every tract j as

[20] See https://salestrip.com/insights/5-ways-businesses-can-avoid-wasting-money-on-travel-expenses/.

fixed effects.

We report our estimates in column (1) of Table 2.2 in the case where there is a single mode of travel. Our estimate of the coefficient on travel time, $\mu_B \theta^B \rho_B$, is approximately equal to -2. As discussed in Section 2.4.2.1, consistent with our interpretation of $e^s_B(j, j_B, m)$ as sampling noise that is mean zero and whose variance approaches zero as the sample size increases in census tract i, the estimates of all parameters in equation (2.23) converge to their corresponding true values as the sample size increases. Alternatively, as discussed in footnote 18, if we assume that the unobserved terms $e^s_B(j, j_B, m, s)$ are mean zero conditional on all observed covariates entering equation (2.23), and weakly correlated across the triplets (j, j_B, m), our estimates of all parameters in equation (2.23) converge to their corresponding true values as the set of destination-transport modes introduced in the analysis grows arbitrarily large.In our baseline estimation approach, we combine our estimate of the composite parameter $\mu_B \theta^B \rho_B$ with our calibrated value of μ_B and the assumption that ρ_B takes the same value as the parameter ρ_C (i.e., we set $\hat{\rho_B} = 1.4$, which corresponds to the estimate of ρ_C reported in column (8) of Table 2.1 to obtain the estimate $\hat{\theta}^B = 14$. We use these parameter estimates to construct model generated shares of workers in j who take business trips to tract j_B via mode m_B given state $s = NB$ from equation (2.47) and use these predicted shares in our quantification.

2.4.2.3 Leisure Travelers

Our model generates an equation for the total number of leisure trips that originate in census tract i, terminate in census tract j_L, and are done using mode of transit m in state s that has very similar structure to that included in equation (2.19) for business trips. Specifically, denoting as $R_L(i, j_L, m_L, s)$ the number of leisure trips from $i$to j_L

Table 2.2: Business and Leisure Trip Estimates

	(1) Business	(2) Leisure
log(travel time)	-2.022***	-1.779***
	(0.0656)	(0.0276)
log(in commuters)	4.019***	-2.242***
	(0.331)	(0.0876)
dest. management share	3.128***	
	(0.196)	
log(OD industry similarity)	0.0405*	
	(0.0175)	
dest. log(distance to beach)		-0.0705***
		(0.0109)
dest. has natl park		1.743***
		(0.123)
dest. hospitality share		6.819***
		(0.169)
dest. log(population)		-0.388***
		(0.0329)
N	9824665	29375132
$\theta^L \cdot \mu^L$	1.5	
$\theta^B \cdot \mu^B$		1.3

Source: California Household Travel Survey, 2010-2012.

using mode of transport m under state s, our model predicts that

$$R_L(i, j_L, m, s) = \psi_0^L(i) \left(\frac{q_L(i, j_L) B(j_L, s)}{d_L(i, j_L, m, s)} \right)^{\mu_L(i)\theta^L} p_L(i, j_L, m)^{-\mu_L(i)\theta^L - 1} \tag{2.23}$$

where $\psi_0^L(i)$ is an unobserved term that is origin specific, $q_L(i, j_L)$ is a shifter (expressed in monetary units) that operates as an amenity value for leisure trips that originate in census tract $i$and have tract j_L as destination, $d_L(i, j_L, m, s)$ is the monetary-equivalent cost of the time spent traveling for leisure from i to j_L through mode m under state s, $B(j_L, s)$ as the amenity value of j_L under state s and $p_L(i, j_L, m)$ is the direct monetary cost of traveling for leisure from i to j_L though mode m. We assume, similarly to the assumptions imposed for commuting and business travel, that

$$d_L(i, j_L, m, s) = D_L(i, m) \tau(i, j_L, m, s)^{\rho^L} \tag{2.24}$$

where $D_L(i,m)$ is an unobserved shifter and, $\tau(i,j_L,m)$ denotes travel time. Similarly to the approach followed when estimating the determinants of business trips in Section 2.4.2.2 we model the impact of the parameters $q_L(i,j_L)$ and $D_L(i,m)$ on the number of leisure trips between tracts i and j_L by mode of transport m as a function of observed covariates; specifically,

$$\theta_L\left(\ln(q_L(i,j_L)\,B(j_L,s)) - \ln(D_L(i,m)\right) = \gamma_L(m)X_L(i,j_L,s) \tag{2.25}$$

where $\gamma_L(m)$ is a model-specific parameter, and $X(i,j_L)$ is a vector that includes the following covariates: (a) an indicator for whether there is a national park in the destination j_l, (b) the distance of the destination j_L to the coastline, (c) the share of employment in the hospitality sector in the destination j_L, (d) the log of the population in the destination j_L and (e) the amenity term $B(j_L,s)$ as defined in (2.52).

For every pair of census tracts i and j_L, every mode of transport m, and every state s, we measure the number of leisure trips $R_L(i,j_L,m,s)$, travel time $\tau(i,j_L,m)$, and commuting costs $p_L(i,j_L,m)$ as indicated in Section 2.4.1. Allowing for sampling error in our measure of $R_L(i,j_L,m_L,s)$ for any given triplet (i,j_L,m_L) in state s, which we denote as $R_L^{obs}(i,j_L,m_L,s)$, we write

$$R_L^{obs}(i,j_L,m_L,s) = R_L(i,j_L,m_L,s) + e_L^s(i,j_L,m_L,s) \tag{2.26}$$

where $e_L^s(j,j_B,m)$ denotes the potential sampling error in our measure of $R_L(j,j_B,m)$. Combining equations (2.23) to (2.26), our resulting estimating equation is analogous to that for business trips in equation (2.23); i.e.,

$$R_L^{obs}(i,j_L,m_L,s) = \exp\big(\psi^L(i) + \gamma_L(m)\mu_L(i)X_L(i,j_L,s) -$$
$$\theta^L\rho_L\mu_L(i)\ln\tau(i,j_L,m_L) - (\mu_L(i)\theta^L+1)p_L(i,j_L,m_L) + e_L^s(i,j_L,m_L)\big)$$

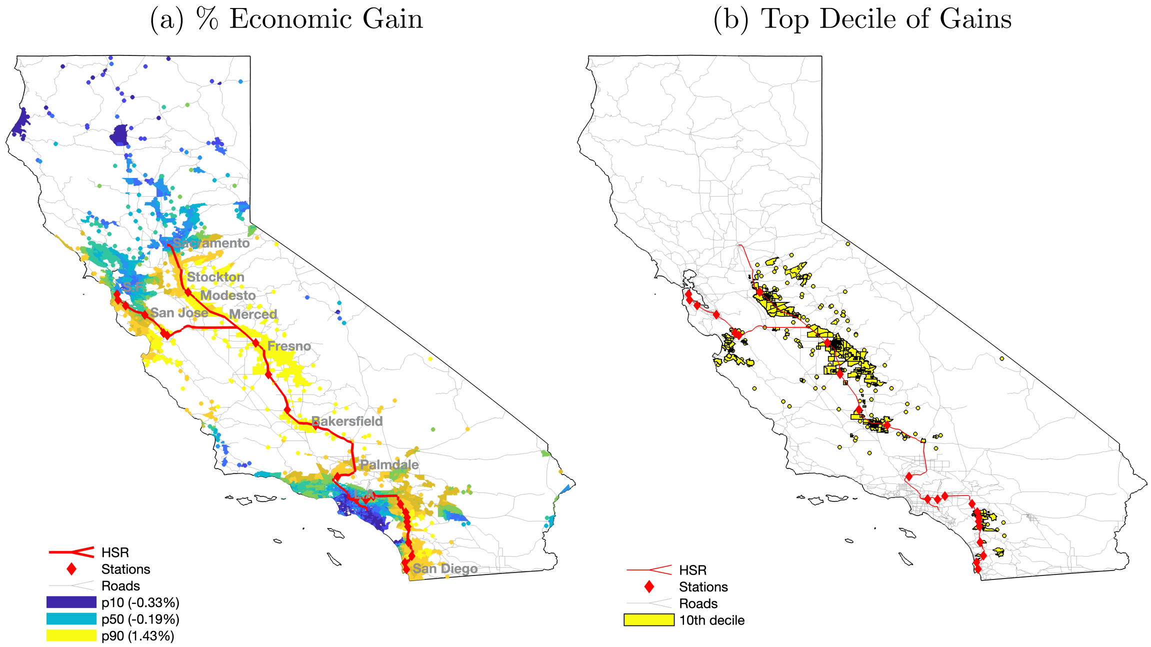
We assume that the share of income spent on leisure travel is constant across tracts, $\mu_L(i) = \mu_L$. Consistent with information from the Bureau of Labor Statistics, which reports average annual spending on travel – including on transportation, food away from home, and lodging – across U.S. households, we calibrate $\mu_L(i) = \mu_L = 0.05$. When we bring equation (2.27) to the data, we make the same restriction as in the business case: consistent with the survey, we exclude destinations that are more than 50 miles from the origin. We report our estimates of equation 2.23 in column (2) of Table 2.2. Our estimate of the composite parameter $\mu_L \theta^L \rho_L$ equals -1.78. Given our calibrated value of μ_L, and setting $\hat{\rho}_L = \hat{\rho}_C = 1.4$, we obtain an estimate of θ^L approximately equal to 26. We use these estimated coefficients to construct model-implied travel flows for leisure trips for each pair of tracts and each mode of travel. We then use our estimated parameters to construct model-implied leisure travel flows between each pair of census tracts as specified in equation (2.46) and we use these model-implied travel flows to measure travel between each pair of tracts (i, j_L) for each mode m_L given state $s = NB$.

2.4.3 Real-Income Effects of High-Speed Rail

Figure 2.4a shows the welfare gains from the HSR across Census tracts. The aggregate gain is 0.4%, on average \$176 per California voter; this figure falls to \$86 in the case that the cost of implementing the full HSR network reaches the current projection, which is about 2.5x the original projection. Only a very small share of commuters save time from the HSR (0.9%), while a much larger share of leisure (12.3%) and business (10%) travelers save time. The median commuter who saves time saves about 12 minutes, while the median long distance traveler who saves time saves about 20 minutes.

The distribution of welfare gains across Census tracts is highly skewed. Tracts in the top decile, ranked by welfare gain and shown in yellow in Figure 2.4b, each gain at least \$1,000 per year. The most significant winner is the Central Valley, where commuters with

Figure 2.4: Real Income Gains from the HSR



initially long commutes traveling north to the San Francisco bay area or south to the LA area benefit considerably. However, 75% of tracts gain less than \$44 per year and many even lose income due to the costs associated with building the HSR.

Of course, the welfare gains from the HSR vary when the cost of building the rail line changes. Our baseline estimation assumed the cost of building as stated in 2008, which was \$40 bn. But even as the cost has grown – the current estimate is around \$105 bn – the real income gains remain positive. Only once the total cost exceeds \$215 bn is net zero economic return from building the HSR.

2.5 Voter Preferences

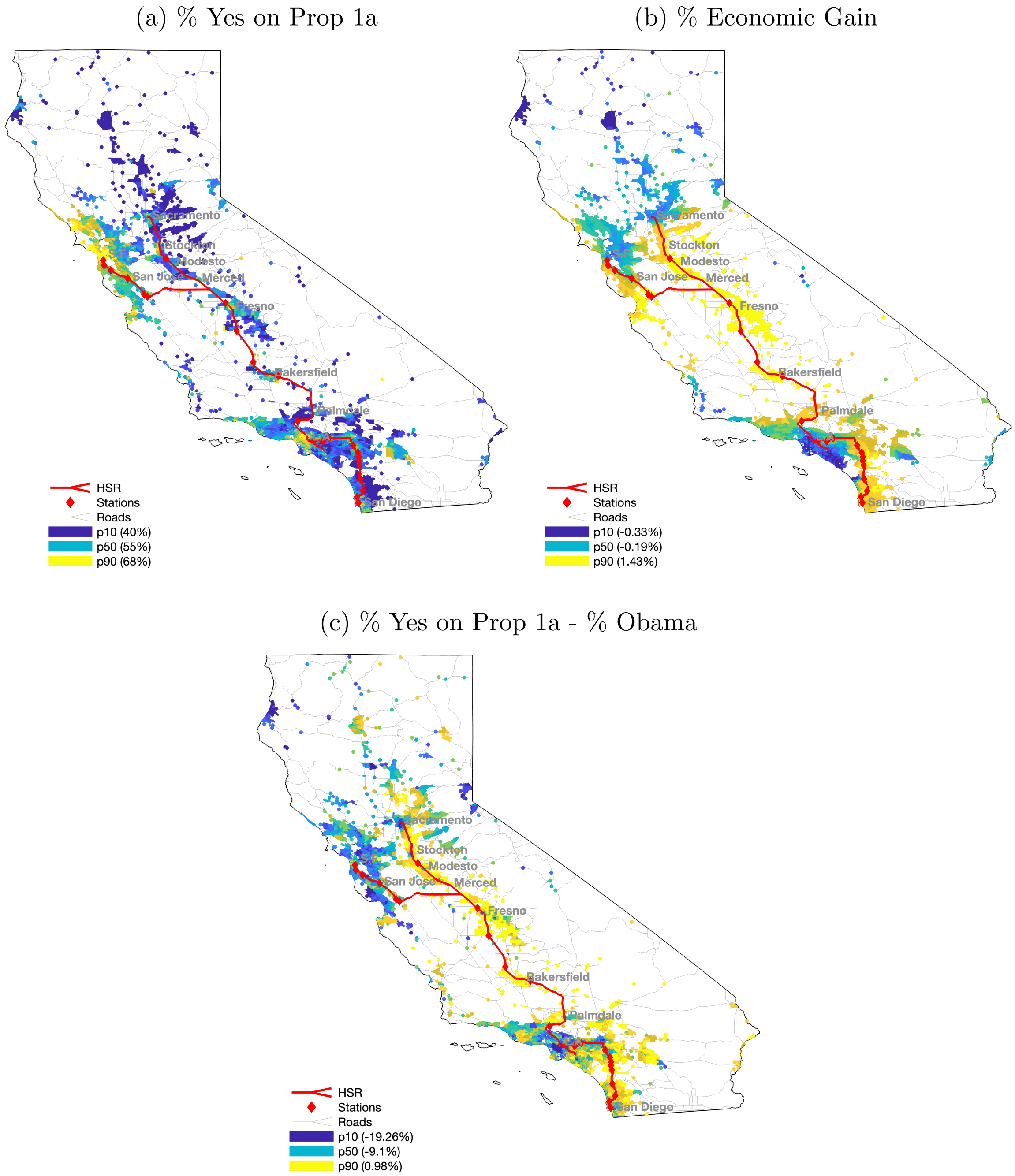
We next describe the procedure we follow to estimate the parameters $(\theta_V, \beta_1, \ldots, \beta_K)$ that enter equation (2.7) and govern the relationship between votes and different components of preferences. Importantly, our goal is to recover estimates of the parameters that may be used to compute the votes on the HSR if the economic or non-economic components of preferences has more or less impact.

Before turning to the model and estimation, we first present some suggestive evidence of the role of both economic and other factors in shaping preferences for infrastructure. Figure 2.5a again shows the share of voters voting in favor of the HSR across tracts. Comparing it with the economic gains as computed in Section 2.3.2 and shown in Figure 2.5b, there appears to be almost no positive correlation between economic gains and votes: if anything, there may be a negative correlation, as areas that gain relatively more economic (i.e., in the Central Valley) seem to be relatively less in favor of the rail line. But then, in Figure 2.5c, we plot, for each tract, the share of voters in each tract voting yes on the HSR *minus* the share of voters voting for Obama. Now, a much clearer relationship between votes and welfare gains appears: areas that were relatively more supportive of the HSR than of Obama were those who stood to gain economically from its creation. Taken together, this set of figures suggests a role for both politics (i.e., the tendency for areas to vote for a given political party) *and* economics in shaping transport preferences. This is precisely the relationship we seek to quantify in this section.

2.5.1 Estimation

As discussed in Section 2.3.1, the error term in equation (2.7) is the sum of two unobserved components. First, a term $\epsilon_W(i)$ that captures the expectational error the representative voter living in location i makes when forecasting the difference in the log of the average expected discounted flow of real-income it will experience depending on whether the high-speed rail, as described in Proposition 1a, is built. Second, a term $\epsilon_a(i)$ that captures all unobserved factors determining the political preferences of residents of census tract i for the high-speed rail.

Figure 2.5: Votes, Economic Gains, and Political Values



The structure of the error term in equation (2.7) implies that defining a procedure that yields consistent estimates of the parameter of interest θ_V requires dealing with two potential sources of endogeneity. First, if voters' expectations are rational, then, across all census tracts, the expectational error $\epsilon_W(i)$ will be correlated with the ex-post realized values of log real-income differences $\ln \hat{W}(i)$. Second, if the political body that decided on the final location of the high-speed railway tracks and stations did so either trying to maximize the overall support for the project or, for any other reason, trying to favor census tracts with particular values of the unobserved term $\epsilon_a(i)$, then this term will be also correlated with the ex-post realized values of log real-income differences $\ln \hat{W}(i)$. While the presence of the expectational error $\epsilon_W(i)$ in the error term will tend to bias the OLS estimate of θ_V towards zero (see Dickstein and Morales, 2018), the presence of the unobserved term $\epsilon_a(i)$ in the regression error term may bias the OLS estimates of θ_V upwards or downwards.

Addressing the omitted variable bias arising from the correlation between $\epsilon_W(i)$ and $\ln \hat{W}(i)$ requires an instrumental variable estimation approach in which we use as a instrument a function of variables belonging to voters' information set at the time of the vote on Proposition 1a (2008). A potential candidate for such instrument is the welfare impact that voters living in any given census tract would have experienced if the fundamentals had remained constant at their 2008 values. Given the assumption that voters' expectations are rational, this instrument would be mean independent of voters' expectational error $\epsilon_W(i)$ as long as voters knew in 2008 the change in real income they would have experienced if the proposed high-speed rail had started to operate at that point.

To address the potential omitted variable bias arising from correlation between $\epsilon_a(i)$ and $\ln \hat{W}(i)$, we exploit information on the technically feasible routes described in Section 2.2 and shown in Figure 2.2. As we discuss in this section, engineers from the HSR authority identified in 1996 three possible routes to connect SF to LA by a high-speed rail.

The determination of these three routes was based on technical feasibility and cost-saving consideration and, thus, was not *directly* impacted by the political preferences that voters in any given census tract had for the high-speed rail project; i.e., was not directly impacted by $\epsilon_a(i)$. We thus would like to build an instrument that exclusively exploits information on these three planned routes, and that incorporates information neither on the route that was finally chosen nor on the actual location of the planned stations, as these choices may have been impacted by the will of political agents to favor particular census tracts.

We construct this instrument by computing the average welfare gain (or loss) of residents of each census tract across 100 simulated high-speed railway networks drawn randomly from the set of all possible networks consistent with the three technically feasible routes connecting SF to LA. Formally, we construct our instrument as

$$z_i = \frac{1}{100} \sum_{n=1}^{100} log(\hat{W_{i,n}^{fs}}) \tag{2.27}$$

where $\hat{W_{i,n}^{fs}}$ is the 2008 change in welfare in location i from a counterfactual high-speed railway network that has 24 stations randomly allocated among all possible locations along the three feasible routes described in Section 2.2. This procedure is described in detail in Appendix 2.9.3.[21]

The instrumental variable z_i defined in equation (2.27) exclusively uses information available to voters in 2008 and, thus, the assumption that voters expectations are rational implies that it will be mean independent of the expectational error term $\epsilon_W(i)$. This instrumental variable also only uses information on a set of possible high-speed railway networks defined exclusively by their technical feasibility and, thus, is not a direct function

[21]Specifically, we first randomly select one of the three feasible routes. Then, we randomly select a set of 24 points along the randomly selected route to serve as stations. We assume that the train travels between these randomly allocated stations at a speed that equals the average travel speed of the actual HSR.

the unobserved political preferences accounted for by the unobserved term $\epsilon_a(i)$. We cannot rule out however that there is a spurious correlation between this unobserved term and the instrument in equation (2.27); this would be the case if, e.g., residents of the census tracts located close to any of the three feasible routes identified in 1996 (shown in Figure 2.2) have on average different values of the term $\epsilon_a(i)$ than those tracts located far away from any of these feasible routes.

We report OLS and IV estimates of the parameter vector $(\theta_V, \beta_1, \ldots, \beta_K)$ in Tables 2.3 and 2.4, respectively. All throughout, we report Conley (1999) standard errors to account for spatial correlation in the data.[22] In the different columns reported in Table 2.3, we progressively increase the number of controls we include in our regression specification. In column (1), we include no controls, obtaining an estimate of θ_V that is negative but not statistically significant. In column (2), to control for differences in political ideology across census tracts, we add a control for the log-odds ratio of the share of voters in each census tract that were registered democrats at the time of the 2008 election. Relative to column (1), adding this controls makes the R^2 increase from 0.07% to nearly 60%, and makes the OLS estimate of θ_V become positive, though still noisily measured. In column (3), we additionally control for a proxy of the extent to which voters in a location value the environment – the log-odds of votes in favor of Proposition 10. [23]

Column (4) additionally controls for a proxy of general support for transportation projects – the log share of votes in support of Proposition 1B. [24] Finally, we introduce in column (5) a series of controls for population density, as the population density of a

[22] In our baseline results, we use a 25km bandwidth. In robustness checks, we use 50km and 100km bandwidths.

[23] Proposition 10 was on the ballot in the same election as the HSR proposition and, if passed, would have allowed the state to issue \$5 billion in bonds for alternative fuel projects (Ballotopedia).

[24] Proposition 1B was on the ballot during the November 2006 Midterm elections, and would have allowed the state to issue \$19.9 billion in bonds for transportation projects (Ballotopedia).

location likely affected whether a station was placed near it, and is likely correlated with the support that Proposition 1a obtained for ideological reasons. Specifically, we control in column (5) for population density in each census tract, which we denote as $popdens_i$, as well as for an inverse-distance weighted average of population density in the surrounding tracts, which we denote as $avgma_i$.[25] In the specification with the largest set of controls, the estimate value of θ_V equals 5.13.

The first column in Table 2.4 reproduces those in column (5) of Table 2.3, to facilitate the comparison of the OLS and IV estimates. Columns (2) and (3) present the first- and second-stage estimates, respectively, for an IV estimator that uses as instrument the welfare impact that voters living in any given census tract would have experienced if the fundamentals had remained constant at their 2008 values. As discussed above, this instrument exclusively address possible endogeneity due to the expectational error term $\epsilon_W(i)$. There is only a small change in the resulting IV estimate of θ_V relative to the OLS estimates. Columns (4) and (5) present the first- and second-stage estimates, respectively, for the IV estimator that uses as an instrument z_i as introduced in equation (2.27). The F-stat associated with the first-stage estimate of the coefficient on the instrument is close to 30, and the resulting IV estimate of θ_V equals 5.69. We use this number as our baseline estimate of θ_V in the subsequent analysis. While these results use the baseline model, Tables 2.7 and 2.8 in Appendix 2.8 presents the same results under the full model.

2.5.2 Political versus Economic Preferences for the HSR

How are important are real income gains from the HSR as compared with other factors, like political preferences, in determining votes for the HSR? To quantify this, we use our model of voting, plus our estimates of each of the components that determine votes including

[25] This is computed as: $avgma_i = \sum_j 1(i \neq i) \cdot \tau(i,j,NB)^{-1} \cdot popdens_j$.

Table 2.3: OLS Estimates of Voting Equation

	(1)	(2)	(3)	(4)	(5)
$\log(\hat{W})$, 2019	-4.617	2.456	2.943	2.834	5.128*
	(4.267)	(2.122)	(1.800)	(1.681)	(2.056)
log-odds Reg Dems Share		0.665***	0.462***	0.455***	0.414***
		(0.0725)	(0.108)	(0.125)	(0.0814)
Environment: log odds Yes on Prop. 10			0.460***	0.430***	0.360***
			(0.0876)	(0.0541)	(0.0519)
Transportation: log odds Yes on Prop. 1b				0.0508	-0.00350
				(0.156)	(0.139)
pop. density					0.0469***
					(0.0103)
Avg. MA Pop. Dens.					0.0457
					(0.127)
Constant	0.203*	0.275***	-0.0165	0.0145	-0.185
	(0.0943)	(0.0470)	(0.100)	(0.0557)	(0.294)
R2	0.00753	0.589	0.658	0.659	0.713
N	7866	7866	7866	7866	7866

Note: This table shows OLS estimates of equation 2.7. Proposition 10 was on the ballot in the same election as the HSR proposition and, if passed, would have allowed the state to issue $5 billion in bonds for alternative fuel projects (Ballotopedia). Proposition 1b, on the ballot in 2006, allocated funds towards transportation projects including highways.

