Essays in Spatial Economics

Jonathan I. Dingel

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

ABSTRACT
Essays in Spatial Economics
Jonathan I. Dingel

A central concern in international economics and urban economics is explaining the distributions of economic assets and activity across space. This dissertation contains three essays examining the pattern of specialization across US cities.

Chapter 1 investigates the determinants of quality specialization within products. A growing literature suggests that high-income countries export high-quality goods. Two hypotheses may explain such specialization, with different implications for welfare, inequality, and trade policy. Fajgelbaum, Grossman, and Helpman (2011) formalize the Linder (1961) conjecture that home demand determines the pattern of specialization and therefore predict that high-income locations export high-quality products. The factor-proportions model also predicts that skill-abundant, high-income locations export skill-intensive, high-quality products (Schott, 2004). Prior empirical evidence does not separate these explanations. I develop a model that nests both hypotheses and employ microdata on US manufacturing plants' shipments and factor inputs to quantify the two mechanisms' roles in quality specialization across US cities. Home-market demand explains at least as much of the relationship between income and quality as differences in factor usage.

In Chapter 2, Donald R. Davis and I develop a theory to jointly explain the distributions of skills, occupations, and industries across cities. Our model incorporates a system of cities, their internal urban structures, and a high-dimensional theory of factor-driven comparative advantage. It predicts that larger cities will be skill-abundant and specialize in skill-intensive activities according to the monotone likelihood ratio property. We test the model using data on 270 US metropolitan areas, 3 to 9 educational categories, 22 occupations, and 21 manufacturing industries. The results provide support for our theory's predictions.

Chapter 3 examines whether larger cities are attractive to consumers. Popular and academic discussions celebrate the virtues of large cities for consumption and leisure. But the standard spatial-equilibrium account says that the consumer attractions of larger cities cannot account for their higher nominal wages and more skilled populations. This chapter revisits that conclusion and shows that the consumption motive can play a first-order role in spatial variation in wage distributions when individuals are heterogeneous. I present a general-equilibrium model in which larger cities offer a greater variety of local goods and services, attracting higher-income individuals who value such variety relatively more. Despite the absence of production-related agglomeration economies, the equilibrium outcomes match a series of facts about spatial variation in wage distributions. I present evidence on the spatial choices of retirees, who consume but do not produce, that is consistent with consumption-driven agglomeration.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I am deeply indebted to Don Davis, who is a great advisor and coauthor. I am very grateful to Amit Khandelwal, Eric Verhoogen, Jonathan Vogel, and David Weinstein for their advice and support throughout my graduate studies.

I also thank many faculty and colleagues for valuable suggestions and encouragement over the years: Nicolas Crouzet, Ron Findlay, Jonathan Fisher, Jessie Handbury, Walker Hanlon, Colin Hottman, Corinne Low, Antonio Miscio, Joan Monras, Eduardo Morales, David Munroe, Suresh Naidu, Peter Neary, Thuy Lan Nguyen, Donald Ngwe, Hyunseung Oh, Paul Piveteau, Keeyoung Rhee, Bernard Salanié, Daniel Sturm, Sébastien Turban, and Reed Walker.

Thanks to my family and friends for unflagging support.

For financial support, I thank the Program for Economic Research at Columbia University and the Institute for Humane Studies.

Chapter 1

The Determinants of Quality Specialization

Jonathan I. Dingel[1]

[1] I am grateful to Don Davis, Amit Khandelwal, Eric Verhoogen, Jonathan Vogel, and David Weinstein for invaluable discussions, guidance, and support. Thanks to Michael Best, Brianna Cardiff, Juan Carlos Hallak, Jessie Handbury, Corinne Low, Kyle Meng, Joan Monras, Eduardo Morales, David Munroe, Paul Piveteau, Bernard Salanié, Daniel Sturm, Allie Tepper, Sebastien Turban, Reed Walker, and Columbia colloquium participants for very helpful comments and suggestions. I am grateful to Rob Feenstra and John Romalis for sharing their parameter estimates. This research was conducted while I was a special sworn status researcher of the US Census Bureau. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the US Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

