Growing Apart: Tradable Services and the Fragmentation of the U.S. Economy*

Fabian Eckert
Yale University

Job Market Paper

[Most recent version here]

November 14, 2018

Abstract

Between 1980 and 2010, the college wage premium in U.S. labor markets with larger initial shares of high-skill service employment grew substantially faster than the nationwide average. I show how this trend can be explained within the context of a Ricardian model of interregional trade, where a reduction in communication costs magnifies regional specialization in high-skill services, raising the skill premium in service-exporting regions and reducing it in service-importing regions. Quantitatively, I show that the decline in communication costs I infer from sectoral trade imbalances can explain a substantial part of the differential skill premium growth across U.S. labor markets in the data. These regional changes aggregate to account for 30 percent of the rise in the overall U.S. college wage premium between 1980 and 2010.

*Contact: fabian.eckert@yale.edu. I owe a great debt of gratitude to my advisors for guidance and support throughout the years: Costas Arkolakis, Samuel Kortum, Giuseppe Moscarini, and Michael Peters. I have also learned and benefitted a lot from conversations with Tatjana Kleineberg, Peter Schott, Conor Walsh, and Fabrizio Zilibotti. For early readings I thank Lorenzo Caliendo, Paul Goldsmith-Pinkham, Brian Greaney, Phil Haile, Xiangliang Li, Maximilian Krahé, Monica Morlacco, Yusuke Narita, and Lucas Zavala. For help with data or the sharing of code I thank David Autor, Nathaniel Baum-Snow, David Dorn, Nicole Fortin, Xavier Giroud, Joachim Hubmer, Jed Kolko, and Nicolas Sheard. I also thank Simon Alder, Treb Allen, Adrien Auclert, Bill Brainard, Rafael Dix-Carneiro, Ilse Lindenlaub, Giovanni Maggi, Guillermo Noguera, and Aleh Tsyvinski for comments that improved the paper. All errors my own.

1 Introduction

Over the last 40 years, the U.S. labor market has experienced a significant, sustained increase in the return to skill. The social and economic inequality accompanying this development has divided U.S. society. A voluminous literature has primarily focussed on two canonical explanations for this aggregate trend: (1) increased exposure to international trade; and (2) skill-biased technological change.

However, the aggregate increase in the college wage premium of about 0.8% annually since 1980 masks substantial and systematic regional variation in its growth. Regions with a large share of high-skill service workers in 1980 saw the premium grow twice as fast as regions with lower shares. The college premium has not only grown over time but has also been *growing apart* across regions.

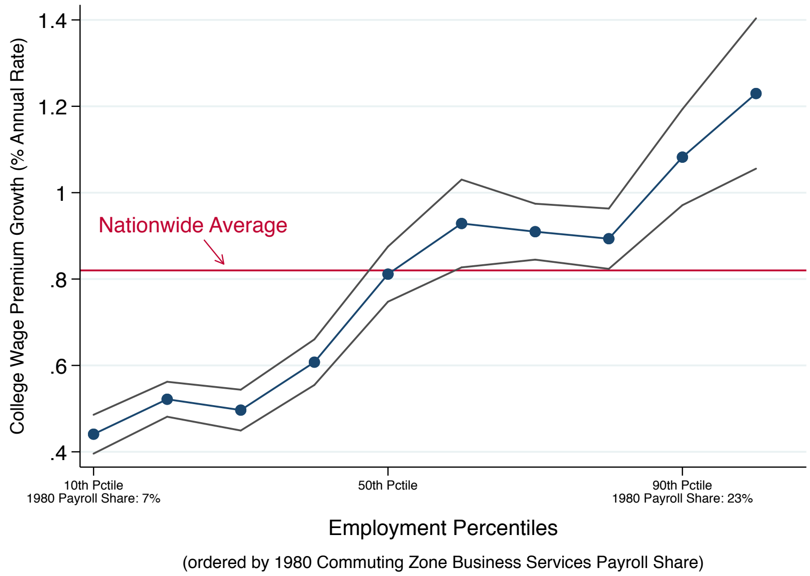
In this paper, I offer a Ricardian explanation for how recent technological change has differentially affected the return to skill across U.S. labor markets. I build on the observation that technological progress has drastically increased labor markets' interconnectedness - a development some commentators have dubbed the "death of distance".[1] With this in mind, I argue that declining trade frictions for high-skill, information-intensive services enabled a small number of local labor markets to provide the most skill-intensive tasks to firms throughout the U.S. economy. A defining characteristic of this ongoing process of specialization is that its gains accrue disproportionately to high-skill workers in exporting labor markets, and to low-skill workers in importing regions. These two effects combine to raise the skilled wage premium in labor markets specialized in such services and to lower it in others.

Consider, for example, the case of Michael Byrd. Byrd founded BakeCrafters, a frozen baked goods company, in 1991 in Chattanooga, TN. In the early 2000s, Byrd started outsourcing his day-to-day customer relationship management to Salesforce, a fast-growing company in San Francisco. A few years later he contracted XERO, a software firm in Denver, to do his accounting. It is likely that communication between Byrd and these service providers occurs via phone calls, email, and occasional in-person meetings. In 2017, Mr. Byrd opened a new distribution center in Lebanon, PA. The company hence raised labor demand for relationship managers and software engineers in San Francisco and Denver (2016 median income: $55k) and low-skill warehouse workers in Lebanon, PA (2016 median income: $27k).[2] All else equal, BakeCrafters helped to raise the college wage premium in Denver and San Francisco and to decrease it in Lebanon.

[1] See the book with the same title by Cairncross (1997) and a similar one by Friedman (2005). Leamer (2007) offers an enlightening review of Friedman (2005) and a discussion of the death of distance hypothesis.

[2] Main Sources: www.FeaturedCustomers.com (2018) and O'Connor (2018).

Figure 1: Skill Premium Growth Across Commuting Zones 1980-2010



Data Source: U.S. Decennial Census (1980-2000) and American Community Survey (2010). 95% Confidence Bands shown. Red Line shows U.S. wide average growth. I define as business services all industries in the NAICS-5 sector except for Waste Management and Motion picture production, distribution, and services. I construct 741 Commuting Zones using the boundaries established in Tolbert and Sizer (1996). I compute the average annual hourly wage growth between 1980-2010 for workers with at least some college and for all others. I then order commuting zones by their 1980 business services payroll share and plot average college wage premium growth rates within employment deciles. Appendix E provides more details.

Throughout the paper, I formalize the notion of "high-skill services" as "business services", a fast-growing class of skill-intensive services mainly used as intermediate inputs that has been recognized by the literature.[3] Examples include management of companies, legal and accounting services, software companies, management consulting, and corporate banking. I refer to frictions inhibiting the trade in business services as "communication costs".[4]

Figure 1 shows high- relative to low-skill wage growth across local labor markets ordered by the fraction of business services in the local payroll in 1980. As the Figure shows, the college premium has risen faster the larger the initial regional payroll share of business

[3]More formally, business services in this paper correspond to the following 2-digit NAICS industries: 51, 52, 53, 54, 55, and 56. Except for waste management and remediation services and which I classify as "local services". See Melvin (1989), Markusen (1989), Fort et al. (2018), and Fort (2017) for recent papers that discuss business/producer services as distinct from consumer services.

[4]Instead of shipping physical output, business service industries communicate information and problem solutions in person or via communication tools, such as computer and cell phones. The so-called "ICT revolution" has been mostly about the rapid progress in developing such communication tools.

services in 1980. I refer to this observation as the *growing apart fact* throughout the paper.

The primary goal of this study is to assess whether changes in communication costs can explain the *slope* of the dotted blue line in Figure 1. A secondary objective is to determine whether changes in communication costs constitute a skill-biased or a skill-neutral form of technological progress, i.e., to gauge their contribution to the change in the aggregate college wage premium since 1980 (the *level* of the horizontal red line in Figure 1).

I begin by documenting two salient features of the business services sector that interact with changes in communication costs to generate differences in the return to skill across labor markets. First, across U.S. labor markets in 1980 the share of business services employment at the 90th relative to the 10th percentile was 1.9 compared to about 1.4 for the goods-producing sector.[5] These numbers hint at marked underlying comparative advantage differences. Second, the business service sector is significantly more skill-intensive than the goods-producing sectors: its college share of employment is more than 2.5 times that of the goods sector, in every decade between 1980 to 2010. A third fact helps to amplify the effect: business services serve as an essential intermediate input into the rest of the economy, with 40% of its output used in goods-production alone.

To illustrate the mechanism, I introduce a simple model of interregional service trade with two regions and two sectors. In its setup, I make assumptions that take the three highlighted empirical properties of business services to their stylized extremes. Out of two regions ("city" and "hinterland"), the city has an exogenous comparative advantage in business service production. The business services sector employs high- and the goods sector low-skill workers. As communication costs fall, the city increases its business services exports, driving out local business service activity in the hinterland. In response, the hinterland increasingly specializes in goods production. Given the differential skill-intensity of the two sectors, these effects combine to raise the skill premium in the city and depress it in the hinterland. If business services are an intermediate input into goods production, the goods sector in the hinterland profits from a decline in input costs, while in the city the same sector suffers from rising input costs, amplifying the effect.

There are two challenges to assessing the quantitative importance of this mechanisms. The first is to infer business service trade flows in the absence of directly observed service shipments between commuting zones. The second is to construct a modeling framework that is flexible enough to be calibrated to match data moments on each of a large number of U.S. local labor markets in 1980 to understand how their distinctive characteristics interact with changes in the trading environment. Models belonging to the recent "Quanti-

[5] In this paper, the goods producing sector comprises all non-service sectors in the economy, i.e., all NAICS industries with 1-digit codes below 5.

tative Spatial Economics" literature (see Redding and Rossi-Hansberg (2017) for a review) can be used to both ends. I embed the simple mechanism into such a quantitative model of interregional trade. In the model, workers with idiosyncratic skills choose their region, sector, and occupation of employment. Additionally, the model features the full set of input-output linkages between sectors.

First I use the structure of the model to infer business service trade flows across regions by building on a simple, yet powerful idea: while in a world without trade, all labor markets have to be self-sufficient, the possibility of trade opens up sectoral deficits and surpluses across regions, reflecting specialization. Drawing on this basic insight, I propose a method that infers service trade flows across regions from local surpluses and deficits that builds on seminal work by Gervais and Jensen (2013). The technique relies on two steps. First I use regional payroll data and information from the input-output tables to construct sectoral deficits and surpluses across regions. Second, I parameterize trade frictions within each sector as a function of distance. I then use the structure of the model to study how sectoral surpluses and deficits change systematically with distance over time to identify changes in trade frictions between 1980 and 2010. The key identification assumption is that for each origin-region productivity is independent of the destination of a shipment, while trade costs depend on the destination in a way that is common across all origins, conditional on the same distance. The estimates suggest that on average delivering a business service input over a distance of 1000 miles has become 70% cheaper between 1980 and 2010.[6]

Ultimately, the effect of a decline in communication costs on skill prices depends on the interplay of regional comparative advantages, sectoral linkages, and the ability of workers to relocate across sectors, occupations, and regions in response. A strength of the quantitative model is its close connection to the data: all parameters on regions' and workers' comparative advantage appear as structural residuals in the model that can be inferred directly from the data by inverting a large set of observable moments. I calibrate the model to match wages by region and education group, and regional employment by sector, occupation, and education group for 741 commuting zones in the United States as well as the aggregate input-output table in 1980.

I use the calibrated model to conduct an exercise aimed at isolating the effect of the decline in communication costs on the spatial distribution of skill prices between 1980 and 2010. In particular, I hold technologies and other parameters fixed at their calibrated levels for 1980, while setting the distance elasticity of service trade, a single parameter, to its

[6]When I apply the same technique to the goods sector, I find that the distance elasticity is roughly constant and of a similar magnitude as existing estimates, in the literature.

estimated 2010 value. The induced effect is as predicted by the simple model: relative to the nationwide trend, the skill premium rises faster in commuting zones with high initial shares of business services employment. The reverse is true for labor markets with larger shares in goods-production in 1980. A simple regression of college wage premium growth on the log payroll share in business services in 1980 run in data and model output reveals that the commuting cost decline can explain about 50% of the positive relationship. I then use auxiliary predictions of the model on differential wage growth rates across sectors and occupations in the different commuting zones to further validate the mechanism.

Regional changes in skill prices and the spatial, sectoral, and occupational relocation of workers combine to generate sizable aggregate effects. The model can explain 30% of the increase in the unconditional aggregate college wage premium between 1980 and 2010. While all skill groups benefit from declining communication costs, high-skill workers see their real wages rise fastest. The college welfare premium increases by 80% over the 30 years of the study, a more substantial increase than nominal wages alone would suggest.

Related Literature Several papers provide evidence that changes in the returns to skill have been spatially unbalanced. Berry and Glaeser (2005), Moretti (2012), Ganong and Shoag (2017), and Giannone (2017) focus on the "end of spatial wage convergence", driven by the fast growth of high-skill wages in a handful of large cities after 1980. Baum-Snow and Pavan (2013) show that the skill premium grew faster in larger metropolitan areas. None of these papers highlights that the extent to which different regions can take advantage of a decline in communication costs appears as an intuitive explanation for these patterns since business services concentrate overwhelmingly in large urban areas.[7]

As part of a nascent literature on domestic trade in services, Atalay et al. (2014) combine restricted-use Census data sets to conclude that, among U.S. establishments, flows of intangibles are likely orders of magnitude more important than flows of physical goods.[8] Fort (2017) exploits a restricted-use survey on U.S. manufacturers' sourcing decisions to provide evidence suggestive of large domestic service trade volumes.[9] Giroud (2013) provides direct causal evidence that reductions in communication costs (flight time decreases) increase investments from headquarters in plants located elsewhere in the US.

[7] Agglomeration spillovers as in Davis and Dingel (2018) and Duranton and Puga (2004) can explain why cities have a static comparative advantage in high-skill activities, (such as business services) but do not highlight mechanisms for faster skill premium *growth* in large cities over the last decades.

[8] The NAICS industry classification treats independent headquarters of vertically integrated companies as a Business service industry ("Management of Companies" (NAICS 55111)).

[9] An early contribution to the literature on international service trade is Griffiths (1975). Other important papers are Deardorff (2001), Hoekman (2006), Francois and Hoekman (2010), and Mattoo et al. (2007).

Jensen and Kletzer (2005) and Jensen and Kletzer (2010) infer "tradability" of service industries from measures of spatial concentration. I build on Gervais and Jensen (2013) who use a multi-sector Armington model to infer sectoral distance elasticities for 1000 sectors in the 2007 Economic Census. Contrary to this paper, they construct a direct proxy for local productivity instead of using additional assumptions on gross trade volumes. Further, they focus on a single cross-section, while changes in the distance elasticity are the focus of the present paper.[10]

Another set of papers highlights the importance of communication costs in changing the spatial organization of production (see Michaels et al. (2018) and Duranton and Puga (2005)). None of these papers estimates a structural model to quantitatively assess the impact of changes in communication costs on local skill prices.[11]

The present paper extends studies of market integration across U.S. regions to the service sector. An early theoretical contribution was Krugman (1991). More recently, Donaldson and Hornbeck (2016) document how the construction of the railroad system facilitated regional agricultural specialization. I argue that declines in business services trade frictions ushered in a similar period of service market integration and focus on its distributional consequences instead of its aggregate gains.

This project also contributes to the literature on skill-biased technological change (see Katz and Murphy (1992) and Krusell et al. (2000)) by highlighting a specific micro-channel through which recent technological change leads to high-skill wage growth and the rise of the aggregate skill premium.

Technically, I combine elements from various papers inspired by Eaton and Kortum (2002): an input-output structure of production and worker relocation across regions as in Caliendo et al. (2015) and Burstein et al. (2017) and occupation and sector choices as in Lee (2015). I also follow these papers in using "hat algebra" (see also Dekle et al. (2007) and Costinot et al. (2012)) to compute counterfactuals.[12]

The remainder of the paper is structured as follows. Section 2 introduces three properties of the business services sector that are important for the mechanism. Section 3 discusses the theory. Sections 4 and 5 explain the calibration strategy and discuss model exercises.

[10] This approach also connects to an early literature on the "regionalization" of Input-Output Tables (see Isard (1953), Moses (1955), Leontief and Strout (1963) and Polenske (1970)).

[11] Strauss-Kahn and Vives (2009) and Aarland et al. (2007) document the increasing spatial concentration of corporate headquarters, suggestive of managerial services shipments back to operating plants.

[12] Burstein and Vogel (2017) and Parro (2013) are recent papers that link *international* trade and rising inequality. Burstein et al. (2015) use techniques from the international trade literature to study aggregate between-group inequality in the United States.

Section 6 concludes.

2 Business Services: Facts

In this section, I introduce three empirical facts on the business services sector that will be central to my analysis. To construct them, I draw on the Public-Use 5% Samples of the U.S. Decennial Census Files and the U.S. input-output tables published by the Bureau of Economic Analysis (BEA), both for various years.[13]

The business services sector in the United States has grown from a mere 5% share of employment in 1950 to an 18% share in 2010. Such growth makes it one of the fastest growing sub-sector of the service economy along with consumer services. Many occupations that have seen large wage gains in recent years are particularly important in the business services sector, e.g., managers, lawyers, and data scientists. For the mechanism in this paper, three properties of the business services sector are of particular import.

Fact 1. Business Services payroll shares differ widely across local labor markets

I compute the distribution of sectoral payroll shares for 741 local labor markets in the United States (see Appendix E for details). Table 8 in the Appendix shows the ratio of the business services payroll shares at the 90th and 10th percentile of the distribution across labor markets. For business services, this ratio is 1.9 in 1980 compared to 1.4 for both the goods and local services sector. Through the lens of a simple Ricardian model, such variation in local specialization is indicative of underlying comparative advantage differences across regions.[14]

Fact 2. Business Services are more skill-intensive than the goods sector

In 1980, only 12% of all goods sector workers had a college degree, whereas 32% of business services workers did. In 2010 these number had risen to 22% and 56%, respectively (see Table 10 in Appendix A). The differential skill intensity of the two sectors implies that sectoral shocks will affect skill group wage averages differently.

[13] I discuss the construction of the underlying sample in detail in Section 4.1 below. In Appendix E.1, I present a full list of sectors subsumed under the label "business services", "goods", and "local services" for the remainder of the paper.

[14] Many papers have elaborated on the reasons behind the concentration of the business services industry in large commuting zones. The quantitative model introduced below has three mechanisms to generate this concentration: (1) technological advantages of regions for business service production, (2) differences in local skill supplies, and (3) "geography" which can make a close neighbor the preferred supplier. For the present analysis, I do not need to distinguish these factors explicitly.

Fact 3. The Goods sector is an important destination for business services

About 40% of overall business services output served as an intermediate input into the goods sector in 1980 (see Appendix A). In contrast, only 1% of goods sector output travelled to the business services sector as an intermediate input in the same year. This sectoral linkage implies that changing trade frictions in the business services sector directly affect input prices in the goods-producing sector.

The central hypothesis advanced in this paper is that the interplay of three forces can explain the systematic variation in the growth rate of the college wage premium across U.S. regions: business services comparative advantage differences across regions (Fact 1); the skill-intensity of the business services sector relative to other tradable sectors in the economy (Fact 2); changes in communication costs. The fact that business services primarily serve as intermediate inputs into goods production (Fact 3) amplifies the mechanism.

The next section introduces a model to formalize the link between these three facts and the impact of a decline in communication costs on returns to skill across regions.

3 Theory

I introduce a general environment and then show two ways of closing it that serve different purposes. The first is a simple way aimed at illustrating the nexus between communication cost declines and the evolution of regional skill premia in a parsimonious setting. The second way is richer and aimed at evaluating the mechanism's quantitative importance. This quantitative version of the model also serves as a framework to infer trade flows in the absence of data on interregional trade in services.

3.1 General Environment

I propose a static theory of trade between a set of locations. There are $r = \{1, ..., R\}$ regions, $s = \{1, ..., S\}$ sectors of production, and $k = \{1, ..., K\}$ worker types. Regions differ in their sector specific productivity, which I denote by A_{rs}. Sectors differ in their use of intermediate inputs from other sectors in the economy and their labor requirements. Worker types differ in the efficiency units of labor they can supply to the different sectors. Throughout, I denote the mass of type k workers in region r by L_{rk}. In general, workers can move across regions and the overall mass of type k workers in the economy is fixed.

Consumer Preferences Workers have Cobb-Douglas preferences over outputs from all S sectors in the economy. They allocate a fraction α_s of their overall consumption expenditure to the sector s commodity. I denote the demand for the sector s commodity of the representative household in region r by C_r^s. I stack sectoral demands within a region into a vector, $\mathbf{C}_r$, and write the utility function as follows:

$$U(\mathbf{C}_r) = \prod_s (C_r^s)^{\alpha_s} \quad \text{where} \quad \sum_s \alpha_s = 1. \tag{1}$$

When taking the model to the data, I will aggregate industries to three sectors s: goods, business services, and local services.

Production with Intermediate Inputs I denote by H_r^s the total efficiency units of labor demanded by sector s in region r. I allow for the full set of input-output linkages and denote the input demand of sector s in region r for sector s' products by $Q_r^{ss'}$. Output in sectors s in region r, Y_r^s, is produced using a Cobb-Douglas technology,

$$Y_r^s = A_{rs}(H_r^s)^{\gamma_s}\left(\prod_{s'}(Q_r^{ss'})^{\gamma_s^{s'}}\right)^{1-\gamma_s} \quad \text{where} \quad \sum_{s'} \gamma_s^{s'} = 1, \tag{2}$$

where γ_s and $\gamma_s^{s'}(1-\gamma_s)$ denote the shares of factor payments going to labor and the intermediate input from sector s', respectively.

Costly Interregional Trade Sectoral outputs are traded across regions subject to a sector-specific trade cost. I assume that trade costs take the usual "iceberg" form: To receive one unit of the sector s output from region r, consumers in region r' need to order $\kappa_{rr'}^s > 1$ units, since $1/\kappa_{rr'}^s$ units "melt" on their way.[15] Without loss of generality, I normalize the cost of shipments within each region-sector to 1, i.e., $\kappa_{rr}^s = 1 \; \forall r$.[16]

3.2 A Simple Model of Service Integration and Skill Premium Growth

In this section, I consider a special case of the general environment that takes the three empirical properties of the business service sector from Section 2 to their stylized extremes. I use this setup to illustrate how these properties interact with declining communication

[15] An implicit assumption is that trade costs are proportional to the value of the shipment.

[16] Sectoral trade costs within a region are not separately identified from region-sector specific productivity terms.

costs to produce the growing apart fact (see Figure 1).[17]

For simplicity, I consider a version of the framework in Section 3.1 with only two regions, two sectors, and two skill groups, i.e., $R = S = K = 2$. I refer to region 1 as city ($r = 1$) and region 2 as hinterland ($r = 2$), sector 1 as business services ($s = b$) and sector 2 as goods ($s = g$), and type 1 workers as high-skill ($k = h$) and type 2 workers as low-skill ($k = l$).

Setup The city is defined by a technological advantage in business service production which I index by $A \equiv A_{1b} > 1$. For simplicity, I set all other productivity terms to 1, i.e., $A_{1g} = A_{2b} = A_{2g} = 1$. To reflect the skill-intensity of the business services relative to the goods sector, I assume that business services firms only demand high-skill and goods sector firms only low-skill labor. I also simplify the input-output structure: both sectors use labor and the goods sector additionally relies on business services as an intermediate input. Given these assumptions the sectoral production technologies in equation 2 simplify as follows:

$$Y_1^b = AH_1^b, \quad Y_2^b = H_2^b, \quad Y_1^g = (H_1^g)^\gamma (Q_1^{gb})^{1-\gamma}, \text{ and } \quad Y_2^g = (H_2^g)^\gamma (Q_2^{gb})^{1-\gamma}.$$

I assume sectoral outputs are homogeneous across regions and goods trade is free, i.e., $\kappa_{rr'}^g = 1 \; \forall r, r'$. Service trade however is costly, i.e., $\kappa_{rr'}^b > 1 \forall r \neq r'$. For simplicity I denote communication costs by $\kappa \equiv \kappa_{rr'}^b \, \forall r \neq r'$. Changes in the communication cost term κ are the central comparative static of this section.

Consumers only demand goods and the utility function in equation 1 simplifies to $U = C_r^g$ accordingly. All workers of a given skill type inelastically supply one unit of efficiency labor to their sector of employment at a wage rate w_r^s. Since the mapping between sectors and skill groups is one-to-one, w_r^b is the high-skill wage, and w_r^g the low-skill wage in this economy.

For analytical clarity, I abstract from factor endowment differences by assuming the ratio of high- to low-skill workers is equal across regions, i.e., $\mu \equiv L_{rh}/L_{rl} \; \forall r$ and workers cannot change their location. If the city had a larger share of high-skill workers than the hinterland, this would add further to its productive advantage in business services production.

[17] The simple model cannot speak to aggregate changes in the college wage premium. Instead, it shows how a decline in communication costs affect regional skill premia differentially. In the quantitative model these regional changes may generate aggregate effects depending on workers' relocation response.