Table 2.4: IV Estimates of Voting Equation without Spillovers

		2008 IV		Random Route IV	
	(1)	(2)	(3)	(4)	(5)
	OLS	FS	IV	FS	IV
$\log(\hat{W})$, 2019	5.128*		4.865*		5.686**
	(2.056)		(2.011)		(2.009)
$\log(\hat{W})$, 2008		0.518***			
		(0.00600)			
$\log(\hat{W})$, Random Route IV, 2008				0.565***	
				(0.103)	
log-odds Reg Dems Share	0.414***	-0.0000745	0.414***	-0.00121	0.414***
	(0.0814)	(0.0000641)	(0.0814)	(0.000741)	(0.0815)
Environment: log odds Yes on Prop. 10	0.360***	0.0000108	0.360***	0.00202	0.359***
	(0.0519)	(0.0000787)	(0.0521)	(0.00136)	(0.0514)
Transportation: log odds Yes on Prop. 1b	-0.00350	0.000105	-0.00318	0.00141	-0.00418
	(0.139)	(0.0000789)	(0.139)	(0.000910)	(0.138)
pop. density	0.0469***	0.00000981	0.0469***	0.0000613	0.0468***
	(0.0103)	(0.00000761)	(0.0103)	(0.000107)	(0.0102)
Avg. MA Pop. Dens.	0.0457	-0.000398*	0.0423	-0.00593*	0.0528
	(0.127)	(0.000160)	(0.127)	(0.00262)	(0.122)
Constant	-0.185	0.00305***	-0.176	0.0146*	-0.204
	(0.294)	(0.000367)	(0.293)	(0.00569)	(0.279)
First stage F-stat			7391.1		29.80
R2	0.713	0.995	0.713	0.560	0.713
N	7866	7866	7866	7866	7866

Note: This table shows IV estimates of equation 2.7. The 2008 IV uses welfare measured using 2008 fundamentals as an instrument for expected welfare, and the random route IV uses an instrument constructed using fake HSRs, as described in detail in Appendix section 2.9.3. Proposition 10 was on the ballot in the same election as the HSR proposition and, if passed, would have allowed the state to issue $5 billion in bonds for alternative fuel projects (Ballotopedia). Proposition 1b, on the ballot in 2006, allocated funds towards transportation projects including highways.

$\hat{\theta}_V$, $\ln(\hat{a}(i))$ and $\hat{W}(i)$ to ask counterfactual questions. First, we quantify the aggregate vote share in favor of the HSR in the case where economic preferences are the only factor influencing votes; in other words, we ask, if people had been voting *only based on the economic effects of the HSR*, ignoring any other preferences, what would the aggregate vote on the HSR have been?

To implement this, we take our baseline estimate of $\hat{\theta}_V = 5.7$ and $\hat{W}(i)$ and set $\ln(\hat{a}(i)) = 1$ and $\epsilon(i) = 0$. Thus, the economic effects will be the only factor affecting votes. In this case, we find that the aggregate vote would have been 50.6%, which is far lower than the actual vote of 52.6% in favor of the HSR. In fact, the economic component can explain less than one-quarter of the aggregate vote in favor of the HSR. A large segment of the vote in favor of the HSR (77%) cannot be explained by economic preferences alone. In a related but slightly different counterfactual, we ask what the aggregate vote share would have looked like if agents had voted only on the non-economic component of preferences. In this case, we set $\hat{W}(i) = 1$ and $\epsilon(i) = 0$ with $\ln(\hat{a}(i)) \neq 1$ and we find that removing the economic component of votes and allowing only the non-economic component to be a factor has a small effect on the aggregate vote share, which falls from 52.6% to 51.8%. In this case, we find a similar result: the non-economic component can explain more than two-thirds of the aggregate vote in favor of the HSR.[26]

We then compute how much of the vote in favor the non-economic preferences component can explain, relative to the economic component. To do this, we compute $v(\hat{a}) - 50 = 1.8$ where $v(\hat{a})$ is the aggregate vote when $\hat{W}(i) = 1$ and $\epsilon(i) = 0$. We then compute the corresponding term for the only-economic component: $v(\hat{w}) - 50 = 0.8$. Dividing these two terms, we find that the non-economic component can explain 3.2 times more of the aggregate vote in favor than the economic-only component can explain.

[26]This figure is computed as: $\frac{51.8-50}{52.6-50} \cdot 100$, and measures the percentage of the aggregate vote share in favor of the HSR (2.6%) is explained by the non-economic component.

We consider these outcomes under several different parameterizations of the model in Table 2.5. In addition to the baseline, no GE case, we consider the case with GE effects. Instead of assuming a power function for travel costs, we instead assume an exponential function. To address some potential for model mispecification, we consider a version where we exclude Census tracts within 5km of the HSR route. In this case, we do not estimate θ_V off of Census tracts where we may have failed to capture forces such as the noise disamenity from living very close to the train tracks. We also consider a "pessimistic" case, where the costs of the HSR are assumed to be higher and the probability of completion is assumed to be lower.

Table 2.5: Determinants of Voter Preferences: Alternative Parameterizations

	θ_V	Average % Gain	Median % Gain	Only $\hat{a}$	Only $\hat{W}$ vote	$\frac{v(\hat{a})-50}{v(\hat{w})-50}$
Baseline	5.69	0.42	0.10	51.84	50.58	3.2
Baseline w/ 5km buffer	5.46	0.42	0.10	53.42	50.56	6.1
Full model	7.63	0.36	0.10	51.73	50.68	2.5
Full model w/ 5km buffer	7.07	0.36	0.10	53.06	50.63	4.9
Baseline+vary ρ	1.53	2.59	0.83	51.50	50.92	1.6
Baseline+no air	12.20	1.12	0.91	48.95	53.53	-0.3
Baseline+pessimistic	10.70	0.21	0.04	51.86	50.55	3.4
Baseline+exponential	2.36	1.80	0.47	51.43	50.98	1.5

Note: This table shows results under different assumptions. "No GE" corresponds to the case without spillovers and without land. "GE" corresponds to the case with spillovers and land. The "5km" buffer rows estimate θ_V excluding tracts within 5km from the HSR tracks.

With the exception of the no air case, in which welfare gains are much larger in magnitude since the initial commute time along many long routes is far higher than when we allow for air travel, the story is very consistent. The HSR would have have passed even if voters put no weight on the economic component of the HSR because their non-economic, including political, preferences, are sufficiently strongly in support of the HSR.

2.6 Conclusion

This paper uses the setting of California's High Speed Rail to study how preferences shape the spatial distribution of transportation infrastructure. To do this, we first build a model of voters' choices over whether or not the rail should be built, where the choice depends on both the economic impacts of the rail and on other factors. To measure the economic gains in each location, we develop a sophisticated mode which captures not only time saved on commuting trips, on leisure trips, and on business trips but also the general equilibrium impacts of the HSR on local development through its impact on land rents and wages. We estimate the model using granular, tract-level data on commuting patterns, leisure and business travel patterns, wages, floor space, and votes on the high speed rail, as well as on other ballot measures. To account for the endogeneity associated with expectational errors as well as the placement of the actual HSR network, we use an instrumental variables strategy.

Using this framework, we show that political and other non-economic preferences play a much larger role than real income gains in determining people's preferences for infrastructure investment. We find that the non-economic component of preferences can explain around three times as much of the variation in votes across locations as much the economic component can explain. Our results have important implications for which transport projects may ultimately be approved, and suggest that, despite the focus on the literature on the real income effects of transport, the projects that are ultimately implemented may not be the ones that maximize real income.

2.7 Model Appendix

We model an economy with a set $\mathcal{J}$ of tracts, each tract i with a fixed resident population $N_R(i)$ and connected to other tracts by various transport modes. Residents consume a traded good, a non-traded service, floor space and leisure trips. They choose where to commute to work, where to take leisure trips, how many such trips to make, and what mode of transportation to use for each of these travel purposes.

Traded good firms produce using labor, floor space, and productivity-enhancing business trips with constant returns to scale; and they choose where to send workers to business trips, how many such trips to make, and what mode of transportation to use. Non-traded services are produced 1-1 from labor. The HSR network is financed with a fixed investment with income taxes. The transport sector of the economy operates constant-returns technologies using the tradeable good as input, with ticket prices covering the price of each trip.

The discrete choices of destinations and modes are governed by idiosyncratic shocks to resident's preferences and tradeable firm's productivity, and by the heteogeneous appeal of destination tracts. Specifically, tracts are heterogeneous in terms of their fundamental productivity, amenities, and stock of floor space. Furthermore, residents living in different tracts are allowed to be heterogeneous in terms of their preferences over modes of travel and for specific leisure destinations. They may also differ in their drivers of income, namely the efficiency units of labor that they supply and their rate of ownership of the local floor space. The implementation of the HSR endows the economy with an option of making faster (and possibly more expensive) trips along some routes compared to the status quo.

In the presentation of the model, variables that are indexed by s may change either endogenously or exogenously based on whether the HSR is built ($s = B$) or not ($s = NB$).

2.7.1 Preferences

When the HSR status is s, the utility U_ω of an individual ω living in tract i who travels to j_C for commuting and to j_L for leisure, by transport modes m_C and m_L respectively, is:

$$U_\omega\left(i,j_C,m_C,j_L,m_L,s\right)=\max_{C,H,S,R_L} B\left(i,s\right)\frac{C^{1-\mu_L(i)-\mu_H(i)-\mu_S(i)}H^{\mu_H(i)}S^{\mu_S(i)}}{d_C\left(i,j_C,m_C,s\right)}\left(\frac{q_L\left(i,j_L\right)B\left(j_L,s\right)}{d_L\left(i,j_L,m_L,s\right)}R_L\right)^{\mu_L(i)}\varepsilon_\omega^C\left(j_C,m_C\right)\varepsilon \tag{2.28}$$

subject to the budget constraint:

$$C+r\left(i,s\right)H+p_S\left(i,s\right)S+p_L\left(i,j_L,m_L,s\right)R_L+p_R\left(i,j_C,m_C,s\right)R_C=\left(1-t\left(s\right)\right)y\left(i,j_C,s\right) \tag{2.29}$$

Expression 2.28 indicates that consumers derive utility from the amenities of their place of residence $B\left(i,s\right)$, as well as from the consumption of tradeable commodities C, housing H, services S, and leisure trips R_L, with Cobb-Douglas shares that we allow to be location-specific. The amenity term may respond endogenously to the local density of economic activity as detailed below. This heterogeneity captures in a reduced-form way the fact that workers in different tracts in California may have different spending patterns on leisure trips and housing, for example due to different demographic characteristics. These workers face a disutility of daily commuting travel $d_C\left(i,j_C,m_C,s\right)$. The utility they derive from leisure trips depends negatively on time travelled $d_L\left(i,j_L,m_L,s\right)$ and positively on the quality of the destination visited, which is a composite of an exogenous origin-destination component $q_L\left(i,j_L\right)$ (capturing, for example, that residents of some locations may on average be more likely to have relatives in some other specific locations) and the destination-specific amenity $B\left(j_L,s\right)$.

In our baseline, the utility cost of travel is a power function of travel time $\tau\left(i,j,s,m\right)$

and depends on travel mode for both commuting ($k = C$) and leisure ($k = L$):

$$d_k\left(i, j, s, m\right) = D\left(i, m\right) \tau\left(i, j, s, m\right)^{\rho^k} \text{ for } k = C, L, \tag{2.30}$$

where ρ^k is the elasticity of travel disutility to travel time, and where $D\left(i, m\right)$ is a location-specific preference for traveling through transport mode m. This term captures in a reduced-form way the fact that workers in different tracts in California may have different tastes for different modes of travel, such as a preference for using cars over public transit.

The last two terms of (2.28), $\varepsilon_\omega^C\left(j_C, m_C\right)$ and $\varepsilon_\omega^L\left(j_L, m_L\right)$, are idiosyncratic preference shocks for commuting and leisure travel to each destination by each travel mode. We assume them to be IID Type-I extreme value distributed:

$$\Pr\left(\varepsilon_\omega^k\left(j_k, m_k\right) < x\right) = e^{-e^{-\theta^k x}} \text{ for } k = C, L, \tag{2.31}$$

where θ^k is the (inverse) of the dispersion of shocks across travel modes and destinations for travel purpose k.

We turn now to the budget constraint 2.29. On the expenditure side, the price per unit of tradeable commodities (C) is normalized to 1 and the cost per unit of floor space for housing (H) is $r\left(i, s\right)$. The monetary cost of traveling from i to j through means m in state s is $p_R\left(i, j_L, m_L, s\right)$, regardless of whether the trip is for commuting or leisure. In each leisure trip, people spend this travel cost and consume c_L units of services at destination, at cost $p_S\left(j_L, s\right)$. Hence, the cost per leisure trip is:

$$p_L(i, j, m, s) = p_R\left(i, j_L, m_L, s\right) + c_L p_S\left(j_L, s\right). \tag{2.32}$$

Regular commuters pay $p_R\left(i, j_C, m_C, s\right)$ per trip from i to j_C through m_C, with the annual cost multiplying this per-trip cost by the annual number of trips (i.e., the working days) R_C.[27] Because non-tradeables are produced 1-1 from labor, the cost of each unit of services is $p_S(i,s) = w(i,s)$. The resulting demand system is quasi-homothetic, with homothetic demand over C, H, S, and R_L after spending $p_C(i, j_C, m_C, s)$ annually on commuting.

The income side of (2.29) equals pre-tax income $y(i, j_C, s)$ net of the tax rate $t(s)$ .The tax equals t if the HSR is approved and 0 otherwise. Pre-tax income is driven by two sources, labor and home ownership:

$$y(i, j_C, s) \equiv e(i)\, w(j_C, s) + \eta(i)\, r(i, s)\,. \tag{2.33}$$

The first term measures labor income. We assume that the returns to labor income equal the efficiency units of residents of tract i, $e(i)$, times the wage per efficiency unit at destination, $w(j_C, s)$. So, within an origin tract, commuters to different destinations earn different wages based on $w(j_C, s)$; and, across origin tracts, commuters to the same destination earn different wages based on $e(i)$. The second term in (2.33) measures returns to home ownership. $\eta(i)$ is the share of the floor space per resident in i that is owned by residents on i. So, an increase in land rents $r(i,s)$ reduces the real income of tract-i residents through the cost of housing, and increases it through the returns to land as a function of $\eta(i)$.

Maximizing out the solutions for consumption C, housing H, services S, and number

[27] This formulation already imposes that, throughout a year, each invidual ω within a tract chooses one destination for work and for leisure. This is consistent with assuming that agents receive idiosyncratic shocks once per year and that the total utility associated with making trips for a given purpose to multiple destinations is additive across destinations, so that only the best destination for each purpose is chosen each year.

of trips R, the solution to (2.28) gives indirect utility:

$$V_\omega\left(i, j_C, m_C, j_L, m_L, s\right) = \frac{B\left(i,s\right)}{r\left(i,s\right)^{\mu_H(i)}\, p_S\left(i,s\right)^{\mu_S(i)}} \frac{I\left(i, j_C, m_C, s\right)}{d_C\left(i, j_C, m_C, s\right)} \left(\frac{q_L\left(i, j_L\right) B\left(j_L, s\right)}{p_L\left(i, j_L, m_L, s\right) d_L\left(i, j_L, m_L, s\right)}\right)^{\mu_L(i)} \varepsilon_\omega^C\left(j_C, m_C\right) \varepsilon \tag{2.34}$$

where

$$I\left(i, j_C, m_C, s\right) \equiv \left(1 - t\left(s\right)\right) y\left(i, j_C, s\right) - p_C\left(i, j_C, m_C, s\right). \tag{2.35}$$

is the disposable income net of taxes and commuting costs.

2.7.2 Welfare

Each resident ω makes discrete choices of destination and transport mode for both commuting and leisure that maximizes indirect utility. These choices are represented by the quadruplet $\{j_C, j_L, m_C, m_L\}$. Destinations are chosen from the set of tracts $\mathcal{J}$ while the set of transport modes for travel purpose $k = L, C$ is $\mathcal{M}_k$.[28] The average yearly real income of tract-i residents is defined as the expected value of indirect utility across the realizations of the ε_ω^C and ε_ω^L preference shocks, that is:

$$V\left(i, s\right) = \mathbb{E}_\omega\left[\max_{\left(j_C, j_L m_C, m_L\right) \in \mathcal{J}^2 \times \mathcal{M}_C \times \mathcal{M}_L} V_\omega\left(i, j_C, m_C, j_L, m_L, s\right)\right]. \tag{2.36}$$

Using standard properties of the extreme-value distributions for the shocks ε_ω^C and ε_ω^L together with (2.34), we can write this average real income simply as:

$$V\left(i, s\right) = \frac{W^C\left(i, s\right)}{P_L\left(i, s\right)^{\mu_L(i)}} \frac{B\left(i, s\right)}{r\left(i, s\right)^{\mu_H(i)}\, P_S\left(i, s\right)^{\mu_S(i)}} \tag{2.37}$$

[28] In our baseline, transport modes for commuting are $\mathcal{M}_C$ = {car,public transit,walking or biking} and for leisure they are $\mathcal{M}_L$ = {car,public transit,air}. Introducing the HSR means an upgrade of the public-transit mode.

where $W^C(i,s)$ captures average income net of commuting costs of residents of i:

$$W^C(i,s) = \left(\sum_{j_C \in \mathcal{J}} \sum_{m_C \in \mathcal{M}_C} \left(\frac{I(i,j_C,m_C,s)}{d_C(i,j_C,m_C,s)} \right)^{\theta_C} \right)^{\frac{1}{\theta_C}}. \tag{2.38}$$

$P_L(i,s)$is akin to a quality-adjusted price index for leisure trips for residents of i, net of travel costs:

$$P_L(i,s) = \left(\sum_{j_L} \sum_{m_L \in \mathcal{M}_L} \left(\frac{p_L(i,j_L,m_L,s)\, d_L(i,j_L,m_L,s)}{q_L(i,j_L)\, B(j_L,s)} \right)^{-\theta_L \mu_L(i)} \right)^{-\frac{1}{\theta_L \mu_L(i)}}, \tag{2.39}$$

Finally, average (amenity-adjusted) real income also depends on local amenities $B(i,s)$, land prices $r(i,s)$, and service prices $P_S(i,s)$.

2.7.3 Tradeble Sector Firms

In the tradeable sector, we assume a measure 1 of firms in each tract. Tracts differ in their productivity $A(j,s)$. Each firm uses foor space H_Y, labor N_Y, and business trips R_B as inputs. A firm ω sending workers on a number R_B of business trips to destination j_B using transport mode m_B produces output according to the Cobb-Douglas production function:

$$Y_\omega(j,R_B,j_B,m_B,s) = A(j,s)\, H_Y^{\mu_{H_Y}(j)} N_Y^{1-\mu_{H_Y}(j)-\mu_B(j)} \left(\frac{q_B(j,j_B)\, A(j_B,s)}{d_B(j,j_B,m_B,s)} R_B \right)^{\mu_B(j)} \varepsilon_\omega^B(j_B,m_B). \tag{2.40}$$

Note first that production functions are allowed to be tract-specific, through heterogeneous Cobb-Douglas shares. Second, business trips are productivity-enhancing, capturing for example that they promote new supplier/customer relationships. Specifically, the returns to business trips depend on the productivity of the destination ($A(j_B,s)$), on an exogenous origin-destination productivity match $q_B(j,j_B)$ (capturing that firms in some locations may on average be more likely to find business partners in some other specific location), and negatively on time traveled (captured by $d_B(j, j_B, m_B, s)$). Finally, the return to business trips also depend on an idiosyncratic productivity shock $\varepsilon_\omega^B(j_B, m_B)$ for the destination and travel mode for these business trips. We assume them to be IID Type-I extreme value distributed:

$$\Pr\left(\varepsilon_\omega^B(j_B, m_B) < x\right) = e^{-e^{-\theta_B x}}, \tag{2.41}$$

where θ_B is the (inverse) of the dispersion of shocks across travel modes and destinations for business travel.

We assume that firms hire labor and floor space before observing realizations of these idiosyncratic business opportunities, and that they choose the business trip destination (from the set of locations $\mathcal{J}$), the mode of transport (from the set of available modes $\mathcal{M}_B$), and the number of trips R_B afterwards. Hence, the firm solves the problem:

$$\Pi = \max_{N_Y, H_Y} \underbrace{\mathbb{E}\left[\max_{(R_B, j_B, m_B) \in (\mathbb{R}^+ \times \mathcal{J} \times \mathcal{M}_B)} Y_\omega(j, R_B, j_B, m_B, s) - p_R(j, j_B, m_B, s)\, R_B\right]}_{\equiv Y(j,s)} - w(j,s)\, N_Y - r(j,s)\, H_Y, \tag{2.42}$$

where $p_R(j, j_B, m_B, s)$ is the monetary cost per business trip. Because conditional on floor space and labor there are decreasing returns to the number of trips, we can solve for the number of trips R_B, plug them back into the expectation and then integrate over realization of idiosyncratic business shocks using standard properties of the extreme value distribution defined in (2.41). After these steps, the expected output net of business costs defined in (2.42) can be written:

$$Y(j,s) \equiv \Omega(j,s)\, H_Y^{\frac{\mu_{H_Y}(j)}{1-\mu_B(j)}} N_Y^{\frac{1-\mu_{H_Y}(j)-\mu_B(j)}{1-\mu_B(i)}}, \tag{2.43}$$

where $\Omega(j,s)$ is an endgenous TFP term that depends on both the TFP of the location

$A(j,s)$ and the distribution of business travel opportunities:

$$\Omega(j,s) \equiv \kappa_B(j) A(j,s)^{\frac{1}{1-\mu_B(j)}} \left(\sum_{j'} \sum_{m \in M_{ij}^B(s)} \left(\frac{q_B(j,j') A(j',s)}{p_B(j,j',m_B,s) d_B(j,j',m_B,s)} \right)^{\theta_B \mu_B(j)} \right)^{\frac{1}{(1-\mu_B(j))\theta_B}}, \tag{2.44}$$

where we have denoted $\kappa_B(j) \equiv \mu_B(j)^{\frac{\mu_B(j)}{1-\mu_B(j)}} (1-\mu_B(j))$.

2.7.4 Travel Choices

The travel decisions of workers and firms imply equations for shares and numbers of trips taken to a given destination that we use to estimate key parameters of the model. Specifically, using standard properties of the extreme-value distributions for the shocks ε_ω^C and ε_ω^L, the solution to (2.36) gives the fraction of residents from i that commute to j_C using transport mode m_C,

$$\lambda^C(i,j_C,m_C,s) = \frac{\left(\frac{I(i,j_C,m_C,s)}{d_C(i,j_C,m_C,s)}\right)^{\theta_C}}{\sum_j \sum_m \left(\frac{I(i,j,m,s)}{d_C(i,j,m,s)}\right)^{\theta_C}}, \tag{2.45}$$

as well as the fraction of residents from i that travel for leisure to j_K through transport mode m_L:

$$\lambda^L(i,j_L,m_L,s) = \frac{\left(\frac{q_L(i,j_L)B(j_L,s)}{d_L(i,j_L,m_L,s)p_L(i,j_L,m_L,s)}\right)^{\mu_L(i)\theta_L}}{\sum_j \sum_m \left(\frac{q_L(i,j)B(j,s)}{d_L(i,j,m,s)p_L(i,j,m,s)}\right)^{\mu_L(i)\theta_L}}. \tag{2.46}$$

Similarly, from the solution to the firm's problem in 2.42 and using standard properties of the extreme value shocks ε_ω^B, the fraction of firms from j sending workers on business trips to j_B takes the same functional form as (2.46):

$$\lambda^B(j, j_B, m_B, s) = \frac{\left(\frac{q_B(j,j_B)A(j_B,s)}{d_B(j,j_B,m,s)p_R(j,j_B,m_B,s)}\right)^{\mu_B(j)\theta_B}}{\sum_{j'}\sum_m \left(\frac{q_B(j,j')A(j',s)}{d_B(j,j',m,s)p_R(j,j',m,s)}\right)^{\mu_B(j)\theta_B}}. \tag{2.47}$$

The last two expressions measure shares of travelers (or firms) making leisure (or business trips) to a destination. In the data, we observe the number of trips to each destination by travel purpose. In the model, the number of leisure trips from i to j_L through means m_L depends not only on the fraction of travelers but also on the intensity of travel. Adding up the optimal choice of number of trips across residents of i, we obtain that the number of leisure trips from i to j_L to m_L is:

$$R_L(i, j_L, m_L, s) = \lambda^L(j, j_L, m_B, s)\,\mu_L(i)\,\frac{N_R(i)\,\overline{I}(i)}{p_L(i, j_L, m_L, s)}, \tag{2.48}$$

where $\overline{I}(i)$ is the average income net of commuting expenditures of location i's residents, itself a function of where they commute for work:

$$\overline{I}(i) \equiv \sum_{j_C}\sum_{m_C}\lambda^C(i, j_C, m_C, s)\,I(i, j_C, m_C, s). \tag{2.49}$$

Similarly, from the solution for R_B from (2.42), the total number of business trips from j to j_B through mode m_B is:

$$R_B(j, j_B, m_B, s) = \lambda^B(j, j_B, m, s)\,\frac{\mu_B(j)}{1-\mu_B(j)}\,\frac{Y(j,s)}{p_B(j, j_B, m_B, s)}. \tag{2.50}$$

2.7.5 Spillovers

We allow for firm productivity and for residential amenities to respond endogeneously to the level of local activity. Specifically, we use similar functional forms as Ahlfeldt et al.

(2015) and assume that spillovers respond to the density of workers in the location and in the surroundings:

$$A(j,s) = Z_A(j)\left(\frac{\tilde{N}_Y(j,s)}{H(j)} + \sum_{k\neq j} e^{-\rho_A^{\text{spillover}}\tau^{min}(j,k,s)}\frac{\tilde{N}_Y(k,s)}{H(k)}\right)^{\gamma_A^{\text{spillover}}} \tag{2.51}$$

$$B(i,s) = Z_B(i)\left(\frac{\tilde{N}_Y(i,s)}{H(i)} + \sum_{k\neq i} e^{-\rho_B^{\text{spillover}}\tau^{min}(j,k,s)}\frac{\tilde{N}_Y(k,s)}{H(k)}\right)^{\gamma_B^{\text{spillover}}} \tag{2.52}$$

where $\tilde{N}_Y(j,s) = \sum_i \lambda^C(i,j,s)\,N_R(i)$ is the number of workers employed in j and $H(j)$ is the total floor space of location j, so that $\frac{\tilde{N}_Y(j,s)}{H(k)}$ is the worker density in j and $\tau^{min}(j,k,s) \equiv \min_m\{\tau(j,k,s,m)\}$ is the fastest travel time across all modes over a given route. In Ahlfeldt et al. (2015), the congestion at residence (denominated B here) depends on how many people live around an area, while the agglomeration at destination (denominated A here) depends on how many people work around an area. Since we assume a fixed number of residents, in our case both spillovers are a function of the endogeneous number of workers in the surrounding areas.