1.1 Introduction

The Linder hypothesis is the oldest theory of product quality in international trade. Linder (1961) posited that profitably exporting a product requires robust demand for that product in the exporter's home market. Since higher-income consumers tend to purchase higher-quality products, Linder conjectured that local consumers' demand causes high-income countries to produce and export high-quality products. This "home-market effect" explanation of quality specialization was recently formalized by Fajgelbaum, Grossman, and Helpman (2011) in a general-equilibrium, monopolistic-competition model. In contrast, the canonical factor-abundance theory of comparative advantage identifies high-income countries' greater supplies of capital and skills as the reason they export high-quality products.[2] These two competing theories have distinct implications for welfare, inequality, and trade policy. Empirical work to date has not identified the importance of each mechanism in quality specialization.

The empirical challenge is that the two theories make the same predictions about country-level trade flows. Each predicts that high-income locations export high-quality products, consistent with the finding that higher-income countries export products at higher prices within narrowly defined product categories (Schott, 2004; Hummels and Klenow, 2005).[3] Similarly, each predicts that high-income locations import high-quality products if preferences are non-homothetic, as indeed they are.[4] Thus, both theories are consistent with the finding that higher-income countries import more from countries exporting products at higher prices (Hallak, 2006). Combining these export and import patterns, both theories predict that countries with more similar incomes trade more intensely with each other, as found by Hallak (2010) and Bernasconi (2013).[5]

[2]For example, Schott (2004, p. 676) suggests that "high-wage countries use their endowment advantage to add features or quality to their varieties that are not present among the varieties emanating from low-wage countries." Linking quality specialization to relative factor supplies dates to at least Falvey (1981).

[3]Throughout this paper, observed "prices" refer to unit values, which are shipments' value-to-quantity ratios. Like international trade data, the data used in this paper describe transactions' values and quantities.

[4]Deaton and Muellbauer (1980, p.144) note that homotheticity "contradicts all known household budget studies, not to mention most of the time-series evidence."

[5]Hallak (2010, p. 459) notes that "several theories can explain a systematic relationship between per capita income and quality production... The prediction of the Linder hypothesis about the direction of trade can be founded on any of these theories."

In this paper, I use theory and data to quantify the roles of the home-market effect and the factor-abundance mechanism in quality specialization across US cities. I develop a model that yields an empirical approach to separate the two mechanisms. It requires plant-level data on shipments and inputs and location-level data on populations and incomes. I implement the empirical strategy using data on US cities and manufacturing plants and find that the home-market effect influences quality specialization at least as much as factor abundance.

To guide my empirical investigation, I introduce a theoretical framework that nests the two mechanisms, each of which has been studied separately. Individuals have non-homothetic preferences over a homogeneous and a differentiated good; higher-income individuals are more likely to consume higher-quality varieties of the differentiated good. This demand assumption makes the model consistent with high-income countries importing high-quality products and generates the home-market effect when trade is costly. Individuals have heterogeneous skills, and goods can be ranked by their skill intensities. This production assumption allows skill-abundant locations to have a comparative advantage in higher qualities when quality is skill-intensive. The model serves two purposes. First, it confirms that each mechanism alone can generate trade flows consistent with the empirical findings described above. Second, the theory identifies a way to separate the two mechanisms using plant-level data. Factor abundance affects specialization exclusively through plants' factor usage. Conditional on plant-level factor intensity, demand alone determines quality specialization. Thus, plant-level data on shipments and inputs can be combined with data on locations' incomes to identify the home-market effect.

To implement this empirical strategy, I use microdata on US manufacturing plants' shipments and inputs from the Commodity Flow Survey and the Census of Manufactures. These sources provide microdata on plants in many cities with different income levels in a single dataset. In contrast, I am not aware of a source containing plant-level shipment and input data from many countries.[6] I document that US cities exhibit the key patterns found in international data. Both outgoing and incoming shipments exhibit higher prices in higher-income cities, and cities with

[6] My empirical approach thus follows the counsel of Krugman (1991a, p.3): "if we want to understand international specialization, a good place to start is with local specialization. The data will be better and pose fewer problems of compatibility, and the underlying economic forces will be less distorted by government policies."

more similar incomes trade more intensely with each other. I therefore proceed to use these data to distinguish between the home-market-effect and factor-abundance hypotheses.