Interregional Trade and the Law of One Price Markets are perfectly competitive and firms price at marginal cost. Since trade in goods is free, the nationwide goods price, p^g, serves as a convenient numeraire:

$$p_1^g = (w_1^g)^\gamma (w_1^b)^{1-\gamma} A^{\gamma-1} = p_2^g = (w_2^g)^\gamma (w_2^b)^{1-\gamma} \equiv p^g = 1 \tag{3}$$

where I suppressed a composite constant.[18]

Since regional sectoral outputs are homogeneous there is no intra-industry trade. Instead, in a trade equilibrium, the city exports services and imports goods, while the hinterland does the opposite.[19] Whenever trade occurs, optimal sourcing behavior of goods firms in the hinterland ensures that the following no-arbitrage relationship holds for business services prices across regions:

$$p_1^b \kappa = p_2^b \quad \Rightarrow \quad \frac{w_1^b}{A}\kappa = w_2^b. \tag{4}$$

It is easy to show that trade takes places and equation 4 holds, as long as $\kappa < \bar{\kappa} \equiv A^\gamma$.[20] I denote by π_{rr}^s the share of spending on sector s in region r directed towards domestic firms, i.e., the "home share". Since regions either export or import within a sector depending on their comparative advantage, $\pi_{11}^b = \pi_{22}^g = 1$ regardless of the value of κ.

Equilibrium Allocations An equilibrium consists of region-sector specific wages $\{w_r^s\}$ and home shares $\{\pi_{rr}^s\}$ that solve four market clearing equations and two no-arbitrage equations.

As a result of the Cobb-Douglas assumption on technologies, a constant fraction of payments in each sector goes to workers and intermediate inputs respectively. The service market clearing equations can then be expressed in terms of sectoral payrolls, factor

[18]Goods prices are multiplied by $\bar{\gamma} = \gamma^{-\gamma}(1-\gamma)^{\gamma-1}$ in both regions.

[19]Since workers cannot move across regions nor sectors, both regions always produce positive quantities in both sectors.

[20]$\bar{\kappa}$ solves the service price no-arbitrage equation in 4 evaluated at the wages that prevail in the autarky equilibrium. The wage level between city and hinterland differs by a factor $A^{1-\gamma}$. Using this in the no-arbitrage equation at the cutoff yields:

$$\frac{w_1^b}{A}\bar{\kappa} = w_2^b \Rightarrow A^{-\gamma}\bar{\kappa} = 1 \Rightarrow \bar{\kappa} = A^\gamma.$$

This cutoff is intuitive: when service trade costs become low relative to productivity differences, service trade occurs.

shares and homes shares alone:

$$w_1^b L_{1h} = \underbrace{\frac{1-\gamma}{\gamma} w_1^g L_{1l}}_{\text{Local Demand}} + \underbrace{(1-\pi_{22}^b)\frac{1-\gamma}{\gamma} w_2^g L_{2l}}_{\text{Exports}} \quad \text{and} \quad w_2^b L_{2h} = \underbrace{\pi_{22}^b \frac{1-\gamma}{\gamma} w_2^g L_{2l}}_{\text{Local Demand}}. \tag{5}$$

In equilibrium, total service sector payroll in each region has to equal the services demand generated by the goods sector. While the city's service producers export services to the hinterland, service producers in the hinterland rely on local demand only.

The goods market clearing equations mirror the business services market clearing equations, but reflect that for goods, the hinterland is an exporter and the city an importer:

$$\frac{w_1^g L_{1l}}{\gamma} = \underbrace{\pi_{11}^g \left[w_1^g L_{1l} + w_1^b L_{1h}\right]}_{\text{Local Demand}} \quad \text{and} \quad \frac{w_2^g L_{2l}}{\gamma} = \underbrace{\left[w_2^g L_{2l} + w_2^b L_{2h}\right]}_{\text{Local Demand}} + \underbrace{(1-\pi_{11}^g)\left[w_1^g L_{1l} + w_1^b L_{1h}\right]}_{\text{Exports}} \tag{6}$$

The left hand side of these equations shows total value of spending on the goods sector and the right hand side highlights its composition.

Together with the no-arbitrage equations 3 and 4 above, equations 5 and 6 can be solved for all equilibrium allocations.

Service Trade and Regional Skill Premia Before considering the intermediate case of $\kappa \in (1, \bar{\kappa})$, two special cases can provide useful intuition. First, in autarky ($\kappa > \bar{\kappa}$) the skilled wage premium across regions is identical:

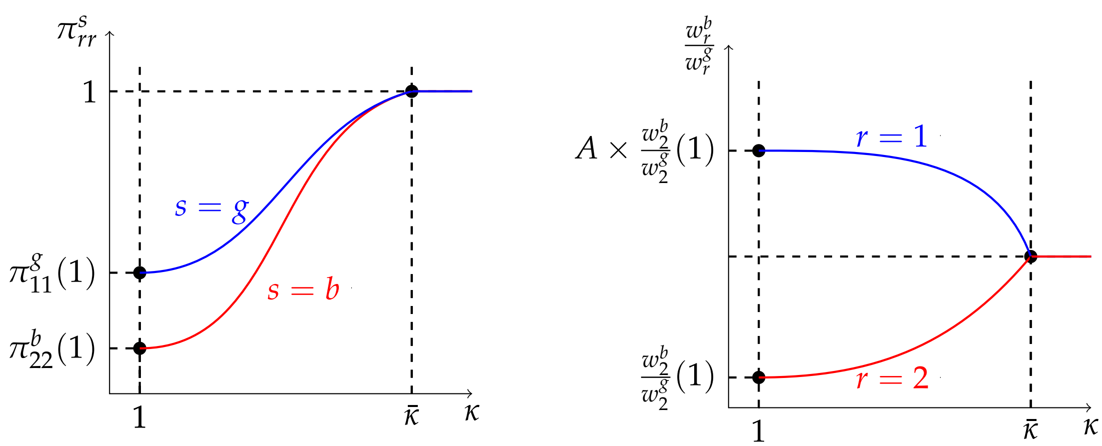
$$\frac{w_1^b}{w_1^g} = \frac{w_2^b}{w_2^g} = \frac{1-\gamma}{\gamma}\mu^{-1}. \tag{7}$$

In the free trade equilibrium ($\kappa = 1$) however, the skill premia across regions differ by the factor A:

$$\frac{w_1^b}{w_1^g} = \frac{1-\gamma}{\gamma}\frac{L_{1l}+L_{2l}}{AL_{1h}+L_{2h}}A > \frac{1-\gamma}{\gamma}\mu^{-1} > \frac{w_2^b}{w_2^g} = \frac{1-\gamma}{\gamma}\frac{L_{1l}+L_{2l}}{AL_{1h}+L_{2h}}. \tag{8}$$

Comparing equation 7 and equation 8 is insightful. In autarky, the city's skill premium is not higher than the hinterland's, despite the local technological advantage in service production. In a world without service trade, high-skill workers depend on local low-skill workers to generate demand for their services. As a result of this mutual dependence, all productive advantages within a location are shared. Relative to the autarky equilibrium,

Figure 2: Service Trade Costs, Regional Specialization, and the Skill Premium



Note: The left panel shows the home shares for the sector in which the respective region does *not* have a comparative advantage plotted as a function of communication costs. For values of κ above $\bar{\kappa}$ these home shares are 1 as both regions are autarkic. As $\kappa \to 1$ specialization occurs and the regions import in the sectors in which they do not have a comparative advantage: goods for the city, business services for the hinterland. Since each region is endowed with a mass of workers in each sector even for $\kappa = 1$, $\pi^s_{rr} > 0$, as regions always produce some output in each sector themselves. The right panel shows the skilled wage premium in both regions as a function of communication costs. For $\kappa > \bar{\kappa}$ the skill premium is constant across regions and unaffected by small movements in κ. As $\kappa \to 1$, the skill premium rises in the city (region 1) and falls in the hinterland (region 2). In the limit, the skill premium in the city is exactly A times the skill premium in the hinterland.

in the free trade equilibrium the skill premium in the city is higher and the skill premium in the hinterland is lower. The difference in the skill premia across regions is given by $A > 1$.[21] Relative *local* factor prices now reflect a location's skill-type specific comparative advantage. Low-skill workers in the city are no longer the essential source of demand for local business services and their relative wages decline accordingly.[22]

Rearranging the equilibrium system, I arrive at closed form expressions for the skill premia in city and hinterland for all $\kappa \in (1, \bar{\kappa})$:

$$\frac{w_1^b}{w_1^g} = \frac{1-\gamma}{\gamma}\left[\mu^{-1} + \frac{L_{2l}}{L_{1h}}(1-\pi_{22}^b)\kappa^{\frac{\gamma-1}{\gamma}}\right] \quad \text{and} \quad \frac{w_2^b}{w_2^g} = \frac{1-\gamma}{\gamma}\mu^{-1}\pi_{22}^b. \tag{9}$$

Equations 9 directly relate regional skill premia to communication costs, κ. The endogenous variable π_{22}^b summarizes the effect of changes in the trading environment on the skill premium in the hinterland, while κ additionally features directly into the expression for the skill premium in the city. Equations 9 show that as regions specialize in accordance with their comparative advantage (i.e., π_{22}^b falls below 1), the skill premium in the city rises, while it falls in the hinterland.

Conveniently, the model also yields an intuitive expression for the business services home share of the hinterland that shows directly how changes in communication costs enable regional specialization:

$$\pi_{22}^b = \mu\frac{\kappa^{\frac{1-\gamma}{\gamma}}L_{1l} + L_{2l}}{A\kappa^{-1}L_{1h} + L_{2h}} \tag{10}$$

Equation 10 shows that as communication costs fall, the hinterland's business service sector loses market share to service firms in the city.

There are two direct effects of a decline in service trade costs on π_{22}^b. First a competition effect captured by the numerator. High-skill workers from the hinterland now compete directly with high-skill workers in the more productive city. The more high skill workers in the city and the more significant the city's productive advantage, the more declines in κ

[21]Notice that if $A = 1$, the skill premium in the free trade equilibrium collapses back to the one in the autarky case, since there would be no incentive to trade.

[22]Equations 7 and 8 highlight another important point: as long as κ is large enough a rise in A will not lead to an increase in the skill premium within a location. The ability to spatially decouple high- and low-skill work is what allows high-skill workers to appropriate the gains from local skill-biased productivity growth. This result helps resolves a tension in some recent work on knowledge spillovers and skill premia in cities: why did knowledge spillovers for high-skill worker not raise the skill premium in cities in earlier decades? The simple model emphasizes that such spillovers only lead to a rise in the local skill premium once business service markets do not have to clear locally. In other words, even though given Cobb-Douglas technologies productivity differences across sectors are factor neutral in autarky, once trade occurs they do appear in relative skill prices within a region.

shift service demand from the hinterland to the city. Second a demand effect captured by the denominator: high local service prices shrink the goods sector in the city, decreasing local demand for services, and pushing the city's service workers to rely more on the hinterland's demand. The strength of this effect depends on the importance of the city for overall business services demand (L_{1l}) and the strength of the input linkage (γ). If business services are not used in goods production, $\gamma \to 1$ and this channel is muted.

The left panel in Figure 2 shows the behavior of the home shares as a function of service trade costs.[23] It shows that as communication costs fall regional specialization increases and city and hinterland increasingly rely on imported goods in the sectors in which they do *not* have a comparative advantage. The right panel of the same Figure shows the concurrent evolution of the regional skill premia. As specialization occurs, the skill premium rises in the city and declines in the hinterland, i.e., the skill premia are *growing apart* across regions.

In fact, if the regions are of equal size, the *ratio* of skill premia takes a particularly intuitive form:

$$\frac{w_1^b/w_1^g}{w_2^b/w_2^g} = \frac{1 + (1 - \pi_{22}^b)\kappa^{\frac{\gamma-1}{\gamma}}}{\pi_{22}^b}$$

In autarky, $\pi_{22}^b = 1$ and the ratio of skill premia is 1, since they are identical across regions. As $\kappa \to 1$, π_{22}^b falls below 1 and the ratio starts to increase until it reaches $A > 1$ in the limit. If there were no comparative advantage differences across regions (i.e., $A = 1$), the ratio would hence be equal to 1 for all values of κ. As $\gamma \to 1$ the ratio increases more slowly with κ reflecting the amplifying role of the input linkage between the business services and the goods sector.

The remainder of the paper consists in assessing the quantitative importance of this simple mechanism in explaining the variation in the growth of the college wage premium across U.S. commuting zones between 1980 and 2010.

3.3 Quantitative Theory

Does the mechanism in Section 3.2 matter quantitatively in explaining the growing apart fact in Figure 1? To answer this question, I enrich the simple model in order to speak to

[23] In the Appendix, I also show that there is an analytical expression for the relationship between the two trade shares in equilibrium:

$$\pi_{11}^g = \left[\gamma + \pi_{22}^b A \kappa^{-\frac{1}{\gamma}}(1-\gamma)\right]^{-1}.$$

So that for $\bar{\kappa} = A^\gamma$ and $\pi_{22}^b = 1$ so that $\pi_{11}^g = 1$ follows directly.

the rich U.S. spatial data. [24]. The expanded model accomplishes three objectives. First, it is flexible enough to be calibrated to moments informative about the underlying regional fundamentals that drove results in the simple model. Second, it allows for rich patterns of worker reallocation in response to a communication cost shock. Third, it serves as a measurement device for changes in sectoral trade frictions across regions, in the absence of data on trade flows. In Appendix C.5, I show how to add structures (land) and capital into the model as additional factors of production, which form part of all our counterfactual exercises.[25]

Differences to Simple Model The general environment is as introduced in Section 3.1. The are three main additions. Regions produce region-specific sectoral varieties. Firms and consumers in all regions consume CES bundles that combine regions' distinct varieties, which introduces a love-of-variety motive for trade. Workers within each skill group make non-degenerate sector-occupation choices to maximize their labor income. Additionally, workers choose their preferred region of employment.

Preferences Agents' preferences over sectoral bundles are as in equation 1 in Section 3.1. I denote by $c^s_{rr'}$ the demand of consumers in region r for the sector s variety from region r' and assume consumers aggregate regional varieties in a CES fashion,

$$C^s_r = \left[\sum_{r'} a^s_{r'} \left(c^s_{rr'}\right)^{\frac{\sigma_s - 1}{\sigma_s}}\right]^{\frac{\sigma_s}{\sigma_s - 1}} \qquad \sigma_s > 1,$$

where $a^s_{r'}$ is a origin-sector specific preference weight. The parameter σ_s controls the substitutability of the region-specific varieties in sector s. As long as $\sigma_s < \infty$ consumers demand sector s varieties from all regions, including their own. The simple model corresponds to the limit case $\sigma_s \to \infty$, in which consumer regard all regional varieties as identical and only buy the cheapest one. The composition of the CES bundle of final consumption, C^s_r, differs across regions in the presence of trade frictions.

In the simple model, there was no need to differentiate the price of a local variety and a sectoral consumption bundle, due to the homogeneity of sectoral outputs. Here, I let p_{rs}

[24]This framework falls into the class of models in the "Quantitative Spatial Economics" (QSE) literature summarized by Redding and Rossi-Hansberg (2017)

[25]I combine several contributions in the QSE literature. The model features many local labor markets and occupation choices as in Burstein et al. (2017). Sectors are linked via an arbitrary set of input-output relationships and there is trade in both intermediate inputs and final goods, as in Caliendo and Parro (2015) and Caliendo et al. (2017). The Roy model component where workers choose sectors and occupations in line with their productive advantage is inspired by Lee (2015).

be the factory-gate price of the region r sector s variety, and P_{rs} the price of the region r sector s CES bundle. Utility maximization yields a CES price index that is an average of the factory gate prices in the different regions, weighted by the bilateral trade frictions:

$$P_r^s = \left[\sum_{r'} (p_{r'}^s \kappa_{r'r}^s)^{1-\sigma_s} \right]^{\frac{1}{1-\sigma_s}}.$$

I denote the fraction of region r spending on sector s that is directed towards varieties produced in region r' by $\pi_{r'r}^s$. These "trade shares", derived from utility maximization, take the familiar form (see eg. Anderson (1979)):

$$\pi_{r'r}^s = (p_{r'}^s)^{1-\sigma_s} (\kappa_{r'r}^s)^{1-\sigma_s} (P_r^s)^{\sigma_s - 1}. \tag{11}$$

Equation 11 generalizes the expression for trade shares from the simple model to an environment where two-way trade occurs for love-of-variety reasons. Note that while individual regional varieties are tradable sectoral CES aggregates are not.

Intermediate Inputs and Sectoral Human Capital The production technology in each region-sector takes the form introduced in equation 2 above. Equation 2 had two components: a bundle of intermediate inputs and a human capital aggregate (value added).

To describe the composition of the intermediate input bundle, I denote by $q_{rr'}^{ss'}$ the demand of region r sector $s$firms for the region r' sector s' variety. Firms aggregate regional varieties within each sector with a constant elasticity of substitution σ_s into a sectoral input bundle $Q_r^{ss'}$,

$$Q_r^{ss'} = \left[\sum_{r'} b_{r'}^s (q_{rr'}^{ss'})^{\frac{\sigma_{s'}-1}{\sigma_{s'}}} \right]^{\frac{\sigma_{s'}}{\sigma_{s'}-1}} \quad \sigma_{s'} > 1,$$

where $b_{r'}^s$ is an origin-sector specific CES weight. The elasticity of substitution is the same as for consumers.[26] As a result, consumers and firms agree on how to allocate a unit of sectoral expenditure across regional varieties, so that equation 11 is also summarizes optimal sourcing decisions for firms.

The value added term in the sectoral production function, H_r^s, consists of the contributions of efficiency units of labor hired in various different occupation o. In particular H_r^s

[26]This assumption is strong, but commonly made in quantitative models with input-output linkages (see Caliendo and Parro (2015) and Caliendo et al. (2017)). Beyond its convenience, a reason for the assumption is the absence of detailed data on interregional trade in inputs versus final consumption goods. This makes it difficult to measure differences between firms' and consumers' willingness/ability to substitute among regional varieties.

is a CES aggregator over the O occupational inputs offered in the economy:

$$H_r^s = \left[\sum_o \mu_{rso} (h_r^{so})^{\frac{\iota-1}{\iota}} \right]^{\frac{\iota}{\iota-1}}$$

where ι parameterizes the ease with which the firm substitutes between human capital in the different occupational categories, h_r^{so}. μ_{rso} is a region, sector, and occupation specific weight reflecting that even within a sector production processes can differ in their occupational requirements across regions.

Individual Location, Sector, and Occupation Choices and their Aggregation Workers obtain two idiosyncratic shocks, revealed one after the other, that drive their location and employment decisions. They first learn their idiosyncratic tastes for locations and choose a location taking into account their expected utility in each potential destination. Once in a location, agents observe local wages in each sector and occupation and learn their sector-occupation specific productivity. They then choose their sector-occupation pair to maximize their income. I start by describing the second choice and then the first, to facilitate exposition.

Agents within a region differ in the number of efficiency units of labor they can supply to the different sectors and occupations. I denote the efficiency units an individual worker i could supply to sector s and occupation o if she chose to work there by ϵ_{so}^i. w_r^{so} is the wage rate per efficiency unit of labor offered in region-sector-occupation pair (r, s, o). Worker i's labor income results from the solution of her sector-occupation choice problem:

$$y_r^i = \max_{s,o} \{ w_r^{so} \times \epsilon_{so}^i \} \tag{12}$$

Note that within a given occupation efficiency units of labor supplied by different types of workers, k, are perfect substitutes.

I denote the expected income of agent i of type k in destination r before making his location choice by $E_k(y_r^i)$. The average indirect utility of a type k agent from moving to location r is then given by:

$$V_{rk}^i = \varrho \times \frac{E_k(y_r^i)}{\prod_s (P_r^s)^{\alpha_s}}$$

where ϱ is a composite constant. Before choosing their location, individuals of type k receive a multiplicative, idiosyncratic preference shock for each region. Individual i's counterfactual level of welfare in region r given her shock η_r^i is given by:

$$\bar{V}_{rk}^i = \varrho \times \frac{E_k(y_r^i)}{\prod_s (P_r^s)^{\alpha_s}} \times \eta_r^i = V_{rk}^i \times \eta_r^i$$

which reflects her type k. Individual i's location choice, r^i, then solves the following problem:

$$r^i = \text{argmax}_r \{\bar{V}^i_{rk}\}$$

To aggregate the individual choices of the mass of heterogeneous agents, I assume that individual heterogeneity within skill groups k is extreme-value distributed. In particular, for the sector-occupation choices, I assume that the distribution of individual efficiency units, ϵ^i_{so}, across individuals i within each skill group k is drawn in an i.i.d. fashion from the following Fréchet distribution:

$$F^{so}_{rk}(\epsilon) = \exp(-T_{rsok}\epsilon^{-\rho_k})$$

where T_{rsok} denotes the mean productivity of type k workers in region r, sector s, and occupation o. While T_{rsok} parameterizes between-group productivity differences, ρ_k determines within-group variations in productivity. This formulation allows for between-group heterogeneity to remain non-parametric, while imposing a parametric assumption on within-group heterogeneity for aggregation purposes.

As is well known, the particular convenience of the Fréchet assumption lies in that it yields closed-form expressions for several endogenous objects of interest. For instance, the fraction of workers of type k in region r, who choose to work in occupation o in sector s, ϕ^{so}_{rk}, is given by:

$$\phi^{so}_{rk} = \frac{T_{rsok}(w^{so}_r)^{\rho_k}}{\sum_{s'}\sum_{o'} T_{rs'o'k}(w^{s'o'}_r)^{\rho_k}}. \quad (13)$$

In a similar way, the average income of a type k workers in region $r$can be expressed in terms of the sector-occupation specific wage rate per efficiency unit, w^{so}_r:

$$w_{rk} = (\sum_s \sum_o (w^{so}_r)^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}$$

Appendix C also shows that there exist intuitive expressions for the total number of efficiency units of labor supplied to a specific region, sector, and occupation at a given wage rate. Appendix C presents all derivations for the results involving the Fréchet distribution.

Note that given the parametric assumption on individual productivities, $E_k(y^i_r) = w_{rk}$. To aggregate the location decisions of individual workers, elements of the idiosyncratic location preference shifter, $\{\eta^i_k\}$, are assumed to be drawn iid from a type specific Fréchet distribution,

$$F_k(\eta) = \exp(-G_{rk}\eta^{-\varkappa}),$$

where G_{rk} is the location- and type-specific mean. G_{rk} can be interpreted as a region r type k specific amenity term. The distributional assumption yields an analytical expression for the share of type k workers in the U.S. economy that chooses to reside in region r given factor prices. In particular:

$$L_{rk} = L_k \times \frac{G_{rk}(V_{rk})^{\varkappa}}{\sum_{r'} G_{r'k}(V_{r'k})^{\varkappa}},$$

where L_k denotes the economy-wide stock of type k workers, which I model as given.

Local Sectoral Sales and Expenditure In order to infer trade flows across regions I require measures of local sectoral demand and supply. The quantitative model provides intuitive equations for these objects in terms of parameters and observable data moments. I denote total regional revenue in sector s by R_r^s and total regional expenditure on sector s by E_r^s. The Cobb-Douglas structure of production implies that each factor receives a fixed share of sectoral revenue. As a result, gross output in a region-sector pair can be expressed in terms of the sectoral payments and the value added share alone,

$$R_r^s = \gamma_s^{-1} \sum_{k,o} w_{rk} L_{rk} \phi_{rk}^{so},$$

where the overall sector s payroll is the sum of the payrolls earned by the different skill groups.

Using this, I can write total expenditure on the sector s bundle in region r, combining final and intermediate input demand:

$$E_r^s = \alpha_s \sum_k L_{rk} w_{rk} + \sum_k R_r^k (1 - \gamma_k) \gamma_k^s. \tag{14}$$

Note that since ϕ_{rk}^{so} and w_{rk} can be expressed as simple functions of w_r^{so}, E_r^s and R_r^s are functions of w_r^{so} only, too.[27]

General Equilibrium The equilibrium is a vector of region-sector-occupation specific skill prices $\{w_r^{so}\}$ and local labor supply by skill type $\{L_{rk}\}$, such that

1. Output markets clear:

$$R_r^s = \sum_{r'} E_{r'}^s \pi_{rr'}^s \tag{15}$$

[27] In Appendix C.3, I show how to use equilibrium conditions to rewrite equation 14 as an eigensystem that is useful in solving the model numerically.

2. Occupation-specific labor markets clear:

$$\mu^{\iota}_{rso}(\frac{w^{so}_r}{w^s_r})^{1-\iota}\gamma_s R^s_r = \sum_k w_{rk} L_{rk} \phi^{so}_{rk} \tag{16}$$

3. Workers make utility maximizing location choices:

$$L_{rk} = L_k \times \frac{G_{rk}(V_{rk})^{\varkappa}}{\sum_{r'} G_{r'k}(V_{r'k})^{\varkappa}}$$

4. And trade is balanced across for all regions,

$$\sum_{r'}\sum_s E^s_{r'}\pi^s_{rr'} = \sum_s E^s_r, \tag{17}$$

   i.e., total expenditure in region r across all sectors is equivalent to the total export revenue of region r across all sectors.