2.7.6 Equilibrium

Equilibrium in the labor market of tract i implies:

$$\underbrace{\sum_i \lambda^C(i,j_C,s)\,e(i)\,N_R(i)}_{N(i,s)} = N_Y(i,s) + \underbrace{\mu_S(i)\frac{N_R(i)\,\overline{I}(i,s)}{w(i,s)} + T_S\sum_j\sum_{m_L} R_L(j,i,m_L,s)}_{N_S(i,s)}. \tag{2.53}$$

where $\lambda^C(i,j_C,s)$ fraction of commuters from i to j_C through any means:

$$\lambda^C(i,j_C,s) \equiv \sum_{m_C}\lambda^C(i,j_C,m_C,s). \tag{2.54}$$

The left-hand side of (2.53) is the supply of efficiency units, $N(i,s)$, to location i. The right hand side is the sum of demand for efficiency units in the tradeable ($N_Y(i,s)$) and non-tradeable ($N_S(i,s)$) sector, with the latter encompassing the demand for local services from location i's residents and from leisure-travelers into i.

Next, using the solution for consumer demand for floor space from (2.28) and for firm's demand for floor space and labor from (2.42), the equilibrium in the housing markets is:

$$N_R(i)\frac{\mu_H(i)\,\overline{I}(i)}{r(i,s)} + N_Y(i,s)\frac{w(i,s)}{r(i,s)}\frac{\mu_{H_Y}(j)}{1-\mu_{H_Y}(j)-\mu_B(j)} = H(i), \tag{2.55}$$

where the first term in the left-hand side is the demand for floor-space from residents of i, the second term is demand from firms located in i, and $H(i)$ is the supply of floor space in i. Finally, since tradeable firms operate subject to constant returns, the zero-profit conditions resulting from (2.42) dictates:

$$w(j)^{1-\mu_B(j)-\mu_{H_Y}(j)}\,r(j)^{\mu_{H_Y}(j)} = \kappa_\Omega(j)\,\Omega(j)^{1-\mu_B(j)} \tag{2.56}$$

for some constant $\kappa_\Omega(j)$ that is a function of $\mu_B(j)$ and $\mu_{H_Y}(j)$.

An equilibrium consists of distributions of land prices $r(j,s)$, wages $w(j,s)$, and supplies of labor into tradeables $N_Y(i,s)$, such that: i) the land market clearing condition 2.55 holds for all tracts; ii) the labor market clearing condition 2.53 holds for all tracts i; and iii) the zero-profit condition 2.56 holds for all tracts j.[29]

[29]These equations include as unknowns the endogenous productivity term $\Omega(j)$, the agglomeration and amenity spillover functions $A(j,s)$ and $B(i,s)$, and the average income $\overline{I}(i)$. Using 2.44, 2.51, 2.52, and 2.49, all these terms can be expressed as functions of the endogenous variables $\left\{r(j,s), w(j,s), N_Y(i,s), \lambda^C(i,j_C,s)\right\}$ which define the equilibrium.

2.7.7 Counterfactuals

We now express the equilibrium value of every endogenous outcome in a scenario where $s = B$ relative to its value in an equilibrium where $s = NB$. The system of equations we derive here is what we implement when running counterfactuals. We let

$$\hat{X}(\cdot) \equiv \frac{X(\cdot, B)}{X(\cdot, NB)}$$

be the ratio of variable X between its equilibrium value when $s = B$ (so that the HSR is built) and when $s = NB$ (not built).

The HSR shock results in new lowest times and monetary travel costs for public-transit travel. So, the shock to the system is given by travel time changes

$$\hat{d}_k(i, j_k, \text{public transit}) = \hat{\tau}(i, j_k, \text{public transit})^{\rho_k}$$

for each travel purpose $k = C, L, B$ (commuting, leisure, or business travel), and to monetary travel cost changes per trip

$$\hat{p}_R(i, j_k, \text{public transit}).$$

On origin-destination pairs where the HSR is chosen, the travel time is faster ($\hat{\tau} < 1$) and the price may increase or decrease depending on the cost of the HSR on that route compared to the pre-HSR; on pairs where it is not chosen, then $\hat{\tau} = \hat{p} = 1$. The HSR shock also entails an increase in the tax to finance the HSR from $t(NB) = 0$ to $t(B) = t$ so that disposable income changes by $1 - t$.

The equilibrium response to $\{\hat{\tau}(i, j, m), \hat{p}_R(i, j, m), 1 - t\}$ consists in changes in land rents $\hat{r}(i)$, wages $\hat{w}(i)$, and labor supplies $\hat{N}_Y(i)$ such that:

i) The land market clears, i.e. (2.55) holds in the counterfactual equilibrium, which implies:

$$\hat{r}(i) = \frac{H_C(i,NB)}{H(i,NB)}\hat{\bar{I}}(i) + \left(1 - \frac{H_C(i,NB)}{H(i,NB)}\right)\hat{w}(i)\hat{N}_Y(i), \tag{2.57}$$

where $H_C \equiv N_R(i)\frac{\mu_H(i)\bar{I}(i)}{r(i,s)}$ is the aggregate housing demand in i and $\hat{\bar{I}}(i)$ is the change in average income of residents of i defined in (2.49),

$$\hat{\bar{I}}(i) = \sum_{j_C}\sum_{m_C}\frac{\lambda^C(i,j_C,m_C,NB)}{\bar{I}(i,NB)}\hat{\lambda}^C(i,j_C,m_C)\hat{I}(i,j_C,m_C), \tag{2.58}$$

where the change in disposable income next of taxes and commuting costs for commuters from i to j_C through m_C is

$$\hat{I}(i,j_C,m_C) = \left(1 + \frac{p_C(i,j_C,m_C,NB)}{I(i,j_C,m_C,NB)}\right)\hat{y}(i,j_C) - \frac{p_C(i,j_C,m_C,NB)}{I(i,j_C,m_C,NB)}\hat{p}_C(i,j_C,m_C), \tag{2.59}$$

the change in pre-tax income is

$$\hat{y}(i,j_C) = \left(1 - \frac{e(i)\,w(j_C,s)}{y(i,j_C,NB)}\right)\hat{\bar{I}}(i) + \frac{e(i)\,w(j_C,s)}{y(i,j_C,NB)}\hat{w}(j), \tag{2.60}$$

and, from (2.45), $\hat{\lambda}^C(i,j_C,m_C)$ is given by:

$$\hat{\lambda}^C(i,j_C,m_C) = \frac{\left(\frac{\hat{I}(i,j,m_C)}{\hat{d_C}(i,j,m_C)}\right)^{\theta^C}}{\sum_{j'}\lambda^C(i,j',m_C,NB)\left(\frac{\hat{I}(i,j',m_C)}{\hat{d_C}(i,j',m_C)}\right)^{\theta^C}} \tag{2.61}$$

ii) the labor market clears, i.e. (2.53) holds in the counterfactual equilibrium, which implies

$$\underbrace{\sum_i\left(\frac{\lambda^C(i,j_C,NB)\,e(i)\,N_R(i)}{\sum_{i'}\lambda^C(i',j_C,NB)\,e(i')\,N_R(i')}\right)\hat{\lambda}^C(i,j)}_{\hat{N}(i,s)} = \frac{N_Y(j,s)}{N(j,s)}\hat{N}_Y(j,s) + \left(1 - \frac{N_Y(j,s)}{N(j,s)}\right)\hat{N}_S(j,s), \tag{2.62}$$

where, in the supply side in the left-hand side of (2.62), $\hat{\lambda}^C(i,j)$ is given by

$$\hat{\lambda}^C(i,j) = \sum_{m_C} \frac{\lambda^C(i,j,m_C,NB)}{\sum \lambda^C(i,j,m'_C,NB)} \hat{\lambda}^C(i,j_C,m_C)$$

and where, in the labor demand side in the right-hand side of (2.62), the demand for labor into services is

$$\hat{N}_S(i) = \frac{N_S^{locals}}{N_S} \frac{\hat{\bar{I}}(i)}{\hat{w}(i)} + \left(1 - \frac{N_S^{locals}}{N_S}\right) \sum_j \sum_{m_L} \left(\frac{R_L(j,i,m_L,NB)}{\sum_{j'} \sum_{m'} R_L(j',i,m',NB)} \right) \hat{R}_L(j,i,m_L,s)$$

where $\frac{N_S^{locals}}{N_S}$ is the share of service labor in the initial equilibrium that satisfies non-traded demand from residents, and where the change in the number of leisure trips is

$$\hat{R}_L(i,j_L,m_L,s) = \hat{\bar{I}}(i,s) \frac{\hat{\lambda}^L(i,j_L,m_L,s)}{\hat{p}_L(i,j_L,m_L,s)},$$

where the change in price per leisure trip is

$$\hat{p}_L(i,j,m) = \frac{c_L p_S(j,NB)}{p_L(i,j,m,NB)} \hat{w}_S(j) + \left(1 - \frac{c_L p_S(j,NB)}{p_L(i,j,m,NB)}\right) \hat{p}_R(i,j,m)$$

and where the change in share of residents from i traveling to j via m for leisure, $\hat{\lambda}^L(i,j_L,m_L,s)$ is

$$\hat{\lambda}^L(i,j_L,m_L,s) = \frac{\left(\hat{p}_L(i,j_L,m_L) \frac{\hat{d}_L(i,j_L,m_L)}{\hat{B}(j_L)}\right)^{-\mu_L(i)\theta^L}}{\sum_{j'} \sum_{m' \in M_{ij}(s)} \lambda^L(i,j',m',NB) \left(\hat{p}_L(i,j',m') \frac{\hat{d}_L(i,j',m')}{\hat{B}(j')}\right)^{-\mu_L(i)\theta^L}}$$

iii) the zero-profit condition (2.56) holds in a counterfactual scenario, i.e.

$$\hat{w}(j)^{1-\mu_B(j)-\mu_{H_Y}(j)} \hat{r}(j)^{\mu_{H_Y}(j)} = \hat{\Omega}(j)^{1-\mu_B(j)}, \tag{2.63}$$

where, from 2.44,

$$\hat{\Omega}(j) = \hat{A}(j)^{\frac{1}{1-\mu_B}} \left(\sum_{j_B} \sum_{m_B} \lambda^B (i, j_B, m_B, NB) \left(\frac{\hat{A}(j_B)}{\hat{d}_B (j, j_B, m_B) \hat{p}_B (j, j_B, m_B)} \right)^{\theta_B \mu_B (j)} \right)^{\frac{1}{(1-\mu_B(j))\theta_B}}, \tag{2.64}$$

and from (2.51),

$$\hat{A}(j) = \left(\sum_k \frac{\frac{\tilde{N_Y}(k,NB)}{H(k)}}{\sum_{k'} e^{-\rho_A \tau^{min}(j,k',NB)} \frac{\tilde{N_Y}(k',NB)}{H(k')}} e^{-\rho_A \tau^{min}(j,k,B)} \hat{\tilde{N}}_Y (k) \right)^{\gamma_A}$$

where the change in the number of workers employed in j_C is:

$$\hat{\tilde{N}}_Y (j_C) = \sum_{i'} \left(\frac{\lambda^C (i', j_C, NB) N_R (i')}{\sum_{i'} \lambda^C (i', j_C, NB) N_R (i')} \right) \hat{\lambda^C} (i', j_C).$$

2.8 Results Appendix

Table 2.6: Commuting Equation Estimates: Exponential Function

	2008				2019			
	(1) PPML	(2) PPML	(3) FS	(4) IV	(5) PPML	(6) PPML	(7) FS	(8) IV
pre-HSR travel time, minutes	-0.102***	-0.105***			-0.106***	-0.109***		
	(0.000276)	(0.000301)			(0.000258)	(0.000261)		
log(workplace earnings, LEHD, 2008)	2.634***			4.876***				
	(0.00895)			(0.165)				
log(earnings, ACS, 2006-2010)			0.347***					
			(0.00870)					
log(workplace earnings, LEHD, 2019)					3.069***			5.107***
					(0.00989)			(0.164)
log(earnings, ACS, 2012-2016)							0.317***	
							(0.00624)	
ρ_C	.039			.021	.034			.021
θ_A	2.63			4.88	3.07			5.11
N	40747886	40747886	7866	7866	40747886	40741069	7865	7865
R2			0.227	0.129			0.291	0.193
F-stat				1600				2600
Destination FE	No	Yes			No	Yes		

Source: LEHD, 2008 and 2019; ACS, 2006-2010; ACS, 2012-2016.

Note: This table shows PPML estimates of the commuting equation 2.18 in the alternative case where $d_C(i,j,m,s) = D_C(i,m)\, exp(\tau(i,j,m,s))^{\rho^C}$.

Table 2.7: OLS Estimates of Voting Equation under Full Model

	(1)	(2)	(3)	(4)	(5)
$\log(\hat{W})$, 2019	-5.580	3.024	3.631	3.518	6.435**
	(5.357)	(2.533)	(2.135)	(2.001)	(2.409)
log-odds Reg Dems Share		0.665***	0.462***	0.455***	0.413***
		(0.0724)	(0.108)	(0.125)	(0.0815)
Environment: log odds Yes on Prop. 10			0.460***	0.430***	0.359***
			(0.0878)	(0.0542)	(0.0519)
Transportation: log odds Yes on Prop. 1b				0.0517	-0.00204
				(0.156)	(0.139)
pop. density					0.0470***
					(0.0103)
Avg. MA Pop. Dens.					0.0482
					(0.127)
Constant	0.204*	0.274***	-0.0175	0.0139	-0.192
	(0.0957)	(0.0471)	(0.101)	(0.0556)	(0.292)
R2	0.00750	0.589	0.658	0.659	0.714
N	7866	7866	7866	7866	7866

Note: This table shows OLS estimates of equation 2.7. Proposition 10 was on the ballot in the same election as the HSR proposition and, if passed, would have allowed the state to issue $5 billion in bonds for alternative fuel projects (Ballotopedia). Proposition 1b, on the ballot in 2006, allocated funds towards transportation projects including highways.

Table 2.8: IV Estimates of Voting Equation under Full Model

		2008 IV		Random Stations IV	
	(1)	(2)	(3)	(4)	(5)
	OLS	FS	IV	FS	IV
$\log(\hat{W})$, 2019	6.435**		6.153**		7.630**
	(2.409)		(2.362)		(2.410)
$\log(\hat{W})$, 2008		0.531***			
		(0.00437)			
$\log(\hat{W})$, Random Route IV, 2008				0.581***	
				(0.105)	
$\log(\hat{W})$, Random Route+Station IV, 2008					
log-odds Reg Dems Share	0.413***	-0.00000456	0.413***	-0.000909	0.413***
	(0.0815)	(0.0000299)	(0.0814)	(0.000626)	(0.0817)
Environment: log odds Yes on Prop. 10	0.359***	-0.0000259	0.359***	0.00167	0.358***
	(0.0519)	(0.0000519)	(0.0520)	(0.00114)	(0.0510)
Transportation: log odds Yes on Prop. 1b	-0.00204	0.0000234	-0.00183	0.00115	-0.00293
	(0.139)	(0.0000389)	(0.139)	(0.000777)	(0.137)
pop. density	0.0470***	0.0000139*	0.0470***	0.0000592	0.0469***
	(0.0103)	(0.00000670)	(0.0103)	(0.0000896)	(0.0102)
Avg. MA Pop. Dens.	0.0482	-0.000250**	0.0452	-0.00524*	0.0608
	(0.127)	(0.0000967)	(0.126)	(0.00224)	(0.119)
Constant	-0.192	0.00264***	-0.184	0.0129**	-0.226
	(0.292)	(0.000223)	(0.292)	(0.00489)	(0.273)
First stage F-stat			14579.2		30.30
R2	0.714	0.997	0.714	0.544	0.713
N	7866	7866	7866	7866	7866

Note: This table shows IV estimates of equation 2.7. The 2008 IV uses welfare measured using 2008 fundamentals as an instrument for expected welfare, and the random route IV uses an instrument constructed using fake HSRs, as described in detail in Appendix section 2.9.3. Proposition 10 was on the ballot in the same election as the HSR proposition and, if passed, would have allowed the state to issue \$5 billion in bonds for alternative fuel projects (Ballotopedia). Proposition 1b, on the ballot in 2006, allocated funds towards transportation projects including highways.

2.9 Data Appendix

2.9.1 Wages

We measure wages at the origin and destination tract levels using data from the LEHD. The LEHD lists, for each tract pair in each year, the number of workers commuting from that origin to that destination and earning below \$1,250 per month, between \$1,250 and \$3,333 per month, or above \$3,333 per month. To construct average wages for each of these three groups, we turn to the American Community Survey. Using the 1-year samples for the corresponding year (either 2008 or 2019), we restrict the data to California and estimate average wages among individuals with monthly earnings in each of the three LEHD categories.[30] We then construct average wages for each origin-destination pair as follows:

$$wage_{od} = \frac{N_{od}^{<\$1250} \cdot wage^{<\$1,250} + N_{od}^{\$1,250-\$3,333} \cdot wage^{\$1,250-\$3,333} + N_{od}^{>\$3,333} \cdot wage^{>\$3,333}}{N_{od}^{<\$1250} + N_{od}^{\$1,250-\$3,333} + N_{od}^{>\$3,333}} \tag{2.65}$$

where N_{od}^{x} is the number of workers who live in o and work in d and have monthly earnings in category x and $wage^{x}$ is the average monthly wage of California workers, conditional on monthly wages being in category x. We then convert monthly earnings to annual earnings.

[30] We do not observe this data at the tract-level for the 1-year ACS samples, so we cannot do this at the tract-level.

2.9.2 Transport Network

Road travel speeds. To measure driving times between all routes, we obtain the shape of the primary and secondary road network in California from the U.S. Census. Then, we randomly select 1,000 routes and obtain the driving travel time along each route from Google Maps using the mode set to road. We do not obtain travel times for all 60+ million routes from Google Maps due to the extraordinary monetary cost associated with doing so, plus the fact that our calibration can closely match observed times. In addition, we do not rely on travel times as reported in the American Community Survey because we do not observe these travel times along routes for which there are no commuters in the data. However, our estimated travel times are fairly well-correlated (0.49) with these ACS reported times. Given the road network, we calibrate road speeds along four different types of roads (primary roads in urban areas, primary roads in non-urban areas, secondary roads in urban areas, and secondary roads in non-urban areas), in addition to a constant term, to match Google driving travel times. An urban road is a road segment that lies in a county with a population of at least 1,000,000. More specifically, we find the parameter vector:

$$\alpha^* = (\alpha^{primary,urban}, \alpha^{primary,non-urban} \alpha^{secondary,urban}, \alpha^{secondary,non-urban}, \alpha^{constant})$$

that solves the following:

$$\alpha^* = \arg\min_{\alpha} \sum_{i}^{1,000} (Est_i(\alpha) - GoogleTime_i)^2$$

where $Est_i(\alpha)$ is the shortest path along route i given the 2008 road network and travel speeds along each type of road given by the vector of parameters α. $GoogleTime_i$ is the Google Maps reported driving time along route i. We can closely match Google travel times, with a correlation between observed and estimated travel times of around 0.99.

Figure 2.6: Calibrated and Observed Road Travel Times

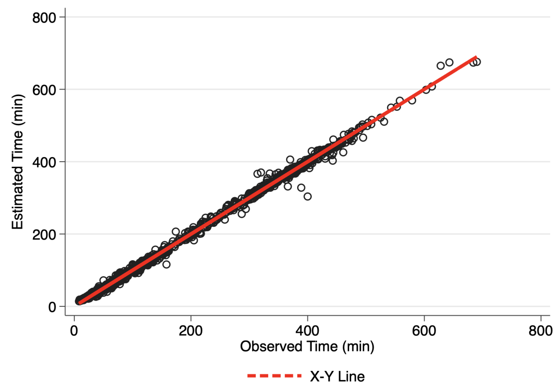


Figure 2.6 shows this correlation; each point lies very close to the 45 degree line.

We then use these parameters and the road network to compute driving times for all of the routes in our sample. To incorporate the air network, we identify commercial air routes between major airports in California as of 2008 from the Bureau of Transportation Statistics. From Google Maps, we obtain flight times between airports and the coordinates of airports. We assume that, when transferring from the road network to the air network, individuals spend 45 minutes. This cost of transferring reflects that people don't simply jump from the highway onto a plane. Then we find the quickest path given the road network and the air network along each route to assign pre-HSR travel speeds along each route.

Public transit network. To estimate travel times and costs via the rail network, we first obtain a shapefile of the location of every rail station in California as of 2013 (the earliest year available) from the California Department of Transportation, accessed via Stanford University. Second, we manually collect time tables on intercity rail service from Capitol Corridor, San Joaquin, Pacific Surfliner, and Amtrak. We do the same for intracity rail service, including BART, Caltrain, ACE, Coaster, Metrolink, and SMART. These timetables are available in PDF form on each rail agency's webpage, and we convert these

PDFs to spreadsheets.[31] From these time tables, we obtain average travel time between each consecutive station. Travel times between stations may vary at different times of the day and depending on whether the train is an express or local; in general, we take an average across all listed times. We then use this information to construct a graph that describes the rail public transit network in California, where edges between stations have weights based on travel times.

To measure the cost of traveling via public transit, we use data from the Capitol Corridor, a passenger train operated by Amtrak in Northern California. From the above procedure, we already observe the travel time between each pair of stations in the Capitol Corridor network. We then obtain fare prices for each pair of stations which we use to estimate equation .[32] Figure 2.7 shows the close relationship between travel fares and travel times for rail travel.

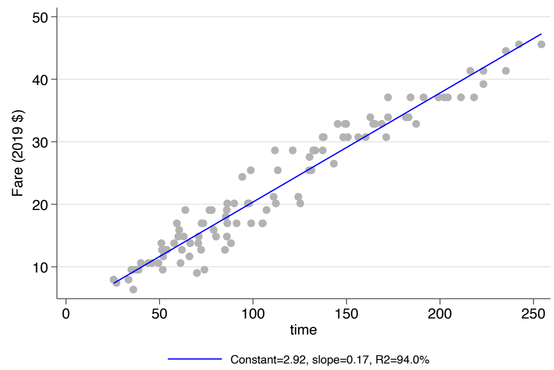
There is far more limited systematic data on the bus network, as it varies widely across cities and counties. Thus, we incorporate the bus network by allowing buses to traverse the entire road network. We calibrate travel speeds for bus using Google Maps using the methodology above and Google travel times where mode is set to bus. As expected, bus travel is much slower than car travel. We obtain data on county-specific costs of bus transport from the American Public Transportation Association, supplemented with data from different counties webpages.

To calibrate transfer costs, we randomly select a set of origin and destinations in the public transit network, which may use any combination of rail and bus. We use Google Maps to obtain travel times along these routes when the mode is set to public transit.

[31] For example, the Pacific Surfliner timetable is available here: https://www.pacificsurfliner.com/globalassets/plan-your-trip/schedules/803091280_pacific-surfliner-timetable_oct-25-2021.pdf/.

[32] This can be found here: https://www.capitolcorridor.org/wp-content/uploads/2016/05/CCJPA_fares_06162016.pdf

Figure 2.7: Capitol Corridor Fares and Travel Times



Source: Capitol Corridor timetables and fares.

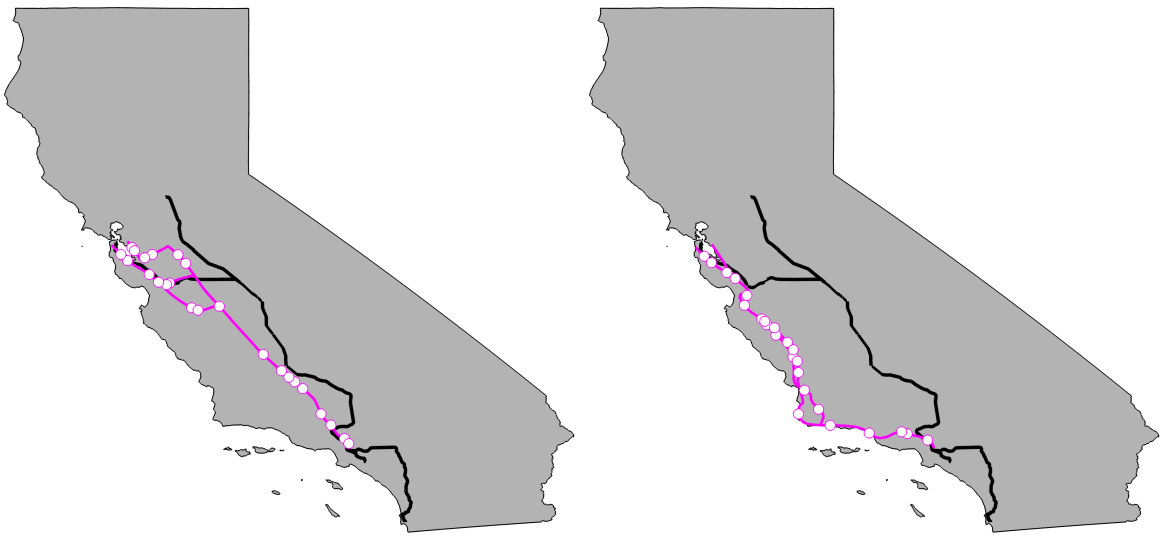
These routes may involve transfers between different rail networks, or between the rail network and the bus network. We then calibrate a single parameter, α^T, which measures the time spent making these transfers. We calibrate it in order to match the observed travel times reported from Google and find a value of $\hat{\alpha}^T = 8.5$.

2.9.3 Construction of the Instrument

The instrument described in Section 2.5 is constructed as follows. We digitized the map shown in Figure 2.2, such that our starting point is a shapefile. For each of the three routes, differentiated by their color in the original map, we divide the line into a series of points by placing a point at uniform increments of 1km along the entire line. To construct the instrument, we first select each of the three routes with probability 1/3. Then, having selected a route, we randomly select 24 points along the line, with equal probability of choosing each point. Each selected point will serve as a station in this alternative high speed rail network. We use 24 as this was the number of stations in the actual, planned HSR network, as well as the number outlined by the California legislature. After choosing any given point, we then exclude points within 5km from that point as possible stations.