Guided by the model, my empirical investigation yields two main results. First, differences in plants' inputs, which may be induced by either mechanism, explain only a minority of the observed specialization across cities. Most of the variation is within-factor-intensity variation. Second, a market-access measure that describes the income composition of proximate potential customers is strongly related to the pattern of within-intensity specialization. Quantitatively, I find that the home-market effect plays at least as large a role as the factor-abundance mechanism in local quality specialization.

More specifically, in my empirical work I infer quality specialization from two empirical measures commonly used in the literature: unit values and estimated demand shifters. The first measure is based on the idea that higher-quality products sell at higher prices and has been widely used in the international trade literature (Hummels and Skiba, 2004; Schott, 2004; Hallak, 2006; Baldwin and Harrigan, 2011). The second measure follows Sutton (1991, 2012), Berry (1994), Hummels and Klenow (2005), Khandelwal (2010), and others in identifying a product as higher-quality when, conditional on price, it has higher market share. When both measures are available in my data, they yield comparable results.

The first empirical finding is that observed factor-usage differences explain a modest share of within-product specialization. Guided by the model, I construct factor-intensity measures using data on plants' employees, equipment, and wages. Between-intensity variation explains about one quarter of the covariance between locations' per capita incomes and outgoing shipment prices. It explains a larger share of the covariance between incomes and estimated demand shifters, but observed factor-usage differences never explain more than half of the specialization by income per capita in any regression specification. Since the factor-abundance mechanism operates only through between-intensity variation, this finding bounds its explanatory power.[7]

The second empirical finding is that the home-market effect plays a quantitatively significant role in quality specialization, at least as large as differences in observed factor usage. Using data

[7]By its nature, this result describes observed factor usage and cannot rule out unobserved inputs.

on cities' incomes and geographic locations, I construct two market-access measures describing the income composition of proximate potential customers. The first omits the residents of the city in which the plant is located, so that it does not reflect any unobserved local supply-side mechanisms. I find that this measure of demand is strongly positively correlated with manufacturing plants' outgoing shipment prices. In fact, this measure explains a larger share of the covariance between income per capita and outgoing shipment prices than plant-level factor usage. The second market-access measure follows the model by including residents in the city of production. This demand measure consistently explains a larger share of the observed specialization across cities than plants' factor inputs. Within-intensity variation in market access explains 54% of the covariance between product prices and incomes per capita, twice that attributable to factor-usage differences.[8] It explains a similar share, 48%, of the covariance between estimated demand shifters and incomes per capita.[9] I conclude that the home-market effect for quality plays a substantial role in the economic geography of US manufacturing.

These findings are important because the two theories have distinct implications. In predicting the quality of a location's exports, one emphasizes its relative factor supplies while the other stresses its relative proximity to high-income customers. These yield very different predictions, for instance, for poor countries that have rich neighbors.[10] To the extent that specializing in producing high-quality goods improves a country's growth prospects, the strong home-market effect found here suggests an advantage of proximity to high-income countries.[11] And since trade policy can affect market access, governments may influence quality specialization.[12]

My empirical strategy of using plant-level data from US cities of different income levels links

---

[8] Using only within-intensity variation is conservative. Unconditionally, variation in market access accounts for 72% of the price-income covariance.

[9] Factor-usage differences explain 46% of the covariance between estimated demand shifters and incomes per capita, so there is considerably smaller residual variation in the decomposition of this measure.

[10] For example, Mexico and Turkey are developing economies that are proximate to high-income customers in the US and EU, respectively. Verhoogen (2008) shows that increased incentive to export caused quality upgrading by Mexican firms.

[11] See Redding (1996), Aghion, Blundell, Griffith, Howitt, and Prantl (2009), and Lederman and Maloney (2012) on quality and growth.