Note that there can be sector-specific deficits within a region. Equation 17 merely imposes balanced trade across all sectors within a region. This is important since I rely on sectoral trade imbalances to infer sectoral trade frictions as discussed in the next Section.

In Appendix C.5, I discuss how to include structures and capital as additional factors of production. There, I assume that land markets clear within each location, introducing an additional form of congestion that raises local prices in response to increased economic activity. Capital markets in turn clear nationally. All workers hold a share proportional to their income in a national real estate and capital portfolio. All counterfactuals allow for these additional margins of adjustment, and I consider them part of the baseline model setup.

Returning to the Simple Model The quantitative framework nests the simple model in Section 3.2. To see this, we can restrict the quantitative framework to two regions $(r = 1, 2)$, two sectors $(s = g, b)$, and two skill groups $(k = h, l)$, and set $\gamma_b = 1$ and $\gamma_g \neq 1, \gamma^b_g \neq 0$. Then we can take three limits: the limit of sectoral trade elasticities, i.e., $\sigma_s \to \infty$ for all sectors, so that varieties are homogeneous across space; the limit of within group heterogeneity so that all agents within a group are equal, $\rho_k \to 0$; the limit of across group productive advantages, i.e., $T_{rhb}/T_{rhg} \to \infty$ and $T_{rlg}/T_{rlb} \to \infty$ for all r, so that high-skill workers always find it income-maximizing to work in business service and low skill workers find it profitable to work in the goods-producing sector.

4 Quantitative Framework: Calibration

In this section, I discuss the quantification of my theory. I take the model to U.S. Public Use Census data, separately for four decades: 1980, 1990, 2000, 2010. Before 1980 some necessary data inputs are available in less detail, and so I restrict the quantitative exploration to the decades from 1980 to 2010.

The first subsection discusses data sources and data construction. The second subsection explains in detail how I construct interregional trade flows and estimate the elasticity of bilateral trade volume to distance for different years and sectors. The remaining sections describe the calibration of the other parameters.

4.1 Data Sources

I use two primary data sources: the U.S. Decennial Census and the Input-Output (IO) Tables for the United States.

Decennial Census I use the 5% public use sample of the Decennial Census files for 1980-2000 and public use sample American Community survey for 2010, both obtained from IPUMS (see King et al. (2010)). The empirical analysis distinguishes individuals along four dimension: sectors, occupations, education, and local labor market. I focus on employed individuals between 16 and 65 years of age for which industry codes, occupation codes, location codes, education codes, income and hours measures are available for a given year. I measure labor supply as annual hours worked, so that L_{rk} denotes the number of hours worked by type k agents in commuting zone r in a given year. I provide further sample selection and construction details in Appendix E.3.

Input-Output Tables I draw on the use tables of the United States in producer value published annually by the Bureau of Economic Analysis (BEA).[28] The use tables contain information about the value of intermediate inputs and value added used to produce one unit of gross output of a given industry. In addition to this information, I extract vectors of final consumption, exports, and import by industry. I make adjustments to final domestic consumption to ensure that gross output by commodity and industry coincide, since my framework abstracts from different industries producing overlapping sets of commodities. Appendix E.2 offers a more detailed discussion.

[28]The BEA recomputed the older use tables using the North American Industry Classification System (NAICS). I use these updated tables.

4.2 Data Construction

The quantitative model has four dimensions of heterogeneity. In the data, I map them to three sectors, four occupations and five skill group, across 741 local markets in the United States. I now discuss these groups in turn.

Commuting Zones A consistent analysis over time requires a fixed geographical definition of what constitutes a "region". Tolbert and Sizer (1996) construct 741 commuting zones by clustering counties based on their 1990 inter-county commuting flows. I use these commuting zones since they provide a constant geography, cover the entire U.S. territory, and provide sufficient spatial detail.[29]

Sectors I group industries into goods (e.g., manufacturing and wholesale trade), business services (e.g., Management of Companies, Legal Services, and Computer systems design) and local services (e.g., Hospitals, Nursing and Accommodation, Restaurants). Appendix E.1.1 offers a complete list.[30]

I map all data sources to the North American Industry Classification System for 2012 (NAICS). More formally, the goods sector comprises all NAICS-1 to NAICS-4 industries, Business Services are all NAICS-5 sector industries, and local services are NAICS-6 to NAICS-8. I treat all three sectors as tradable and infer the degree of their tradability from the calibrated model as discussed below. Results confirm that the NAICS-6 to NAICS-8 sectors are the least tradable of the three coarse sectoral groups.[31]

[29] Autor and Dorn (2013) provide a useful crosswalk from Census county groups (1980) and Public Use Micro-data Areas (PUMAs) to the Tolbert and Sizer (1996) commuting zones. They also provide an adjustment to sample weights in cases where Census spatial units are split into several commuting zones. Autor and Dorn (2013) is the first study to use this definition of local labor markets in the economics literature. Eckert and Peters (2018) and Burstein et al. (2017) are more recent studies that use the same delineations.

[30] The input-output tables of the United States show international trade in all sectors, and since my calibration strategy relies on matching the aggregate input-output relationships exactly, all sectors are allowed to be tradable. What I refer to as non-tradable services are local services like restaurants, education, hospitals, janitors and the like whose trade volumes are likely less affected by recent technological change than the information-intensive business services

[31] It may seem counterintuitive that local services, e.g., restaurant meals, are tradable. However, the input-output tables record positive trade flows for them. For example, visitors from other countries or other commuting zones consume restaurant meals in New York (in person travel constitutes an important form of service trade). In Appendix A.1, I state the four forms of service trade defined in the General Agreement on Tariffs and Trade (GATT) by the WTO (Uruguay Round of Negotiations).

Education Types I group individuals into five groups based on their educational attainment: less than high school, high school, some college, college, five and more years of college. Recall that the quantitative model involved a parametric assumption on within skill group heterogeneity. Choosing a large number of skill groups then allows for more realistic patterns of adjustment, since the model allows for between-group heterogeneity to remain non-parametric.

Occupations There are approximately 320 occupational groups in the Decennial Census Files. I organize these occupations into groups that exhibit qualities that are likely to be important in the current setting. Many occupations in the business services sector require specialized skills and are highly tradable, i.e., they can often serve their function from a distance. As an example consider a manager in a headquarter who can use telephone and internet to instruct workers. Many occupations in the local service sector instead require personal contact between worker and customer, e.g., a bank teller (and so are "non-tradable"). Another distinction within non-tradable occupations is their skill content. The occupations of physicians and janitors differ substantially in their educational requirements.

To capture these qualitative differences, I categorize occupations into four broad categories based on their skill intensity and their tradability. The measure of skill-intensity, I use a measure of "abstract task intensity" constructed in Autor et al. (2003) from the Dictionary of Occupational Titles, published by the U.S. Department of Labor in 1977. To measure tradability, I use the offshorability measure employed in Autor and Dorn (2013).[32]

In the 1980 Decennial Census data, I compute aggregate annual labor supply by occupation. I then order all occupations ranked by their "abstract task intensity" and split them to form two groups each accounting for 50% of annual labor supply and ranked by their "abstract task intensity". I repeat this exercise with the tradability measure. Then I group occupations above the median in terms of abstractness and tradability together and call them "abstract-tradable" (AT). Similarly, I create three more groups called "non-abstract tradable" (NAT), "abstract non-tradable" (ANT) and "non-abstract non-tradable" (NANT). As an example, managerial professions are in AT, assembly line workers and phone operators in NAT, doctors, and teachers in ANT and cooks in NANT. I then

[32] Fortin et al. (2011) construct the ingredients for a measure of offshorability of occupations, which I interpret as measuring tradability, from O*NET data. Autor and Dorn (2013) use a simple average of two of the Fortin et al. (2011) measures - "face-to-face contact" and "on-site job" - to measure offshorability, and I adopt their index. Burstein et al. (2017) is another paper that uses an offshorability measure as an index for tradability. They employ the offshorability measure constructed in Blinder and Krueger (2013).

hold these groups fixed for all remaining decades. Appendix E.1.2 lists example occupations for each of these four groups and Appendix E.5 provides further detail on their construction.

Summary For each decade, I add up annual hours within each commuting zone, sector, occupation, and skill group to obtain a measure of labor supply. The model structure imposes that workers of the same skill earn, on average, the same wage within every sector and occupation within a region. Dividing total income by total hours within each region-education bin yields a measure of the average hourly wage by skill group within each location, w_{rk}.

From the IO tables, I derive vectors of final domestic consumption, imports, exports and gross output by sector for all four years of my analysis.

4.3 Inferring Service Trade Flows

Information on interregional trade flows within the United States is very sparse. Research on intranational trade in goods has drawn on a single nationally-representative source: the Commodity Flow Survey (CFS) conducted by the U.S. Census Bureau.[33] However, the CFS is limited in a crucial dimension: it does not contain information on services. This calls for a calibration strategy of service trade frictions, that does not rely on detailed, repeated cross-sections of sectoral trade flows. In this section, I propose a methodology that builds on work by Gervais and Jensen (2013) to infer service trade flows between U.S. commuting zones. The technique relies on three ingredients: detailed regional data on sectoral payrolls and hours worked, the aggregate input-output tables, and the gravity equation for interregional trade implied by the quantitative model. I present the procedure in three steps and implement it separately for each decade between 1980 and 2010.

Step 1: Constructing Local Sectoral Output and Expenditure I use the structure of the quantitative model in Section 3.3 to construct measures of local sectoral output and expenditure. I introduce the rest of the world (ROW) into the analysis as a 742nd region for reasons made explicit below. Recall the expressions for regional sales,

$$R_r^s = \gamma_s^{-1} \sum_k w_{rk} L_{rk} \phi_{rk}^s, \tag{18}$$

and expenditure,

[33] Prominent studies that use the CFS data include Hillberry and Hummels (2008), Allen and Arkolakis (2014), Duranton et al. (2014), and Dingel (2016).

$$E_r^s = \alpha_s \sum_k L_{rk} w_{rk} (1+\omega_r) + \sum_{s'} R_r^{s'} (1-\gamma_{s'}) \gamma_{s'}^s. \tag{19}$$

Here, ω_r is an exogenous subsidy to U.S. consumer paid for by ROW workers. Introducing such exogenous transfers is a simple way to rationalize the large U.S. trade deficit implicit in the input-output data.[34] I obtain the technical coefficients, γ_s and γ_s^k, and the utility function parameters, α_s, directly from the input-output tables of the respective year. These parameters, together with the data on sectoral payrolls, $\{w_{rk} L_{rk} \phi_{rk}^s\}$, imply a vector of regional sales and expenditure, $\{R_r^s, E_r^s\}$ for every region, sector, and decade. $\{E_r^s\}$ combines two sources of demand for the sector s output in region r: intermediate input demands and final consumption. Summing R_r^s across regions yields the gross sectoral output from the input-output tables in that year. Likewise, summing E_r^s across regions yields total domestic expenditure on sector s output in the US. For the rest of this section I treat $\{R_r^s, E_r^s\}$ as data.

Step 2: Constructing a Gross Sectoral Export Measure Subtracting regional demand from regional supply implies a measure of local net sectoral exports,

$$NX_r^s = R_r^s - E_r^s.$$

If $NX_r^s > 0$ region r produces more output than its firms and consumers absorb. In this case NX_{rs} provides a lower bound on the value of sector s output region r exports. It only yields a lower bound since regions are likely to import some output even in sectors in which they are net exporters. Similarly, net importing regions are likely to export some output. The phenomenon of gross trade flows in excess of net flows is referred to as inter-industry trade in the literature (see Helpman (1987) and Hummels and Levinsohn (1995) for classic discussions). The accounting identity for sectoral trade volume (V_r^s) clarifies the relationship between the inter-industry component and the net component in generating

[34] Figure 8 in Appendix A.5 shows sectoral trade balances over time. The overall trade balance is negative for every decade covered by the study. I match this by imposing a transfer from ROW to U.S. regions, from which every consumer benefits in proportion to their wage, i.e., $\omega_r = \omega > 0 \ \forall r \neq ROW$ and $\omega_{ROW} < 0$. To rationalize it through the lens of the model, I assume there is a subsidy ω that is distributed to U.S. consumers in proportion to their labor income and financed by a tax on ROW workers, denoted ω_{ROW}. In Appendix D.5, I discuss the details of inferring ω, ω_{ROW} for every decade so as to match the U.S. trade deficit in the IO tables for the United States for the respective year.

overall volume:

$$
\begin{aligned}
V_r^s &\equiv EX_r^s + IM_r^s & (20)\\
&= \mid EX_r^s - IM_r^s \mid + [EX_r^s + IM_r^s - \mid EX_r^s - IM_r^s \mid] & (21)\\
&= \mid NX_r^s \mid + \underbrace{IIT_r^s}_{\text{Interindustry Trade}} & (22)
\end{aligned}
$$

where $\{EX_r^s, IM_r^s\}_{rs}$ denotes region-sector specific exports and imports.

Equation 20 highlights that in the absence of information on total trade volume, any amount of interindustry trade is consistent with observed net balances. Unfortunately, there exists no data on gross business services trade flows that could inform IIT_r^s directly.

In the next step, I parameterize sectoral trade frictions as a function of distance. I then show that a measure of *aggregate* trade volume, combined with the structure of the model and data on regional output and demand $\{R_r^s, E_r^s\}_{rs}$ is sufficient to infer the unique matrix of interregional trade flows. I now discuss how to construct a proxy for the gross aggregate trade volume by sector.

There is an intuitive lower bound on nationwide inter-industry trade IIT^s. I observe interindustry trade between the sum of all U.S. regions and the rest of the world (ROW) in the input-output tables. The minimum amount of interindustry trade flows across the 742 regions in the model is equal to the interindustry trade between U.S. regions and ROW:

$$IIT^s_{min} = V^s_{US} - \mid NX^s_{US} \mid$$

where US abbreviates "United States". This allows me to construct a measure of total gross exports of U.S. regions by sector:

$$EX^s = IIT^s_{min}/2 + \sum_{r \neq ROW} \mathbf{1}(NX_{rs} > 0) NX_{rs}. \tag{23}$$

Equation 23 combines the interindustry trade that the U.S. regions engage in with ROW with the minimum exports necessary to cover net balances. I use this measure of sectoral trade volume in the baseline calibration. There is also an upper bound on the total value of gross exports: the total value of production, which I observe. In Appendix F, I describe robustness exercises that set total sectoral trade volume to values between their upper and lower bound.[35] I discuss these exercises in more detail at the end of this section.

[35] There are reasons to believe that net flows are closer to gross flows in the business services sector compared to the goods sector. First, U.S. business service exports are very close to net exports as evidenced by Figure 8 in the Appendix, especially in earlier decades. Second, small- or mid-sized regions specialize in

Step 3: Inferring Trade Flows In this section, I show how to use the demand and supply measures constructed in Step 1, $\{R_r^s, E_r^s\}$, and the sectoral export volume constructed in step 2, EX^s, to impute interregional trade flows.

The following lemma is useful to construct a formal identification argument.

Lemma. *Consider a mapping of the form:*

$$A_i = \sum_{j=1,...,N} B_j \frac{\lambda_i K_{ij}}{\sum_k \lambda_k K_{kj}} \quad \forall i = 1, ..., N. \tag{24}$$

For any strictly positive vectors $\{A_i\} \gg 0$ and $\{B_i\} \gg 0$, such that $\sum_i A_i = \sum_i B_i$, and any strictly positive matrix $\mathbf{K} \gg 0$ there exists a unique (to scale), strictly positive vector $\{\lambda_i\} \gg 0$.

Proof. See Appendix D.1. □

The market clearing equation for sector s implied by the quantitative model (see equation 15) yields equation 24 by defining some composite terms:

$$R_r^s = \sum_{r'} E_{r'}^s \frac{(p_r^s)^{1-\sigma_s}(\kappa_{r'r}^s)^{1-\sigma_s}}{\sum_{r''}(p_{r''}^s)^{1-\sigma_s}(\kappa_{r''r'}^s)^{1-\sigma_s}} \equiv \sum_{r'} E_{r'}^s \frac{\lambda_{rs} K_{rr'}^s}{\sum_{r''} \lambda_{r''s} K_{r''r'}^s} \tag{25}$$

In order to use the Lemma in practice, I require $\sum_r R_r^s = \sum_r E_r^s$. As a result of the U.S. trade deficit with ROW, the equality does not hold across U.S. commuting zones alone. It does, however, hold for the U.S. regions and the ROW region together.

The Lemma implies the existence of a large set of different trade flow matrices that all generate the same regional *net* export balance, i.e., match $\{R_r^s, E_r^s\}$. There is exactly one such set for every trade cost matrix $\mathbf{K}^s$. To pin down $\mathbf{K}^s$, I parameterize its entries as a function of bilateral distance, $d_{rr'}$,

$$K_{rr'}^s \equiv d_{rr'}^{(1-\sigma_s)\bar{\delta}^s} \equiv d_{rr'}^{\delta^s}.$$

I refer to δ^s as the sector specific distance elasticity of trade.[36] Identifying the product $\delta^s \equiv (1-\sigma_s)\bar{\delta}^s$ is sufficient to infer interregional trade flows. This reduces the quantification of the trade cost matrix to the calibration of a single composite parameter, δ^s.

the production of particular goods that exhibit strong returns to scale, say tractor, and hence need to import many other goods that form part of the consumption bundle. The business services sector is more narrow. For example, it seems unlikely that JP Morgan in Manhattan would import legal services from a small town when it has world-class legal services all around it in New York.

[36]Other papers that parameterize interregional trade flows as a function of distance are Allen and Donaldson (2018),Ahlfeldt et al. (2015), and Monte et al. (2015). I use the centroid-to-centroid distance in miles.

The moment I use to identify δ^s is the measure of gross exports constructed in Step 2, EX^s. I subtract the shipments from region r to itself from the market clearing equation 25 above to obtain an expression for region r sector s gross exports, EX_r^s:

$$EX_r^s = \sum_{r' \neq r} E_{r'}^s \frac{\lambda_{rs} K_{rr'}^s}{\sum_{r''} \lambda_{r''s} K_{r''r'}^s} \tag{26}$$

For a given δ^s, equation 25 can be solved for a unique vector of location fixed effects, $\{\lambda_{rs}\}$, that rationalize E_r^s and R_r^s. These fixed effects then imply a vector of region-sector specific exports via equation 26. Summing these exports across all regions in the United States yields a model implied measure of EX^s.[37] I show in Appendix that equation 26 is strictly increasing in δ^s for a given vector $\{\lambda_{rs}\}$, so that there exists a value of δ^s that minimizes the following criterion function:

$$\Omega(\delta^s) = \mid \log \frac{EX_{DATA}^s}{\sum_r EX_{rMODEL}^s} \mid .$$

For each sector and year I find $\delta^s = \text{argmin}_{\delta^s} \Omega(\delta^s)$. I describe the algorithm in more detail in Appendix D.3. There I also explain how to use observed international imports and exports from the IO tables to calibrate $\lambda_{ROW,s}$ so as to match imports and exports (i.e., the only bilateral set of flows I observe in the data) exactly. I also show that this implies values for E_{ROW}^s and R_{ROW}^s.

Table 1 presents the results. It suggests that trade frictions for goods have been fairly stable throughout the period $1980-2010$. However, business services trade frictions have declined significantly.[38]

The fact that trade frictions for goods appear constant, justifies the use of a distance elasticity, δ^g, obtained from running a gravity equation in the CFS data, which is only available since 1992. In particular, since goods trade frictions are not the focus of this paper, I

[37] Gervais and Jensen (2013) use the 2007 cross-section of the Economic Census to infer industry-specific distance elasticities. They proxy directly for the vector of fixed effects $\{\lambda_{rs}\}$ using sales per worker over payroll per worker in each region-sector. Such a procedure has the disadvantage of not matching E_r^s and R_r^s exactly and requires data that distinguishes sales and payroll by sector-region, which is not widely available.

[38] The former finding is in line with the finding in Allen and Arkolakis (2018), who show that the distance coefficient in a gravity equation for goods trade among U.S. states is virtually unchanged between 2007 and 2012 (using Commodity Flow Survey (CFS) Data). They find $\delta^g = -1$. Monte et al. (2015) estimate the same elasticity using CFS Data for 123 CFS regions within the United States and find $\delta^g = -1.29$.

Eaton and Kortum (2018), using 2010 international bilateral trade data, find a distance elasticity of -1.4 on Professional Services and on Administrative Services using OLS log-log regression. Their Pseudo Maximum Likelihood estimates yields slightly lower estimates.

Table 1: Estimates Transformed Distance Elasticities, δ^s

Year	Goods Sector	Business Services
1980	-1.6	-2.1
1990	-1.6	-1.8
2000	-1.6	-1.6
2010	-1.6	-1.5
$\Delta_{80\to10}\%$	0%	-28%

Notes: The table shows estimates of sectoral elasticities of trade costs with respect to distance, δ^s: $K^s_{rr'} = d^{\delta^s}_{rr'}$. Here $d_{rr'}$ is the distance in miles between the centroids of commuting zones r and r' based on the demarkation in Tolbert and Sizer (1996). The data underlying the estimates is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Additionally, data from the input-output use tables in producer prices published by the Bureau of Economic Analysis is used.

use the estimates from Monte et al. (2015) who find $\delta^g = -1.29$ using the 2012 CFS data.[39]

I assume throughout the paper that σ_s is constant so that changes in δ^s reflect changes in $\bar{\delta}^s$. Suppose for example $\sigma_b = 4$ then the change in business service trade frictions between 1980 and 2010 is given by:

$$\frac{K^b_{rr',2010}}{K^b_{rr',1980}} = (d_{rr'})^{\frac{-1.51+2.1}{-3}} = (d_{rr'})^{-0.2}$$

For a distance of a 1000 miles this corresponds to a 75% decrease. For comparison, in a similar framework Lee (2015) estimates that the accession of China into the WTO lowered bilateral trade frictions between the U.S. and China by 26%.

The procedure outlined above also relates to an older literature on the "regionalization" of input-output tables with noteworthy contributions by Isard (1953), Moses (1955), Leontief and Strout (1963) and Polenske (1970).

Discussion and Summary The separate identification of the local productivity terms and the trade elasticity relies on the following argument. For any origin-region unobserved productivity (λ_{rs}) is independent of the destination of a shipment. However, trade costs depend on the destination in a way common to the whole country, given the same distance, independently of where the origin-region is. Information on total exports and relative productivities across regions (as measured by NX_{rs}) then imply a unique set of trade flows. The Lemma in Step 3 above formalizes this intuition.

[39] Note that I normalized trade friction *within* each region to 1 in each decade. Changes in within-region trade frictions are not separately identified from regional productivity changes. As a result, what I really measure are changes in the cost of shipping services out of a region *relative* to shipping them within.

In the Appendix, I provide details on a range of robustness exercises. First I consider an alternative specification of the trade cost matrix. One feature of phone calls or emails is that the physical distance between sender and receiver is almost always irrelevant for the cost of communication. However, there is a fixed cost of moving from in-person communication to communication mediated by an electronic device. I hence propose an alternative parameterization of trade costs, where shipping business services to any other commuting zone incurs the same fixed cost. I then estimate this fixed cost over time and find it, too, is declining. Second, I use an alternative calibration of the *change* of business service trade frictions. I calibrate the 1980 trade flows for business services as described above. However, I infer the 2010 distance elasticity by postulating that the total trade volume is either 50 percent of all business service output in 2010. This trade volume implies a decline in the distance elasticity relative to 1980. In Appendix F, I discuss more details on these robustness exercises and replicate all main results for these alternative specifications.

The output of this section is a set of calibrated distance elasticities $\{\delta^s\}$, a set of bilateral trade share matrices $\{\pi^s_{rr'}\}$, and vectors of regional output and demand $\{E^s_r, R^s_r\}$ for every decade between 1980 and 2010.

4.4 Other Parameters

This section discusses the calibration of the remaining parameters.

Factor Shares and ρ_k I obtain the Cobb-Douglas coefficients in the production function and the utility function directly from the input-output table of the respective year as listed in Table 2. I conduct all counterfactuals relative to the 1980 cross-section, holding factor shares and utility function parameters fixed at their 1980 values.

Calibrating the factors shares for the rest of the world is more involved since I do not observe ROW input-output coefficients in the U.S. input-output table. I infer these coefficients to be consistent with ROW sectoral sales, expenditures, and the U.S. trade deficit. I provide details in Appendix D.5.

Table 2: Technical Coefficients over Time

Sector (s)	γ_s^g	γ_s^b	γ_s^{ls}	γ	α
Goods (g)	.82	.12	.06	.42	.52
Business Services (b)	.22	.57	.21	.68	.06
Local Services (ls)	.64	.15	.21	.71	.42

Notes: The table shows factor shares for three aggregate sectors obtained from the input-output Use tables in producer prices published by the Bureau of Economic Analysis. The shares computed as the fractions of sectoral payments directed towards the various purposes.