Figure 2.8: Example Fake HSRs for Instrument

(a) Fake Network Along I-5 (b) Fake Network Along Coast



Note: This figure shows two "fake" HSRs constructed for the instrument, with the route chosen with equal probability from proposed routes as in Figure 2.2 and 24 points, serving as stations, chosen with equal probability along the route.

Figure 2.8a shows an example of a HSR constructed along the I-5 route, and Figure 2.8b shows an example along the Coast.

We repeat this procedure 100 times, and obtain 100 fake HSR networks as a result. In order to estimate the welfare gains under each fake HSR, we need to assign travel speeds to convert rail distances to rail travel times. We do this by assuming that the average speeds along the I-5 route and the Central Valley route are 264 km per hour, and slightly slower at 240 km per hour for the Coastal Route. These times reflect system averages reported in CHSR documents. Then, for each fake network, we construct the implied welfare gains. Ultimately, our instrument is the average welfare gain implied over these 100 fake HSR networks.

CHAPTER 3

Trade Networks in Latin America: Spatial Inefficiencies and Optimal Expansions

(with Elena Ianchovichina)

.

3.1 Introduction

Differences in labor income across regions and municipalities are expected to diminish with development as barriers to factors and goods mobility decline and technology spreads within countries. Yet, spatial inequality has remained high throughout Latin America (Acemoglu and Dell (2010)). One potential reason for this may be insufficient and spatially misallocated road networks, which are an important determinant of trade costs (Limao and Venables (2001); Atkin and Donaldson (2015)). The findings of transport studies lend support to this hypothesis. Road density in Latin America is comparable to that in Sub-Saharan Africa and lower than elsewhere in the world (Ndulu (2007)). While geography and other physical characteristics of Latin American countries may explain the low road density, road occupancy rates are very high and large areas remain inaccessible (Fay et al. (2017)). Railway networks are also underdeveloped, implying that railway services are neither an effective substitute for, nor complement to, road transport.

We explore the extent to which trade connectivity issues affect the efficient spatial distribution of economic activity within and across countries in Latin America either because the existing road infrastructure is spatially misallocated or because it is insufficient.[1] Using data from multiple sources and the general equilibrium spatial framework of Fajgelbaum and Schaal (2020), we construct optimal transport networks and optimal expansions to existing networks in most Latin American countries, as well as within MERCOSUR and the Andean Community. Comparisons between the existing and optimal networks or between existing networks and their optimal expansions allow us to identify inefficiencies in the existing networks and their associated welfare effects.

We find that in several LAC countries, including Argentina, Brazil, and Peru, transport networks are relatively inefficient. Misallocation of roads in these countries is associated with annual welfare losses of between 1.5% and 2.4%. Therefore, the model results suggest that inefficiencies in existing road networks generate significant trade frictions in the largest and most populous countries on the continent. At the other end of the spectrum are a few countries with relatively efficient transport networks, including Guatemala, Costa Rica, and El Salvador. These countries will not gain much if the social planner had the power to lift existing networks and place them optimally. As is the case for Europe presented in Fajgelbaum and Schaal (2020), the loss associated with road misallocation in Latin America is around 1% in simple average terms and 1.6% when countries are weighted according to population, indicating that losses are larger in the region's more populous countries. In most countries, new road expansions can deliver welfare increases comparable to the gains from replacing existing networks with optimal ones. These gains are largest in the countries with the largest inefficiencies in existing networks, including Argentina, Brazil, and Bolivia. The model-implied optimal investments correlate relatively well with World

[1] Since it is infeasible to build the optimal road network, its comparison with the existing network provides a sense of the losses due to the misallocation of existing roads.

Bank road investments because both the model and the World Bank prioritize projects in high population areas.

Trade frictions also arise due to inefficiencies in optimal transnational road networks within MERCOSUR and the Andean Community. Spatial misallocation of transnational road networks is associated with average annual welfare losses of 1.8% in MERCOSUR and 1.5% in the Andean Community. These losses can be remedied with road investments that improve and expand the existing road network. In the case of MERCOSUR, expansion yields an annual welfare increase of 1.9%, while in the case of the Andean Community, the gain is 1.5%. The model suggests improvements in connectivity between the largest cities within MERCOSUR and between the largest cities in each member country. Given the location of these cities, these are improvements mostly along the coastal highways. Most of the investments occur in Brazil (71%) and in Argentina (22%), with the remainder split equally between Uruguay and Paraguay, but the highest welfare gain accrues to the least developed country in the trade bloc – Paraguay (3.3%). Within the Andean community, half of the infrastructure growth occurs in Colombia, a quarter each in Peru and Ecuador, and only 2% in Bolivia. These investments are again pro-poor benefiting the most the poorest member of the bloc – Bolivia – where welfare increases by slightly more than 5%. Optimal investments improve connectivity between La Paz in Bolivia, along the coast of Peru to Lima and through Quito to Medellin.

The findings in this paper are important because they are informative as to the optimal spatial distribution of road infrastructure projects in Latin America. Policy makers and other stakeholders can use these findings to better target their road investments, with the objective of lowering trade costs and getting a bigger growth boost per dollar spent.

Our paper is related to the literature on the aggregate effects of misallocation (Restuccia and Rogerson (2008); Hsieh and Klenow (2009)), particularly the body of work on spatial resource misallocation due to frictions or government policies (Desmet and Rossi-

Hansberg (2013); Fajgelbaum et al. (2018); and Hsieh and Moretti (2019)). It contributes directly to the large and growing literature on evaluating the economic returns to improving infrastructure systems. Some studies in this literature use historical data to measure the welfare effects of transport infrastructure, including the colonial railways in India (Donaldson (2018)), the railway network in 19th century U.S. (Donaldson and Hornbeck (2016*a*)) and in late 19th century Argentina (Fajgelbaum and Redding (2022)), China's national trunk highway system (Faber (2014)) and the U.S. highways (Allen, Arkolakis and Takahashi (2019)). Other studies explore the impact of transport infrastructure on different aspects of development. Baum-Snow (2007) assesses the effect of highways on suburbanization in the US, while Sotelo (2016) evaluates the impact of highway investments on agricultural productivity in Peru. More recently, Bird, Lebrand and Venables (2019) study the effects of the Road Belt Initiative on economic development in Central Asia.

Another part of this literature looks at the role of trade costs in different geographic settings. Eaton and Kortum (2002) and Anderson and Van Wincoop (2004) pioneered an approach that fits standard quantitative trade models to data on the geographic distribution of economic activity across countries; they use this framework to evaluate the economic and welfare effects of exogenous shocks to trade costs. Caliendo et al. (2017) and Ramondo, Rodríguez-Clare and Saborío-Rodríguez (2016) explore the role of trade frictions within countries in the presence of factor mobility. Many papers study how actual changes in transport costs shape domestic economic activity, including the impact of road expansion on productivity across US industries (e.g. Fernald (1999)) and the impact of US highways on regional economic outcomes (e.g. Chandra and Thompson (2000) and Duranton, Morrow and Turner (2014)).

This paper is most closely linked to applications of the framework of Fajgelbaum and Schaal (2020) to Europe (Fajgelbaum and Schaal (2020)) and Africa (Graff (2019)). Ours is the first study to assess the spatial inefficiencies of existing transport networks and the effects of optimal expansions to existing transport networks for all Latin American countries, except those in the Caribbean[2], using the model of Fajgelbaum and Schaal (2020). Building on their methodology, we develop a new discretization procedure to deal with the large spatial scale of Brazil. This procedure also allows us to study road connections between different groups of Latin American countries.

Misallocation in the transport network is measured by the wedge between the optimal level of investment along each link in the network that the social planner would choose relative to the investment observed in the data. The social planner jointly solves the optimal transport problem and the optimal allocation problem to determine the optimal investment in the transport network, the optimal allocation of production and consumption across locations, and the gross trade flows between locations, given fundamentals including endowments and transport costs. In contrast to the standard models in the literature, transport costs in this model are endogenous as they depend on how much is invested in each link of the road network. The level of investment in each road link and consumption and production in each location are also endogenous and chosen by the social planner to maximize welfare. New road investments reduce transport costs and influence through general equilibrium forces the prices at which goods are produced and sold in each location and the quantities traded between locations. Other things equal, trade flows of a good between a pair of locations is higher when the road infrastructure linking the two locations is better. Trade flows decline with congestion along the link and increase with the price differentials between the two locations.

Like other models with an optimal transport problem (Alder and Kondo (2020), Allen, Arkolakis and Takahashi (2019)), the model of Fajgelbaum and Gaubert (2018) requires choosing least-cost routes across pairs of locations. However, it offers some distinct advan-

[2]The Caribbean countries are excluded due to their small size and lack of land access to South and Central America.

tages compared to other methodologies. First, given the large size of the road networks in many countries in South America, the main benefit is the substantial savings in computation time due to a reduction of the search space in the special convex case with congestion. Such savings are possible because in this case optimal infrastructure investments can be represented as functions of optimal prices, avoiding a direct search in the network space.

Second, congestion enables the social planner to focus on the optimization over the transport network itself and conduct a search for the global optimum. In this case, consumption and production in each location are not fixed but respond to the general equilibrium forces in the model and are therefore endogenous. By contrast, in the case of no congestion, which is the typical case in other studies and a special case in Fajgelbaum and Gaubert (2018), the optimal transport problem can be solved independently from the general equilibrium outcome by mapping sources with fixed supply to destinations with fixed demand. Only in this special non-convex case without congestion does the global solution in Fajgelbaum and Gaubert (2018) match closely with the least cost route optimization solutions in the literature. A third advantage is the flexibility of the framework, which can easily be matched to data on actual transport networks around the world. The model's fundamentals can be calibrated such that the solution to the planner's problem matches spatially disaggregated data on population and economic activity given an observed transport network. Counterfactuals can be considered assuming a specific technology to build infrastructure, a specific technology to produce goods, and specific consumer preferences.

The paper is structured as follows. Section 3.2 discusses the main features of the model, the data used to calibrate the model, and the representation of existing road networks in the model. Section 3.3 presents the baseline results of optimal road networks and expansions and respectively the associated spatial inefficiencies and gains from optimal expansions. Section 3.4 explores the robustness of the results. Section 3.5 presents optimal expansions of trans-national road networks within several groups of countries. We offer concluding remarks in Section 3.6.

3.2 Methodology

This section briefly describes the model of Fajgelbaum and Schaal (2020), the data, and the discretized road network representation used to implement the model.

3.2.1 Model

The social planner solves a triple-nested optimization problem:

$$\max_{I_{jk}} \max_{Q^n_{jk}} \max_{c_j, h_j, D^n_j, L^n_j, V^n_j, X^n_j} \sum_j L_j \cdot U(c_j, h_j) \tag{3.1}$$

subject to several constraints including the availability of traded and non-traded goods, the balanced-flows constraint requiring that in each location demand for a good should not exceed its supply net of exports to other locations, the network building constraint, the local factor market clearing conditions, and conditions requiring that consumption, trade flows and factor use are always non-negative. The objective of the social planner is to maximize the sum of population-weighted welfare across locations within each country.[3]

The innermost utility maximization problem in equation 3.1 is a standard allocation problem of choosing per-capita consumption of a non-traded good h_j and a composite traded good c_j in location j, bundling together the outputs D^n_j of a discrete set of N tradable sectors. Total labor across each sector n in each location j is $L_j = \sum_n L^n_j$. There is a fixed supply $V^n_j = (V^{1n}_j, ..., V^{Mn}_j)$ of M primary factors that are immobile across

[3] We focus on this objective function because this paper studies optimal national and transnational transport networks. This framework could also be used to consider transport networks that minimize inequality in a country, or weigh the utility of some locations more heavily than others, but these cases are beyond the scope of this paper.

locations but mobile across sectors within each location. Finally, $X_j^n = (X_j^{1n}, \ldots, X_j^{Nn})$ are intermediate inputs, and thus represent the quantity of each sector's output allocated to producing sector n's good in location j.

The second nested problem in 3.1 is the optimal flow problem. It determines the gross flow of goods, Q_{jk}^n, traded along the road link between locations j and k, regardless of where the goods were produced. Goods transit through the network until they reach the destination where they are consumed or used as an intermediate input. Transporting each unit of good n from j to k requires τ_{jk}^n units of good n where the per unit cost is a function of the total quantity of good n shipped along link jk, Q_{jk}^n, and the level of infrastructure on this link, I_{jk}:

$$\tau_{jk}^n = \tau_{jk}(Q_{jk}^n, I_{jk}) \tag{3.2}$$

where the per-unit cost of shipping is increasing in the quantity of good n shipped, implying decreasing returns or the presence of congestion:

$$\frac{\partial \tau_{jk}(Q, I)}{\partial Q} \geq 0$$

Congestion reflects the fact that increased shipping activity increases marginal transport costs due to road damage, road accidents, longer travel times, and other potential reasons. In contrast, increased road investments in link jk, which either improve the road surface, widen the road and/or increase the number of road lanes, reduce marginal transport costs:

$$\frac{\partial \tau_{jk}(Q, I)}{\partial I} \leq 0$$

The transport technology captures variation in geographic characteristics such as ruggedness. If elevation is higher in node j than k and it is cheaper to transport goods from j to k, then $\tau_{jk} < \tau_{kj}$. The optimal flow problem combines (i) an optimal transport problem of how to map production sources to destinations and (ii) a least-cost route problem with congestion. These two problems must be solved jointly because determining the least-cost routes requires information about the flows and the supply and demand of each good, which are endogenously determined in the solution to the allocation problem. The transport technology in 3.2 can be represented as a constant-elasticity transport function:

$$\tau_{jk}(Q, I) = \delta^{\tau}_{jk} Q^{\beta} / I^{\gamma} \tag{3.3}$$

Parameter β assumes non-negative values. If $\beta > 0$, the marginal cost of shipping is increasing in the shipped quantity. If $\beta = 0$ the marginal cost of shipping is invariant to the quantity shipped as in the standard iceberg case. Parameter γ represents the elasticity of the per-unit transport cost to infrastructure investment. Finally, the scalar δ^{τ}_{jk} captures the geographic frictions that affect per-unit transport costs, given the quantity shipped and infrastructure I. Finally, the outer-most nested problem in 3.1 is the optimal network problem. Its solution determines the optimal infrastructure investment I_{jk} in each link jk in the road network. The building of infrastructure requires a resource (i.e. asphalt and/or other materials) whose supply K is fixed in the aggregate within a country. Resource K cannot be used for any other purpose. This assumption implies that the opportunity cost of building infrastructure in any location is only foregoing infrastructure elsewhere. Building infrastructure I_{jk} in link jk requires an investment of $\delta^{I}_{jk} I_{jk}$ units of K. Then the network building constraint is given as:

$$\sum_{j} \sum_{j \in N(j)} \delta^{I}_{jk} \cdot I_{jk} \leq K \tag{3.4}$$

where $N(j)$ is the set of neighboring locations to location j and $j \in N(j)$. From location j, which can be interpreted as a county, goods can be shipped to locations other than $N(j)$, but in this case they must transit through a sequence of connected locations. Thus, total spending on the existing road network is measured as K: the total cost of investment across all links in the network. Investment takes place in a link jk when there is some minimum infrastructure $\underline{I}_{jk}$ in this road link. This investment may be limited by an upper bound, $\overline{I}_{jk}$, imposed by geographic constraints on the capacity to build on the link.

Given 3.3, the optimal level of infrastructure in a link is:

$$I_{jk} = \min\left(\max(I^*_{jk}, \underline{I}_{jk}), \overline{I}_{jk}\right)$$

where I^*_{jk} is the optimal infrastructure in link jk in the unconstrained optimal network problem (i.e., $\underline{I}_{jk} = 0, \overline{I}_{jk} = \infty$).

$$I^*_{jk} = \left(\frac{\gamma}{P_K}\frac{\delta^\tau_{jk}}{\delta^I_{jk}}\left(\sum_n P^n_j (Q^n_{jk}(P^n_j, P^n_k))^{1+\beta}\right)\right)^{(1/(1+\gamma))}$$

The network problem is tractable because in this setup optimal infrastructure in a link is only a function of prices in each location. Given the relatively small set of prices, the model is solved link by link, instead of searching in the very large space of all networks. Optimal infrastructure increases with gross flows shipped along link jk and their prices at origin. It decreases with the price of building material, P_K, as it increases the cost of building infrastructure; it increases with δ^τ_{jk} reflecting that infrastructure investments offset geographic frictions; and decreases with the marginal cost of building infrastructure δ^I_{jk}. Note that infrastructure affects only trade of goods, not the movement of people.

Fajgelbaum and Schaal (2020) show how, using a no arbitrage condition governing prices across space from the first order conditions of the social planner's problem, gross flows of each commodity shipped along link jk can be expressed as follows:

$$Q^n_{jk} = \left(\frac{1}{1+\beta}\frac{I^\gamma_{jk}}{\delta^\tau_{jk}}\max\left(\frac{P^n_k}{P^n_j} - 1, 0\right)\right)^{1/\beta}$$

All else equal, as noted by Fajgelbaum and Schaal (2020), trade flows of a commodity between a pair of locations will be higher when infrastructure quality is higher. Similarly, trade flows will be higher when there is more to gain from trade; that is, when the price of the commodity in the destination is much higher than the price in the origin. This result significantly simplifies the data required to implement the model since intranational trade data, from which we could observe Q^n_{jk}, can be hard to obtain in many settings.[4] Given this result, computing trade flows between locations for each commodity requires a measure of geographic trade frictions δ^{τ}_{jk} the price of the commodity in the destination k relative to the origin j, $\frac{P^n_k)}{P^n_j}$, a measure of the infrastructure linking j and k, I_{jk}, and parameter values for β and γ.

3.2.2 Data

To implement the model, we must first construct a grid describing the spatial distribution of economic activity and population within each country and a discretized road network describing existing road network connections between each pair of locations in each country. Following Fajgelbaum and Schaal (2020), we divide each country into 1-arc degree cells (in most cases), 0.5 arc-degree cells or 0.25 arc-degree cells (in some cases), depending on the surface area of the country.[5] Table 3.1 shows, for each country, the size of cells that comprise that country's grid. Brazil, however, is too large even for a grid of 1 arc-degree cells as in this case the grid is made up of more than 800 cells. When the number of grid cells exceeds approximately 275 cells, the calibration exercise becomes prohibitively computationally intense. Thus, in the case of Brazil, we construct grid cells based on the

[4]This is of particular concern in our setting where reliable data on within country trade is severely lacking.

[5]The countries in our setting are quite a bit larger in size relative to those in Europe, which is why we use 1 arc-degree cells in most cases as compared with 0.5 arc-degree cells used in Fajgelbaum and Schaal (2020).

boundaries of mesoregions, which are subdivisions of Brazilian states used by the Brazilian Institute of Geography and Statistics.

Table 3.1: Road Network Summary Statistics

	Actual Network			Discretized Network		
Country	**Network Length (km)**	**Avg. Lanes per km**	**Cell Size**	**# Cells**	**Network Length (km)**	**Avg. Infrastructure**
Argentina	855,713	1.67	1	294	118,406	0.73
Bolivia	121,098	1.98	1	101	51,248	0.56
Brazil	727,041	2	-	137	109,479	1.27
Chile	118,456	1.92	1	69	18,226	0.88
Colombia	410,003	1.8	1	108	55,227	0.52
Costa Rica	20,254	1.96	0.5	22	4,996	1.41
Ecuador	87,531	1.91	0.5	87	23,216	0.66
El Salvador	30,313	1.93	0.25	32	6,489	1.2
Guatemala	42,981	1.92	0.5	44	10,018	0.83
Mexico	796,553	2.09	1	201	102,400	1.36
Nicaragua	26,952	1.95	0.5	47	12,276	0.99
Panama	11,196	1.96	0.5	33	4,255	1.07
Paraguay	49,434	2.01	1	42	12,911	0.78
Peru	226,267	2.03	1	85	35,269	0.97
Uruguay	81,089	1.97	0.5	75	10,082	1.19
Venezuela, RB	123,567	1.67	1	84	42,280	0.63

Note: The average infrastructure index is the distance-weighted, road-type weighted average number of lanes connecting two different grid cells.

To measure population and value added within each grid cell, we follow Fajgelbaum and Schaal (2020) and obtain population data from NASA-SEDAC's Gridded Population of the World (GPW) v.4 and value-added data from Yale's G-Econ 4.0; both datasets refer to 2005.[6] The GPW population data are reported for 30 arc-second cells (approximately 1 km) and the G-Econ value-added data are reported for 1 arc-degree cells (appr. 100km). In Panel (a) of Figure 3.1 we show the distribution of population across grid cells in Argentina. Here, brighter colored cells represent areas with relatively large populations while areas with darker colored cells represent areas with smaller populations. In the case of Brazil, since our grid cells are consistent with those used by statistical agencies, we use mesoregion level data on population and income as provided by the Brazilian government. In Panel (a) of Figure 3.2 we show the distribution of the population across grid cells in Brazil.

Some countries in our sample have unique geographies that pose challenges to road building; we handle these situations on a case-by-case basis. In Peru, Iquitos is surrounded by a natural reserve area and is inaccessible for road building, while other areas of the country are covered by forest. Southern Chile is similarly covered by natural reserve areas and has extremely difficult-to-build geography. Finally, a large lake prevents road building in southeastern Nicaragua. In each of these cases we adjust the baseline grid to accommodate these unique conditions.

In Peru, we measure unbuildable areas using the University of Maryland's Global tree cover dataset, and consider grid cells to be unbuildable if at least 80% of the grid cell is covered by tree canopy. Using this methodology, we exclude the Northeastern part of the

[6] In a robustness check, we instead use population data from WorldPop. We use data corresponding to 2005 as this is the latest year for which G-Econ data are available, and it is important to use population and value-added data from the same year. Another option would be to use Nighttime lights (NTL) data to measure economic activity in each grid cell. However, since rural commodity producing areas are common in our setting, relying on NTL data to measure economic activity may underestimate true output in these areas.

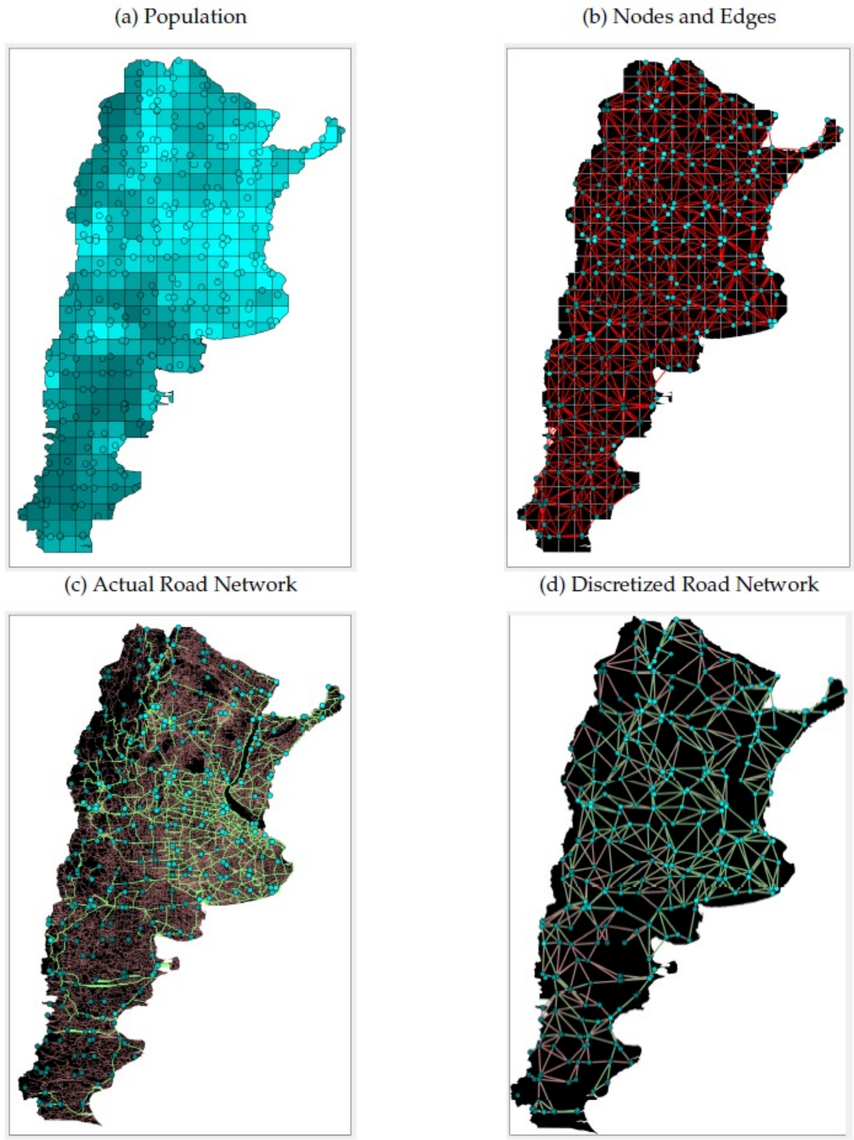
country from the grid since roads cannot be built through the reserve areas. In the case of Chile, we focus on the Northern part of the country and exclude the southern portion of the country based on the distribution of reserve areas as measured by the Chilean government. In Nicaragua, we obtain data on the locations of inland bodies of water from the International Steering Committee for Global Mapping, accessed via New York University. We use this information to restrict the grid to cells that are not majority water.

3.2.3 Discretization process

The data described above are used to construct the locations of nodes as well as the links in the model's graph for each country. The GPW data is used to locate the population centroid of each cell; these nodes are typically very close to a node on the road network. We then define the links between nodes in contiguous cells. Since cells are square-shaped, this includes up to 4 nodes connecting horizontal and vertical neighbors and up to 4 nodes along the diagonals.