[12] Helpman and Krugman (1989, p.2): "It is clear that changing one's view of why trade happens, and how international markets work, ought to change one's view of what kind of trade policy is appropriate."

my results to a number of findings in urban and regional economics. I provide the first characterization of production specialization within product categories across cities. Previous empirical work describing variation in manufacturing across US cities has focused on inter-industry specialization (Henderson, 1991; Holmes and Stevens, 2004; Davis and Dingel, 2013) or described the products available to retail consumers without tracking production locations (Handbury and Weinstein, 2011). The finding that the geography of demand plays a major role in specialization complements a nascent literature describing the consumption benefits of living in cities with high-income populations (Glaeser, Kolko, and Saiz, 2001; Diamond, 2012; Handbury, 2012).

The paper is organized as follows. Section 1.2 describes the two competing hypotheses. Section 1.3 introduces a model nesting both and shows how to separate them using plant-level data. Section 1.4 describes the US microdata and pattern of specialization and exchange. Section 1.5 reports the empirical results. Section 1.6 concludes.

1.2 Background

Linder (1961) started from the proposition that home demand is essential to developing an exportable product. Since higher-income consumers tend to purchase higher-quality products, Linder suggested that demand composition causes higher-income locations to produce higher-quality products. A novel implication was that countries with more similar incomes would trade more intensely with each other. Despite its informal theoretical underpinnings, this trade-flow prediction motivated many empirical investigations (Deardorff, 1984).

Krugman (1980) formalized how economies of scale and trade costs generate a "home-market effect" in which the country with a larger home market for a product is the net exporter of that good. Demand differences determine trade because economies of scale and costly transport mean that a larger home market is a competitive advantage. First, economies of scale cause each product to be produced in a single location and sold to many markets. Second, producing in the larger market minimizes transportation costs. The Krugman (1980) model featured homothetic preferences and two products produced by different industries, omitting the roles of non-homothetic preferences and product quality emphasized by Linder.

Fajgelbaum, Grossman, and Helpman (2011) recently formalized how income differences can determine the pattern of quality specialization and trade in a general-equilibrium, monopolistic-competition model. They describe a world economy without traditional supply-side determinants of the pattern of trade. The composition of income in a location determines the composition of demand, since higher-income households are more likely to purchase a higher-quality variety. Plants produce higher qualities in higher-income locations because it is more profitable to produce in the larger home market. These mechanics are consonant with the story suggested by Linder (1961).[13] In equilibrium, high-income locations disproportionately produce, export, and import high-quality products.

The canonical factor-abundance theory of comparative advantage yields the same set of predictions when preferences are non-homothetic. An early example is Markusen (1986), in which the income elasticity of demand for capital-intensive manufactures is greater than one, so that high-income, capital-abundant countries specialize in manufactures that are exported to other high-income countries.[14] Many other models make analogous assumptions about the alignment of comparative advantage and relative demand.[15] In these theories, high-income countries both demand higher-quality products and have Ricardian or factor-abundance-driven comparative advantage in producing them, so that "tastes and capabilities are correlated" but not causally linked (Murphy and Shleifer, 1997, p. 6).

Thus, both theories are consistent with the growing body of empirical evidence suggesting that high-income countries export and import high-quality products. Schott (2004) shows that unit val-

[13]Linder's informal narrative focused on the role of entrepreneurial discovery in bringing products to market. He emphasized the informational costs of distance more than transportation costs and did not explicitly address economics of scale (Linder, 1961, p.89-90).

[14]Markusen (1986, p. 1003) obtains "a Linder-type trading pattern based on a Linder-type demand assumption" with no trade costs and thus no home-market effect. See also Bergstrand (1990). Strictly speaking, these are general-equilibrium models of intersectoral specialization. Falvey (1981) introduced a partial-equilibrium model of within-industry specialization across qualities by capital intensity consonant with the within-product interpretation of factor-abundance theory suggested by Schott (2004).