I show in Appendix D.4 that individual log income within each skill group, location and sector is Gumbel distributed. A convenient implication is that the variance of log income within a region-sector-occupation-type bin is only a function of ρ_k. Drawing on this insight I calibrated ρ_k to match the average variance of log income within these bins in the data. The results from this procedure are listed in Table 3. The estimates imply that more educated workers are more similar in their human capital holdings than the least educated group.[40]

Table 3: Estimates of ρ_k

Skill Type (k)	1	2	3	4	5
ρ_k	1.14	1.46	1.41	1.47	1.47

Notes: The table presents estimates of the elasticity of labor supply to wages per efficiency unit in a given region, sector, and occupation for the $K = 5$ different education types considered in the baseline calibration of the model. The parameters are estimated in the micro-data of the 5% sample of the U.S. Decennial Census (1980-2000) and the American Community Survey (2010).

Regional Fundamentals I refer to parameters indexed by region r as "regional fundamentals". The regional fundamentals in the quantitative framework are: $\{T_{rsok}, A_{rs}, G_{rk}, a_r^s, b_r^s\}$.

[40] ρ_k is an important parameter since it regulates the type specific response to changes in efficiency wages. Given the formula for labor supply of type k to sector s it is easy to see the role of ρ_k

$$\phi_{rk}^{so} = \frac{T_{rsok}(w_r^s)^{\rho_k}}{\sum_{s'} T_{rsok}(w_r^{s'})^{\rho_k}} \Rightarrow \frac{d\log\phi_{rk}^{so}}{d\log w_r^{so}} = \rho_k(1-\phi_{rk}^{so}) > 0$$

ϕ_{rk}^{so} is a summary statistic for the mean productivity of individuals of type k in sector s. Intuitively, if ϕ_{rk}^{so} is large fraction of k types will already have sorted into s reducing the scope for more to follow if w_r^{so} increases. A higher ρ_k implies that the labor supply response is larger, since individuals are more similar in terms of productivity meaning a larger mass of agents are indifferent between sector-occupation pairs and will move given small movements in the skill price.

Conveniently, I do not need to calibrate these regional fundamentals explicitly to conduct counterfactual exercises. In Appendix C.4, I show how to rewrite the equilibrium system in changes as in Dekle et al. (2007) or Costinot and Rodríguez-Clare (2014). The "in-changes" technique allows me to replace all expressions involving regional fundamentals with regional data moments informative about them. I provide more details when describing the model exercises and in Appendix C.4.

Parameters from the Literature Goos et al. (2014) estimate the elasticity of substitution between different occupations to be $\iota = 0.9$. The estimate implies that occupations are complements in the production. I use their estimate in my baseline exercise. In Appendix F, I offer a robustness check and instead follow Burstein et al. (2017) in setting $\iota = 1.93$, which makes occupations substitutes. The alternative value for ι does not affect results substantially.

Gervais and Jensen (2013) estimate the elasticity of substitution between regional varieties from an Armington model with trade across metropolitan areas. This implies $\sigma_g = 5.5, \sigma_b = 5$ and $\sigma_{ls} = 6$. The number for traded goods is also in line with estimates from Caliendo and Parro (2015) using international trade between the United States and other countries. Appendix F considers a set of robustness exercises involving different value of σ_s, results do not change appreciably.

$\varkappa$ is the elasticity of local labor supply to local real wages. I assume this elasticity to be identical across skill groups and set it to $\varkappa = 1.5$, which is roughly in the middle of the range of values used in the literature on geographic mobility as reviewed in Fajgelbaum et al. (2018).

4.5 Summary of Calibration

Table 4 provides an overview of the baseline calibration of the model.

Table 4: Overview of Parameterization of Model

	Value	Description	Strategy	Source
δ^s	(-1.5)-(-2.1)	Distance Elasticity for Service Sectors	Estimated	IO Tables, Local Data, Armington Structure
ρ_k	1.14-1.47	Labor Supply Elasticity	Estimated	Within Group Variance of Earnings
α_s	0.52, 0.6, 0.42	Cobb-Douglas Coefficients in Utility Function	Calibrated	IO Table
γ_s, γ_s^k	...	Factor Shares in Production	Calibrated	IO Tables
ι	0.9	Elasitcity of Substitution between Occupations	Literature	Goos et al. (2014)
σ_s	5.5,5,6	Elasitcity of Substitution between Regional Varieties	Literature	Gervais and Jensen (2013)
$\varkappa$	1.5	Spatial Labor Supply Elasticity	Literature	Fajgelbaum et al (2018)
δ^g	-1.23	Distance Elasticity of Goods Trade Costs	Literature	Monte et al. (2018)

Notes: The table summarizes the parameterization of the model used for all baseline counterfactual exercises shown in the body of the paper.

5 Quantitative Framework: Counterfactual Exercises

In this section, I use the theoretical framework introduced above to isolate the effect of a decrease in communication costs on the U.S. economy in 1980. In particular, I calibrate the model to the 1980 data and change a *single* parameter: I move the distance elasticity of business services trade frictions from its 1980 to its (lower) 2010 value, i.e., I set $\delta^b = \delta^b_{2010}$.

5.1 The Quantitative Model in Changes

I solve for the counterfactual equilibrium in changes (see Costinot and Rodríguez-Clare (2014) and Dekle et al. (2007) for a discussion). For a given variable or parameter x, $\hat{x}$ denotes x'/x where x' is the value of the variable or parameter in the $\delta^b = \hat{\delta}^b_{2010}$ equilibrium, while x is its value in the $\delta^b = \hat{\delta}^b_{1980}$ equilibrium. Rewriting the model in changes is useful since it implies that instead of calibrating regional fundamentals ($\{T_{rsok}, A_{rs}, G_{rk}, a_r^s, b_r^s\}$), I can instead use data moments informative about these objects in 1980 to compute the counterfactual equilibrium.

As an example consider equation 11 from above expressed both in levels and changes:

$$\pi^s_{r'r} = \frac{(p^s_r)^{1-\sigma_s}(\kappa^s_{r'r})^{1-\sigma_s}}{\sum_{r''}(p^s_{r''})^{1-\sigma_s}(\kappa^s_{r''r'})^{1-\sigma_s}} \Rightarrow \hat{\pi}^s_{r'r} = \frac{\hat{p}^{1-\sigma_s}_{rs}(\hat{\kappa}^s_{r'r})^{1-\sigma_s}}{\sum_{r''}\hat{p}^{1-\sigma_s}_{r''s}(\hat{\kappa}^s_{r''r'})^{1-\sigma_s}\pi^s_{r''r'}} \tag{27}$$

All endogenous variables in Equation 27 are now expressed in changes, $\hat{\pi}^s_{r'r}$ and $(\hat{p}^s_r)^{1-\sigma_s}$, and I solve for them instead of their counterparts in levels. For $s = b$ I insert

$$\hat{\kappa}^b_{r'r} = d^{\delta^b_{2010}-\delta^b_{1980}}_{r'r},$$

while for all other sectors, I keep trade costs at their 1980 level, i.e., $\hat{\kappa}^s_{r'r} = 1$. Finally, the $\pi^s_{r''r'}$ in the denominator is a data object. It is the fraction of sector s expenditure in region r' spent on the region r'' in 1980, a result of above imputation procedure. Equation 27 is a good example of the different objects that appear in the equilibrium system written in changes: region specific parameters that change (e.g., $\hat{\kappa}^b_{r'r} \neq 1$), others that do not (e.g., $\hat{\kappa}^g_{r'r} = 1$), endogenous variables in changes (e.g., $(\hat{p}^s_r)^{1-\sigma_s}$), and data objects in 1980 that include information about regional fundamentals (e.g., $\pi^s_{r''r'}$). I show the full model rewritten in changes in Appendix C.4.

5.2 The Distributional Impact of Communication Cost Changes

In this section, I explore whether the Growing Apart channel highlighted by the simple model is quantitatively important in explaining Figure 1. I first discuss the degree to which the mechanism can explain the slope of the line and then discuss the nationwide effect on the skill premium in the next subsection.

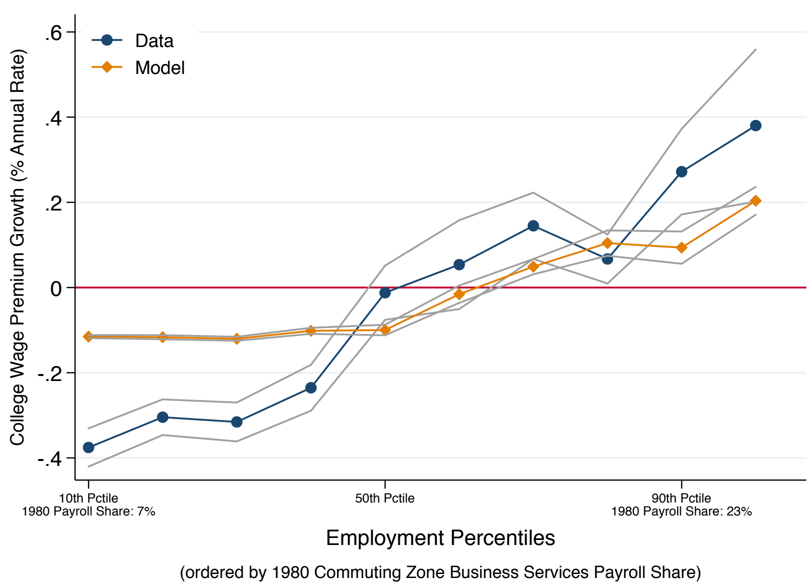
Figure 3 replicates Figure 1 from the introduction, with the average nationwide skill premium growth subtracted out. The blue line depicts the data, while the orange line shows the model generated college wage premium growth. I compute the college wage premium identically in both model and data. Both lines are based on the same cross-section of 1980 wages and employment counts, but the orange line then draws on the wage rate and employment counts predicted by the model for 2010, while the blue lines relies on the 2010 cross-section in the data. Each dot or diamond denotes the average college wage premium growth within a decile of national employment.

The measured the decline in communication costs, indeed induces systematic regional growth in the college wage premium that is in line with that observed in the data. The growing apart mechanism is active and can explain a significant part of the systematic relationship between initial specialization in business services and the subsequent growth in the local college premium.

The reason the model predicts similar growth for the first four deciles of employment in commuting zones with the lowest business service payroll shares in 1980 is that these commuting zones do not differ much in the size of their local business services sector in 1980. As a result the model infers similar comparative advantages for these commuting zones vis-a-vis the rest of the economy and predicts a similar impact of declining communication costs on local labor demand.

The grey lines denote 95% confidence intervals around the means taken within deciles of national employment across commuting zones ordered by their 1980 business services payroll share. The fact that this interval is much tighter for means to the left relative to the right of the graph reflects that commuting zones with small business services payrolls in 1980 also are less populous on average, so that there are more of them in a decile of national employment. The intervals are tighter in the model than in the data, reflecting that the models' mechanism interacts directly with the local business services share in 1980, while there may be additional forces at work in the data.

Figure 3: The Growing Apart Effect, 1980-2010

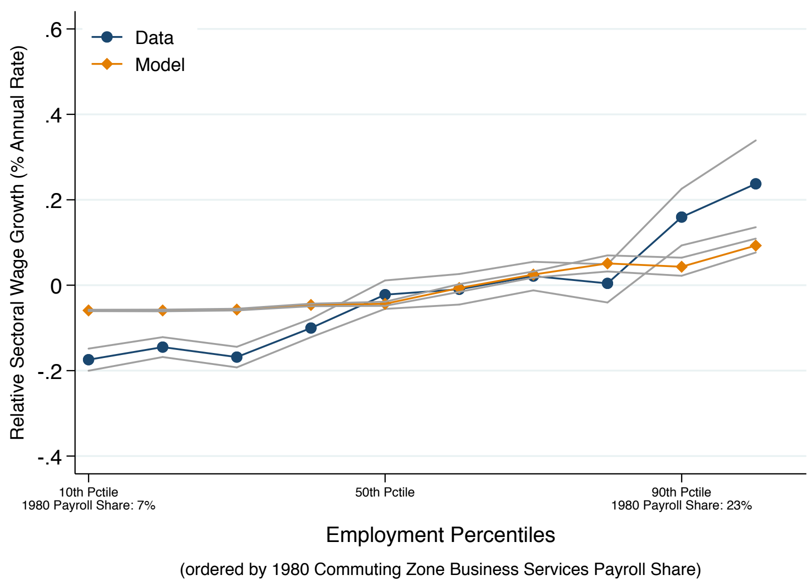


Note: The blue line replicates the Figure from the introduction (see Figure Figure 1) except with the nationwide average growth of the college wage premium subtracted out. Grey lines show 95% confidence intervals. The orange line is compute analogously using the 1980 data and the wages and employment counts implied by the model for 2010.

In Appendix F, I show that this finding is robust to using different elasticities of substitution between regional varieties (σ_s) and between occupations (ι). I also, show that a specification for communication costs that does not depend on distance but instead a simple fixed costs of shipping services beyond the home region that is calibrated to the same data on regional net balances, generates a very similar graph. Lastly, I show that a larger decline in business service trade costs would bring the model generated data yet more in line with the actual data.[41] In Appendix A.6, I additionally decompose the response of the economy to the change in communication costs into different margins of adjustment. I switch off spatial reallocation altogether and also show results where spatial reallocation occurs but prices for structures and capital do not adjust.

To understand the driving forces behind this result I draw on additional predictions about sector- and occupation-specific wage growth.

Figure 4: Business Services versus Goods Sector Wage Growth, 1980-2010



Note: The blue line replicates the Figure from the introduction (see Figure Figure 1) with two differences. First, I aggregate wages to the sector level and compute relative wage growth within the business services relative to the goods sector. Second, subtract out the nationwide average growth of the wage ratio. Grey lines show 95% confidence intervals. The orange line is compute analogously using 1980 data and model implied wages and employment counts for 2010.

Sectors Figure 4 shows Figure 3 with the y-axis depicting the wage growth differential between the business services sector and the goods sector instead of growth in the college wage premium. In line with the anticipated regional specialization, business services sector wages rise faster than goods sector wages in regions that initially specialized in

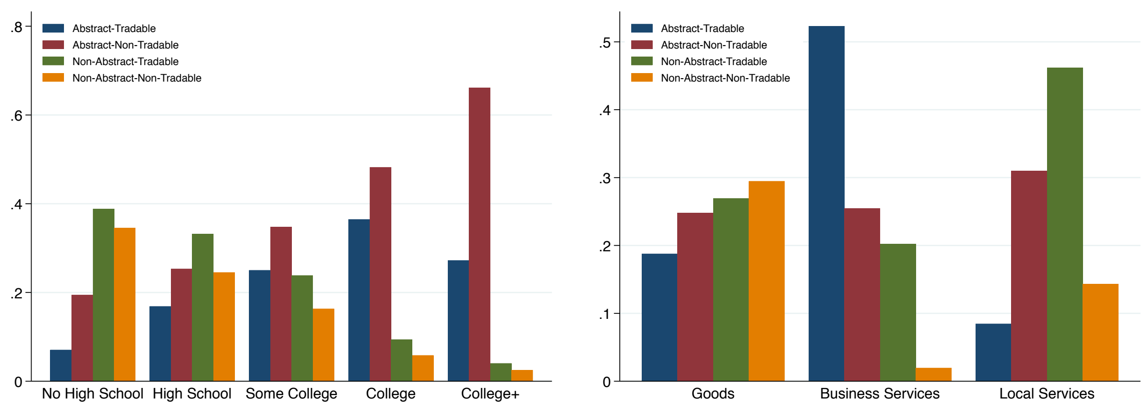
[41] In the Appendix, I assume that in 2010 50% of total business services sales in the United States are traded across commuting zones. This assumption implies a value of δ^b_{2010} that is substantially lower than the one estimated above.

business services production. The growth differential is smaller across sectors than it is across education groups, reflecting that in model and data workers of all education types are found in all sectors. However, as Table 10 in the Appendix shows the college share in the business services sector is around 2.5 times that of the goods sector. Sectoral wage changes as shown in Figure 4 hence translate into relatively higher wage gains for college educated workers in business service specialized regions, and relatively higher gains for less educated workers in goods producing regions.

Occupations A more subtle implication of the decrease in communication costs is its impact on occupational returns. Figure 5a shows how workers of different skill sort across the four occupation groups in 1980. Instead, Figure 5b depicts occupational employment shares across the three sectors present in the calibrated model.

Figure 5: Occupational Employment Across Skill Groups and Sectors

(a) Occ. Empl. Shares by Skill Group, 1980 (b) Occ. Empl. Shares by Sector, 1980



Note: All Data from 5% Publics Use Samples of U.S. Decennial Census Files for 1980 (obtained via IPUMS, see King et al. (2010)). I compute total hourly labor supply within each education group and within each sector, I then compute the fraction of this labor that is supplied to one of four occupation categories: AT: Abstract-Tradable, ANT: Abstract-Non-Tradable, NAT: Non-Abstract-Tradable, NANT: Non-Abstract-Non-Tradable Occupations.

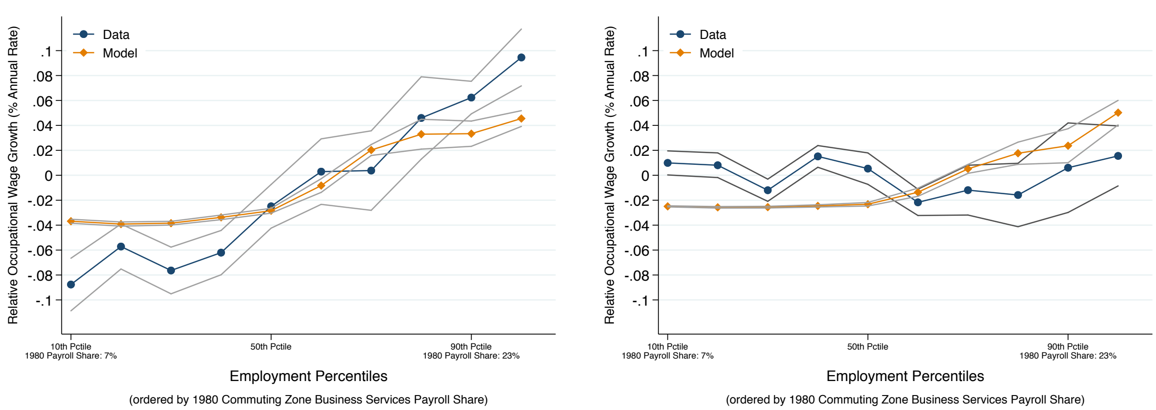
Changes in communication costs enable regional *sectoral* specialization which changes sectoral wages across regions as shown in Figure 4. Occupations link changes in sectoral returns to changes in the return to skill via the intensity of the different occupational inputs used by sectors and the sorting of workers into occupations in line with their abilities. Recall that in Section 4.2 above, I grouped the 320 occupations in the U.S. census files into four occupational groups based on their attachment to a particular location ("tradability") and their cognitive requirements ("abstractness").

The left panel of Figure 6 shows the relative wage growth of the two occupational groups that are above the median in terms of their abstractness, but one is above the median, the other below in terms of tradability. Managers, architects, and lawyers are examples of occupations in the first group, while dentists, psychologist, and secondary school teachers are professions in the abstract-non-tradable bin. The model replicates the data well. The decline in communication costs entails substantial wage growth for tradable-abstract relative to non-tradable abstract occupations. Figure 5, explains the model's success in predicting occupation group specific wage growth profiles across commuting zones. The business services sector relies heavily on abstract-tradable workers. The reason these workers benefit from a decline in communication cost is precisely that their "output", e.g. strategic direction, is not tied to a particular location the way a teacher is tied to the class present in his classroom. Instead, strategic advice by managers in Denver can decisively affect production processes in locations throughout the United States. A decline in communication costs then has an effect on these occupations that is reminiscent of the "Superstar Effect" discussed by Rosen (1981): it allows workers specialized in these occupations to extend the spatial reach of their output vastly.

Figure 6: Communication Costs and Relative Occupational Wage Growth

(a) AT vs ANT Occupations

(b) NAT vs NANT Occupations



Note: AT: Abstract-Tradable, ANT: Abstract-Non-Tradable, NAT: Non-Abstract-Tradable, NANT: Non-Abstract-Non-Tradable Occupations. These two Figures show relative annualized hourly wage premium growth for two different occupation at a time across commuting zones between 1980 and 2010 in the data (blue line) and the model (orange). The data is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Wages are computes as unconditional average hourly labor income for workers with at least some college education and workers with only high school education or less. To compute the lines in the Figure, I compute the average growth rate of the wage ratio (occupation 1 to occupation 2) within deciles of employment across commuting zones ordered by their business services payroll share in 1980. The Figure shows 95% Confidence Bands on these within-decile averages.

The right panel of Figure 6 shows the wage growth difference between non-abstract tradable and non-abstract-non-tradable professions. The growth difference for these two occupations is much less pronounced than those for the abstract occupations. As Figure 5b shows the sorting of these two occupational groups across sectors is much less striking than it was for the abstract occupational groups. The reason is that the task intensity measured use to construct these groupings in Section 4.2 do not do well in distinguishing tradable and non-tradable non-abstract occupations well. Barkeepers, for example, are classified as tradable when in reality their services are very much tied to a particular gastronomical venue limiting the spatial reach of their activity. With a more precise occupational grouping, I would have expected non-abstract tradable occupations to see faster labor demand growth vis-a-vis non-abstract, non-tradable occupations in regions not specialized in business services production.

5.3 Communication Costs and the Aggregate College Wage Premium

In this section, I consider the aggregate implications of the decline in communication costs. The first column of table 5 groups workers with some college, college and more than college education together and compares their relative wage growth to that of the two remaining education groups. In the data, the unconditional college wage premium increased by about 27% between 1980 and 2010. The full model can explain more than 30 percent of this increase. The inclusion of structures and capital is important in generating this result. As Table 5 reveals, without structures (and capital) the model explains about 17 percent of the increase in the data.

Table 5: Changes in the Aggregate College Wage and Welfare Premium

	Δ% Wage Premium
Data	27.6%
Model without Spatial Reallocation	5.7%
Model without Structures	4.3%
Full Model	10.0%

Note: The model implies commuting zone level changes in the average wage for each type of worker. For each commuting zone, I use the hourly labor supply from the original (1980) equilibrium by skill group to compute the average wage growth across all commuting zones of workers with at least some college and high-school or less. The table present the log ratio of these growth rates.

To understand this recall that Figure 3 above showed that the college wage premium rises most markedly in the top two deciles of employment. The large local expansion not only increases the demand of the business services sector but also generates substantial demand spillovers so that demand for structures of the local goods and local services sector also increases. Since structures are in limited supply, this raises local prices considerably. Furthermore, since business services are skill-intensive, the ratio of college to high-school educated workers rises substantially in these commuting zones. In spatial equilibrium, the nominal wage increase necessary to attract high-skill workers into these commuting zones to work in the business services sector, increase the aggregate nominal college wage premium in the aggregate substantially. Restricted housing supply is hence a potent amplifying force for the effect of communication cost declines on the measured aggregate college wage premium.

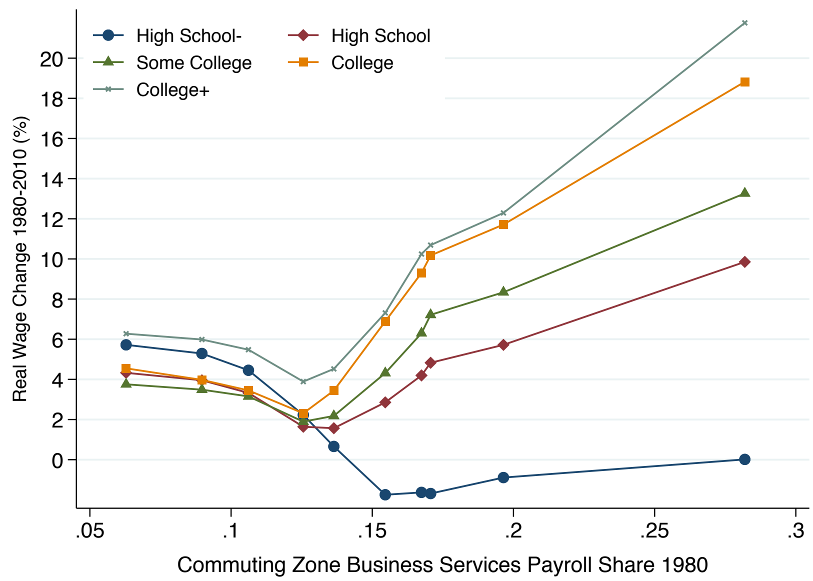
5.4 Growing Apart in Real Wages

In this section, I turn to the predictions of the model for real wage growth of workers of different education levels across commuting zones. In the data, real wages are hard to measure. Accordingly, I view the following results a more speculative. Figure 7 shows average welfare growth for workers of all education groups in the model across commuting zones with different initial business services payroll shares. The Figure is constructed by computing the average real wage growth in each commuting zone of workers of different education types between 1980 and 2010. I then weight these growth rates using the 1980 education group specific employment counts for each commuting zones to construct averages within deciles of employment, ordered by the 1980 business services payroll share of the respective commuting zone. Each dot in the Figure signifies the average real wage growth within a decile. The unequal spacing of the points reflects that deciles differ to varying degrees in the average business services payroll share of the commuting zones that go into their construction.