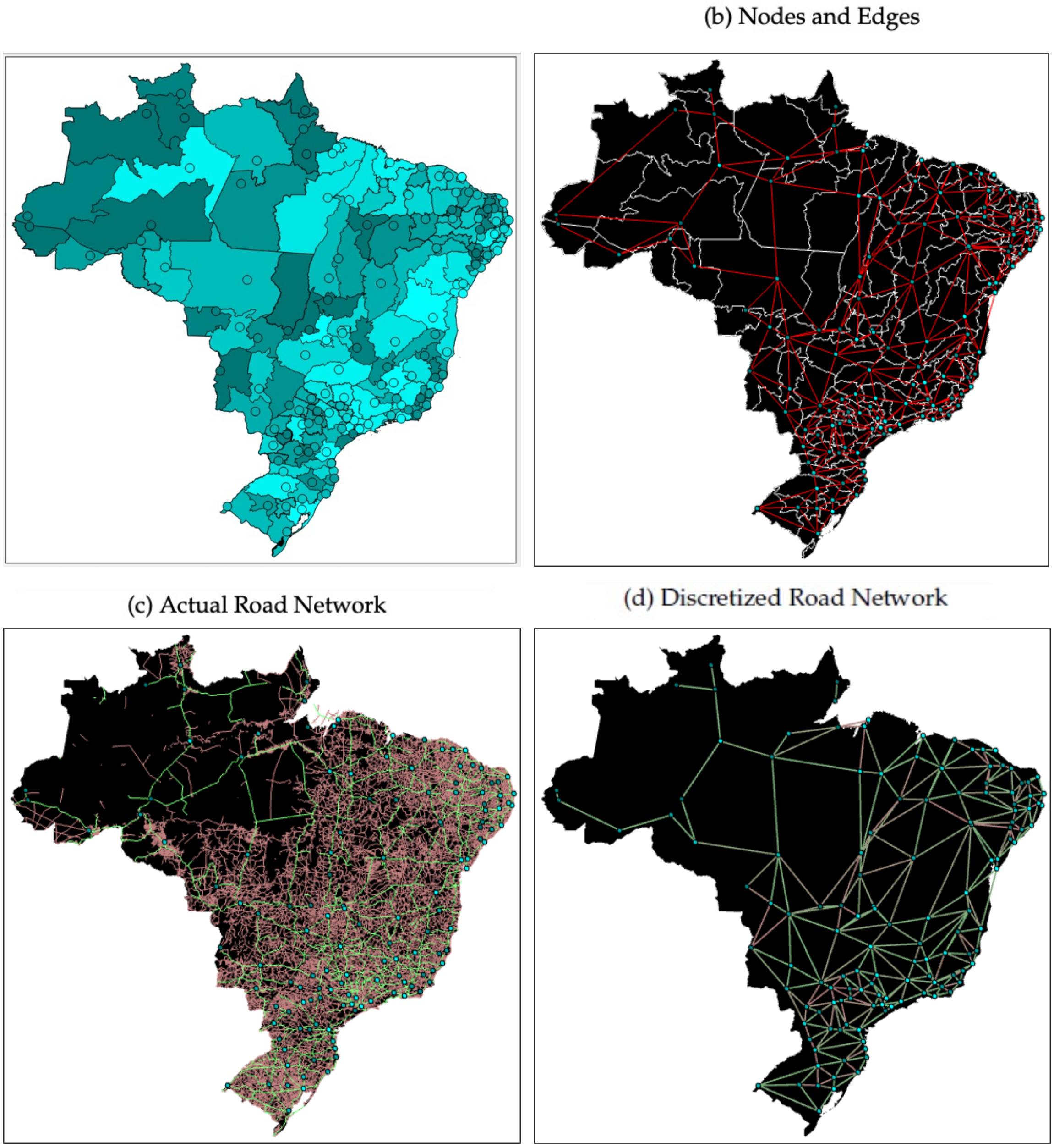
Panel (b) of Figure 3.1 shows the nodes and edges that comprise the largest possible graph in Argentina. The color of each node corresponds with its relative population, with brighter nodes indicating more populous areas. In the case of Brazil, where our grid cells are neither square in shape nor of uniform size, we define the edges of the underlying graph as follows: two centroids are connected via an edge if they share a border, and are within 800 km from each other. The additional distance criteria are required to avoid linking mesoregions that are technically neighboring but may be very far apart in space and thus not truly directly connected since the distribution of cell sizes (mesoregion areas) varies widely. This graphical representation of connections between locations within each country serves as each country's baseline grid, from which we build the discretized road network. Panel (b) of Figure 3.2 shows the nodes and edges for the graph in Brazil.

Figure 3.1: Discretization of Argentina's Road Network



Note: In Panel (a), brightly colored cells indicate areas with higher levels of population while darker cells indicate areas with lower levels of population. In Panel (b), brightly colored nodes are more populous. In Panel (c), green lines indicate primary roads while red segments indicate non-primary roads along the actual road network; same in Panel (d), along the discretized road network.

Figure 3.2: Discretization of Brazil's Road Network



Note: In Panel (a), brightly colored cells indicate areas with higher levels of population while darker cells indicate areas with lower levels of population. In Panel (b), brightly colored nodes are more populous. In Panel (c), green lines indicate primary roads while red segments indicate non-primary roads along the actual road network; same in Panel (d), along the discretized road network.

We then convert each country's actual road network into a discretized road network. The actual and discretized road networks of Argentina and Brazil are shown on Panels (c) and (d) of Figures 3.1 and 3.2, respectively. Our objective is to measure the quality of road network connections between each grid cell within each country. To this end, we obtain data on road networks for each country in our sample from the Global Roads Inventory Project (GRIP), accessed via the World Bank. For each road segment in each country, GRIP provides data on the type of road (motorway, trunk, primary, secondary, tertiary, local), the type of surface (paved, unpaved, asphalt, ground), and the number of lanes along that road segment. However, the data on road quality in GRIP is not as complete as the data available for European or other developed countries. For example, the number of lanes for each road segment listed by GRIP is highly incomplete; in most cases, information on this attribute is missing.

We thus supplement the GRIP road network data with road network data from OpenStreetMap (OSM). For each country in our sample, we calculate the average number of lanes by road type in the OSM data and use these values as our measure of the number of lanes for the corresponding country and road type in the GRIP data. For example, we compute the average number of lanes on primary roads in Colombia and assume that all primary roads in Colombia have this number of lanes. One downside to this approach is that there is little variation in the number of lanes as we fail to capture roads with very many or very few lanes. Still, we use GRIP as our primary data source because it has much more complete information on the type of road and road surface than the OSM data. In a robustness check, detailed in Section 4.1, we instead use road speeds to measure road quality, as in Graff (2019) and obtain similar results.

In our baseline results, we use the measure of infrastructure quality in Fajgelbaum and Schaal (2020). We use road network attributes from each country's road network to construct a measure of infrastructure quality along each road segment in each country's discretized road network. To do this, we first compute the shortest path along the observed road network between each set of nodes in the network. The optimal path on the road network between j and k is $P(j,k)$. We then measure the average number of lanes and average road type along a link from j to k, respectively, as:

$$lanes_{jk} = \sum_{s \in P(j,k)} \omega_{jk}(s) \cdot lanes(s)$$

$$nat_{jk} = \sum_{s \in P(j,k)} \omega_{jk}(s) \cdot nat(s)$$

where $s \in P(j,k)$ is each road network segment in the optimal route between j and k and $\omega_{jk}(s) = \frac{length(s)}{\sum_{s \in P(j,k)} length(s')}$ is the length traveled along segment s relative to the length of the journey along the road network from j to k. Then, we measure average infrastructure quality along each link as $I_{jk}^{obs} = lanes_{jk} \times K^{1-nat_{jk}}$. Following Fajgelbaum and Schaal (2020), we set $K = 1/5$, which is in line with the ratio of construction and maintenance costs along trunk roads relative to highways as reported by Doll et al. (2008). Note that for routes traveled completely along national routes, the average infrastructure index will be equal to the average number of lanes along the route. We impose that infrastructure quality is symmetric, so that $I_{jk}^{obs} = I_{kj}^{obs}$. In some cases, the optimal road network route deviates considerably from the cells containing j or k. We classify these cases as $P(j,k) = \emptyset$, indicating that there is no direct link between these locations. In panel (c) of Figure 3.1, we show Argentina's actual road network. Here, primary roads are colored in green and secondary roads are colored in red. In panel (d), we show the discretized road network where we have converted the nodes and edges as shown in Panel (b) into a discretized version of the actual road network. Figure 3.2 shows the same set of figures for Brazil, where we take a different approach and focus on connections between mesoregions, to accommodate Brazil's large size.

Table 3.1 displays the characteristics of the actual and discretized networks for the countries in our sample. In every case, the discretized network is much smaller in distance than the actual road network since the length of the discretized network reflects the shortest path between each set of nodes in the discretized network, while the actual road network comprises thousands of segments connecting thousands of nodes throughout the country. In other words, the discretized network summarizes the actual road network's connections between the population centers of each grid cell. Average infrastructure, while correlated with the average number of lanes per kilometer, is often lower in magnitude than the average number of lanes in each country. This is because we adjust the average infrastructure index to account for travel along primary versus non-primary roads. For example, as shown in Panels (c) and (d) of Figure 3.1, much of southern Argentina can only be traversed along non-primary roads; thus, average infrastructure quality across discretized edges in Argentina is considerably lower than the average number of lanes, since travel between a considerable number of nodes involves travel along non-primary roads.

3.2.4 Assumptions and Calibration

Our baseline results assume fixed labor, such that people cannot reallocate across space as infrastructure changes, and the convex case of the parameterization on congestion as described in Section 2.1. We set β, the parameter governing congestion, to 0.13 and γ, the parameter governing returns to infrastructure, to 0.10, as calibrated by Fajgelbaum and Schaal (2020) to match existing empirical estimates. We maintain the assumptions on preferences, production, and the values of parameters as outlined in Fajgelbaum and Schaal (2020). We assume that individuals have Cobb-Douglas preferences over traded and non-traded goods with the parameter α governing the share of non-traded goods in consumption. Traded goods enter the utility function through a CES aggregator across goods produced in each location with elasticity of substitution σ. We set $\alpha = 0.4$ and σ, the demand elasticity, to 5.

Production is a linear function of productivity and labor, where n is location n in country j:

$$Y_j^n = z_j^n L_j^n$$

Following Fajgelbaum and Schaal (2020), we assume that each location within a country produces a single tradable good. In each country, we allow for 10 differentiated goods plus a homogenous good. Each differentiated good is produced by a unique producer; we assume that the 10 differentiated producers are located in the 10 most populous grid cells in each country. The homogenous good is produced in the remaining cells. Finally, we assume that geographic trade frictions, which enter the transport technology in equation (5), are $\delta_{jk}^{\tau} = \delta_0^{\tau} dist_{jk}$. In lieu of data on interregional trade in Latin America or any existing estimates of this parameter from the region from which to draw, we use the value of δ_0^{τ} calibrated by Fajgelbaum and Schaal (2020) to match the level of intra-regional trade in Spain.

Given the observed discretized road network in each country, data on population and value added in the underlying grid of each country, and the parameterization described above, we calibrate each grid cell's productivity to match observed value added. In Figure 3.3, which pools all countries in our sample together, we show that with this calibration the model can closely match the distribution of income in the data.

Finally, implementing counterfactuals also requires assumptions on the cost of building along each link in the discretized road network. For this, we maintain the functional form and parameterization of the building cost function used in Fajgelbaum and Schaal (2020):

$$\ln\left(\frac{\delta_{jk}^{I,GEO}}{dist_{jk}}\right) = \ln\left(\delta_0^I\right) - 0.11 \times 1(dist_{jk} > 50km) + 0.12 \times \ln(ruggedness_{jk})$$

where ruggedness along an edge between j and k is measured as:

$$ruggedness_{jk} = \frac{1}{2}\left(ruggedness_k + ruggedness_j\right)$$

and ruggedness for each grid cell is constructed as the standard deviation of the change in elevation across neighboring grid cells. We use elevation data from the ETOPO1 Global Relief Model, which is provided at a finer granularity than our grid cells, to compute:

$$ruggedness_n = \left(\sum_{i\in J(n)} \sum_{k\in N(i)} (elev_i - elev_k)^2\right)^{\frac{1}{2}}$$

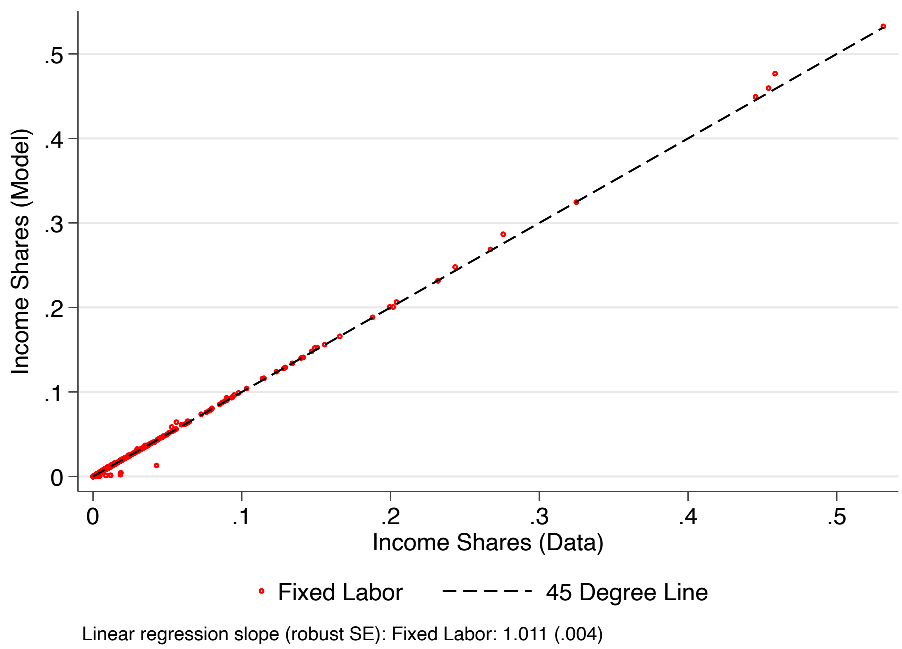
where $J(n)$ is the set of ETOPO1 grid cells contained in country's discretized grid cell n and N(i) is the set of eight neighboring cells to each cell i in ETOPO1. Ruggedness will be higher in places with large changes in elevation. Given this cost function, the model recognizes that it is more costly to build in places with large changes in elevation (i.e. mountainous areas), and less costly per kilometer to build along longer links.

Note that with this methodology, we measure the cost of building roads along a link up to scale in each country. To implement counterfactual analyses, we will set $K = 1$ in the network building constraint of equation 3.4, so we re-scale each δ^I_{jk} to satisfy this constraint. As a result, the cost of building in each link is measured as a share of the initial network size. We do not assign a dollar value cost to building along each link, and thus we cannot perform a cost-benefit analysis of any simulated infrastructure expansions.

3.3 Baseline Results

In our main results, we consider two counterfactual scenarios: a reallocation of the existing road network and a 50% expansion of the existing road network. The reallocation counter-

Figure 3.3: Calibrated Model Fit



factual studies what each country's road network would look like if a social planner took all the resources used to build the country's existing road network, as measured by K in the network building constraint of equation (6), and changed the spatial distribution of the investments made with these resources in order to create the optimal network. Considering this counterfactual allows us to understand the extent to which the existing road network differs from the optimal one, holding constant the resources used to build the network. The expansion counterfactual, on the other hand, asks where the social planner would invest in infrastructure if the resources used to build the existing road network, as measured by K in the network building constraint, were to grow by 50%.

Figures 3.4 and 3.5 shows the location of infrastructure investments in each scenario for Argentina, Bolivia, Brazil, Colombia, Mexico, Paraguay, and Peru. Figures showing results for the remaining countries in our sample are included in the Appendix. In these figures, brighter green, thicker segments represent links with larger levels of infrastructure growth. Infrastructure growth could mean any investments that improve road network connections between nodes; for example, improving road quality, adding new lanes, or creating new roads. In the reallocation scenarios, results for which are shown in the right-hand side panel in each row, we also have red links which represent areas where infrastructure would be reduced if roads were to be reallocated towards the optimal network. Note that in the reallocation case, reduction in infrastructure along a link (i.e., a link colored in red) does not indicate that that link should not have been built; instead, given a fixed level of resources, the level of infrastructure along the link is higher than the social planner would allocate to that link. Similarly, green lines reflect links to which a social planner would allocate more resources.

In Argentina, under both scenarios, new infrastructure investments are made to enhance roads radiating from Buenos Aires toward urban centers in Entre Rios and Santa Fe. Some large investments are also made in poorer areas in the North and in provinces east of the province of Buenos Aires. The optimal network in Argentina indicates overinvestment in roads in the sparsely populated South, areas in the far East and North. In Colombia, new investments center on Bogota, connecting cities in the North, Northwest, and Southwest. The optimal network points to overinvestment in roads east of Bogotá and in the North of the country. In Mexico, the expansion of the network covers the eastern part of Mexico, starting at the border city of Cuidad Juarez, extending to Monterrey and the densely populated core at the center of which is Mexico City. The model suggests overinvestment in trunk road infrastructure in the Western parts of the country and the Yukatan Peninsula. The model suggests optimal expansions along the coast of Peru.

Turning to Bolivia, most investments are in the interior of the country, connected La Paz and El Alto to Santa Cruz. In Paraguay, the optimal investments improve connections between La Asunción and the surrounding cities. In Brazil, the optimal network improves connections along the coast, including between São Paolo, Rio De Janeiro, and the surrounding coastal cities, as well as links connecting northern and southern cities, via Brasilia. In Chile, shown in the Appendix, new investments are concentrated in the central and more populous parts of the country, while the optimal network suggests major overinvestment in the North of Chile (Figure 3.17). Results for the other Latin American countries are also available in the Appendix Figures 3.18 through 3.20. In the following two sections, we study the factors that influence the distribution of infrastructure investments as well as the welfare gains obtained by each country under each scenario.

3.3.1 Drivers of Infrastructure Growth

What drives infrastructure expansion in some places within a country as compared to others? To understand the location of the optimal expansion and reallocation, we pool all the countries together and then estimate the following regression where gI_{ic} is infrastructure growth:

Figure 3.4: Counterfactual Results

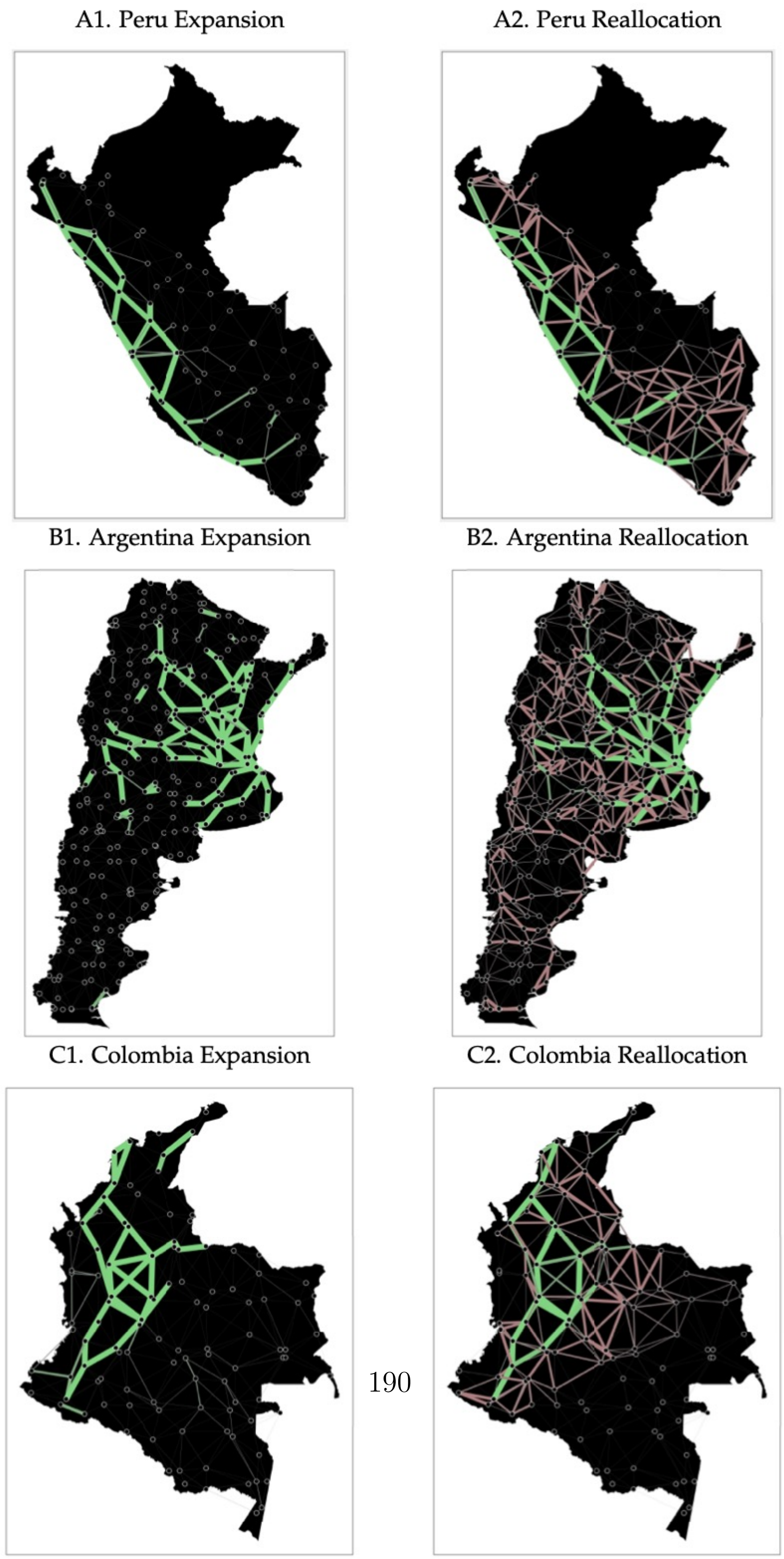


Figure 3.5: Counterfactual Results (cont'd)

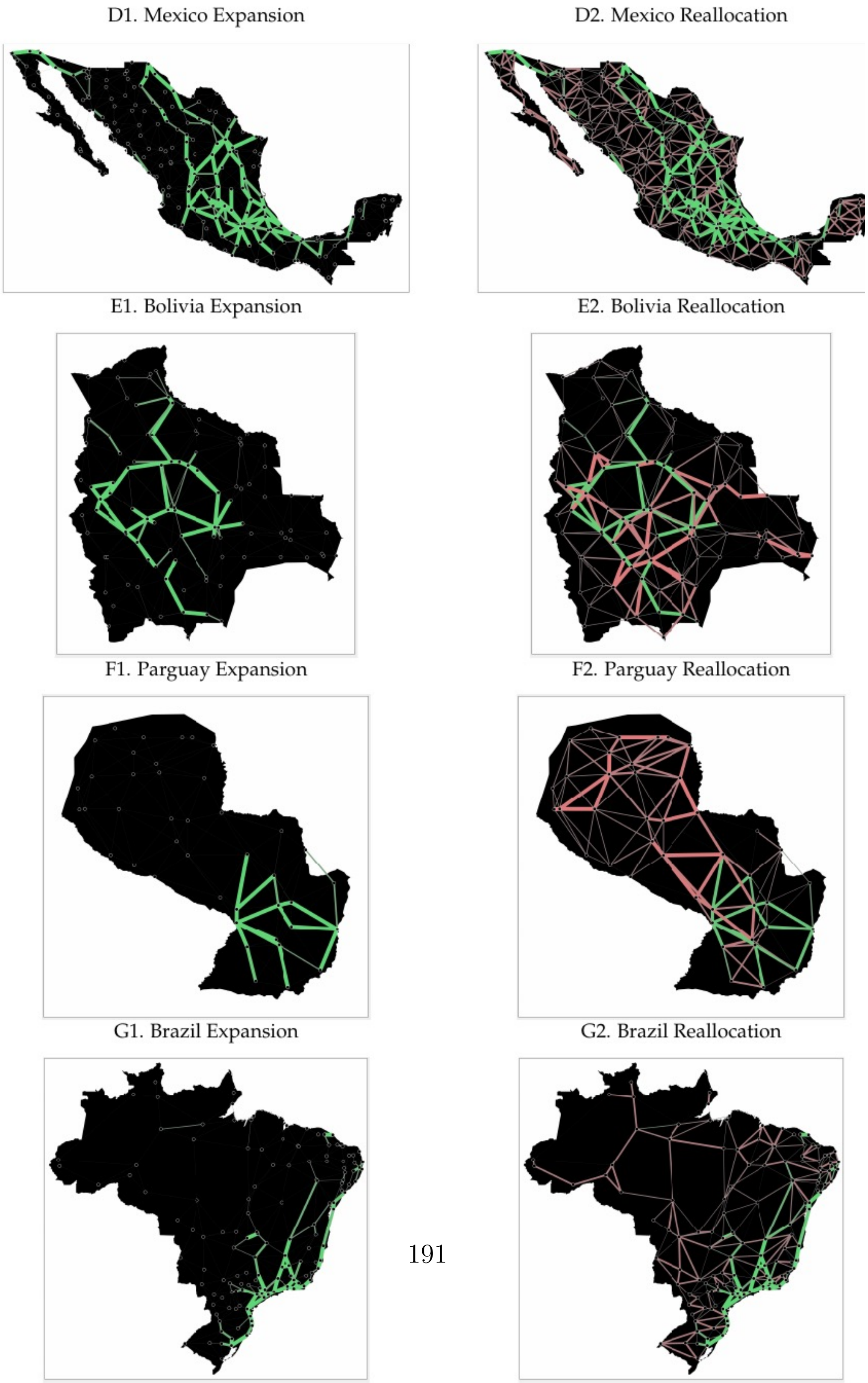


Table 3.2: Drivers of Infrastructure Growth

	Expansion			Reallocation		
	1	2	3	4	5	6
Infrastructure	-0.151***	-0.219***	-0.226***	-0.207***	-0.396***	-0.410***
	-0.00921	-0.0108	-0.00989	-0.0243	-0.0257	-0.0233
Population		0.170***	0.174***		0.471***	0.481***
		-0.0159	-0.0102		-0.0378	-0.021
Tradeable Income per capita			0.109*			0.221*
			-0.0387			-0.092
Differentiated producer			0.109**			0.205***
			-0.0309			-0.0498
Country FE	X	X	X	X	X	X
N	1461	1461	1461	1461	1461	1461
R2	0.484	0.659	0.673	0.27	0.582	0.595

$$gI_{ic} = \alpha + \beta_1 \cdot \log(L_{ic}) + \beta_2 \cdot log(I_{ic}) + \beta_3 \log(y_{ic}) + \beta_4 \cdot + \gamma_c + \epsilon_{ic} \tag{3.5}$$

where where an observation is a cell i in each country c, γ_c is a country fixed effect, y_{ic} is tradable income per capita, L_{ic} is population, and standard errors are clustered at the country level. In Table 3.2 below we show results from estimating equation 3.5 across all LAC countries. Columns (1) through (3) correspond to the case of the 50% expansion in road investment while columns (4) through (6) correspond to the reallocation case. In columns (3) and (6) we also include an indicator for whether a grid cell is a differentiated producer.