[15]Flam and Helpman (1987) focus on a setting in which the high-wage country, which demands higher qualities, has Ricardian comparative advantage in producing higher-quality varieties. Using aggregate trade flows, Fieler (2011) estimates a two-sector version of the Eaton and Kortum (2002) Ricardian model. She finds that the industry with a greater income elasticity has greater dispersion in idiosyncratic productivities, causing higher-TFP countries to have comparative advantage in these luxuries. Examining variation across 56 broad sectors, Caron, Fally, and Markusen (2012) find a positive correlation between industries' income elasticities of demand and skill intensities.

ues in product-level US import data are higher for higher-income, more capital- and skill-abundant exporting countries; Hummels and Klenow (2005) find a positive relationship between unit values and exporter income per capita using data from more than 50 importing countries.[16] Khandelwal (2010) estimates demand shifters using US import data and finds that they are positively related to exporting countries' GDP per capita and capital abundance. Feenstra and Romalis (2012) and Hallak and Schott (2011), using other methods, also report that higher-income countries export products inferred to be higher quality. High-income countries import narrowly defined products at higher prices (Hallak, 2006), and higher moments of the income and import price distributions are similarly related (Choi, Hummels, and Xiang, 2009).

These common predictions for country-level trade flows motivate this paper's use of plant-level data to separate the two mechanisms. In short, the challenge prior work has faced is that customers and workers are the same people in country-level data.[17] As the model demonstrates, assessing the factor-abundance hypothesis requires looking at the factors of production employed by exporting plants. A series of studies using firm-level data have shown that exporters and firms producing higher-quality products use more capital-intensive and skill-intensive production. Verhoogen (2008) describes exporting-induced quality upgrading by demonstrating that the Mexican peso crisis induced initially more productive plants to become exporters, increase their average wages, and raise their capital-labor ratio. Hallak and Sivadasan (2013) show that, conditional on size, exporting firms in Chile, Colombia, India, and the United States are more capital-intensive and pay higher wages. These firm-level findings are consistent with the factor-abundance explanation of quality specialization. But they do not provide evidence that differences in factor abundance relate to differences in output across locations, since they describe establishments in a single location.[18]

As a result, there is no prior empirical evidence distinguishing the home-market effect for

[16] Torstensson (1996) reports similar results using Swedish imports and more aggregated product categories.

[17] In addition to looking at country-level capital abundance, Schott (2004) shows that the unit values of exported products are positively correlated with the capital-labor ratio of the relevant three-digit ISIC industry in the exporting country. However, much of the variation reflects cross-country differences in capital abundance, a fact noted by Dollar, Wolff, and Baumol (1988, p. 33). The mean pairwise correlation between any two of the 28 industries' capital-labor ratios across the 34 countries in the Schott (2003) data is 0.5. Moreover, industry data necessarily aggregate heterogeneous plants and may not represent exporters' factor intensities.

[18] These data describe plants in many cities within a single country, but the authors did not exploit cross-city variation.

quality from factor-abundance-determined quality specialization. There is a large literature on the Krugman (1980) home-market effect, in which a larger home market causes specialization in the industry with greater economies of scale.[19] This empirical work has identified the economies-of-scale home-market effect by using observable sectoral characteristics, such as transport costs and elasticities of substitution (Hanson and Xiang, 2004). But these cross-industry sources of variation are unavailable when considering quality specialization within products. Moreover, since the composition of income and the composition of human capital are closely related, both across countries and cities, it is empirically difficult to distinguish the home-market effect for quality from factor-abundance theories of comparative advantage using aggregate data.

I proceed to introduce a theoretical framework that incorporates both of these mechanisms and their interaction in equilibrium. This allows me to derive an empirical strategy that relies on observing plants' inputs and outputs.

1.3 Theory

I introduce a theoretical framework describing an economy in which both the home-market effect and relative factor abundance may influence the pattern of production and exchange. I use a high-dimensional framework with many locations, qualities, and skills.[20] It nests a version of the Fajgelbaum, Grossman, and Helpman (2011) home-market-effect model and a traditional factor-abundance model as special cases. Nesting the two mechanisms within a single framework allows me to analyze each in isolation and their interaction.

The theory delivers two results that are key to the empirical investigation. First, it confirms that quality specialization is overdetermined. Each mechanism alone is sufficient to cause high-income locations to disproportionately produce, export, and import high-quality varieties in equilibrium. Second, the theory identifies an important distinction between the two mechanisms. Conditional

[19]On identifying the Krugman (1980) home-market effect, see Davis and Weinstein (1999), Davis and Weinstein (2003), and Hanson and Xiang (2004).