As can be seen, the most educated workers (college and college plus) experience the fastest real wage growth in regions specialized in business services. These regions are also, on average, populous regions. Likewise, the communication cost shock introduces real wage growth for low skill workers in regions that specialized in goods production in 1980. Since utility maximizing agents choose their location to maximize their real wage, Figure 7 suggests that communication cost declines "pull" high- and low-skill workers in different "directions" in space: while high-skill workers see more substantial welfare gains in, on average, large commuting zones, low-skill workers are increasingly better off in small places.[42]

[42]Note that through the lens of a model with a homothetic utility function, Figure 1 in the introduction already is informative about differential welfare growth for high- and low skill workers within a commuting zone, since the price index in the denominator cancels out when taking the wage ratio.

Figure 7: Real Wage Changes Across U.S. Commuting Zones, 1980-2010



Note: The model implies commuting zone level changes in the real wage (welfare) for each type of worker. For each commuting zone, I use the hourly labor supply from the original (1980) equilibrium by skill group to compute the average real wage growth across within each commuting zones of workers by education group. I then order all commuting zones by their 1980 business services payroll share and compute average welfare increases by education group for each decile of national employment. Each symbol in the Figure represents such an average. Real Wage Growth rates are computed for the time period between 1980 to 2010.

Interestingly, workers with post-graduate degrees (see Figure 5above) see substantially faster wage growth than college-educated workers in locations not very specialized in business services in 1980. The reason is that the local services sector includes high-skill services, which employs many workers with post-graduate degrees, in particular doctors. This compositional reason explains why the wages of these workers grow faster than college worker wages in regions with a low business services payroll shares in 1980, which, on average, are also small regions: in these commuting zones increased goods-production generates demand spillovers into the local sectors.

An interesting feature of Figure 7 is that it shows the plight of mid-sized labor markets without a clear competitive advantage: in these regions, all workers experience hardly any real wage growth. These regions are not specialized enough in business services to compete with the large, very specialized local labor markets. At the same time, the local goods sectors benefit less from cheaper business services imports since they have a non-negligible local presence of business service providers. These changes in real wages suggestion incentives for spatial sorting, whereby more high-skill types sort into larger metropolitan areas and less skilled types into smaller regions.

Table 6: Changes in the Aggregate College Welfare

	Δ% Welfare Premium
No Spatial Reallocation	43.1%
No Land	46.2%
Full	82.7%

Note: The model implies commuting zone level changes in the real wage (welfare) for each type of worker. I use the hourly labor supply from the original (1980) equilibrium by skill group to compute the average real wage growth across all commuting zones of workers with at least some college and high-school or less. The table present the log ratio of these growth rates.

Overall, high skill workers experience faster real wage growth than low skill workers across all regions making them the chief benefactors of the communication cost decline. Not surprisingly Table 6 then reveals that the college welfare premium has increased substantially faster than the college wage premium alone would suggest. Part of the reason why real wages for high- relative to low-skill workers increase faster than their nominal equivalents is that the share of business services in the final consumption bundle is almost insignificant (see Table 2). As Figure 10 in the Appendix shows, the increase in local prices of business services in business service intensive regions does not lead to a substantial increase in the local consumer price index (CPI). At the same time, the local goods sector becomes less competitive, hence exports less, putting further downward pressure on local goods sector prices. Furthermore, imported goods become cheaper since goods producers in other regions have cheaper access to business services. The fate of goods-producing regions is different: there, all else equal, goods prices rise due to increased exports, even though they decrease overall through access to much cheaper business services (see Figure 10). These effects combine to explain the more substantial increase of college welfare premium as compared to the college wage premium.

6 Conclusion

The rise in income inequality since the 1980s has had a marked impact on the political, social, and economic cohesion of the United States. Geography plays an essential role in these developments. Increasingly, high-skill, well-educated workers concentrate in a handful of large labor markets plugged into the global marketplace. At the same time, many parts of the United States seem increasingly decoupled from the fast-moving, skill-hungry global economy.

Why is it no longer the case that high- and low-skill workers experience equally shared wage and welfare gains in the same locations? In this study, I applied an understanding of one of the critical features of recent technological progress to suggest an answer to this question. I argue that the recent technological advances have fundamentally altered the spatial linkages that connect U.S. local labor markets. These changes have enabled a spatial fragmentation of high- and low-skill activities that is unprecedented in human history. Today lawyers in New York, in a single day, can advise clients throughout the country in video calls as if they were locally present. At the same time, firms can use the internet to find and interact with the foremost experts to whatever problem they confront without ever meeting face-to-face.

In the present paper, I argued that a distinctive feature of this development is that it generates labor demand for low- and high-skill workers in different localities. As a result, the skill premium rises in some labor markets and declines in others, with the aggregate effect determined by the ease with which workers relocate across occupations, sectors, and regions. I presented a method to quantify how much easier it has become to trade services across space since the 1980s by drawing on data on regional trade imbalances and a structural model of interregional trade. The estimated change is substantial and explains a large part of the unequal growth of the skill premium across U.S. commuting zones between 1980 and 2010. It also generates a substantial increase in the aggregate college wage gap and an even more substantial increase in the college welfare premium. These results suggest that the ongoing spatial re-organization of the production structure of the United States plays an important role in understanding the rise in inequality.

Throughout this study, the absence of data on service trade flows presented a key challenge. While my imputation technique allows me to deliver some first insights on the effect of changes in service trade frictions on the U.S. economy, it is vital that our measurement of the service economy improves. A final report by the MIT working group on service offshoring (see Sturgeon (2006)) highlighted the asymmetry in current data collection efforts: while services account for more than 85% of U.S. private sector GDP, there exists very little information on the services that are bought and sold by companies. The Census collects data on 6,000 physical products used as inputs by U.S. firms, but only classifies fewer than 100 separate service inputs. In future work, I plan on drawing on micro-datasets from other countries to answer the many remaining questions on how increasing domestic service trade affects contemporary developed economies.

References

AARLAND, K., J. C. DAVIS, J. V. HENDERSON, AND Y. ONO (2007): "Spatial organization of firms: The decision to split production and administration," *The RAND Journal of Economics*, 38, 480–494.

AHLFELDT, G. M., S. J. REDDING, D. M. STURM, AND N. WOLF (2015): "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 83, 2127–2189.

ALLEN, T. AND C. ARKOLAKIS (2014): "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 129, 1085–1140.

——— (2018): "Modern Spatial Economics: A Primer," *Yale mimeo*.

ALLEN, T., C. ARKOLAKIS, AND X. LI (2015): "On the Existence and Uniqueness of Trade Equilibria," .

ALLEN, T. AND D. DONALDSON (2018): "The Geography of Path Dependence," Tech. rep.

ANDERSON, J. E. (1979): "A theoretical foundation for the gravity equation," *American Economic Review*, 69, 106–116.

ATALAY, E., A. HORTAÇSU, AND C. SYVERSON (2014): "Vertical integration and input flows," *American Economic Review*, 104, 1120–48.

AUTOR, D. AND D. DORN (2013): "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 103, 1553–97.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The skill content of recent technological change: An empirical exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

BAUM-SNOW, N. AND R. PAVAN (2013): "Inequality and city size," *Review of Economics and Statistics*, 95, 1535–1548.

BERRY, C. R. AND E. L. GLAESER (2005): "The divergence of human capital levels across cities," *Papers in Regional Science*, 84, 407–444.

BLINDER, A. S. AND A. B. KRUEGER (2013): "Alternative measures of offshorability: a survey approach," *Journal of Labor Economics*, 31, S97–S128.

BURSTEIN, A., G. HANSON, L. TIAN, AND J. VOGEL (2017): "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence from the US," Tech. rep., National Bureau of Economic Research.

BURSTEIN, A., E. MORALES, AND J. VOGEL (2015): "Accounting for changes in between-group inequality," Tech. rep., National Bureau of Economic Research.

BURSTEIN, A. AND J. VOGEL (2017): "International trade, technology, and the skill premium," *Journal of Political Economy*, 125, 1356–1412.

CAIRNCROSS, F. (1997): *The Death of Distance : How the Communications Revolution Will Change our Lives*, Harvard Business School Press Boston, Mass.

CALIENDO, L., M. A. DVORKIN, AND F. PARRO (2015): "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," .

CALIENDO, L. AND F. PARRO (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *The Review of Economic Studies*, 82, 1–44.

CALIENDO, L., F. PARRO, E. ROSSI-HANSBERG, AND P.-D. SARTE (2017): "The impact of regional and sectoral productivity changes on the US economy," *The Review of Economic Studies*, 85, 2042–2096.

COSTINOT, A. AND A. RODRÍGUEZ-CLARE (2014): "Trade Theory with Numbers: Quantifying the Consequences of Globalization," in *Handbook of International Economics*, Elsevier, vol. 4, 197–261.

COSTINOT, A., J. VOGEL, AND S. WANG (2012): "An Elementary Theory of Global Supply Chains," *Review of Economic Studies*, 80, 109–144.

DAVIS, D. R. AND J. I. DINGEL (2018): "A Spatial Knowledge Economy," Tech. rep., National Bureau of Economic Research.

DEARDORFF, A. V. (2001): "International provision of trade services, trade, and fragmentation," *Review of International Economics*, 9, 233–248.

DEKLE, R., J. EATON, AND S. KORTUM (2007): "Unbalanced trade," *American Economic Review*, 97, 351–355.

DINGEL, J. I. (2016): "The Determinants of Quality Specialization," *The Review of Economic Studies*, 84, 1551–1582.

DONALDSON, D. AND R. HORNBECK (2016): "Railroads and American economic growth: A market access approach," *The Quarterly Journal of Economics*, 131, 799–858.

DORN, D. (2009): "Essays on inequality, spatial interaction, and the demand for skills," Ph.D. thesis.

DURANTON, G., P. M. MORROW, AND M. A. TURNER (2014): "Roads and Trade: Evidence from the US," *Review of Economic Studies*, 81, 681–724.

DURANTON, G. AND D. PUGA (2004): "Micro-foundations of urban agglomeration economies," in *Handbook of regional and urban economics*, Elsevier, vol. 4, 2063–2117.

——— (2005): "From sectoral to functional urban specialisation," *Journal of Urban Economics*, 57, 343–370.

EATON, J. AND S. KORTUM (2002): "Technology, geography, and trade," *Econometrica*, 70, 1741–1779.

——— (2018): "Trade in Goods and Services," *Yale mimeo*.

ECKERT, F. AND M. PETERS (2018): "Spatial Structural Change," .

FAJGELBAUM, P. D., E. MORALES, J. C. SUÃ¡REZ SERRATO, AND O. ZIDAR (2018): "State Taxes and Spatial Misallocation," *The Review of Economic Studies*, rdy050.

FORT, T. C. (2017): "Technology and production fragmentation: Domestic versus foreign sourcing," *The Review of Economic Studies*, 84, 650–687.

FORT, T. C., J. R. PIERCE, AND P. K. SCHOTT (2018): "New perspectives on the decline of us manufacturing employment," *Journal of Economic Perspectives*, 32, 47–72.

FORTIN, N., T. LEMIEUX, AND S. FIRPO (2011): "Decomposition methods in economics," in *Handbook of labor economics*, Elsevier, vol. 4, 1–102.

FRANCOIS, J. AND B. HOEKMAN (2010): "Services trade and policy," *Journal of economic literature*, 48, 642–92.

FRIEDMAN, T. L. (2005): *The world is flat: A brief history of the twenty-first century*, Macmillan.

GANONG, P. AND D. SHOAG (2017): "Why has regional income convergence in the US declined?" *Journal of Urban Economics*, 102, 76–90.

GERVAIS, A. AND J. B. JENSEN (2013): "The tradability of services: Geographic concentration and trade costs," Tech. rep., National Bureau of Economic Research.

GIANNONE, E. (2017): "Skilled-biased technical change and regional convergence," Tech. rep.

GIROUD, X. (2013): "Proximity and investment: Evidence from plant-level data," *The Quarterly Journal of Economics*, 128, 861–915.

GOOS, M., A. MANNING, AND A. SALOMONS (2014): "Explaining job polarization: Routine-biased technological change and offshoring," *American Economic Review*, 104, 2509–26.

GRIFFITHS, B. (1975): "Invisible Boarders to Invisible Trade," .

HELPMAN, E. (1987): "Imperfect competition and international trade: evidence from fourteen industrial countries," *Journal of the Japanese and international economies*, 1, 62–81.

HILLBERRY, R. AND D. HUMMELS (2008): "Trade responses to geographic frictions: A decomposition using micro-data," *European Economic Review*, 52, 527–550.

HOEKMAN, B. (2006): *Liberalizing trade in services: a survey*, The World Bank.

HUBMER, J. (2018): "The Race Between Preferences and Technology," .

HUMMELS, D. AND J. LEVINSOHN (1995): "Monopolistic competition and international trade: reconsidering the evidence," *The Quarterly Journal of Economics*, 110, 799–836.

ISARD, W. (1953): "Regional commodity balances and interregional commodity flows," *American Economic Review*, 43, 167–180.

JENSEN, J. B. AND L. G. KLETZER (2005): "Tradable services: Understanding the scope and impact of services offshoring," in *Brookings trade forum*, JSTOR, 75–133.

——— (2010): "Measuring tradable services and the task content of offshorable services jobs," in *Labor in the new economy*, University of Chicago Press, 309–335.

KATZ, L. F. AND K. M. MURPHY (1992): "Changes in relative wages, 1963–1987: supply and demand factors," *The Quarterly Journal of Economics*, 107, 35–78.

KING, M., S. RUGGLES, J. T. ALEXANDER, S. FLOOD, K. GENADEK, M. B. SCHROEDER, B. TRAMPE, AND R. VICK (2010): "Integrated public use microdata series, current population survey: Version 3.0.[machine-readable database]," *Minneapolis: University of Minnesota*, 20.

KRUGMAN, P. (1991): "Increasing returns and economic geography," *Journal of Political Economy*, 99, 483–499.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 68, 1029–1053.

LEAMER, E. E. (2007): "A flat world, a level playing field, a small world after all, or none of the above? A review of Thomas L Friedman's The World is Flat," *Journal of Economic Literature*, 45, 83–126.

LEE, E. (2015): "Trade, inequality, and the endogenous sorting of heterogeneous workers," .

LEONTIEF, W. AND A. STROUT (1963): "Multiregional input-output analysis," in *Structural interdependence and economic development*, Springer, 119–150.

MACPHERSON, A. (2008): "Producer Service Linkages and Industrial Innovation: Results of a Twelve-Year Tracking Study of New York State Manufacturers," *Growth and Change*, 39, 1–23.

MARKUSEN, J. R. (1989): "Trade in producer services and in other specialized intermediate inputs," *The American Economic Review*, 85–95.

MATTOO, A., R. M. STERN, AND G. ZANINI (2007): *A handbook of international trade in services*, The World Bank.

MELVIN, J. R. (1989): "Trade in producer services: a Heckscher-Ohlin approach," *Journal of Political Economy*, 97, 1180–1196.

MICHAELS, G., F. RAUCH, AND S. J. REDDING (2018): "Task Specialization in US Cities from 1880 to 2000," *Journal of the European Economic Association*.

MONTE, F., S. J. REDDING, AND E. ROSSI-HANSBERG (2015): "Commuting, migration and local employment elasticities," Tech. rep., National Bureau of Economic Research.

MORETTI, E. (2012): *The new geography of jobs*, Houghton Mifflin Harcourt.

MOSES, L. N. (1955): "The stability of interregional trading patterns and input-output analysis," *The American Economic Review*, 45, 803–826.

O'CONNOR, D. (2018): "Lebanon distribution center prepares to open," .

PARRO, F. (2013): "Capital-skill complementarity and the skill premium in a quantitative model of trade," *American Economic Journal: Macroeconomics*, 5, 72–117.

POLENSKE, K. R. (1970): "An empirical test of interregional input-output models: Estimation of 1963 Japanese production," *The American Economic Review*, 60, 76–82.

REDDING, S. J. AND E. ROSSI-HANSBERG (2017): "Quantitative spatial economics," *Annual Review of Economics*, 9, 21–58.

ROSEN, S. (1981): "The Economics of Superstars," *American Economic Review*, 71, 845–858.

STRAUSS-KAHN, V. AND X. VIVES (2009): "Why and where do headquarters move?" *Regional Science and Urban Economics*, 39, 168–186.

STURGEON, T. (2006): "Services offshoring working group final report," *MIT Service Offshoring Working Group Final Report*.

TOLBERT, C. M. AND M. SIZER (1996): "US Commuting Zones and Labor Market Areas: A 1990 Update," *ERS Staff Paper*.

WWW.FEATUREDCUSTOMERS.COM (2018): "Featured Customers Review," .

A Additional Figures and Regressions

A.1 Modes of Service Trade

The members of the World Trade Organization (WTO) signed a General Agreement on Trade and Services (GATS) as part of its Uruguay round of negotiations. As part of this agreement, WTO members agreed on a now widely accepted definition of what constitutes trade in services. Table 7 lists the four modes os service trade defined in the GATS.

Table 7: Modes of Service Trade as defined by the WTO

Mode	Criteria
Cross-border supply	Service delivered within the territory of the Member, from the territory of another Member
Consumption abroad	Service delivered outside the territory of the Member, in the territory of another Member, to a service consumer of the Member
Commercial presence	Service delivered within the territory of the Member, through the commercial presence of the supplier
Presence of a natural person	Service delivered within the territory of the Member, with supplier present as a natural person

Note: The Table shows the four modes of service trade as defined in the General Agreement on Service Trade (GATS) by the World Trade Organization that entered into force in 1995 as a result of the Uruguay round of negotiations.

A.2 Concentration Measures of Sectoral Production

Table 8 shows two measures of concentration of sectoral production. For 741 local labor markets in the United States (see Tolbert and Sizer (1996)) I subdivide the local economies into three sectors and construct the share of each in the total local payroll. The sectors are goods-producing sectors, business services sectors, and local services (see E.1.1 for more detail on the grouping). Then I compute the payroll shares at three percentile of the distribution of payroll shares across regions: 10th, 50th, and 90th percentile. Table 8 present rations of the sectoral payroll share at the 90th relative to the 10th percentile and relative to the median of the distribution across commuting zones. As can be seen business services employment became substantially more concentrated.

Table 8: Concentration Measures of Sectoral Production

	p90/p10			p90/p50		
Year	Goods Sectors	Business Services	Local Services	Goods Sectors	Business Services	Local Services
1980	1.41	1.90	1.45	1.17	1.50	1.19
1990	1.45	1.94	1.40	1.17	1.51	1.19
2000	1.46	2.20	1.38	1.17	1.64	1.19
2010	1.51	2.27	1.31	1.18	1.67	1.16

Note: All Data from 5% Publics Use Samples of U.S. Decennial Census Files from 1980 to 2000, and from the American Community Survey for 2010 (obtained via IPUMS, see King et al. (2010)). Using the PUMA identifiers in the data I construct the 741 commuting zones from Tolbert and Sizer (1996). Then I compute the share of total local hours worked in a commuting zone in a given year that are worked in one of three aggregate sectors: goods, business services, and local services. I then consider the distribution of these employment shares across region for each year and sector separately and compute the employment share of the commuting zone at the 10th, 50th, and 90th percentile. Using these statistics, I compute the p90/p50 and p90/p10 ratios as shown.

A.3 Uses of Business Services in the Economy

Table 9 shows a collapsed version of the Input-Output tables from the Bureau of Economic Activity for the year 1980.

Table 9: The Use of Business Services in the Economy, 1980

	Percentage of Output used as				
	Intermediate Inputs			Final Use	Total
Sector	Goods Sectors	Business Services	Local Services		
Goods Sectors	48	1	8	43	100
Business Services	39	18	11	32	100
Local Services	9	3	6	82	100

Note: The table is based on data from the input-output Use tables in producer prices for 1980 published by the Bureau of Economic Analysis.

A.4 The College Share in the Business Services and Goods Sectors

Table 10 shows the college share of employment (measured in hours) for the business services and goods-producing sector as defined in Appendix E.1. The table is based on data from the 5% Public Use Decennial Census Files from 1950-2000 and the 5% Public Use Sample from the American Community Survey (see King et al. (2010)).

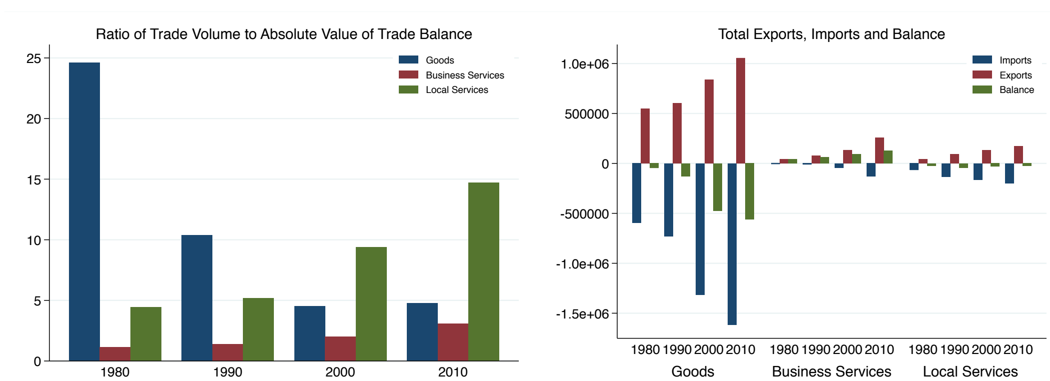
Table 10: College Share of Employment in Goods and Business Service Sector

	1980	1990	2000	2010
Goods	.12	.16	.19	.22
Business Services	.32	.41	.49	.56

Note: All Data from 5% Publics Use Samples of U.S. Decennial Census Files from 1980 to 2000, and from the American Community Survey for 2010 (obtained via IPUMS, see King et al. (2010)). I compute total annual hours worked in each sector for all workers with at least some college and the total hours supplied to a given sector. The Table shows the share of hours supplied to a given sector in a given year that is attributed to college educated workers.

A.5 Aggregate Sector Trade Balances in the United States

Figure 8: Gross Flows to Net Flows in U.S. Imports and Exports

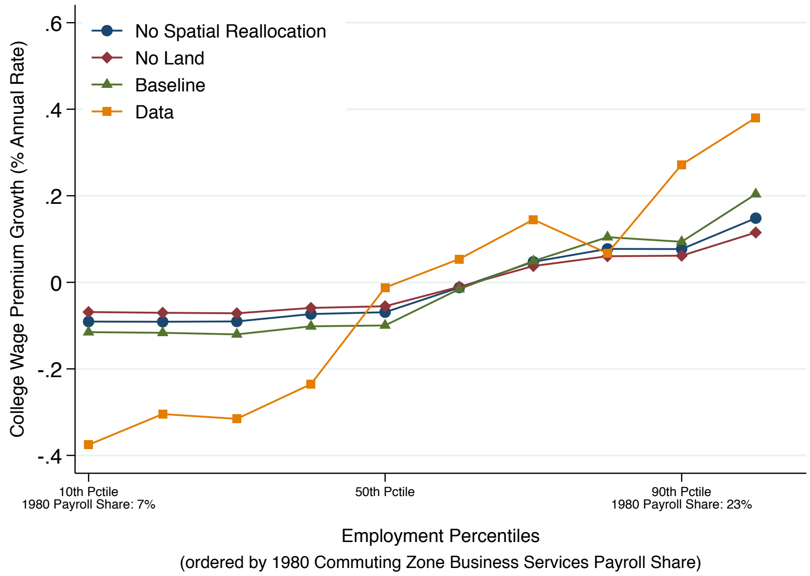


Note: The first Figure shows the ratio of total sector trade volume (the combined absolute values of exports and imports) to the absolute value of the net sectoral trade balance. The second Figure shows the total volume of sectoral exports, imports, and the trade balance for various years. All underlying data stems from the input-output tables of the Bureau of Economic Analysis for the respective year.

A.6 Decomposition of Growing Apart Effect Into Margins of Adjustment

Figure 9 shows the impact of different model margins of adjustment in generating the differential growth in the skill premium. The blue line shows results for a version of the model in which workers cannot migrate across regions. The red line shows results for the specification of the model in which workers can relocate across space, but without structures as an input in production - so that the presence of regional varieties is the only force of congestion. Lastly, the green line shows the performance of the full model in which workers reallocate across regions, sectors, and occupations and both structures and capital are inputs into production. As can be seen the possibility of spatial reallocation alone dampens the effect. This is intuitive: high skill workers with a strong comparative advantage in business services move into business service specialized regions and hence high skill wage growth there is reigned in. Likewise, low-skill workers relocate to goods-producing regions. Adding structures, i.e., a congestion force, into the model amplifies the effect. As high-skill workers crowd into business services specialized regions to work in business services, local goods prices increase faster, as the local factor - structures - gets more expensive. In spatial equilibrium, this entails higher nominal wages for high-skill workers in business service exporting regions as a compensating differential fort the high cost of living. Adding a form of agglomeration forces, whereby local business services productivity depends on the fraction of high-skill workers, is likely to further amplify the effect.