The results in columns (1) and (4) show that higher levels of initial infrastructure are associated with smaller increases in infrastructure in both counterfactual scenarios. Initial infrastructure alone can explain almost 50% of the variation in infrastructure growth across grid cells under the expansion scenario and almost 30% under the reallocation scenario. Adding in population as a covariate in columns (2) and (5), we find that grid cells with larger populations experience larger gains in infrastructure; population explains an additional 17% of the variation in infrastructure growth under the expansion scenario and an additional 31% under the reallocation scenario. Finally, in columns (3) and (6), we add income per capita and an indicator for whether a grid cell is a differentiated producer. Grid cells with higher per capita incomes, which are more productive, and cells with differentiated producers, also see higher levels of infrastructure growth. Although differentiated producers are allocated based on the most populous grid cells, these areas become relatively more connected even after controlling for population. In both scenarios, these four covariates explain roughly 60%-70% of the variation in infrastructure growth, with most of the variation being explained by the initial level of infrastructure and the population in the cell.

3.3.2 Welfare Gains

We measure per-capita consumption in each grid cell in the calibrated model, as well as the per-capita consumption in each grid cell resulting from each counterfactual. Given our parametrization of the utility function and measures of population in each grid cell, we can construct aggregate welfare in each country in the initial equilibrium, and again under each counterfactual. Figure 3.6a below shows average welfare gains, measured as a percentage of annual consumption, that each country could expect to obtain under a 50% expansion of the existing road network. For example, assuming that consumption is 70% of GDP, a 2.2% welfare gain in Argentina would mean an annual increase in the country's income of about $6.86 billion. We find that Argentina, Brazil, Peru, Bolivia, and Mexico have the most to gain from an expansion of the existing road network, while Panama, Uruguay, Costa Rica, Guatemala, and El Salvador have relatively less to gain. On average, under the expansion counterfactual, countries in Latin America would gain about 1.6% of consumption, weighing each country by its total population, or around 1% giving each country an equal weight.

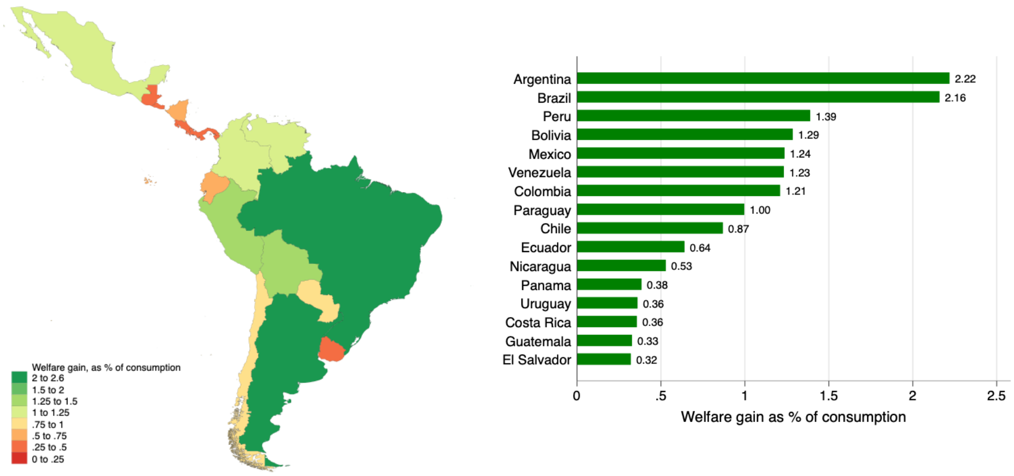
Figure 3.6b below shows welfare gains that countries could expect to obtain from a reallocation of the existing road network. The gains are comparable to those obtained by a 50% expansion of the road network. On average, Latin American countries would gain about 1.6% of consumption weighing each country by its total population, or around 1% unweighted. The road networks of Argentina, Brazil, Peru, and Mexico are relatively more misallocated, with gains from a reallocation ranging from 1.3%-2.4%, while El Salvador, Guatemala, Costa Rica, Uruguay, and Panama are relative less misallocated, with gains below 0.5%. These results are comparable in magnitude to the average gain from network reallocation of 1.7% obtained for Europe by Fajgelbaum and Schaal (2020), as well as to the average gain from network reallocation of 1.1% obtained for Africa by Graff (2019). While the welfare gains vary across countries, we find that as in the European case larger countries in terms of population tend to gain more. In addition, as in Fajgelbaum and Schaal (2020), the welfare gains from optimal networks are substantially larger under non-convex parametrization without congestion and in this case the solution closely matches the least-cost-route optimization solutions in the literature. For example, in the non-convex parametrization case, the welfare gain of Colombia rises from 1.2% to 2.3% and under the expansion case, from 1.2% to 2.9%. This suggests that our estimates of welfare change are relatively conservative.

3.3.2.1 Gains within Countries

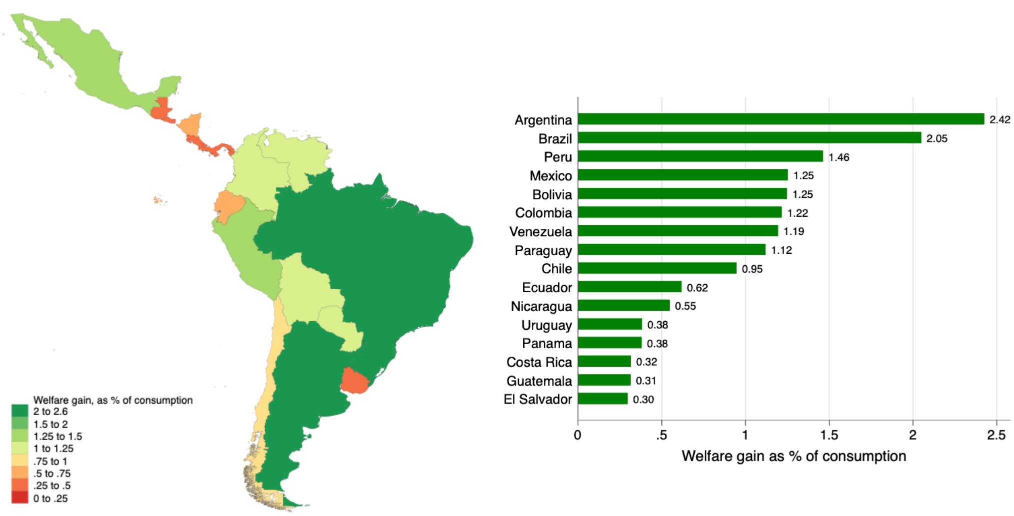
In Figure 3.7, we plot the percentage change in consumption against the log of the initial income per capita for each grid cell in all the countries in our sample, showing separately the data points corresponding to Argentina, Mexico, and Brazil. The results in this figure suggest that the areas that gain the most from the expansion are those that had relatively lower levels of per capita income before the expansion. In all countries, we find a significant negative relationship between the initial level of per capita income in a location and the percentage change in consumption experienced in that location after the implementation of the optimal 50% expansion. As noted by Fajgelbaum and Schaal (2020), this is consistent

Figure 3.6: Welfare gains

(a) 50% Expansion

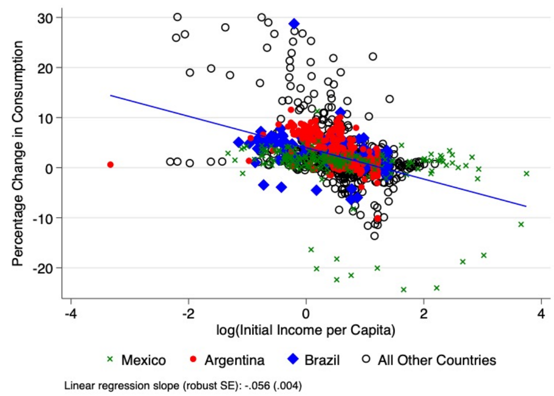


(b) Reallocation



Note: . This figure shows, in panel (a), the welfare gain across countries from the 50% expansion counterfactual, and in panel (b), the welfare gain across countries from the reallocation counterfactual.

Figure 3.7: Correlation of Gains and Initial Income Across Countries



Note: This figure shows the relationship between initial income per capita in each cell within each country, and the percentage change in consumption implied by the model under the 50% expansion case.

with the social planner's desire to equalize the marginal utility of consumption across locations. In this sense, the expansions we study here can reduce inequalities across space by increasing consumption in the locations with initially low levels.

3.3.3 Infrastructure Growth and World Bank Investments

How do the optimal infrastructure investments proposed by the model correlate with investments currently being made in these countries? To study this, we use a list of World Bank Transportation projects in each LAC country between 2005 and 2020 obtained from the World Bank Projects API. Each project is associated with at least one set of latitude and longitude coordinates, and some projects are also associated with a lending amount. The spatial distribution of these projects is shown in Figure 3.8 below. Many of these investments are located close to cities; for example, several investments are clustered around Asunción in Paraguay, Buenos Aires in Argentina, Lima, Peru, and along the Coast of Brazil. In Appendix Figure 3.21, we plot the within-country correlation between population and World Bank investments, where both population and World Bank investments are measured for each grid cells within each country. They are well-correlated in all cases, except in the case of Costa Rica.

For each grid cell in each country, we compute the correlation between the model-implied level of infrastructure growth in the 50% expansion counterfactual first with the total amount of dollars spent on infrastructure projects in that grid cell since 2005 and then with the total number of projects in that grid cell. Figure 3.9 shows these correlations across countries. While this is an imperfect comparison since we do not observe exactly the links that World Bank investment projects are targeting with these projects, some interesting patterns appear. In most countries, the correlation is quite high; for example, in Paraguay, Guatemala, and Colombia. The reason for this is that within most countries there is a high correlation between population and the level of World Bank investment. This relationship is shown in Appendix Figure 3.22: countries with a high level of correlation between population and World Bank investments also have a high level of correlation between model-implied infrastructure growth and World Bank investments. Since the model prioritizes investments in high population areas, we find that World Bank and model-implied investments are generally well-correlated. But this is not, for example, the case in Costa Rica, where most World Bank projects are along the coast while the model highlights investment in the interior, near San Jose.

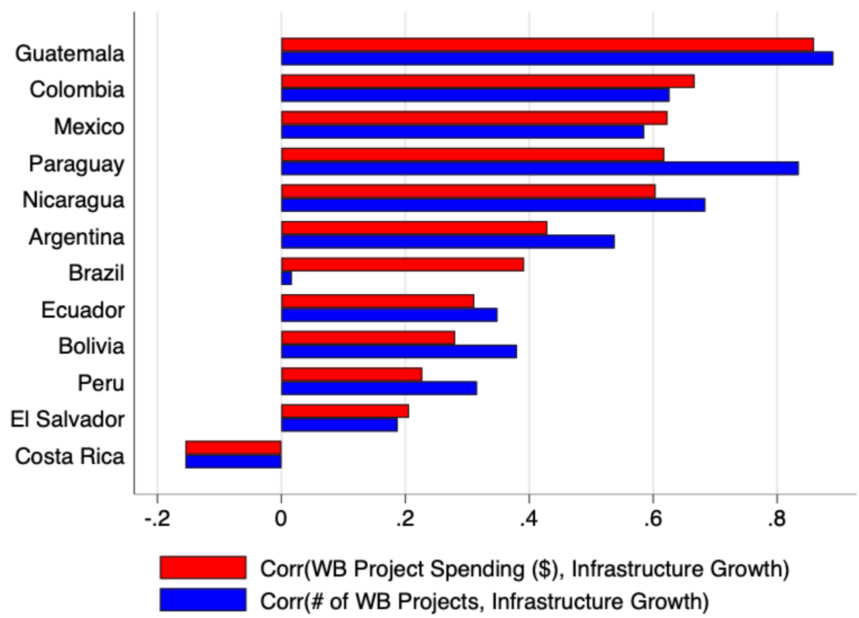
3.4 Robustness

In this section, we consider the robustness of our results to alternative data sources, measures of infrastructure quality, grid cell sizes, and size of the expansion.

Figure 3.8: Distribution of World Bank Transportation Projects

Source: World Bank Projects API. Includes transportation sector projects in each country between 2005 and 2020. Note: Each red dot shows the location of a project; dots are sized by the relative funding of the project. For projects associated with more than one location, funding is assumed to be uniformly distributed across locations.

Figure 3.9: Correlations Between Road Network Investments & World Bank Projects



Note: The blue bars show the country-level correlation between the number of World Bank infrastructure projects within each grid cell and the amount of model predicted infrastructure investment under the 50% expansion counterfactual. The red bars show the country-level correlation between amount of spending on World Bank infrastructure projects within each grid cell and the amount of model predicted infrastructure investment, under the 50% expansion counterfactual. Only countries with more than one observed World Bank Project are included.

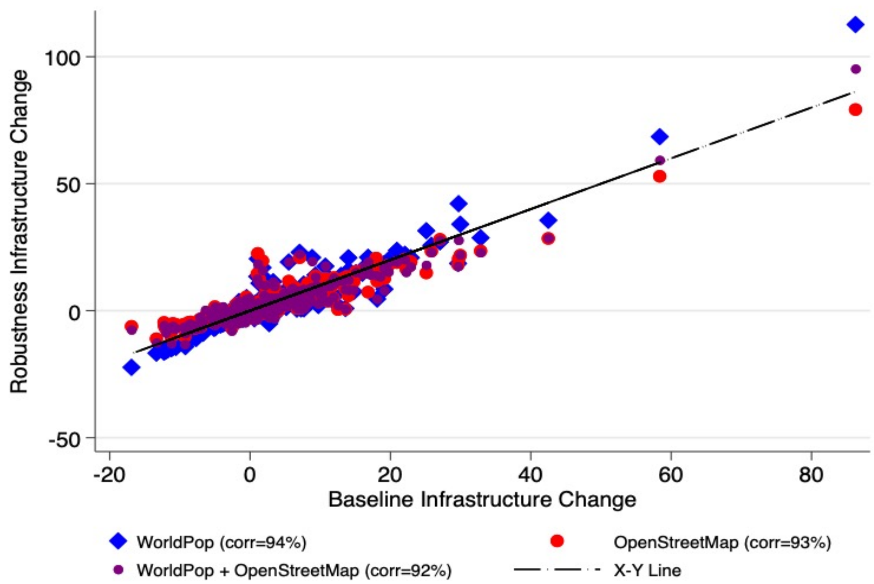
3.4.1 Alternative Data Sources

We undertake three robustness checks related to our data sources. First, instead of using SEDAC's GPW dataset to measure the population in each grid cell, we use data from WorldPop. Second, following Graff (2019), we measure infrastructure quality along a link using the average speed that a car could travel between nodes in the network, as computed via OpenStreetMaps. Finally, we combine both data sources by including WorldPop population data and OSM travel speeds simultaneously. Figure 3.10 shows correlations between changes in infrastructure in each grid cell in each country under the baseline case and under each of the three robustness checks. On the x-axis, we show the change in infrastructure in each grid cell across countries and counterfactuals, relative to the mean. On the y-axis, we show the change in infrastructure in that grid cell in each of these three robustness checks. The correlations are very high, above 90 in all three cases, suggesting our results are not sensitive to these different sources of data. In Appendix Table 3.3, we show the welfare gains for each of these robustness checks across both scenarios in each country. In general, the numbers are very similar across specifications.

3.4.2 Rural Road Quality

One concern with our approach to measuring infrastructure quality is that the data we use may overstate the quality of roads in rural areas. We therefore explore the robustness of our results to this measure by reducing travel speeds in very isolated areas, where road quality may be lower than measured. We focus here on Brazil and Bolivia, which both include large, remote, difficult-to-access areas where this problem may be especially acute. We identify which grid cells are "rural" in each country based on whether population density is below 20 people per square kilometer. In Brazil, we find that about 16% of the population lives in areas that we identify as ultra-low density; in Bolivia, that figure is

Figure 3.10: Robustness Check Road Network Investment Correlations



Note: This figure shows infrastructure changes across locations for each robustness check. The "Baseline" is our baseline specification, using GRIP road network quality data and SEDAC-GPW population data. The "WorldPop" points use WorldPop data on populations of grid cells in lieu of SEDAC-GPW data, and the "OpenStreetMap" uses travel speeds as computed with OpenStreetMap to measure infrastructure quality in lieu of GRIP measures of road segment quality. Finally, "WorldPop + OpenStreetMap" uses WorldPop data on the populations of grid cells and OSM data on travel speeds to measure infrastructure quality.

29%. Given this set of rural grid cells within each country, we use the OSM version of our results as described in Section 4.1 and reduce speeds by 20% along edges with a rural destination and a rural origin, and by 10% if the origin or destination grid cell is rural.

We re-estimate the expansion counterfactual and find that this change has almost no effect in the case of either country. In Brazil, welfare falls by about 0.06 percentage points relative to our baseline results, and the correlation in the location of investments with this rural version and our baseline version is nearly one. Since this change to the quality of rural roads affects only a small share of the population, it is not too surprising our aggregate welfare results do not change much. In the case of Bolivia, where this adjustment to road speeds affects a larger share of the population, we find a similarly very small change in welfare relative to the baseline case of about 0.01 percentage points and a correlation in investments' location of 99%. Because the model prioritizes high-population locations, changing the initial level of road quality in remote areas does not affect the results very much. Thus, we are reassured that our results are robust to changes in our measurement of rural road quality.

3.4.3 10% Expansion

In our main analysis, we considered a 50% expansion counterfactual. However, many countries may face competing priorities and limited fiscal space such that a major infrastructure push may not be possible at this time. Thus, in this section we focus on Brazil and consider a 10% expansion. The welfare gain from a 10% expansion is 1% of annual consumption. This result highlights the non-linearity of gains from improving infrastructure: though we reduce the size of the expansion by 80%, the welfare gains obtained under this much smaller expansion are 44% of the gains obtained under the larger, 50% expansion. In terms of infrastructure placement, the reduction in investment in the 10% case as compared with the 50% case is very uniform across links. The grid-cell level change in investment has a correlation coefficient of 98%; the main difference in the two counterfactuals is the total level of investment allocated to each link which is reduced under the smaller expansion.

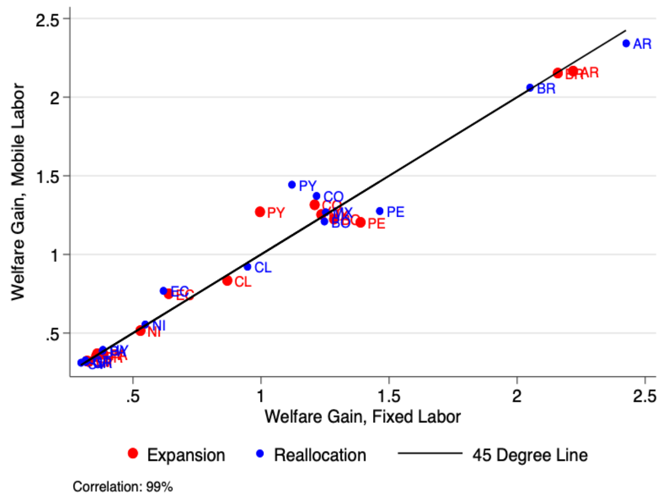
3.4.4 Labor Mobility

In this section, we relax the assumption of immobile labor and allow for perfect labor mobility across space within countries. In this version, we calibrate the model to match not only income per capita in each location, as in the case of immobile labor, but also to match the distribution of the population. Then, we study the effects of the expansion and reallocation counterfactuals. Figure 3.11 shows the correlation between welfare gains in each country under the assumption of fixed labor and under the assumption of perfect labor mobility, for both the reallocation and expansion counterfactuals. In the latter case, workers can reallocate across space given the change in infrastructure and income in each location implied by each counterfactual. In turn, the distribution of investment outcomes factors in that workers are mobile. The correlation between the welfare gains under these two different assumptions is very high, around 99%.

3.5 Transnational Trade Networks

In this section, we evaluate optimal expansions and reallocation in two transnational road networks. First, we consider road connectivity within the group of countries that are signatories to the MERCOSUR free trade agreement (FTA) – Argentina, Brazil, Paraguay, and Uruguay. Second, we explore the road networks connecting countries in the Andean Community, a free trade area including Bolivia, Ecuador, Peru, and Colombia. With this analysis we would like to assess the extent to which road infrastructure plays a role in inflating trade costs and thus limiting the gains from these two regional free trade

Figure 3.11: Welfare Gains Under Labor Mobility Assumptions



Note: This figure shows the correlation in welfare gains across countries in the case where labor is mobile versus in the case where it is immobile.

agreements.[7]

3.5.1 Discretization

To study optimal expansions and reallocations among groups of countries, we deviate from the original methodology in Fajgelbaum and Schaal (2020) and follow a discretization procedure like the one used in the case of Brazil. Although in theory we could implement exactly the same process on groups of countries as we did in earlier sections, in practice the countries in our sample are too large when grouped together for this to be computationally feasible.[8] Thus, for our transnational analyses, we do not rely on square, uniformly-sized

[7] We do not include any border frictions or tariffs that would affect trade of goods across countries in this analysis.

[8] Fajgelbaum and Schaal (2020) adopt this approach for the union of 24 European countries, but the countries in our setting are far too large for this approach to be feasible.

grid cells but instead use the Level 1 administrative borders (province or state-level) to construct "grid cells" in each country. One challenge with this approach is that it can result in a wide range of cell sizes. For example, Paraguay's Level 1 regions are much smaller both in surface area and population compared with those in Argentina and Brazil. Thus, in the cases of Paraguay and Uruguay, we combine small provinces together. For example, in our MERCOSUR analysis, Paraguay consists of three regions and Uruguay consists of four regions.

As in our single-country analysis, we use SEDAC's GPW to identify the most populous place in each cell; these points become the nodes corresponding to the state or province in which they lie. Thus, the total number of nodes in the connected network will be the total number of states across the group of countries. Edges between nodes are determined based on which grid cells are neighbors, meaning that they share a border. This set of nodes and edges forms the basis for our discretized graph. We compute the total population in each state using SEDAC's GPW and total value added in each state using G-ECON.[9] Since we are no longer working with rectangular cells and we observe population at a very fine granularity, we estimate value added in each grid cell as the population-weighted value added. Figure 3.12 shows the grid that we work with, as well as the relative levels of population in each cell. Purple lines indicate country borders. To measure distances between links and infrastructure quality across links, we use OpenStreetMap to compute travel distances and travel speeds between each node in the network.[10] As in our OSM robustness check, we define infrastructure quality along a link as the average travel speed along that link. We also use OSM to identify whether links exist in the real road network:

[9]Another option is to use official statistics for population and/or income as we did in the case of Brazil. However, in this cross-country case, we prefer to use a uniform source because we would like to avoid combining slightly different data on countries within each connected group.

[10]This is the same methodology used in one of our robustness checks. The road networks of each country are extremely complex, and this approach eases the computational burden associated with combining them together.