[20]The problems at hand necessitate such an approach. Matching the facts that both outgoing and incoming shipment prices are increasing in average income necessitates a many-location model. Making comparisons across and within qualities of different factor intensities, which is at the heart of my empirical strategy, necessitates many quality levels.

on plant-level skill intensity, the correlation between local income and plants' output quality is due solely to the home-market effect. This result is the basis of my empirical approach.

In the model, there are K locations indexed by k. Location k has a population of size N_k made up of heterogeneous individuals whose skills, indexed by ω, are distributed according to the density $f(\omega, k)$. I take skill distributions as exogenously determined. This is a standard assumption in models of international trade and innocuous for the purpose of distinguishing the roles of the factor-abundance and home-market effect mechanisms.[21] I assume that locations can be ranked by their skill abundance in the likelihood-ratio sense. The skill distribution $f(\omega, k)$ is strictly log-supermodular, so high-k locations are skill-abundant.[22]

1.3.1 Preferences

Consumer preferences are non-homothetic, so the income distribution influences the composition of demand. As in Fajgelbaum, Grossman, and Helpman (2011), individuals consume a differentiated good and a homogeneous good. Varieties of the differentiated good are indexed by j, and J_q denotes the set of varieties with quality q. For individual h, the utility of consuming z units of the homogeneous good and a unit of variety $j \in J_q$ of the differentiated good is

$$u_{hj} = zq + \epsilon_{hj}, \tag{1.1}$$

where ϵ_{hj} is the individual's idiosyncratic valuation of the variety.[23] An individual's vector of idiosyncratic valuations, ϵ_h, is drawn from the generalized extreme value distribution, $G_\epsilon(\epsilon) = \exp\left[-\sum_{q\in Q}(\sum_{j\in J_q}\exp(-\epsilon_j/\theta_q))^{\theta_q}\right]$, where Q denotes the set of qualities and θ_q governs the strength of idiosyncratic differences among varieties with quality q. This specification yields a

[21] Factor mobility would be relevant in considering counterfactuals, since individuals may migrate across cities in response to economic changes.

[22] My theoretical approach makes extensive use of log-supermodularity as an analytical tool. See Costinot (2009) for an introduction to log-supermodularity in the context of trade theory. In $\mathbb{R}^2$, a function $f(\omega, k)$ is log-supermodular if $\omega > \omega'$, $k > k' \Rightarrow f(\omega, k)f(\omega', k') \geq f(\omega, k')f(\omega', k)$ and strictly log-supermodular when the inequality is strict. Davis and Dingel (2013) provide evidence that cities' skill distributions are consistent with the log-supermodularity assumption.

[23] Varieties are thus both vertically and horizontally differentiated (Beath and Katsoulacos, 1991, p.4-6). Conditional on ϵ_h, all consumers prefer higher-q varieties. If all varieties were the same price, consumers would not be unanimous in their ranking of them due to ϵ_h, so these products are horizontally differentiated.

nested-logit demand system (McFadden, 1978).

An individual chooses variety j and quantity z to maximize utility. The homogeneous good is the numeraire, and variety j is available at price p_j. A consumer with income y_h therefore chooses the variety j that maximizes $(y_h - p_j)q + \epsilon_{hj}$, where $z = y_h - p_j$ is the amount of the homogeneous good purchased after buying a single unit of differentiated variety j. As Fajgelbaum, Grossman, and Helpman (2011) show, the fraction of individuals with income y who demand variety j of quality q is

$$\rho_j(y) = \rho_j(q) \cdot \rho_q(y) = \frac{\exp(-p_j q/\theta_q)}{\sum_{j' \in J_q} \exp(-p_{j'} q/\theta_q)} \frac{\left[\sum_{j' \in J_q} \exp((y - p_{j'})q/\theta_q)\right]^{\theta_q}}{\sum_{q'} \left[\sum_{j' \in J_{q'}} \exp((y - p_{j'})q'/\theta_{q'})\right]^{\theta_{q'}}}.$$