Figure 9: Decomposing the Margins of Adjustment

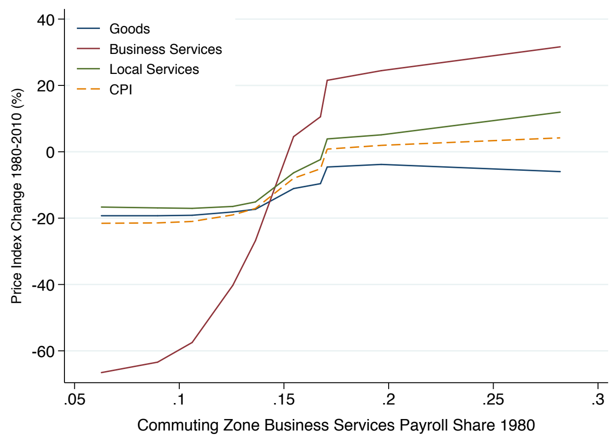


Note: This Figure shows annualized college wage premium growth across commuting zones between 1980 and 2010 in the data (orange line) and the various specifications of the model (other colors). The data is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Wages are computes as unconditional average hourly labor income for workers with at least some college education and workers with only high school education or less. To compute the lines in the Figure, I compute the average growth rate of the wage ratio (college to high-school) within deciles of employment across commuting zones ordered by their business services payroll share in 1980. The Figure shows 95% Confidence Bands on these within-decile averages. In this Figure, the model implied college premium growth is shown for the following three specifications of the model: (1) the model without spatial relocation of workers, structures, and capital (blue); (2) the model with spatial relocation but without structures, and capital; (3) the full model with spatial relocation and structures and capital in value added.

A.7 Sectoral Price Indices Across Regions

Figure 10 shows changes sectoral price index changes induces by the communication cost decline.

Figure 10: Sectoral Price Indices Changes Across Commuting Zones, 1980-2010



Notes: This Figure shows model implied changes in sectoral CES price indices between 1980 and 2010 across commuting zones in the United States. Growth Rates are over 30 years. It also shows changes in the resulting consumer price index, which are computed as the following function of sectoral price index changes, $\hat{P}_{rs}$:

$$\hat{P}_{CPI,r} = \prod_s \hat{P}_{rs}^{\alpha_s}$$

where α_s is the share of sector s in final consumption. $\hat{x} = x'/x$ where x' is the value of the variable in the "counterfactual" equilibrium and x the value of the variable in the "initial" equilibrium.

B Derivations of the Simple Model

In this section, I derive show how to derive the expressions for equilibrium objects shown in Section 3.

Autarky Equilibrium Consider the equilibrium as stated in the main part of the paper. An equilibrium in the simple model consists of a vector of region-sector specific wages $\{w_r^s\}$ and a vector of trade shares $\{\pi^s\}$ that solve the following four market clearing equations and two non-arbitrage equations.

Business services market clearing in each region:

$$w_1^b L_{1h} = \frac{1-\gamma}{\gamma} w_1^g L_{1l} + (1-\pi_{22}^b)\frac{1-\gamma}{\gamma} w_2^g L_{2l} \quad \text{and} \quad w_2^b L_{1h} = \pi_{22}^b \frac{1-\gamma}{\gamma} w_2^g L_{1l}$$

Goods market clearing in each region:

$$w_1^g L_{1l} = \pi_{11}^g \gamma \left[w_1^g L_{1l} + w_1^b L_{1h} \right] \quad \text{and} \quad w_2^g L_{2l} = \gamma \left[w_2^g L_{2l} + w_2^b L_{2h} \right] + (1-\pi_{11}^g)\gamma \left[w_1^g L_{1l} + w_1^b L_{1h} \right]$$

No-arbitrage equations for sectoral prices:

$$(p_1^b - p_2^b)(1 - \pi_{22}^b) = (\frac{w_1^b}{A_{1b}}\kappa - w_2^b)(1 - \pi_{22}^b) = 0 \tag{28}$$

and:

$$(p_1^g - p_2^g)(1 - \pi_{11}^g) = ((w_1^g)^\gamma (w_1^b)^{1-\gamma} A_{1b}^{\gamma-1} - (w_2^g)^\gamma (w_1^b)^{1-\gamma})(1 - \pi_{11}^g) = 0$$

Since trade in goods is costless, it is convenient to choose the nationwide price of the good as numeraire,

$$p_1^g = p_2^g = p^g = 1.$$

Autarky Equilibrium In autarky, homes shares are one, $\pi_{11}^g = \pi_{22}^b = 1$, and the system reduces to the following four equations:

Business services market clearing in each region:

$$w_1^b L_{1h} = \frac{1-\gamma}{\gamma} w_1^g L_{1l} \quad \text{and} \quad w_2^b L_{1h} = \frac{1-\gamma}{\gamma} w_2^g L_{1l}$$

Goods market clearing in each region:

$$w_1^g L_{1l} = \gamma \left[w_1^g L_{1l} + w_1^b L_{1h} \right] \quad \text{and} \quad w_2^g L_{2l} = \gamma \left[w_2^g L_{2l} + w_2^b L_{2h} \right]$$

Along with the normalization of the goods price,

$$p_1^g = p_2^g = p^g = 1.$$

From the normalization of the goods price I obtain the following:

$$w_1^g = (w_1^b)^{\frac{\gamma-1}{\gamma}} A_{1b}^{\frac{1-\gamma}{\gamma}} \quad \text{and} \quad w_{2g} = (w_2^b)^{\frac{\gamma-1}{\gamma}}$$

Plugging this into the respective service market clearing equations produces:

$$w_1^b = \mu^{-\gamma} A_{1b}^{1-\gamma} (\frac{1-\gamma}{\gamma})^\gamma \quad w_1^g = \mu^{1-\gamma} A_{1b}^{1-\gamma} (\frac{\gamma}{1-\gamma})^{1-\gamma} \quad w_2^b = \mu^{-\gamma} (\frac{1-\gamma}{\gamma})^\gamma \quad w_1^g = \mu^{1-\gamma} (\frac{\gamma}{1-\gamma})^{1-\gamma}.$$

But then taking ratios yields the result in the body of the paper:

$$\frac{w_r^b}{w_r^g} = (\frac{1-\gamma}{\gamma})\mu^{-1}.$$

Also relative wages across region in autarky are given by:

$$\frac{w_1^s}{w_2^s} = A_{1b}^{1-\gamma}.$$

The Cutoff Condition Service trade occurs when goods-producing firms in the hinterland find it profitable to purchase service from the city at autarky wage levels. This occurs if and only if:

$$p_1^b \kappa \leq p_2^b$$

At autarky prices this inequality has to hold with equality at the cutoff value for service trade costs, $\bar{\kappa}$:

$$\frac{w_1^b}{A_{1b}}\bar{\kappa} = w_2^b \Rightarrow \frac{\mu^{-\gamma} A_{1b}^{1-\gamma} (\frac{1-\gamma}{\gamma})^{\gamma}}{A_{1b}}\bar{\kappa} = \mu^{-\gamma}(\frac{1-\gamma}{\gamma})^{\gamma}$$

Solving this equation for $\bar{\kappa}$ yields $\bar{\kappa} = A_{1b}^{\gamma}$.

Service Trade Equilibrium For all $\kappa < \bar{\kappa}$, trade shares are no longer one, $\pi_{11}^g, \pi_{22}^b \neq 1$. The equilibrium system can then be written:

Business services market clearing in each region:

$$w_1^b L_{1h} = \frac{1-\gamma}{\gamma} w_1^g L_{1l} + (1-\pi_{22}^b)\frac{1-\gamma}{\gamma} w_2^g L_{2l} \quad \text{and} \quad w_2^b L_{1h} = \pi_{22}^b \frac{1-\gamma}{\gamma} w_2^b L_{1l}$$

Goods market clearing in each region:

$$w_1^g L_{1l} = \pi_{11}^g \gamma \left[w_1^g L_{1l} + w_1^b L_{1h} \right] \quad \text{and} \quad w_2^g L_{2l} = \gamma \left[w_2^g L_{2l} + w_2^b L_{2h} \right] + (1-\pi_{11}^g)\gamma \left[w_1^g L_{1l} + w_1^b L_{1h} \right]$$

No-arbitrage equations for sectoral prices:

$$\frac{w_1^b}{A_{1b}}\kappa = w_2^b \quad \text{and} \quad w_{1g}^{\gamma} w_{1b}^{1-\gamma} A_b^{\gamma-1} = w_{2g}^{\gamma} w_{2b}^{1-\gamma} \tag{29}$$

Since trade in goods is costless, it is convenient to choose the nationwide price of the good as numeraire,

$$p_{1g} = p_{2g} = p_g = 1.$$

From the normalization of the goods price I obtain again:

$$w_{1g} = (w_1^b)^{\frac{\gamma-1}{\gamma}} A_{1b}^{\frac{1-\gamma}{\gamma}} \quad \text{and} \quad w_{2g} = (w_2^b)^{\frac{\gamma-1}{\gamma}} \tag{30}$$

Substituting out π^b from the two business services market clearing equations yields another equation just in wages:

$$w_1^b L_{1h} + w_2^b L_{1h} = \frac{1-\gamma}{\gamma} \left[w_1^g L_{1l} + w_2^g L_{2l} \right] \tag{31}$$

Equations 30 and 31 together with the non-arbitrage equation for business services prices, $\frac{w_1^b}{A_b}\kappa = w_2^b$, are four equations in four unknowns that can be solved for the region-sector wage vector:

$$w_1^b = A_{1b}\kappa^{-1}w_2^b \quad w_1^g = (w_2^b)^{\frac{\gamma-1}{\gamma}}\kappa^{\frac{1-\gamma}{\gamma}} \quad w_2^b = [\frac{1-\gamma}{\gamma}]^{\gamma}\left[\frac{[L_{1l}\kappa^{\frac{1-\gamma}{\gamma}} + L_{2l}]}{[L_{1h}A_b\kappa^{-1} + L_{2h}]}\right]^{\gamma} \quad w_2^g = (w_2^b)^{\frac{\gamma-1}{\gamma}}. \tag{32}$$

Regional skill premia in the free trade equilibrium are given by:

$$\frac{w_1^b}{w_1^g} = A_{1b}(w_2^b)^{\frac{1}{\gamma}} = [\frac{1-\gamma}{\gamma}]\left[\frac{[L_{1l}\kappa^{\frac{1-\gamma}{\gamma}} + L_{2l}]}{[L_{1h}A_b\kappa^{-1} + L_{2h}]}\right]A_{1b} \quad \frac{w_2^b}{w_2^g} = (w_2^b)^{\frac{1}{\gamma}} = [\frac{1-\gamma}{\gamma}]\left[\frac{[L_{1l}\kappa^{\frac{1-\gamma}{\gamma}} + L_{2l}]}{[L_{1h}A_b\kappa^{-1} + L_{2h}]}\right]$$

I can use the business service market clearing equation in the hinterland to solve for the domestic trade share π^b:

$$\pi_{22}^b = \frac{\gamma}{1-\gamma}\frac{w_2^b L_{1h}}{w_2^g L_{1l}} = \frac{\gamma}{1-\gamma}\frac{L_{1h}}{L_{1l}}(w_2^b)^{\frac{1}{\gamma}} = \mu\frac{L_{1l}\kappa^{\frac{1-\gamma}{\gamma}} + L_{2l}}{L_{1h}A_b\kappa^{-1} + L_{2h}}$$

In a similar fashion, I can find an expression for the city's domestic trade share in goods:

$$\pi_{11}^g = \frac{1}{\gamma}\frac{w_1^g L_{1l}}{w_1^g L_{1l} + w_1^b L_{1h}} = \frac{1}{\gamma}\frac{1}{1+\mu\frac{w_1^b}{w_1^g}} = \frac{1}{\gamma}\frac{1}{1+\mu A_{1b}\kappa^{-\frac{1}{\gamma}}(w_2^b)^{\frac{1}{\gamma}}} = \frac{1}{\gamma + \pi_{22}^b A_{1b}\kappa^{-\frac{1}{\gamma}}(1-\gamma)}$$

The expressions for wages in the hinterland in equations 32 can be re-written using this expression for π^b:

$$w_2^b = [\frac{1-\gamma}{\gamma}]^{\gamma}L_{1h}^{-\gamma}(L_{1l}\pi_{22}^b)^{\gamma} \equiv \bar{\gamma}_b(\mu^{-1}\pi_{22}^b)^{\gamma} \quad w_2^g = [\frac{1-\gamma}{\gamma}]^{\gamma-1}L_{1h}^{1-\gamma}\left[L_{1l}\pi_{22}^b\right]^{\gamma-1} \equiv \bar{\gamma}_g\left[\mu^{-1}\pi_{22}^b\right]^{\gamma-1} = ($$

From business services market clearing in the city I can express wages in the city as a function of π^b and κ alone:

$$w_1^b L_{1h} = \frac{1-\gamma}{\gamma}(w_2^b)^{\frac{\gamma-1}{\gamma}}\kappa^{\frac{1-\gamma}{\gamma}}L_{1l} + (1-\pi_{22}^b)\frac{1-\gamma}{\gamma}(w_2^b)^{\frac{\gamma-1}{\gamma}}L_{2l}$$

But then:

$$w_1^b = L_{1h}^{-\gamma}A_{1b}^{1-\gamma}(\frac{1-\gamma}{\gamma})^{\gamma}\left[L_{1l} + (1-\pi_{22}^b)\kappa^{\frac{\gamma-1}{\gamma}}L_{2l}\right]^{\gamma} \equiv \gamma_b A_{1b}^{1-\gamma}\left[\mu^{-1} + (1-\pi_{22}^b)\kappa^{\frac{\gamma-1}{\gamma}}\frac{L_{2l}}{L_{1h}}\right]^{\gamma}$$

Using the expression in equations 32 for the goods wage in the city:

$$w_1^g = A_{1b}^{1-\gamma}\gamma_g \left[\mu^{-1} + (1-\pi_{22}^b)\kappa^{\frac{\gamma-1}{\gamma}} \frac{L_{2l}}{L_{1h}}\right]^{\gamma-1} \equiv A_{1b}^{1-\gamma}\gamma_g \left[\mu^{-1} + (1-\pi_{22}^b)\kappa^{\frac{\gamma-1}{\gamma}} \frac{L_{2l}}{L_{1h}}\right]^{\gamma-1} = (p_1^b)^{\frac{\gamma-1}{\gamma}}$$

Consider again the no-arbitrage equations in the service trade equilibrium. Rearranging the one for business services yields:

$$\frac{w_1^b}{w_2^b} = A_{1b}\kappa^{-1} \propto \kappa^{-1}$$

Rearranging the one in the goods sector yields:

$$\frac{w_1^g}{w_2^g} = \left(\frac{w_1^b}{w_2^b}\right)^{\frac{\gamma-1}{\gamma}} A_{1b}^{\frac{1-\gamma}{\gamma}} = \left(A_{1b}\kappa^{-1}\right)^{\frac{\gamma-1}{\gamma}} A_{1b}^{\frac{1-\gamma}{\gamma}} = \kappa^{\frac{1-\gamma}{\gamma}} \propto \kappa^{\frac{1-\gamma}{\gamma}}$$

C Derivations of the Quantitative Model

This section contains all derivations for the quantitative model omitted in the main part of the paper. It also presents an extension to include land prices and capital in the computations and other materials pertaining to the quantitative model.

C.1 Notation Rules

Throughout the paper, I follow a simple set of rules on how to index exogenous parameters, regional fundamentals, and endogenous variables.

- All Location-specific parameters are indexed with subscripts in the order r, s, o, k, but not separated by a comma.
- If several different regions appear in one equation they are differentiated as follows: r, r', r''. Similarly for s, o, k.
- Trade shares and trade costs are an exception with two subscripts in that order denoting origin and destination region (rr') and a superscript denoting the sector.
- Endogenous variables are indexed with regions and skill groups in subscript (the two "immutable" indices in the baseline model), and sectors and occupation variables in superscript (the two "choice" indices).

- Fundamental parameters not indexed by region are always indexed by subscripts, and by an additional superscript if there are two indices, e.g. $\gamma_s^{s'}$.
- Where attributes of individual workers appear they are indexed by superscript i.

C.2 Aggregation Results

In the baseline model individuals draw sector-occupation specific productivities from the following Fréchet distribution:

$$F(\epsilon) = \exp(-T_{rsok}\epsilon^{-\rho_k})$$

Income of individual i is given by

$$y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\}$$

First compute

$$\begin{aligned}
P(y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\}) &= \int_0^{\infty} -T_{rsok}\rho k^{-\rho_k-1}(w_r^{so})^{\rho_k}\exp(-k^{-\rho_k}\sum_{s'}\sum_{o'} T_{rs'o'k}(w_r^{s'o'})^{\rho_k})dk \\
&= \int_0^{\infty} -T_{rsok}\rho_k k^{-\rho_k-1}(w_r^{so})^{\rho_k}\exp(-k^{-\rho_k}\sum_{s'}\sum_{o'} T_{rs'o'k}(w_r^{s'o'})^{\rho_k})dk \\
&= \frac{T_{rsok}(w_r^{so})^{\rho_k}}{\sum_{s'}\sum_{o'} T_{rs'o'k}(w_r^{s'o'})^{\rho_k}} = \phi_{rk}^{so}
\end{aligned}$$

I compute the probability density of income, conditional on choosing a particular $s - o$ combination being chosen:

$$P(y < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\}) = \frac{1}{P(y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\})}\int_0^k -T_{rsok}\rho_k k^{-\rho_k-1}(w_r^{so})^{\rho_k}\exp(-k^{-\rho_k}\sum_s \sum$$

$$P(y < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\}) = \frac{\sum_s\sum_o (w_r^{so})^{\rho_k} T_{rsok}}{(w_r^{so})^{\rho_k} T_{rsok}}\int_0^k -T_{rsok}\rho_k k^{-\rho_k-1}(w_r^{so})^{\rho_k}\exp(-k^{-\rho_k}\sum_s\sum_o (w_r^{so})^{\rho_k} T$$

$$P(y < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\}) = \int_0^k -\rho_k k^{-\rho_k-1}\sum_s\sum_o (w_r^{so})^{\rho_k} T_{rsok}\exp(-k^{-\rho_k}\sum_s\sum_o (w_r^{so})^{\rho_k} T_{rsok})dk$$

$$(y < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon_{rso}^i\}) = \int_0^k -\rho_k k^{-\rho_k-1}\sum_s\sum_o (w_r^{so})^{\rho_k} T_{rsok}\exp(-\left[\frac{k}{(\sum_s\sum_o (w_r^{so})^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}}\right]^{-\rho_k})dk$$

This is a again a Fréchet Distribution. But then the average wage of a type k worker in commuting zone r is given by the mean of this distribution:

$$\bar{y} = \Gamma(1 - \frac{1}{\rho}) \times (\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}$$

where $\Gamma(.)$ denotes the Gamma function.

Next I derive the expected effective labor supply by a type k worker conditional on choosing a sector occupation pair:

$$P(\epsilon < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\}) = P(\epsilon < \frac{\bar{k}}{w_r^{so}} \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\})$$
$$P(\epsilon < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\}) = P(y < \bar{k} \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\})$$

$$P(\epsilon < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\}) = \int_0^{\bar{k}} -\rho_k \bar{k}^{-\rho_k - 1} \sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok} \exp(- \left[\frac{\bar{k}}{(\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}} \right]^{\rho_k}) d\bar{k}$$

$$P(\epsilon < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\}) = \int_0^{k} -\rho_k k^{-\rho_k - 1} (w_r^{so})^{-\rho_k - 1} \sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok} \exp(- \left[\frac{\bar{k}}{(\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rso}} \right.$$

$$P(\epsilon < k \mid y^i = \max_{s,o}\{w_r^{so} \times \epsilon^i_{rso}\}) = \int_0^{k} -\rho_k k^{-\rho_k - 1} (w_r^{so})^{-\rho_k} \sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok} \exp(- \left[\frac{k}{(w_r^{so})^{-1} (\sum_s \sum_o (w_r^{so}}\right.$$

This is again a Fréchet distribution. so the mean human capital supplied by a worker condition on choosing the sector-occupation combination $s - o$ is:

$$\bar{\epsilon} = \Gamma(1 - \frac{1}{\rho_k}) \times (w_r^{so})^{-1} (\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}$$

But then average income of a type k worker in sector s, occupation o is:

$$w_{rk} = w_r^{so} \bar{\epsilon} = (\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}$$

Total efficiency labor supply to region r, sector s, occupation o:

$$h_r^{so} = \sum_k L_{rk} \phi_{rk}^{so} \Gamma(1 - \frac{1}{\rho_k}) \times (w_r^{so})^{-1} (\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok})^{\frac{1}{\rho_k}}$$

I can then also derive the occupational market clearing condition from the main part of the paper:

$$h_r^{so} w_r^{so} = \mu_{rso}^{\iota} (\frac{w_r^{so}}{w_r^s})^{1-\iota} \gamma_s R_r^s$$

$$\sum_k L_{rk} \phi_{rk}^{so} w_{rk} = \mu_{rso}^{\iota} (\frac{w_r^{so}}{w_r^s})^{1-\iota} \gamma_s R_r^s$$

Lastly, I derive the labor supply expressions. The expected indirect utility of individual i in skill group k if she moved to location r is given by:

$$\bar{V}_{rk} = \varrho \times \frac{w_{rk}}{\prod_s (P_r^s)^{\alpha_s}} \times \eta_r^i = V_{rk} \times \eta_r^i$$

But then:

$$\begin{aligned} P(\bar{V}_{rk} = \max_{r'}\{\bar{V}_{r'k}\}) &= \int_0^{\infty} -G_{rk} \varkappa_k k^{-\varkappa_k - 1} V_{rk}^{\varkappa_k} \exp(-k^{-\varkappa_k} \sum_{r'} G_{r'k} (V_{r'k})^{\varkappa_k}) dk \\ &= G_{rk} V_{rk}^{\varkappa_k} \int_0^{\infty} -\varkappa_k k^{-\varkappa_k - 1} \frac{\sum_{r'} G_{r'k} (V_{r'k})^{\varkappa_k}}{\sum_{r'} G_{r'k} (V_{r'k})^{\varkappa_k}} \exp(-k^{-\varkappa_k} \sum_{r'} G_{r'k} (V_{r'k})^{\varkappa_k}) dk \\ &= \frac{G_{rk} V_{rk}^{\varkappa_k}}{\sum_{r'} G_{r'k} (V_{r'k})^{\varkappa_k}} = \frac{L_{rk}}{L_k} \end{aligned}$$

which completes the derivations of the results involving the Fréchet distribution in Section 3.3.

C.3 A Useful Eigensystem

I introduce a feature of the model that is convenient in computing its equilibria numerically. Using the goods market clearing equation 15, I can rewrite equation 14 as follows:

$$\begin{aligned} E_r^s &= \alpha_s \sum_{s'} \sum_{r'} E_{r'}^{s'} \pi_{rr'}^{s'} \gamma_{s'} + \sum_{s'} (\sum_{r'} E_{r'}^{s'} \pi_{rr'}^{s'}) (1 - \gamma_{s'}) \gamma_{s'}^s \\ &= \sum_{s'} \sum_{r'} E_{r'}^{s'} \pi_{rr'}^{s'} \left[\alpha_s \gamma_{s'} + (1 - \gamma_{s'}) \gamma_{s'}^s \right] \end{aligned} \tag{33}$$

Equation 33 is an eigensystem. The eigenvector corresponding to the eigenvalue 1, is equal to a scaled version of the vector $\{E_r^s\}_{rs}$, denote $\{\lambda E_r^s\}_{rs}$. A normalization of GDP to 1 then pins down λ:

$$GDP = \sum_r \sum_s \sum_{r'} \lambda E^s_{r'} \pi^s_{rr'} \gamma_s \Rightarrow \lambda = \left[\sum_r \sum_s \sum_{r'} E^s_{r'} \pi^s_{rr'} \gamma_s \right]^{-1}$$

I rely heavily on this result when solving the model numerically.

C.4 Baseline Model in Changes

The baseline model can be written as a set of equations expressed in differences. Here $\hat{x} = \frac{x'}{x}$ where x denotes the endogenous variable in the original equilibrium and x' its value in the counterfactual equilibrium. This approach allows me to replace many parameters with objects taken directly from the data. I only allow the service sector trade cost parameter to change. For all other parameters indexed by a region I set $\hat{x} = 1$, i.e., assume they remain unchanged. I now list the model equations re-written in changes in the order that they are used in the algorithm to compute counterfactual. In this algorithm I started with a guess for the vector $\hat{w}^{so}_r$ and then compute all objects below in order starting with 34 and finally updating the guess using equation 35.