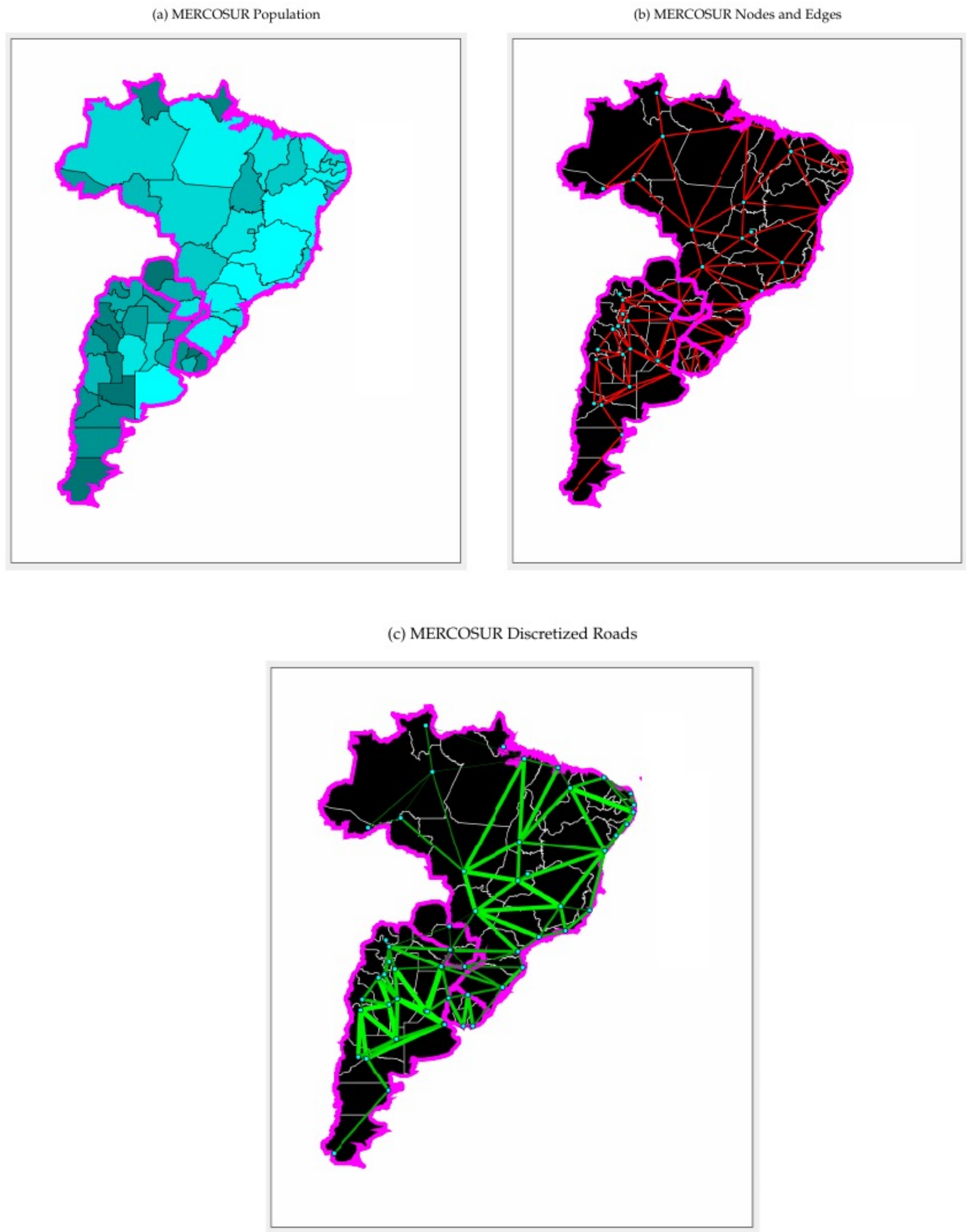
if there is no road network path between two points, we exclude the edge connecting those two points from our discretized road network.

Figure 3.12, panel (c) shows the discretized road network for the MERCOSUR group of countries; in this case, all links are colored in green because there is no differentiation between primary versus secondary roads. Thicker, brighter links are those with higher quality infrastructure as measured by high travel speeds, while thinner, dimmer lines are those with lower infrastructure quality. Figure 3.13 shows the same set of figures for the Andean Community countries. Maintaining the same assumptions on preferences and technology described in Section 3.2, we then calibrate the fundamentals of the model as in our single-country analysis.

Following Fajgelbaum and Schaal's application to transnational road networks within Europe, we assume that each country produces a country-specific differentiated product, in addition to a homogeneous good and use the same parameters as in the benchmark case. We assume that the largest locations in terms of observed population within each country produce the differentiated product of that country, while the remaining locations produce the homogeneous product. In the case of MERCOSUR, we set the number of differentiated producers to 7 in Brazil (the largest and most populous country in this setting), 5 in Argentina, and one each in Paraguay and Uruguay. In the case of the Andean Community, we set the number of differentiated producers to 3 in each country except Bolivia, where we assume only one location produces a differentiated product.[11]

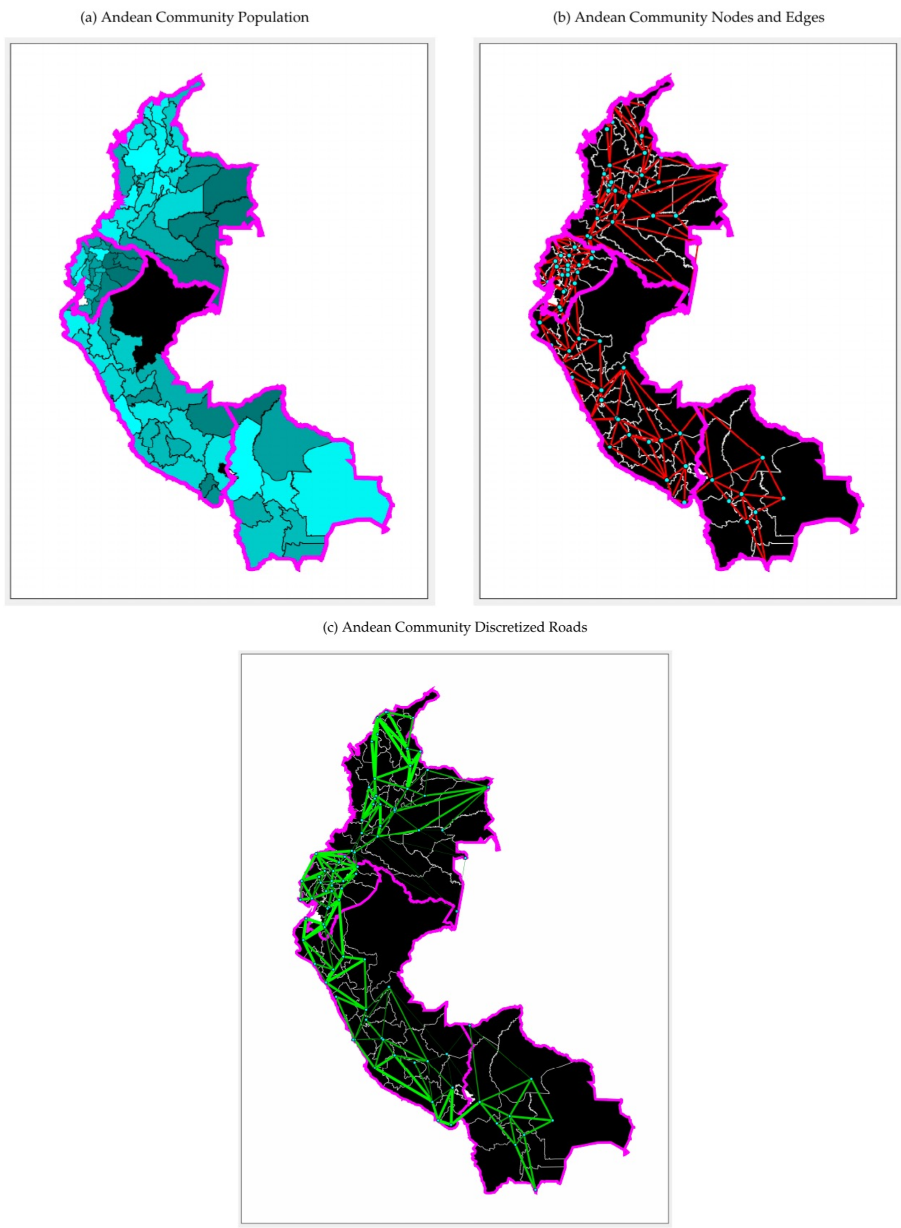
[11] Choices governing the number of differentiated producers in each country are motivated by the number of cells in each country's grid; for example, Bolivia has only 9 cells compared with Colombia where there are 33 cells.

Figure 3.12: MERCOSUR Discretization



Note: Brighter cells in panel (a) show cells with larger population. Thicker brighter links show higher quality infrastructure in panel (c). Magenta lines show country borders.

Figure 3.13: Andean Community Discretization



Note: Brighter cells in panel (a) show cells with larger population. Magenta lines show country borders.

3.5.2 Results

We study the same 50% expansion and reallocation counterfactuals examined on a country-by-country basis in Section 3.3. First, we consider these counterfactuals in the case of the MERCOSUR countries. Panels (a) and (b) of Figure 3.14 show the results for the optimal 50% expansion and the optimal reallocation for this group of countries, respectively. In the expansion case, the model tells us that it is best to improve the roads connecting the largest cities of each member country (which in our model are also the locations producing the differentiated good). Given the location of these cities, these are improvements mostly along the coastal highways.

Most of the investments, as measured by the percentage of total infrastructure growth, are in Brazil (71%) and Argentina (22%), while the remainder is split between Uruguay and Paraguay. The optimal reallocation produces similar results, with resources allocated away from the less populous areas of Brazil and Argentina to finance the expansion. We find that the expansion yields an annual welfare increase of 1.91%, while the reallocation would yield a welfare gain of 1.75%. The results indicate that deficiencies in the road network connecting major cities in MERCOSUR member countries increase trade costs and limit the gains from regional trade.

Second, we consider these counterfactual scenarios for countries in the Andean Community. The results are shown in Figure 3.15. We find that the optimal expansion yields a welfare gain of 1.47% while the optimal reallocation yields a welfare gain of 1.54%. Both the optimal expansion and reallocation improve connections between La Paz in Bolivia, along the coast of Peru to Lima, and through Quito to Medellin. In the case of the reallocation scenario, resources for this investment are drawn from the interior of each country. In the optimal expansion, 50% of the infrastructure growth is in Colombia, 25% is in Peru, 23% in Ecuador, and the remainder in Bolivia.

Figure 3.14: MERCOSUR Counterfactual Networks

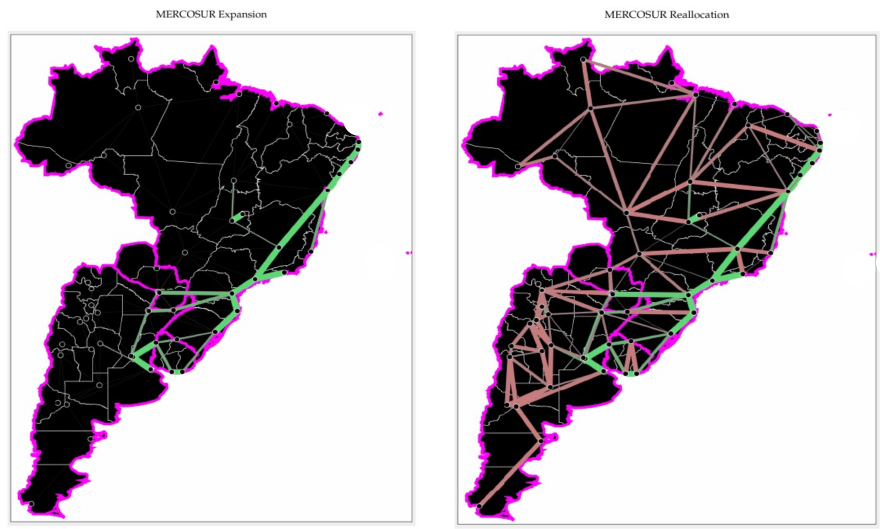
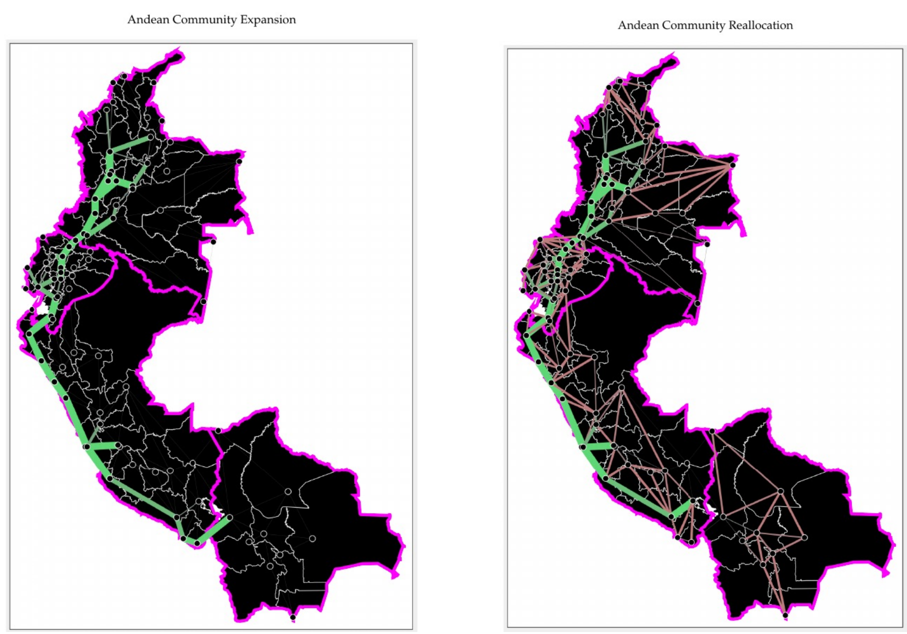


Figure 3.15: Andean Community Counterfactual Networks



Welfare gains under both counterfactual scenarios vary across countries. In Figure 3.16, we show the welfare gains obtained by each country within each grouping under both counterfactual scenarios. Figure 3.16a shows country-level welfare gains for MERCOSUR and Figure 3.16b shows the same for the countries in the Andean Community. Among the MERCOSUR countries, we find that Paraguay (+3.3%) and Brazil (+2.3%) experience the largest gains; Argentina experiences the smallest gains and a small decline in welfare in the reallocation scenario (-0.29%). Among the Andean Community, gains are largest in Bolivia (+5.67%) and smallest in Peru (+0.67%).

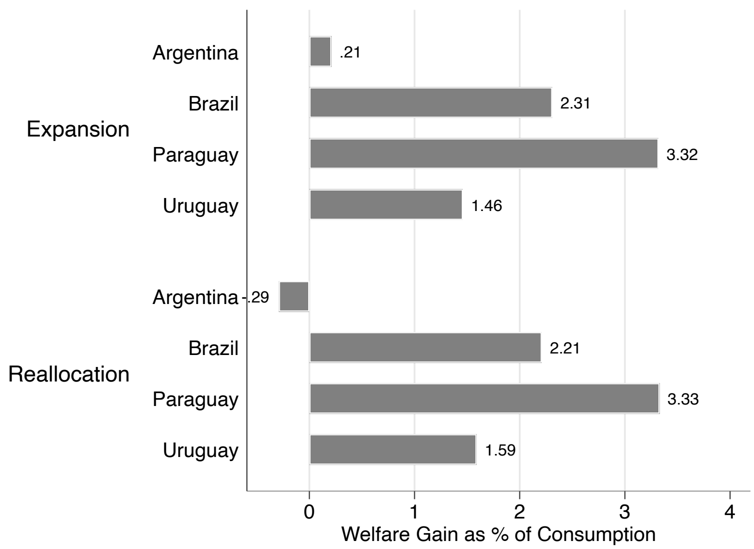
3.6 Conclusion

We explore the extent to which road connectivity issues affect the efficient spatial distribution of economic activity within and across countries in Latin America either because the existing road infrastructure is spatially misallocated or because it is insufficient. Using the general equilibrium spatial framework of Fajgelbaum and Schaal (2020) and data from multiple sources, we construct optimal transport networks and optimal expansions to existing networks in most Latin American countries, as well as within MERCOSUR and the Andean Community. We assess the average annual welfare losses due to inefficient domestic road networks in Latin America at 1.6%, if weighted by population, and just below 1% in simple average terms. Spatial inefficiencies are highest in Argentina, Brazil, and Peru, where the welfare losses due to such inefficiencies average 2.4%, 2.0% and 1.5%, respectively, and lowest in El Salvador, Guatemala, and Costa Rica where the losses are 0.3%. These results are robust to changes in data sources and model assumptions and suggest that domestic trade costs associated with inefficient road networks are sizable in the most populous economies in Latin America.

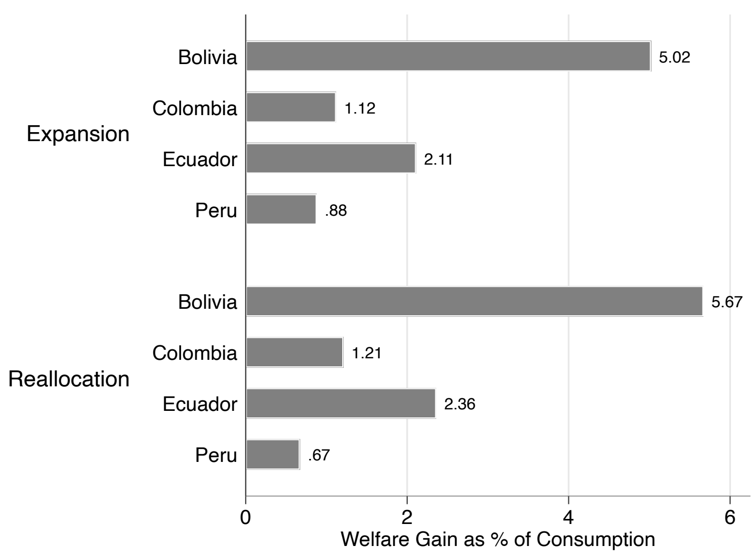
We identify optimal expansions to existing networks that can correct these inefficiencies

Figure 3.16: Transnational Welfare Gains

(a) MERCOSUR



(b) Andean Community



Note: This figure shows welfare gains across countries for both MERCOSUR and the Andean community.

and we show that these investments tend to reduce spatial inequality. The average regional welfare gains of these investments are on par with those assessed by Fajgelbaum and Schaal (2020) for Europe. Model-implied optimal investments in improving and expanding existing networks correlate relatively well with World Bank road investments because both the model and the World Bank prioritize projects in high population areas. Within countries, we show that the optimal road infrastructure investments tend to boost consumption in areas with lower levels of per capita income before the expansion. This is consistent with the model's objective to equalize the marginal utility of consumption across locations.

Identifying the optimal transnational road networks for the countries that are signatories to MERCOSUR and the Andean Community allows us to determine the extent to which trade costs deter regional trade. We find that spatial misallocation of transnational road networks is associated with average annual welfare losses of 1.8% in MERCOSUR and 1.5% in the Andean Community. These losses can be remedied with road investments that improve and expand the existing road networks. In the case of MERCOSUR, expansion yields an annual welfare increase of 1.9%, while in the case of the Andean Community, the gain is 1.5%. In both cases, the transnational expansions benefit the most the poorest country in the trade bloc. The model improves connectivity between the largest cities within MERCOSUR and between the largest cities in each member country. Given the location of these cities, these are improvements mostly along the coastal highways. Most of the investments occur in Brazil (71%) and in Argentina (22%), with the remainder split equally between Uruguay and Paraguay. Within the Andean community, half of the infrastructure growth occurs in Colombia, a quarter each in Peru and Ecuador, and only 2% in Bolivia. Optimal investments improve connectivity between La Paz in Bolivia, along the coast of Peru to Lima, and through Quito to Medellín.

It is important to keep in mind the following caveat. The 50% expansion of the road network depicts a scenario equivalent to a major road infrastructure push. However, the paper does not factor in the financing costs of increasing the size of the infrastructure budget. If resources are raised by increasing taxes or pulling resources from other public investments, the welfare gains of road infrastructure investments would be smaller. Regardless, the findings of the paper are useful and timely because they are indicative of the optimal spatial distribution of road infrastructure projects. By optimally locating their road projects, governments can lower trade costs and achieve a bigger growth boost per dollar spent.

3.7 Appendix

Table 3.3: Robustness Check Welfare Gains

	Expansion				Reallocation			
Country	**Base**	**WorldPop**	**OSM**	**WorldPop + OSM**	**Base**	**WorldPop**	**OSM**	**WorldPop + OSM**
Bolivia	1.3	1.8	1.5	1.5	1.2	1.8	1.5	1.5
Chile	0.9	0.8	0.7	0.7	0.9	0.9	0.7	0.7
Costa Rica	0.4	0.4	0.3	0.3	0.3	0.4	0.3	0.3
Ecuador	0.6	0.7	0.5	0.6	0.6	0.7	0.5	0.6
El Salvador	0.3	0.2	0.1	0.1	0.3	0.2	0.1	0.1
Guatemala	0.3	0.3	0.2	0.2	0.3	0.3	0.2	0.2
Nicaragua	0.5	0.6	0.4	0.5	0.5	0.7	0.4	0.5
Panama	0.4	0.6	0.6	0.6	0.4	0.6	0.6	0.6
Paraguay	1.0	1.0	0.6	0.6	1.1	1.1	0.7	0.7
Uruguay	0.4	0.4	0.3	0.3	0.4	0.4	0.3	0.3
Venezuela	1.2	1.2	1.0	1.1	1.2	1.2	1.0	1.1
Bolivia	1.3	1.8	1.5	1.5	1.2	1.8	1.5	1.5
Chile	0.9	0.8	0.7	0.7	0.9	0.9	0.7	0.7
Costa Rica	0.4	0.4	0.3	0.3	0.3	0.4	0.3	0.3

Note: This table shows welfare gains across countries under robustness checks. The "Base" column lists our baseline welfare estimate, using GRIP road network quality data and SEDAC-GPW population data. The "WorldPop" column uses WorldPop data on populations of grid cells in lieu of SEDAC-GPW data, and the "OSM" uses travel speeds as computed with OpenStreetMap to measure infrastructure quality in lieu of GRIP measures of road segment quality. Finally, "WorldPop + OSM" uses WorldPop data on the populations of grid cells and OSM data on travel speeds to measure infrastructure quality.