1. Labor supply by type k to occupation o in sector s can be written:

$$\hat{\phi}^{so}_{rk} = \frac{(\hat{w}^{so}_r)^{\rho_k}}{\sum_{s'} \sum_{o'} (\hat{w}^{s'o'}_r)^{\rho_k} \phi^{s'o'}_{rk}} \tag{34}$$

2. Type-level average wage:

$$\hat{w}_{rk} = \left[\sum_s \sum_o (\hat{w}^{so}_r)^{\rho_k} \phi^{so}_{rk} \right]^{\frac{1}{\rho_k}}$$

3. Cost per unit of value added:

$$\hat{w}_{rs} = \left[\sum_o (\hat{w}^{so}_r)^{1-\iota} X^{so}_r \right]^{\frac{1}{1-\iota}}$$

with $X^{so}_r = \frac{(\mu_{rso})^{\iota} (w^{so}_r)^{1-\iota}}{\sum_o (\mu_{rso})^{\iota} (w^{so}_r)^{1-\iota}} = \frac{\sum_k w_{rk} L_{rk} \phi^{so}_{rk}}{\sum_o \sum_k w_{rk} L_{rk} \phi^{so}_{rk}}$.

4. Industry commuting zone level price index:

$$\hat{P}^s_r = \left(\sum_{r'} (\hat{p}^s_{r'} \hat{\kappa}^s_{r'r})^{1-\sigma_s} \pi^s_{r'r} \right)^{\frac{1}{1-\sigma_s}}$$

5. Local factory gate prices:

$$\hat{p}_r^s = (\hat{w}_r^s)^{\gamma_s} \times \left(\prod_{s'} \hat{P}_r^{s'} \gamma_s^{s'}\right)^{1-\gamma_s}$$

6. Expenditure shares on goods from elsewhere:

$$\hat{\pi}_{rr'}^s = \frac{(\hat{p}_r^s \hat{\kappa}_{rr'}^s)^{1-\sigma_s}}{\sum_k (\hat{p}_{r''}^s \hat{\kappa}_{r''r'}^s)^{1-\sigma_s} \pi_{r''r'}^s}$$

7. Occupation market clearing:

$$(\frac{\hat{w}_r^{so}}{\hat{w}_r^s})^{1-\iota} \hat{R}_r^s = \sum_k \hat{w}_{rk} \hat{L}_{rk} \hat{\phi}_{rk}^{so} X_{rk}^o \tag{35}$$

where $X_{rk}^o = \frac{w_{rk} L_{rk} \phi_{rk}^{so}}{\sum_k w_{rk} L_{rk} \phi_{rk}^{so}}$ is the payroll share of type k in location r in occupation o. I also use:

$$\hat{R}_r^s = \sum_{r'} \hat{E}_{r'}^s \hat{\pi}_{rr'}^s \frac{E_{r'}^s \pi_{rr'}^s}{\sum_{r'} E_{r'}^s \pi_{rr'}^s} = \sum_{r'} \hat{E}_{r'}^s \hat{\pi}_{rr'}^s x_{rr'}^s$$

where I defined $x_{rr'}^s \equiv \frac{E_{r'}^s \pi_{rr'}^s}{\sum_{r'} E_{r'}^s \pi_{rr'}^s}$. Also notice that I can write

$$\hat{E}_r^s = \sum_{s'} \sum_{r'} \hat{E}_{r'}^{s'} \hat{\pi}_{rr'}^{s'} Z_{rr's'}^s$$

this is the "useful" eigensystem introduced in the main text with $\hat{E}$ a scaled version of the eigenvector corresponding to the eigenvalue 1. Also note that I defined:

$$Z_{rr's'}^s = \frac{E_{r'}^{s'} \pi_{rr'}^{s'} \left[\alpha_s (1+\omega) \gamma_{s'} + (1-\gamma_{s'}) \gamma_{s'}^s\right]}{\sum_{r'} \sum_{s'} E_{r'}^{s'} \pi_{rr'}^{s'} \left[\alpha_s (1+\omega) \gamma_{s'} + (1-\gamma_{s'}) \gamma_{s'}^s\right]}$$

The normalization of US GDP to 1000 pins down the scaling factor for the eigenvalue:

$$\sum_{r \neq ROW} \sum_s R_r^s \gamma_s = 1000$$

Substituting in the market clearing equation and writing the system in changes yields:

$$\sum_{r \neq ROW} \sum_s \left[\sum_{r'} \lambda E_{r's} \pi_{rr'}^s\right] \gamma_s = 1 \Rightarrow \lambda = \left[\sum_{r \neq ROW} \sum_s \sum_{r'} \hat{E}_{r'}^s \hat{\pi}_{rr'}^s z_{rr'}^s\right]^{-1}$$

where I have defined:

$$z_{rr'}^s = \frac{E_{r'}^s \pi_{rr'}^s \gamma_s}{\sum_{r \neq ROW} \sum_s \sum_{r'} E_{r'}^s \pi_{rr'}^s \gamma_s}$$

8. And a spatial equilibrium condition for each worker type k

$$\hat{L}_{rk} = \frac{\left(\frac{\hat{w}_{rk}}{\prod_s (\hat{P}_r^s)^{\alpha_s}}\right)^{\varkappa_k}}{\sum_{r'} \left(\frac{\hat{w}_{r'k}}{\prod_s (\hat{P}_{r'}^s)^{\alpha_s}}\right)^{\varkappa_k} L_{rk}} \tag{36}$$

in the code I need to additionally make ensure that $\hat{L}_{ROW,k} = 1$ always, since ROW workers cannot move. Importantly enforcing the adding up constraint is also necessary:

$$\sum_r L_{rk} = L_k \Rightarrow \sum_r \hat{L}_{rk} \frac{L_{rk}}{\sum_r L_{rk}} = 1$$

C.5 The Model in Changes with Migration, Capital, and Structures

In the baseline model the value added share in production was γ_s and it was composed entirely of labor. In this section, I decompose the value added bundle into three components: a share β_s spent on labor, a share δ_s spent on structure and $1 - \beta_s - \delta_s$ spent on capital. I need to adjust the data construction and amend the equilibrium system in changes as follows.

I use capital and structures shares in value added from Hubmer (2018), who obtains them directly from more disaggregated IO tables for the U.S.. The value I used are listed in Table 11.

Table 11: Structures and Capital Shares in Value Added

Sector (s)	Year	δ_s	β_s
Goods	1980	0.12	0.69
Business Services	1980	0.19	0.63
Local Services	1980	0.39	0.55
Goods	2010	0.14	0.50
Business Services	2010	0.24	0.57
Local Services	2010	0.31	0.65

Notes: The Table shows the shares of structures (δ_s) and labor (β_s) in value sectoral value added for three aggregate sectors in every decade from 1980 and 2010. The shares are drawn from Hubmer (2018) and based on the extended input-output tables of the Bureau of Economic Analysis for the respective decade. The 1980 numbers are in fact from 1982 and the 2010 numbers from 2007, since the extended input output files are only published every five years.

I abstract from structures in final use. Each commuting zone has a fixed endowment of structures, H_r, units of which are rented out at rental rate r_r.

The land market clearing equation in region r is given by:

$$\sum_s \gamma_s \delta_s R_r^s = H_r r_r$$

I can express total sectoral sales in terms of the local payroll,

$$R_r^s = \frac{\sum_o \sum_k L_{rk} w_{rk} \phi_{rk}^{so}}{\beta_s \gamma_s},$$

and plug the resulting expression into the land market clearing equation:

$$\sum_s \frac{\gamma_s \delta_s}{\beta_s \gamma_s} \sum_o \sum_k L_{rk} w_{rk} \phi_{rk}^{so} = H_r r_r$$

This equation can easily be rewritten in changes, in terms of endogenous objects in changes and data already constructed:

$$\sum_s \frac{\delta_s}{\beta_s} \sum_o \sum_k \hat{L}_{rk} \hat{w}_{rk} \hat{\phi}_{rk}^{so} \frac{L_{rk} w_{rk} \phi_{rk}^{so}}{\sum_s \frac{\delta_s}{\beta_s} \sum_{o,k} L_{rk} w_{rk} \phi_{rk}^{so}} = \hat{r}_r$$

Capital markets are assumed to clear on the national level. There is a fixed national capital stock K units of which are rented out at rate R.

The rest of the world region does not use capital nor structures in production. Capital market clearing across U.S. regions is then given by:

$$\sum_{r \neq ROW} \sum_s \gamma_s (1 - \delta_s - \beta_s) R_r^s = K \times R$$

where R is the rental rate of capital, which is endogenous. K is the exogenous stock of capital. Expressing the same equation in terms of local payrolls (which are observable):

$$\sum_{r \neq ROW} \sum_s \gamma_s (1 - \delta_s - \beta_s) \frac{\sum_o \sum_k L_{rk} w_{rk} \phi_{rk}^{so}}{\beta_s \gamma_s} = K \times R$$

Expressing this equation in changes:

$$\sum_{r \neq ROW} \sum_s \frac{(1 - \delta_s - \beta_s)}{\beta_s} \sum_o \sum_k \hat{L}_{rk} \hat{w}_{rk} \hat{\phi}_{rk}^{so} \frac{L_{rk} w_{rk} \phi_{rk}^{so}}{\sum_{r \neq ROW} \sum_s \frac{(1-\delta_s-\beta_s)}{\beta_s} \sum_o \sum_k L_{rk} w_{rk} \phi_{rk}^{so}} = \hat{R}$$

I assume that all land holdings and the entire capital stock is held in a national portfolio in which every citizen holds a share proportional to his or her income. I then just need to solve for the factor ζ by which everyone's income gets scaled up as a result of the capital returns earned on this portfolio.

Total rental and capital income in the economy is:

$$\Theta = \sum_r \sum_s \gamma_s (1-\delta_s-\beta_s) R_r^s + \sum_r \sum_s \gamma_s \delta_s R_r^s = \sum_r \sum_s \gamma_s (1-\delta_s-\beta_s) \frac{\sum_o \sum_k L_{rk} w_{rk} \phi_{rk}^{so}}{\beta_s \gamma_s} + \sum_r \sum_s \delta_s \frac{\sum_o \sum_k L_{rk} w}{\beta_s}$$

which can be expressed more concisely as:

$$\Theta = \sum_r \sum_s \sum_o \sum_k \left[\gamma_s (1-\delta_s-\beta_s) + \gamma_s \delta_s\right] \frac{L_{rk} w_{rk} \phi_{rk}^{so}}{\gamma_s \beta_s}$$

But then the factor of proportionality by which every citizen's pre subsidy wage income gets scaled up as a result of her shares in the national portfolio is given by:

$$\zeta = \frac{\Theta}{\sum_k \sum_r w_{rk} L_{rk}}$$

Also I ca compute the changes in the value of the national portfolio,

$$\hat{\Theta} = \sum_r \sum_s \sum_o \sum_k \hat{L}_{rk} \hat{w}_{rk} \hat{\phi}_{rk}^{so} \frac{\left[\gamma_s (1-\delta_s-\beta_s) + \gamma_s \delta_s\right] R_r^s}{\sum_r \sum_s \left[\gamma_s (1-\delta_s-\beta_s) + \gamma_s \delta_s\right] R_r^s},$$

and use it to construct changes in the factor of proportionality:

$$\hat{\zeta} = \frac{\hat{\Theta}}{\sum_k \sum_r \hat{w}_{rk} \hat{L}_{rk} \frac{w_{rk} L_{rk}}{\sum_k \sum_r w_{rk} L_{rk}}}.$$

The adjustment for the deficit is now slightly changed. In particular the subsidy now adjusts across counterfactuals. Each consumer in the economy gets his income scaled up (down) as an exogenous transfer to match the nationwide deficit:

$$\omega \sum_r \sum_k L_{rk} w_{rk} (1+\zeta) = D \Rightarrow \omega = \frac{D}{\sum_r \sum_k L_{rk} w_{rk} (1+\zeta)}$$

Changes in the subsidy are then given by:

$$\hat{\omega} = \left[\frac{(1+\zeta')}{(1+\zeta)} \sum_r \sum_k \hat{w}_{rk} \hat{L}_{rk} \frac{w_{rk} L_{rk}}{\sum_k \sum_r w_{rk} L_{rk}}\right]^{-1}$$

As before the tax collected from ROW citizens adjusts. The equation that pins down the ROW tax share of income is:

$$\omega_{ROW} \sum_s \gamma_s R^s_{ROW} = D \Rightarrow \omega_{ROW} = \frac{D}{\sum_k w_{ROWk} L_{ROWk}}.$$

Rewriting this equation in changes:

$$\hat{\omega}_{ROW} = \frac{1}{\sum_k \hat{w}_{ROWk} \frac{w_{ROWk} L_{ROWk}}{\sum_k w_{ROWk} L_{ROWk}}}$$

The system written in changes is again very similar to the baseline with migration case studied in Section **??**. Here I just highlight equations that are added or changed as a result of introducing capital and structures into the model.

With endogenous location choices for some or all types I rewrite the new system of equations in changes:

1. Local factory gate prices now include payments to all factors used in production:

$$\hat{p}^s_r = \hat{w}_{rs}^{\gamma_s \beta_s} \times \hat{r}_r^{\gamma_s \delta_s} \times \hat{R}^{\gamma_s(1-\beta_s-\delta_s)} \times \left(\prod_{s'} \hat{P}_r^{s' \gamma_s^{s'}} \right)^{1-\gamma_s}$$

2. The occupation market clearing equation itself has not changed, but the objects that enter it:

$$(\frac{\hat{w}^{so}_r}{\hat{w}^s_r})^{1-\iota} \hat{R}^s_r = \sum_\tau \hat{w}_{rk} \hat{L}_{rk} \hat{\phi}^{so}_{rk} X^o_{rk} \tag{37}$$

where $X^o_{rk} = \frac{w_{rk} L_{rk} \phi^{so}_{rk}}{\sum_k w_{rk} L_{rk} \phi^{so}_{rk}}$ is the payroll share of type k in location r in occupation o. I also use:

$$\hat{R}^s_r = \sum_{r'} \hat{E}^s_{r'} \hat{\pi}^s_{rr'} \frac{E^s_{r'} \pi^s_{rr'}}{\sum_{r'} E^s_{r'} \pi^s_{rr'}} = \sum_{r'} \hat{E}^s_{r'} \hat{\pi}^s_{rr'} x^s_{rr'}$$

where the expression for $\hat{E}^s_r$ has changed due to the endogenous subsidy and national asset portfolio return:

$$\hat{E}^s_r = \sum_{s'} \sum_{r'} E^{s'}_{r'} \pi^{s'}_{rr'} \left[\alpha_s \gamma_{s'} \delta_{s'} (1+\omega')(1+\zeta') + (1-\gamma_{s'}) \gamma^s_{s'} \right] \frac{E^{s'}_{r'} \pi^{s'}_{rr'}}{\sum_{s'} \sum_{r'} E^{s'}_{r'} \pi^{s'}_{rr'} \left[\alpha_s \gamma_{s'} \delta_{s'} (1+\omega)(1+\zeta) + \right.}$$

note how this equation cannot fully be expressed in changes and features the new national asset portfolio share, ζ', and the new subsidy, ω'. Instead of GDP I now normalize total labor value added to 1000:

$$\sum_{r \neq ROW} \sum_s R^s_r \gamma_s \delta_s = 1000$$

$$\sum_{r \neq ROW} \sum_{s} \left[\sum_{r'} \lambda E_{r'}^{s} \pi_{rr'}^{s} \right] \gamma_s = 1 \Rightarrow \lambda = \left[\sum_{r \neq ROW} \sum_{s} \sum_{r'} \hat{E}_{r'}^{s} \hat{\pi}_{rr'}^{s} z_{rr's} \right]^{-1}$$

where I defined

$$z_{rr'}^{s} = \frac{E_{r'}^{s} \pi_{rr'}^{s} \gamma_s \delta_s}{\sum_{r \neq ROW} \sum_s \sum_{r'} E_{r'}^{s} \pi_{rr'}^{s} \gamma_s \delta_s}$$

D Calibration Details

D.1 A useful Lemma

Lemma. *For any strictly positive vectors $\{A_i\} \gg 0$ and $\{B_i\} \gg 0$, such that $\sum_i A_i = \sum_i B_i$, and any strictly positive matrix $\mathbf{K} \gg 0$ there exists a unique (to scale), strictly positive vector $\{\lambda_i\} \gg 0$.*

Proof. Define $\sum_k \lambda_k K_{kj}$ as μ_j, then rewrite the above equation as two equations:

$$\lambda_i^{-1} = \sum_{j=1,\ldots,N} \mu_j^{-1} R_i^{-1} E_j K_{ij} \quad \text{and} \quad \mu_j = \sum_k \lambda_k K_{kj}$$

The result is then a direct corollary of results in Allen et al. (2015). □

D.2 Monotonicity of local Exports as a function of δ^s

Consider the expression for local exports in region r, sectors s, given in equation 26:

$$EX_r^s = \sum_{r' \neq r} E_{r'}^s \frac{\lambda_{rs} K_{rr'}^s}{\sum_{r''} \lambda_{r''s} K_{r''r'}^s}$$

An alternative way to write total exports in region r, sectors s, is to subtract shipments to itself from total local output in sector s:

$$EX_r^s = R_r^s - E_r^s \pi_{rr}^s = R_r^s - E_r^s \frac{\lambda_{rs}}{\sum_{r'} \lambda_{r's} K_{r'r}^s}$$

But then

$$\frac{d \log EX_r^s}{d \log \delta^s} = E_r^s \pi_{rr}^s \frac{\sum_{r'} \lambda_{r's} K_{r'r}^s \log d_{r'r}}{\sum_{r'} \lambda_{r's} K_{r't}^s} > 0$$

So that for a given vector $\{\lambda_{rs}\}$, increasing δ^s strictly reduces gross exports in all regions and hence overall exports. By implication, for any $\{\lambda_{rs}\}$, there is a unique δ^s to update $\mathbf{K}^s$.

D.3 Details on Trade Frictions Estimation

Here I describe how I estimate δ^s in a given decade and sector.

Step 1: Choose a distance between ROW and U.S. regions.[43] I assume this distance is twice the maximum distance in the continental US. Guess a value for δ^s and construct the matrix $\mathbf{K}^s$ of dimension $N+1 \times N+1$. Guess a vector λ_s of dimension $N+1$ that sums to 1.

Step 2: I now adjust the vector λ entry for ROW, so as to match observed exports from all U.S. regions to ROW. Using the market clearing equation to solve for λ^s_{ROW} in terms of observed ROW exports:

$$EXP^s_{ROW} = \sum_{r \neq ROW} E^s_r \frac{\lambda^s_{ROW} K^s_{ROWr}}{\sum_{r'} \lambda^s_{r'} K^s_{r'r}} \Rightarrow \lambda^s_{ROW} = \frac{EXP^s_{ROW}}{\sum_{r \neq ROW} E^s_r \frac{K^s_{ROWr}}{\sum_{r'} \lambda^s_{r'} K^s_{r'r}}} \tag{38}$$

Note that I do not need ROW expenditure, E^s_{ROW}, to infer λ^s_{ROW}. The current λ vector with the $n+1th$ entry replaced by λ^s_{ROW} would produce exports of ROW that match the data. However, the normalization for the λ no longer holds, so I normalize the λ vector to sum to 1 and go back to equation 38. I iterate on this expression until I found the vector λ that solve equation 38 and sums to 1.

Next I use U.S. exports to ROW (i.e., ROW imports) to impute expenditure in ROW:

$$IMP^s_{ROW} = \sum_{r \neq ROW} E^s_{ROW} \frac{\lambda^s_r K^s_{iROW}}{\sum_{r'} \lambda^s_{r'} K^s_{r'ROW}} \Rightarrow E^s_{ROW} = \frac{IMP^s_{ROW}}{\sum_{r \neq ROW} \frac{\lambda^s_r K^s_{iROW}}{\sum_{r'} \lambda^s_{r'} K^s_{r'ROW}}}$$

This I can simply calculate - no need for iteration here. But then I can also calculate ROW total sales in sector s by using market clearing:

$$R^s_{ROW} = \sum_r E^s_r \frac{\lambda^s_{ROW} K^s_{ROWr}}{\sum_{r'} \lambda^s_{r'} K^s_{r'r}}$$

[43] In order to implement this strategy an assumption is needed on the distance between U.S. regions and ROW. I set this distance equal to the largest distance in the continental U.S. (about 3000 miles). In practice this parameter turns out to have no bearing on the magnitude of the estimated distance elasticity since more than 80% of the trade volume occurs within the United States. It does affect the implied ROW output $R_{ROW,s}$. Intuitively, given observed bilateral flows between the U.S. and ROW, if the distance between the two is large, the model imputes that ROW must be very productive given the size of its shipments to the US, relative to what commuting zones ship among one another. This implies that it is possible to calibrate this distance so as to match the U.S. to ROW GDP ratio.

This number is easy to obtain in practice: U.S. GDP relative to world GDP is between 30-40%. This allows to construct ratio of U.S. to ROW GDP which can be used as calibration target. Intuitively, if ROW is further away, to match the observed imports and exports of the U.S. to and from ROW, ROW needs more productive and hence richer. So as we increase the average distance the ratio of U.S. to ROW GDP increases.

Lastly, I then update λ to be such market clearing holds for all regions using the following mapping.

$$R_r^s = \sum_{r'} E_{r'}^s \frac{\lambda_r^s K_{rr'}^s}{\sum_{r''} \lambda_{r''}^s K_{r''r'}^s} \Rightarrow \lambda_r^s = \frac{R_r^s}{\sum_{r'} E_{r'}^s \frac{K_{rr'}^s}{\sum_{r''} \lambda_{r''}^s K_{r''r'}^s}}$$

I then ensure the normalization holds and then again adjust λ_{ROW}^s so that foreign imports are exactly met etc. I then use the converged λ^s and the measures for ROW expenditure to compute gross exports by sector for every commuting zone in the United States:

$$EXP_{r,Model}^s = \sum_{r' \neq r} E_{r'}^s \frac{\lambda_r^s K_{rr'}^s}{\sum_{r''} \lambda_{r''}^s K_{r''r'}^s}$$

Step 3: I now use the measure for gross exports on the region-sector level to evaluate the following criterion function:

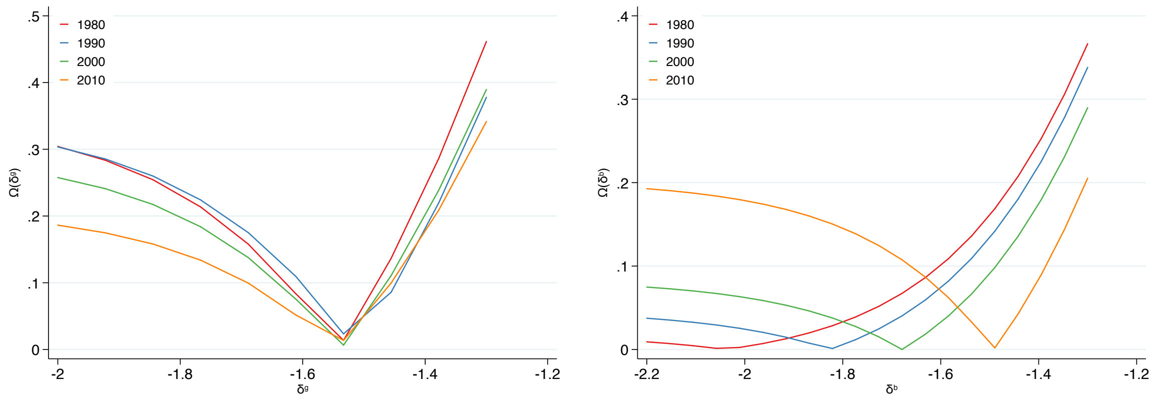
$$\Omega(\delta^s) = \mid \log\left(\frac{(\sum_{r \neq ROW} EXP_{r,Model}^s)}{EXP_{''DATA''}^s}\right) \mid \tag{39}$$

Step 4: For each sector and decade, I repeat steps 1 to 3 for a large number of values of δ^s to identify the value that minimizes equation 39.

Figure 11 below shows the criterion function, $\Omega(\delta^s)$, evaluated over a grid of δ^s points for the four decades in my sample.

Figure 11: Criterion Functions for Sectoral Trade Elasticities

(a) Value of Criterion: Goods Trade　　(b) Value of Criterion: Business Services



Notes: The two graphs show the criterion function,

$$\Omega(\delta^s) = \left| \log\left(\frac{(\sum_{r \neq ROW} EXP^s_{r,Model})}{EXP^s_{\text{"}DATA\text{"}}}\right) \right|,$$

graphed over a grid of values for δ^s. For each value of δ^s, I compute implied interregional trade flows and total regional gross exports and then compute the criterion using the total gross exports implied by regional trade imbalances and the international trade inferred from the input-output tables of the respective year. The Figure shows that for each year and sector there is a unique value δ^s minimizing the criterion function.