Figure 3.17: Expansions and Reallocations

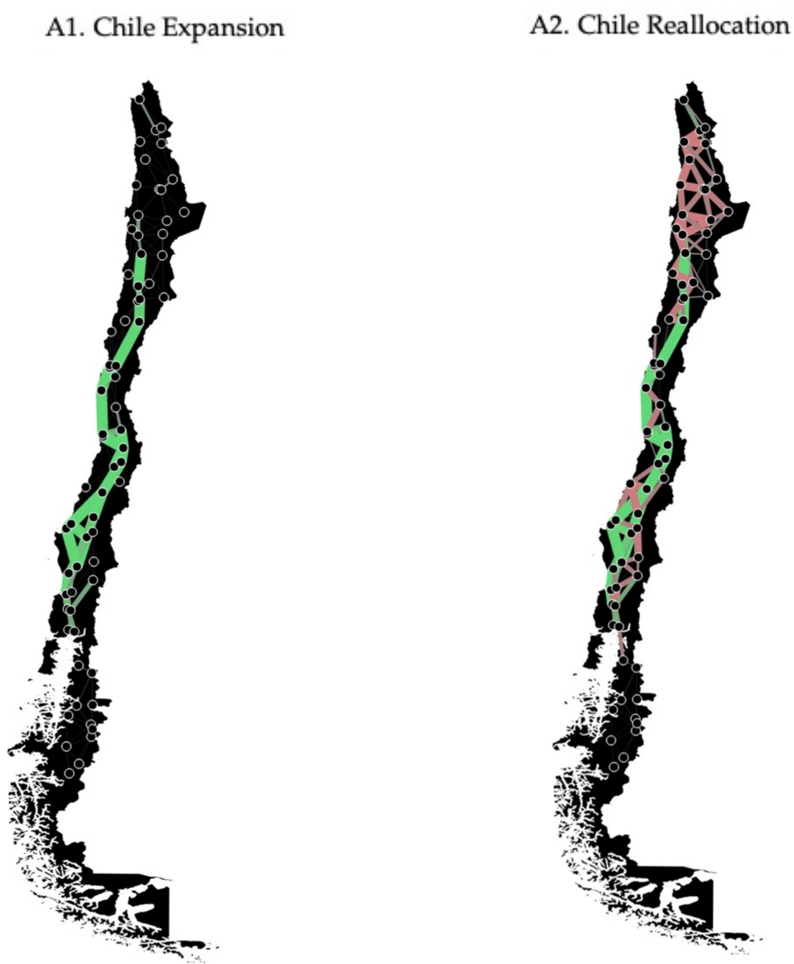


Figure 3.18: Expansions and Reallocations

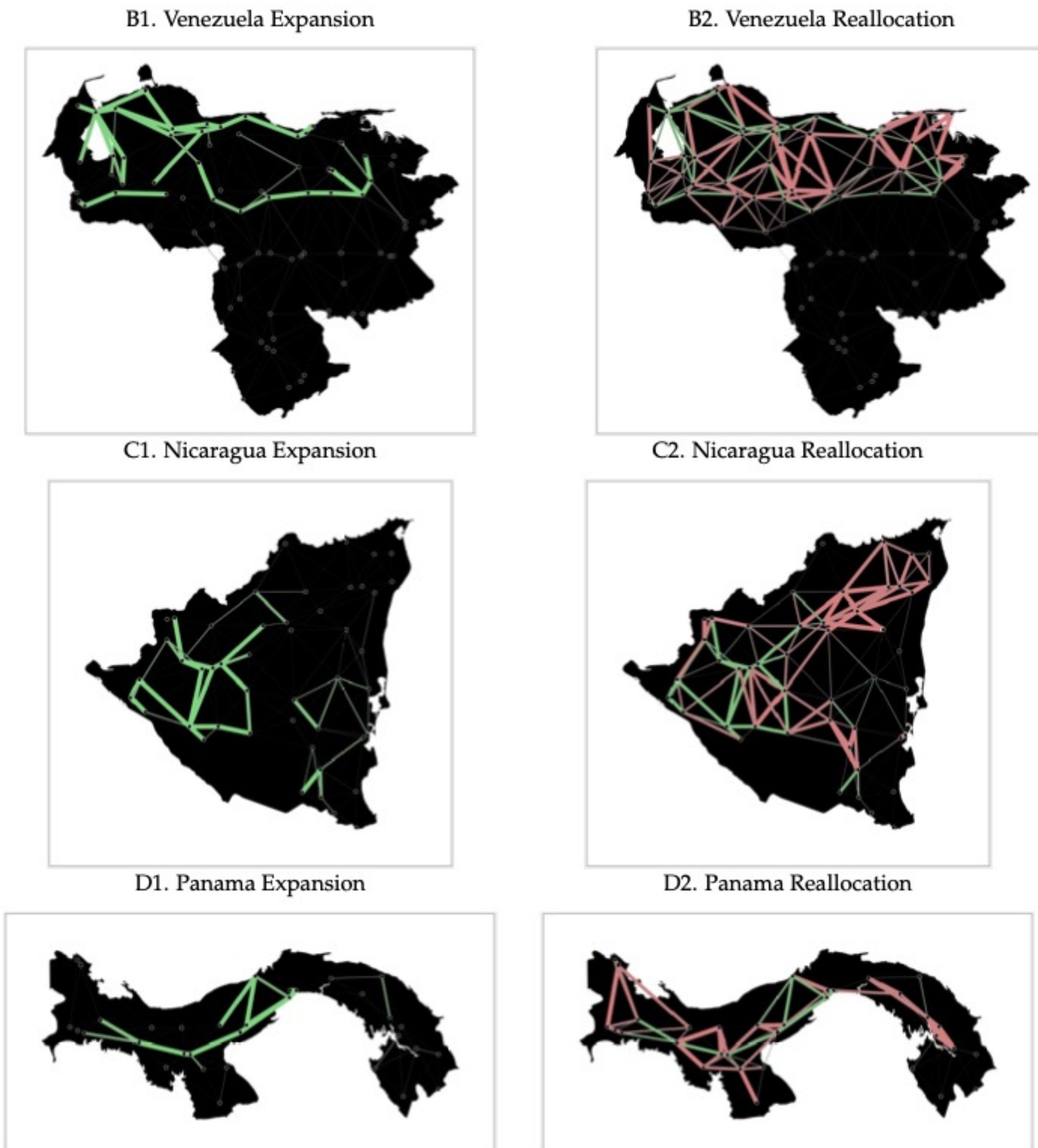


Figure 3.19: Expansions and Reallocations

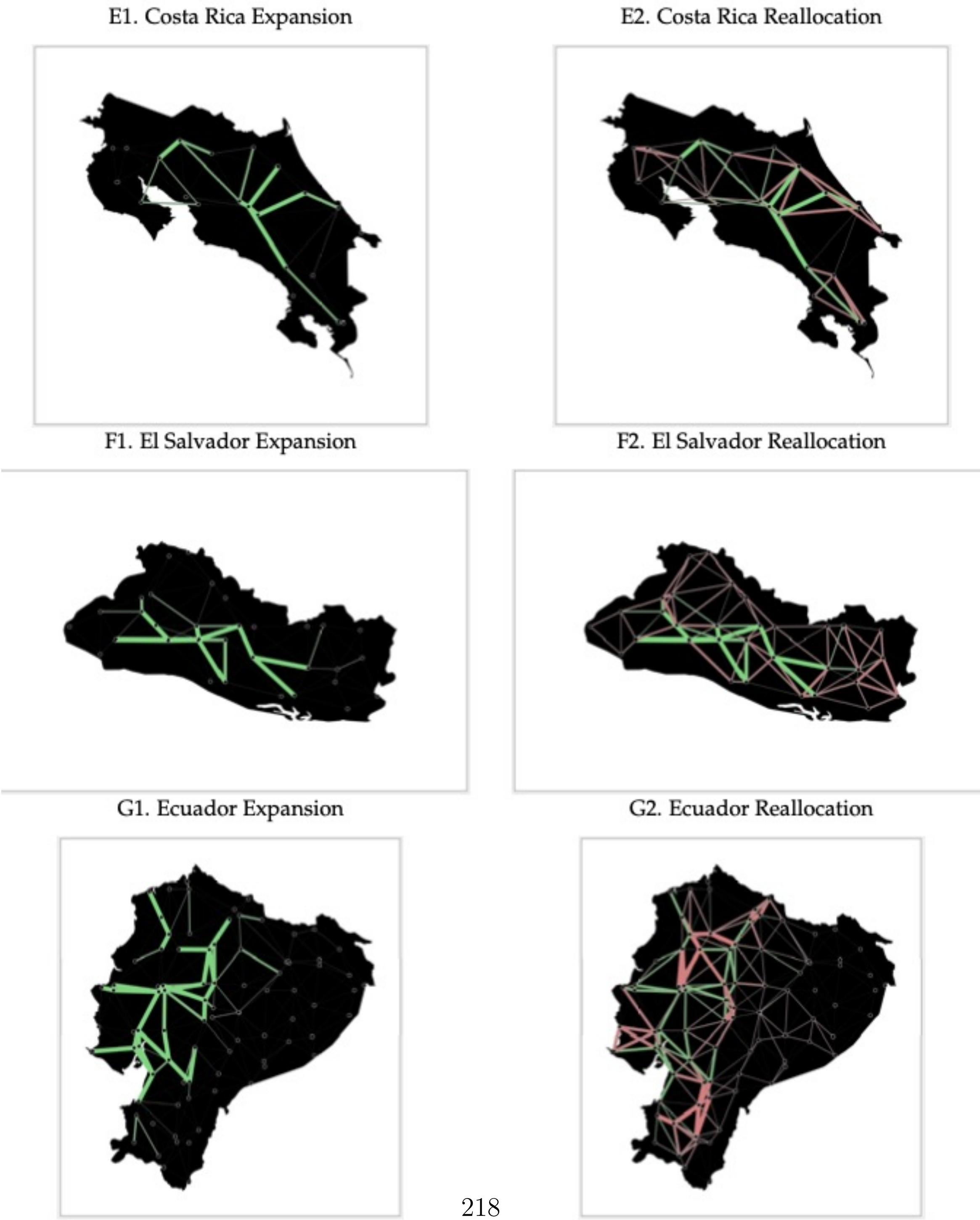


Figure 3.20: Expansions and Reallocations

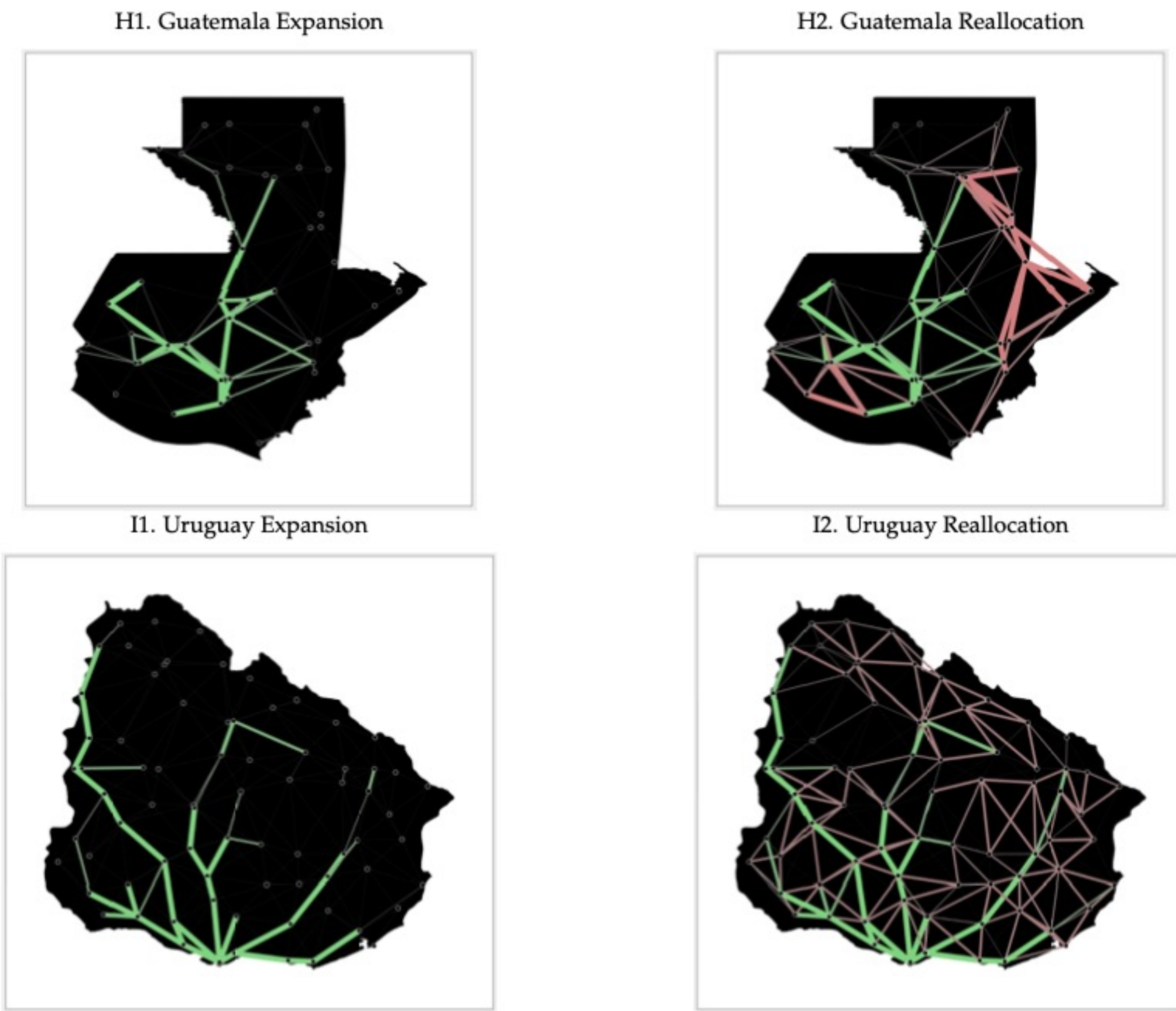
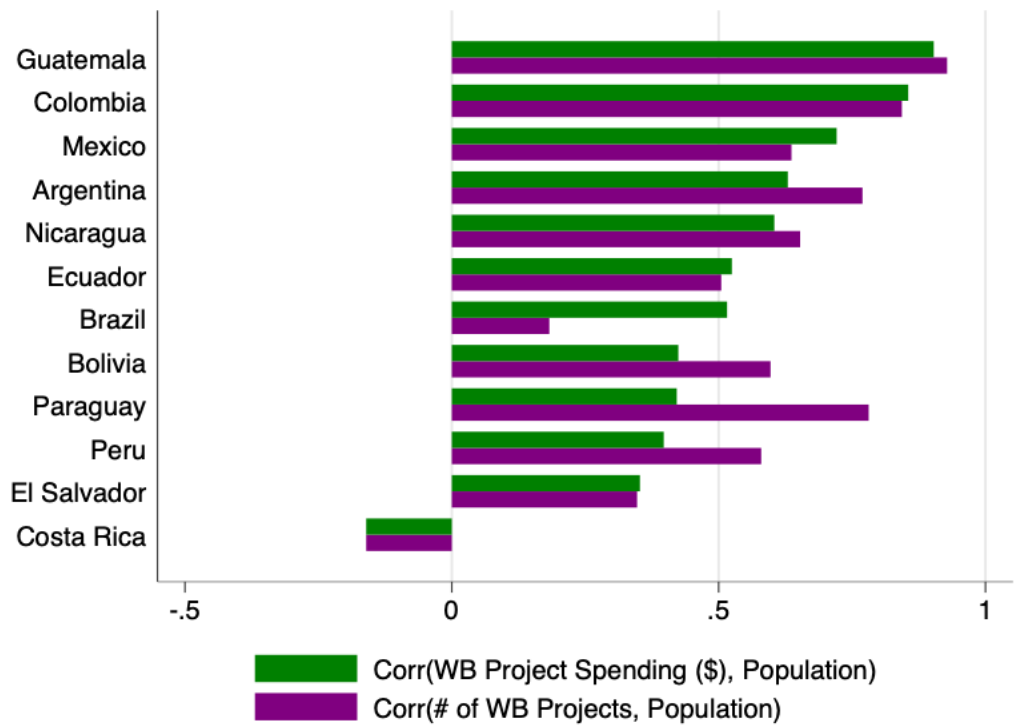
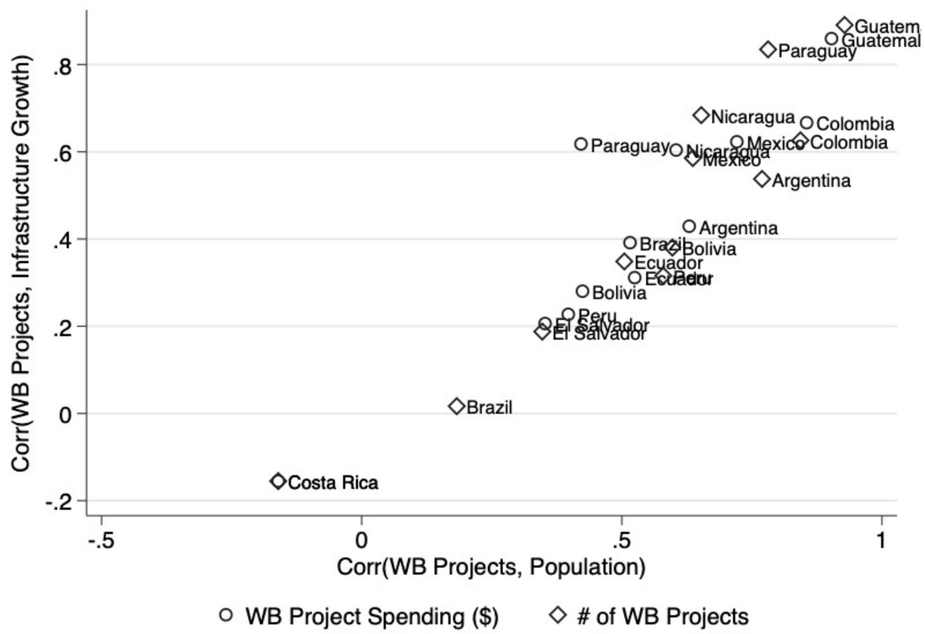


Figure 3.21: Correlations Between Population and World Bank Investments



Note: These are correlations across grid cells within a country between population and the level of World Bank infrastructure projects, as measured by the number of projects (purple bars) or the total amount of spending (green bars). Only countries with more than one observed World Bank Project are included.

Figure 3.22: Population, Model-Implied Infrastructure, and World Bank Investments



Note: This figure plots, for each country, the correlation between World Bank projects and population on the x-axis and the correlation between World Bank projects and the model-implied infrastructure growth under the expansion counterfactual on the y-axis.

BIBLIOGRAPHY

Acemoglu, Daron, and Melissa Dell. 2010. "Productivity differences between and within countries." *American Economic Journal: Macroeconomics*, 2(1): 169–88.

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. 2015. "The economics of density: Evidence from the Berlin Wall." *Econometrica*, 83(6): 2127–2189.

Alder, Simeon, David Lagakos, and Lee Ohanian. 2014. "Competitive pressure and the decline of the Rust Belt: A macroeconomic analysis." National Bureau of Economic Research.

Alder, Simon, and Illenin Kondo. 2020. "Political Distortions and Infrastructure Networks in China: A Quantitative Spatial Equilibrium Analysis."

Allen, Treb, Costas Arkolakis, and Yuta Takahashi. 2019. "Universal gravity." *Journal of Political Economy*.

Anderson, James E, and Eric Van Wincoop. 2004. "Trade costs." *Journal of Economic literature*, 42(3): 691–751.

Antras, Pol, and Elhanan Helpman. 2004. "Global sourcing." *Journal of political Economy*, 112(3): 552–580.

Antras, Pol, Teresa Fort, and Felix Tintelnot. 2022. "Trade Policy and Global Sourcing: An Efficiency Rationale for Tariff Escalation." *University of Chicago, Becker Friedman Institute for Economics Working Paper*, , (2022-89).

Atkin, David, and Dave Donaldson. 2015. "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs." National Bureau of Economic Research.

Autor, David H, David Dorn, Gordon H Hanson, and Jae Song. 2014. "Trade adjustment: Worker-level evidence." *The Quarterly Journal of Economics*, 129(4): 1799–1860.

Babcock, Michael. 1976. "An Empirical Locational Analysis Of The Wheat Flour Milling Industry." *The Review of Regional Studies*, 6(3): 11–18.

Babcock, Michael W, Gail L Cramer, and William A Nelson. 1985. "The impact of transportation rates on the location of the wheat flour milling industry." *Agribusiness*, 1(1): 61–71.

Bartelme, Dominick G, Arnaud Costinot, Dave Donaldson, and Andres Rodriguez-Clare. 2019. "The textbook case for industrial policy: Theory meets data." National Bureau of Economic Research.

Baum-Snow, Nathaniel. 2007. "Did highways cause suburbanization?" *The Quarterly Journal of Economics*, 775–805.

Bernard, Andrew B, Andreas Moxnes, and Yukiko U Saito. 2019. "Production networks, geography, and firm performance." *Journal of Political Economy*, 127(2): 639–688.

Bird, J, M Lebrand, and A Venables. 2019. "The Belt and Road Initiative Reshaping Economic Geography in Central Asia?/World Bank Policy Research Working Paper№ 8807-Washington DC."

Brueckner, Jan K, and Harris Selod. 2006. "The political economy of urban transport-system choice." *Journal of Public Economics*, 90(6-7): 983–1005.

Caliendo, Lorenzo, and Fernando Parro. 2015. "Estimates of the Trade and Welfare Effects of NAFTA." *The Review of Economic Studies*, 82(1): 1–44.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte. 2017. "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy." *The Review of Economic Studies*, 85(4): 2024–2096.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte. 2018. "The impact of regional and sectoral productivity changes on the US economy." *The Review of economic studies*, 85(4): 2042–2096.

Caselli, Francesco, and Wilbur John Coleman II. 2001. "The US structural transformation and regional convergence: A reinterpretation." *Journal of political Economy*, 109(3): 584–616.

Chandra, Amitabh, and Eric Thompson. 2000. "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system." *Regional Science and Urban Economics*, 30(4): 457–490.

Chaney, Thomas. 2008. "Distorted gravity: the intensive and extensive margins of international trade." *American Economic Review*, 98(4): 1707–21.

Costinot, Arnaud, and Dave Donaldson. 2016. "How large are the gains from economic integration? Theory and evidence from US agriculture, 1880-1997." National Bureau of Economic Research.

Cox, Lydia. 2021. "The Long-Term Impact of Steel Tariffs on US Manufacturing." mimeo, Harvard University.

Desmet, Klaus, and Esteban Rossi-Hansberg. 2013. "Urban Accounting and Welfare." *American Economic Review*, 103(6): 2296–2327.

Dingel, Jonathan I, and Felix Tintelnot. 2020. "Spatial economics for granular settings." National Bureau of Economic Research.

Dixit, Avinash, and John Londregan. 1996. "The determinants of success of special interests in redistributive politics." *the Journal of Politics*, 58(4): 1132–1155.

Donaldson, Dave. 2016. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, 108(4): 899–934.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review*, 108(4-5): 899–934.

Donaldson, Dave, and Richard Hornbeck. 2016*a*. "Railroads and American Economic Growth: A Market Access Approach." *The Quarterly Journal of Economics*, qjw002.

Donaldson, Dave, and Richard Hornbeck. 2016*b*. "Railroads and American Growth: A "Market Access" Approach." *Quarterly Journal of Economics*, 799–858.

Duranton, Gilles, Peter M Morrow, and Matthew A Turner. 2014. "Roads and Trade: Evidence from the US." *The Review of Economic Studies*, 81(2): 681–724.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, geography, and trade." *Econometrica*, 70(5): 1741–1779.

Eckert, Fabian, and Michael Peters. 2018. "Spatial structural change." *Unpublished Manuscript.*

Faber, Benjamin. 2014. "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System." *The Review of Economic Studies*, rdu010.

Fajgelbaum, Pablo, and Cecile Gaubert. 2018. "Optimal spatial policies, geography and sorting." National Bureau of Economic Research.

Fajgelbaum, Pablo, and Stephen Redding. 2022. "Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914." *Journal of Political Economy*, 130(5).

Fajgelbaum, Pablo D, and Edouard Schaal. 2020. "Optimal transport networks in spatial equilibrium." *Econometrica*, 88(4): 1411–1452.

Fajgelbaum, Pablo D, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar. 2018. "State taxes and spatial misallocation." *The Review of Economic Studies*, 86(1): 333–376.

Fay, Marianne, Luis Alberto Andres, Charles Fox, Ulf Narloch, and Michael Slawson. 2017. *Rethinking infrastructure in Latin America and the Caribbean: Spending better to achieve more.* World Bank Publications.

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

Glaeser, Edward L, and Giacomo AM Ponzetto. 2018. "The political economy of transportation investment." *Economics of Transportation*, 13: 4–26.

Glaeser, Edward L, and Janet E Kohlhase. 2004. "Cities, regions and the decline of transport costs." In *fifty years of regional science*. 197–228. Springer.

Gollin, Douglas, and Richard Rogerson. 2010. "Agriculture, Roads, and Economic Development in Uganda." *NBER Working Paper Series.*

Graff, Tilman. 2019. "Spatial Inefficiencies in Africa's Trade Network." National Bureau of Economic Research.

Grossman, Gene, and Elhanan Helpman. 2021. "When Tariffs Disrupt Global Supply Chains."

Harwood, Joy. 1991. "U.S. Flour Milling on the Rise." *Food Research and Policy*, 34–38.

Held, Walter G. 1979. "U.S. Wheat Industry." *USDA Agricultural Economic Report*, , (32).

Holian, Matthew J, and Matthew E Kahn. 2015. "Household demand for low carbon policies: Evidence from California." *Journal of the Association of Environmental and Resource Economists*, 2(2): 205–234.

Hornbeck, Richard. 2012. "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe." *American Economic Review*, 102(4): 1477–1507.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing constraints and spatial misallocation." *American Economic Journal: Macroeconomics*, 11(2): 1–39.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly journal of economics*, 124(4): 1403–1448.

Hyslop, J.D., and R. P. Dahl. 1964. "The Effect of Changes in Transportation Costs on Minnesota Wheat Flour Millers and Oilseed Processors." *University of Minnesota Agricultural Experiment Station*, 3–11.

Kahn, Matthew E, and John G Matsusaka. 1997. "Demand for environmental goods: Evidence from voting patterns on California initiatives." *The Journal of Law and Economics*, 40(1): 137–174.

Kim, CS, C Hallahan, Glenn Schaible, and Gerald Schluter. 2001. "Economic analysis of the changing structure of the US flour milling industry." *Agribusiness*, 17(1): 161–171.

Kim, Sukkoo. 1995. "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U. S. Regional Manufacturing Structure, 1860–1987*." *The Quarterly Journal of Economics*, 110(4): 881–908.

Krugman, Paul. 1980. "Scale economies, product differentiation, and the pattern of trade." *The American Economic Review*, 70(5): 950–959.

Limao, Nuno, and Anthony J Venables. 2001. "Infrastructure, geographical disadvantage, transport costs, and trade." *The world bank economic review*, 15(3): 451–479.

Long, Jason, and Henry Siu. 2018. "Refugees from Dust and Shrinking Land: Tracking the Dust Bowl Migrants." *The Journal of Economic History*, 78(4): 1001–1033.

Melitz, Marc. 2003. "The Impact of Transportation Rates on the Location of the Wheat Flour Milling Industry." *Econometrica*, 71(6): 1695–1725.

Meyer, John R., and Alexander L. Morton. 1975. "The U.S. Railroad Industry in the Post-World War II Period: A Profile." In *Explorations in Economic Research, Volume 2, number 4. NBER Chapters*, 449–501. National Bureau of Economic Research, Inc.

Minneapolis Mills Fight for Life, Blame Transit Rates. 1965. "Minneapolis Mills Fight for Life, Blame Transit Rates." *Minneapolis Tribune.*

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, migration, and local employment elasticities." *American Economic Review*, 108(12): 3855–90.

Ndulu, Benno J. 2007. *Challenges of African growth: Opportunities, constraints, and strategic directions.* World Bank Publications.

Nightingale, Edmund. 1967. "Some Effects of Recent Changes in the Railway Grain-Rate Structure on Interregional Competition and Regional Development." In *Transportation Problems and Policies in the Trans-Missouri West.* , ed. Jack Ottoson and Howard Davidson, 105–168. University of Nebraska Press.

Ramondo, Natalia, Andrés Rodríguez-Clare, and Milagro Saborío-Rodríguez. 2016. "Trade, domestic frictions, and scale effects." *American Economic Review*, 106(10): 3159–3184.

Redding, Stephen J, and Matthew A Turner. 2015. "Transportation Costs and the Spatial Organization of Economic Activity, vol. 5 of Handbook of Regional and Urban Economics."

Restuccia, Diego, and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic dynamics*, 11(4): 707–720.

Severen, Christopher. 2019. "Commuting, labor, and housing market effects of mass transportation: Welfare and identification." *The Review of Economics and Statistics*, 1–99.

Sotelo, Sebastian. 2016. "Domestic Trade Frictions and Agriculture." *Manuscript, University of Michigan.*

The Northwestern Miller. 1959. "List of U.S. Flour Mills." *The Northwestern Miller.*

The Northwestern Miller. 1961. "List of U.S. Flour Mills." *The Northwestern Miller.*

The Northwestern Miller. 1967. "List of U.S. Flour Mills." *The Northwestern Miller.*

The Northwestern Miller. 1975. "List of U.S. Flour Mills." *The Northwestern Miller.*

The Northwestern Miller. 1985. "List of U.S. Flour Mills." *The Northwestern Miller.*

The Northwestern Miller. 1990. "List of U.S. Flour Mills." *The Northwestern Miller.*

The Southwestern Miller. 1955. "The Southwestern Miller."

The Southwestern Miller. 1963. "Kansas Millers Plea Against Peril to Trade." *The Southwestern Miller.*

Tsivanidis, Nick. 2019. "Evaluating the impact of urban transit infrastructure: Evidence from bogota's transmilenio."

United States Court of Appeals, Fifth District. 1962. "Arrow Transportation Co. v. Southern Ry. Co."

USDA. 1964. *The Farm Index: 1964 World Report.* Economic Research Service.

Van Patten, Diana, and Esteban Méndez. 2022. "Voting on a Trade Agreement: Firm Networks and Attitudes Toward Openness." National Bureau of Economic Research.

Wilson, Steven. 2009. "Population Dynamics of the Great Plains: 1950 to 2007." *Current Population Reports.*

Wishart, David J. 2004. *Encyclopedia of the great plains.* U of Nebraska Press.

World Trade Organization. 2004. *World Trade Report.*