D.4 Calibrating ρ_k

I follow the strategy outlined in Eckert and Peters (2018) for calibrating ρ_k. Note that in the model earnings of an individual i in region r who chose to work in sector s and occupation o are:

$$y^i = w_r^{so} \times \epsilon_s^i$$

Since ϵ_i^s is Fréchet distributed with shape parameter ρ_k, realized income is distributed according to a Fréchet distribution, too:

$$F(y) = \exp(-y^{-\rho_k}(\sum_s (w_r^s)^{\rho_k} T_{rsok}))$$

But then it is easy to show that log realized income follows a Gumbel distribution:

$$P(\log y < k) = P(y < \exp(k)) = \exp(-\exp(-\rho_k(k - \frac{1}{\rho_k}\log(\sum_s \sum_o (w_r^{so})^{\rho_k} T_{rsok}))))$$

The variance of log income is given by:

$$\text{var}(\log y) = \frac{\pi^2}{6}(\frac{1}{\rho_k})^2$$

So that the variance of log income conditional on r, s, k, o is just a function of ρ_k. This provides an intuitive way of estimating ρ_k. I estimate the following regression in the micro data underlying the estimation data set:

$$\log w^i = \delta_{r,s,o} + u^i_{rs}$$

separately for each k. u^i_{rs} denotes an unexplained residual. Then I compute:

$$\hat{\rho}_k = \sqrt{\frac{\pi^2}{6}\frac{1}{\text{var}(\hat{u}^i_{rs})}}$$

The results from this procedure are listed in Table 3. These estimates imply that more educated workers are more similar in their human capital holdings than the least educated group.

D.5 Calibrating Factor Shares

Recall that the imputation procedure for trade flows implies a $E_{ROW,s}$ and a R_{ROW}, s for every sector and decade. I can then write the following six equations for the ROW:

$$E^s_{ROW} = \alpha_s(w_{ROW}L_{ROW} \times (1-\omega)) + \sum_k R^k_{ROW}(1-\gamma_k)\gamma^s_k \quad \forall s$$

$$R^s_{ROW} = \gamma_s^{-1}\mu^s_{ROW}w_{ROW}L_{ROW} \quad \forall s$$

I assume that the average wage in foreign is the same for all skill types and equal to 1. I also assume that the employment share in all three sectors is equal to 1/3. I then need to calibrate L_{ROW} and $\alpha_s, \gamma_s, \gamma^k_s$. I choose values for these 12 parameters so as to ensure that the above six equations hold. In practice changing these parameters has no impact on my outcomes of interest which is the wage distribution within the United States.

I treat the trade deficit of ROW with the United States as an exogenous constant denoted D_t. I then solve for a subsidy in the United States that is funded through a tax in ROW so as to rationalize D_t. I assume taxes and subsidies are proportional to labor income and the same for all types. Denoting the subsidy in the United States by ω and the tax in ROW ω_{ROW} I need to solve the following two equations:

$$\omega \sum_r \sum_s L_r^s w_r^s = D \Rightarrow \omega = \frac{D}{\sum_r \sum_s L_r^s w_r^s}$$

$$\omega_{ROW} \sum_s \gamma_s R_{ROW}^s = D \Rightarrow \omega_{ROW} = \frac{D}{\sum_s \gamma_s R_{ROW}^s}$$

Given that I normalize GDP $\sum_r \sum_k L_{rk} w_{rk} = 1000$, $\omega = D/1000$ is fixed and remains constant across counterfactuals. ω_{ROW} is endogenous and allowed to adjust across counterfactuals as R_{ROW}^s changes.

E Data Appendix

E.1 Definition of Sectors and Occupations

E.1.1 Sectoral Groupings

The three sectors used in the calibration of the model consist of the following sub-industries in the 2010 IO tables (BEA Naics codes in brackets):

- **Goods Sector:** Farms (111CA), Forestry, fishing, and related activities (113FF), Mining (21), Utilities (22), Construction (23), Manufacturing (31G), Wholesale trade (42), Retail trade (44RT), Transportation and warehousing (48TW)
- **Business Services Sector:** Information (51), Finance, insurance, real estate, rental, and leasing (FIRE), Professional and business services (PROF); except Real Estate and Waste management and remediation services
- **Local Services Sector:** Real Estate (531), Waste management and remediation services (562), Educational services, health care, and social assistance (6), Arts, entertainment, recreation, accommodation, and food services (7), Other services, except government (81), Government (G)

E.1.2 Occupational Groupings

There are approximately 320 occupations in each decennial census file used in this paper. A complete list of the 320 occupations is available from the author on request. Here I list some examples of occupations falling into each of the four groups used in the paper:

- **Abstract-Tradable:** Managers and specialists in marketing, advertising, and public relations; Legislators; Operations and systems researchers and analysts; Purchasing managers, agents and buyers, n.e.c.; Financial managers; Lawyers; Architects; Computer software developers; Statisticians; Human resources and labor relations managers;

- **Abstract-Non-Tradable:** Speech therapists; Earth, environmental, and marine science instructors; Repairers of data processing equipment; Psychologists; Respiratory therapists; Clinical laboratory technologies and technicians; Secondary school teachers; Veterinarians; Dentists; Police, detectives, and private investigators

- **Non-Abstract-Tradable:** Knitters, loopers, and toppers textile operatives; Butchers and meat cutters; Administrative support jobs, n.e.c.; Welders and metal cutters; Telephone operators; Dancers; Metal platers; General office clerks

- **Non-Abstract-Non-Tradable:** Bakers; Protective services, n.e.c.; Miners; Heating, air conditioning, and refrigeration mechanics; Parking lot attendants; Dental assistants; Pest control occupations; Funeral directors; Bus drivers; Baker

E.2 Collapsing the IO Tables: Creating the IO Data Used

In my calibration procedure I match the input-output tables of every year exactly. To enable the model to be calibrated in this way I collapse the IO tables to just three sectors and make some minor adjustments.

I add the rows for government and scrap value into the intermediate input industry "Other Services", which is later collapsed to the "Local Services" sector in my calibration. Similarly, I add all government production columns to the "Other Services" column. The USE tables I use are by construction Industry times Commodity tables, where the number of industries is equal to the number of commodities, but a given industry can produce more than one commodity. This implies that gross output by sector is not equal to gross output by commodity. In my model commodities and sectors coincide and hence I need to make an adjustment. I adjust final consumption within each sector (net of imports and exports) such that gross output by commodity (last column of the table) is brought in line with gross output by industry (last row of the table). The changes to final consumption are not large and can be thought of as adjustments to inventories in that year. After this adjustments the last row of the IO table and the last column coincide.

E.3 Constructing Labor Supply And Wages

For the work with the Census data files, I follow the sample selection procedure in Autor and Dorn (2013) with some minor modifications.

The sample of workers considers individuals who were between 16 and 64 who were employed in the year of the census. I drop workers with missing occupation codes, education codes, industry codes or county-group/PUMA identifier. I calibrate the model to sectoral hours worked in each commuting zone and hence rely on a measure for usual hours worked per year. As in Autor and Dorn (2013) I impute hours worked per week or missing weeks by taking averages within occupation-education groups for which the respective variable is not missing and then replacing the missing value with this average. I then multiply total weeks worked, usual hours worked per week and the sample weight together and collapse the data by commuting zone, sector, occupation and skill group used in this paper.

To construct the hourly wage measure I again consider individuals who were between 16 and 64 who were employed in the year of the census. I drop workers with missing occupation codes, education codes, industry codes or county-group/PUMA identifier. In addition, I also drop all individuals with missing income, with imputed income, or with farm or business income. I also drop workers who are self-employed and who have missing values for hours or weeks worked. I also restrict my sample to individuals who work at least 35 hours a week and 40 weeks per year. I inflate all wages to the year 2004 using the Personal Consumption Expenditure Index obtained from the FRED database at the St. Louis Fed. Lastly, I multiply top coded yearly earnings by 1.5 times the top coded value and assign the average earnings at the 1st percentile of the earnings distribution to workers earning less than that. I then divide total yearly earnings by total hours worked per year within each commuting zone, sector, occupation and skill group bin used in this paper to obtain a measure for yearly hourly wages.

E.4 Constructing Industry Groups

For industries I proceed as follows. First I construct a crosswalk between the ind1990 variable, which is consistently available in the Census data, and the Naics 2012 coding system. I construct weights using the 2000 cross-section employment counts for cases where one ind1990 code maps into the several NAICS 2012 codes. Next I concord the NAICS2012 codes with the modified NAICS codes used in the BEA IO tables in each year. Lastly, I concord the two waves of BEA IO Tables, aggregating the more recent table to the 65 industries used in the earlier table. Then I group industries into goods, business services, and local services according to the classification in Section E.1.1.

E.5 Constructing Occupational Groups

I take the hourly labor supply data set that results from the procedure described in Section E.3. The dataset contains total hours supplied within each commuting zone, sector, occupation and skill group bin used in this paper to obtain a measure for yearly hourly wages. I merge several crosswalks and additional data onto these files.

First, I use the occupation codes in the Census data ("occ") to merge in the occ1990dd codes introduced in Dorn (2009). This allows me to merge in two measures of task intensity of occupations used in Autor and Dorn (2013): offshorability and abstractness.

Fortin et al. (2011) derive the ingredients for the measure of offshorability of occupations, which I interpret as tradability, from O*NET data. Autor and Dorn (2013) use a simple average of two aggregate variables: "face-to-face contact" and "on-site job" and I adopt their measure. Fortin et al. (2011) define "face-to-face contact" as the average value of the O*NET variables "face-to-face discussions," "establishing and maintaining interpersonal relationships," "assisting and caring for others," "performing for or working directly with the public," and "coach- ing and developing others." "on-site job" is the average of the O*Net variables "inspecting equipment, structures, or material," "handling and moving objects," "operating vehicles, mechanized devices, or equipment," and the mean of "repairing and maintaining mechanical equipment" and "repairing and maintaining electronic equipment."

The measure of abstractness is constructed in Autor et al. (2003) from the Dictionary of Occupational titles published by the U.S. Department of Labor in 1977. The "abstractness" measure of an occupation is the average of two DOT variables: "direction control and planning," measuring managerial and interactive tasks, and "GED Math," measuring mathematical and formal reasoning requirements (see Autor et al. (2003) for details).

In the 1980 cross-section, I then sum up total labor supply by occ1990dd (there are about 320 such occupations). I order all occupations by their "abstractness score" and then find the cutoff occupation such that approximately half of the hours supplied in the U.S. in 1980 are in occupations that are less abstract than the cutoff occupation. I do the same for the offshorability measure, which I interpret as a measure of tradability. Then I create four occupation categories: Abstract-Tradable, which contains occupations with an abstractness and tradability score above the median, and Abstract-Non-Tradable, Non-Abstract-Tradable, and Non-Abstract-Non-Tradable similarly defined. These are the four occupation categories used in this paper.

E.6 Adjusting hourly labor supply to match value added shares

The Census data, in combination with the structure of the model implies a level of value added for each sector. The IO tables also give a level of value added for each sector. The IO data is almost certainly more accurate. Consequently, I adjust the Census data to be in line with the IO data on the coarse sector level. I do so without distorting relative average wage levels of occupations, sectors and commuting zones by changing the labor supply (in hours) instead. I make these adjustments in the aggregate and then reallocate hours in proportion to region-sector hours in the data across regions and sectors. Table 12 shows the value added shares by sector in the IO tables and the payroll shares by sector in the Census data.

Consider a given year. From the IO tables I obtain a value added share for each of the three coarse sectors, $\mu^{IO}_{VA,s}$.

The data consists of hourly wages by commuting zone, detailed industry, occupation and education type. I denote the hourly supply within these bins by $L_{r,s,o,k}$ and the corresponding hourly wage by $w_{r,s,o,k}$. If I scale GDP in the data, the sectoral value added share is simply:

$$\mu^{Census}_{VA,s} = \sum_{r,o,k} L_{r,s,o,k} w_{i,s,o,k}$$

But then

$$\mu^{IO}_{VA,s} = \sum_{i,o,k} \frac{\mu^{IO}_{VA,s}}{\mu^{Census}_{VA,s}} L_{i,s,o,k} \bar{w}_{i,s,o,k}$$

So that all I have to do is scale the labor supply count in sector s by the adjustment factor $\frac{\mu^{IO}_{VA,s}}{\mu^{Census}_{VA,s}}$. In the case with structures and capital the adjustment is similar.[44]

[44]In fact, reassuringly, when including capital and structures the payroll share in the Census and the labor share in value added are even more similar than they when structures and capital are disregarded.

Table 12: Value Added Shares in IO-Tables and Payroll Shares in Census Files

Sector (s)	Year	Census Files	IO-Tables
Goods	1980	.55	.50
Business Services	1980	.14	.15
Local Services	1980	.32	.36
Goods	2010	.35	.35
Business Services	2010	.20	.21
Local Services	2010	.45	.43

Notes: The Table presents the aggregate sectoral value added shares implied by the 5% Sample of the U.S. Decennial Census Files (1980) and the American Community Survey (2010) and the same quantities obtained from the Input-Output Use Tables in Producer Prices from the Bureau of Economic Analysis for 1980 and 2010.

F Robustness

In this section, I discuss a number of robustness checks on the results in Section 5.

F.1 The Functional Form of Business Service Trade Cost

In the body of the text, I parameterized business services trade costs as a function of distance. This was motivated by a literature on international services trade which finds that flows do decline with distance in a way similar to goods (see e.g. Eaton and Kortum (2018)). The absence of data on trade flows for (Business) Services *within* the United States makes it impossible to assess whether this assumption is supported in the data (some small-sample surveys suggest it is, e.g. Macpherson (2008)). While our intuition for the flow-distance relationship in goods trade is firmly rooted in the fact that transporting goods over longer distances is more costly and hence less done, the rise of digital forms of transmitting information suggests that this could be different for business services. This suggests another formulation of business services trade costs as a fixed cost, that is equal regardless of which commuting zone a service is shipped to:

$$K^s_{rr'} \equiv \kappa_{rr'}^{(\sigma_s-1)\bar{\delta}^s} \equiv \kappa^{\delta^s} > 1 \forall r \neq r'.$$

I can repeat the calibration exercise for business services trade costs from above by setting $\kappa = 2$, without loss of generality, and calibrating δ^s to match the same targets as in Section 4.3. The resulting "trade cost elasticities" are given by:

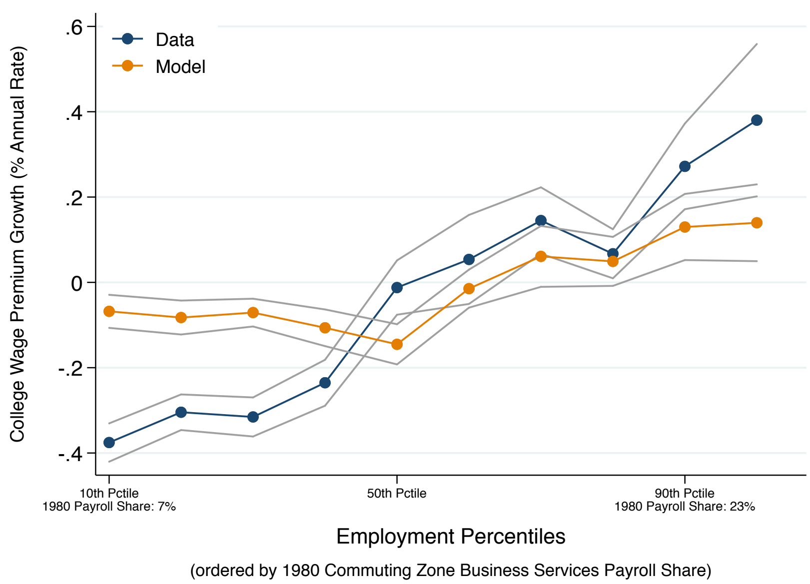
Table 13: Estimates Transformed Distance Elasticities, δ^s

Year	Business Services
1980	-17.21
1990	-15.63
2000	-14.58
2010	-13.26

Notes: The table present estimates of δ^b in the following specification for business services trade costs: $K^s_{rr'} \equiv 2^{\delta^s} > 1 \forall r \neq r'$. In this specification there are only fixed costs to business services trade and costs do not vary with distance.

I now reproduce Figure 3 using this specification for business services trade costs, while leaving trade frictions for goods and local services as calibrated above.

Figure 12: The Growing Apart Effect, 1980-2010

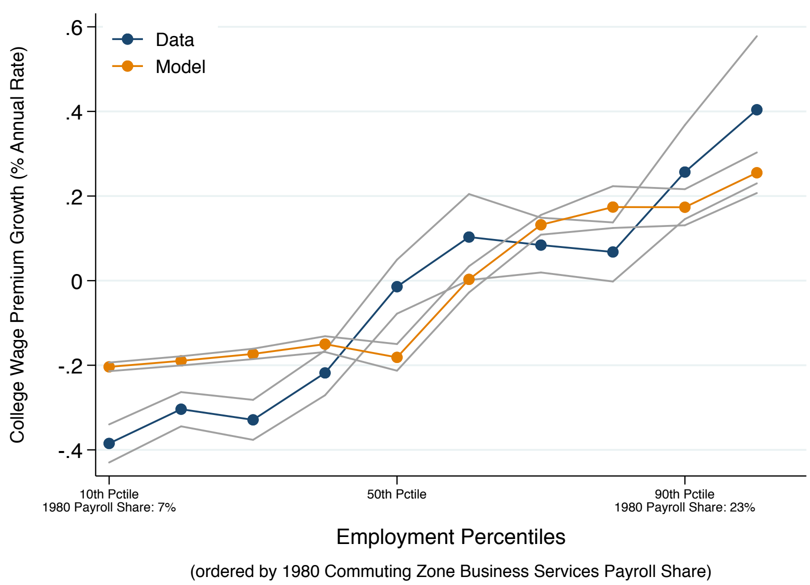


Notes: This Figure show college wage premium growth across commuting zones between 1980 and 2010 in the data (blue line) and the model (orange). The data is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Wages are computes as unconditional average hourly labor income for workers with at least some college education and workers with only high school education or less. To compute the lines in the Figure, I compute the average growth rate of the wage ratio (college to high-school) within deciles of employment across commuting zones ordered by their business services payroll share in 1980. The Figure shows 95% Confidence Bands on these within-decile averages. In this Figure, the model line shown is for the baseline calibration of the model but with the distance elasticity of business services estimated under the assumption of a fixed cost of service trade only.

F.2 Alternative Calibration of Sectoral Trade Costs

As indicated in the text, the assumption that there is no interindustry trade among regions in the United States is very strong. I relax this assumption here. For 1980, I calibrate the distance elasticity for the 1980 cross-section as before. This corresponds to $\approx$ 13 percent of gross business services output being trade across U.S. commuting zones in 1980. The estimate in the main body of the text corresponds to an increase to $\approx$ 15 percent by 2010. As discussed in the text, these numbers are lower bounds on the actual gross volume traded. Accordingly, here I calibrate δ^b_{2010} by assuming a much larger fraction of business service revenue is traded across U.S. regions in 2010: 50%. Figure 13 shows the resulting growth of the college wage premium across regions. As can be seen the model fit improves significantly. This suggests that if the decline in business services trade costs was yet more severe than estimated, the channel highlighted in this paper grows yet more potent to explain the fact in Figure 1 from the introduction.

Figure 13: College Wage Premium Growth Across Commuting Zones, 1980-2010



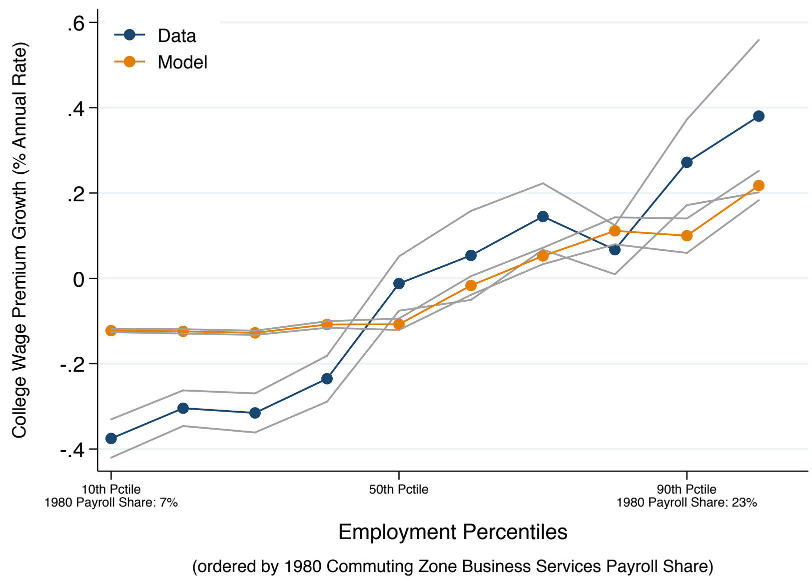
Notes: This Figure show college wage premium growth across commuting zones between 1980 and 2010 in the data (blue line) and the model (orange). The data is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Wages are computes as unconditional average hourly labor income for workers with at least some college education and workers with only high school education or less. To compute the lines in the Figure, I compute the average growth rate of the wage ratio (college to high-school) within deciles of employment across commuting zones ordered by their business services payroll share in 1980. The Figure shows 95% Confidence Bands on these within-decile averages. In this Figure, the model line shown is for the baseline calibration of the model but with the distance elasticity of business services set to the value implied by a trade volume of 50% of total business service sales in the 2010 calibration of the model.

F.3 Alternative Elasticity of Substitution between Occupations

In the main body of the text I used $\iota = 0.9$, which is drawn from Goos et al. (2014), and makes occupations complements.Burstein et al. (2017), using an alternative estimating strategy, obtain the estimate $\iota = 1.93$ so that occupations are substitutes. Here, I reproduce the exercises from the main part of the paper using the Burstein et al. (2017) estimate.

Figure 14: College Wage Premium Growth Across Commuting Zones, 1980-2010

(a) $\iota = 1.92$



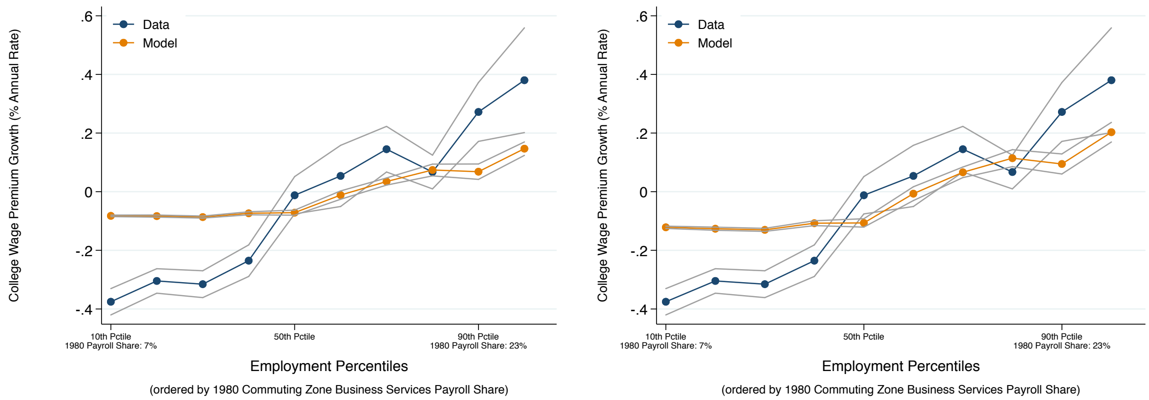
Notes: This Figure show college wage premium growth across commuting zones between 1980 and 2010 in the data (blue line) and the model (orange). The data is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Wages are computes as unconditional average hourly labor income for workers with at least some college education and workers with only high school education or less. To compute the lines in the Figure, I compute the average growth rate of the wage ratio (college to high-school) within deciles of employment across commuting zones ordered by their business services payroll share in 1980. The Figure shows 95% Confidence Bands on these within-decile averages. In this Figure, the model line shown is for the baseline calibration of the model but with the elasticity of substitution between different occupational inputs set to $\iota = 1.92$.

F.4 Alternative Elasticity of Substitution between Regional Varieties

In the main part of the paper, I relied on estimates for σ_s drawn from two sources: Caliendo and Parro (2015) for goods and Gervais and Jensen (2013) for services. These estimates are quite similar across sectors, so that $\sigma_s \approx 6\ \forall s$. In the present section I offer some robustness with regards to this estimate. I consider two alternatives: $\sigma_s = 3\ \forall s$ and $\sigma_s = 9\ \forall s$.

Figure 15: College Wage Premium Growth Across Commuting Zones, 1980-2010

(a) $\sigma = 9$ (b) $\sigma = 3$



Notes: These two Figures show college wage premium growth across commuting zones between 1980 and 2010 in the data (blue line) and the model (orange). The data is constructed from the 5% sample of the U.S. Decennial Census (1980-2000) and American Community Survey (2010). Wages are computes as unconditional average hourly labor income for workers with at least some college education and workers with only high school education or less. To compute the lines in the Figure, I compute the average growth rate of the wage ratio (college to high-school) within deciles of employment across commuting zones ordered by their business services payroll share in 1980. The Figure shows 95% Confidence Bands on these within-decile averages. In this Figure, the model line shown is for the baseline calibration of the model but with the trade elasticity σ set to 3 and 9 respectively, for all sectors.