This demand system has two important properties. First, consumers' incomes are systematically related to the quality of the variety they purchase. The market share of variety j of quality q varies with income according to $\frac{1}{\rho_j(y)} \frac{\partial \rho_j(y)}{\partial y} = q - q_a(y)$, where $q_a(y)$ is the average quality consumed by individuals with income y.[24] The fraction of individuals purchasing variety j rises with income if and only if its quality exceeds $q_a(y)$. Second, the elasticity of demand takes a simple form: holding the terms within summations fixed, $\frac{\partial \rho_j(y)}{\partial p_j} \frac{p_j}{\rho_j(y)} = -\frac{q}{\theta_q} p_j$. This property will cause producers of quality q to charge a constant additive markup, $\frac{\theta_q}{q}$.[25]

1.3.2 Production

Production involves employing workers of heterogeneous skills, so relative factor supplies may be a source of comparative advantage. Both the homogeneous good and the differentiated good are produced using a constant-elasticity-of-substitution technology. The homogeneous good is freely traded, produced by perfectly competitive firms, and used as the numeraire. Varieties of the differentiated good are produced by monopolistically competitive firms.

[24] $q_a(y) \equiv \sum_{q \in Q} q \frac{\left[\sum_{j' \in J_q} \exp((y-p_{j'})q/\theta_q)\right]^{\theta_q}}{\sum_{q'} \left[\sum_{j' \in J_\omega} \exp((y-p_{j'})\omega/\theta_{q'})\right]^{\theta_{q'}}}$.

[25] I use the nested-logit demand system in part because the constant-additive-markup property makes the model analytically tractable. Only the former property, that high-income consumers are more likely to purchase high-quality varieties, is necessary for the home-market effect to influence the pattern of specialization. A broad class of non-homothetic preferences exhibit this property.

1.3.2.1 Producing the homogeneous good

The freely traded homogeneous good is produced using a continuum of labor inputs, with skill ω available in location k at wage $w(\omega, k)$. Production exhibits constant returns to scale, so the total cost of producing quantity $x(z, k)$ at unit cost $c(z, k)$ is $x(z, k)c(z, k)$.[26] The unit cost resulting from hiring $\ell(\omega)$ units of skill ω per unit of output is

$$c(z, k) = \min_{\ell(\omega)} \int_{\omega \in \Omega} \ell(\omega) w(\omega, k) d\omega \quad \text{s.t.} \quad \left(\int_{\omega \in \Omega} b(\omega, z) \ell(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \geq 1.$$

The technological coefficients $b(\omega, z)$ describe the contribution of each skill type in production and therefore characterize the homogeneous good's skill intensity. The elasticity of substitution across inputs σ is greater than one and finite. Cost minimization yields per-unit input demands $\ell(\omega, z, k) = w(\omega, k)^{-\sigma} b(\omega, z)^{\sigma}$ wherever $x(z, k) > 0$.

1.3.2.2 Producing varieties of the differentiated good

Firms may enter into the differentiated-good sector by choosing a quality level q that is produced by incurring fixed cost f_q, paid in units of the numeraire. The constant marginal cost of producing units of quality q in location k is

$$c(q, k) = \min_{\ell(\omega)} \int_{\omega \in \Omega} \ell(\omega) w(\omega, k) d\omega \quad \text{s.t.} \quad \left(\int_{\omega \in \Omega} b(\omega, q) \ell(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}} \geq 1.$$

The resulting input demands are $\ell(\omega, q, k) = w(\omega, k)^{-\sigma} b(\omega, q)^{\sigma} c(q, k)^{\sigma}$, with marginal cost

$$c(q, k) = \left(\int_{\omega \in \Omega} b(\omega, q)^{\sigma} w(\omega, k)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}.$$

A firm producing $x(q, k)$ units of quality q in location k hires $x(q, k)\ell(\omega, q, k)$ of skill ω.

Firms producing a differentiated variety in location k can export one unit to destination k' at marginal cost $c(q, k) + \tau_{qkk'}$, where the trade cost $\tau_{qkk'}$ is incurred in units of the numeraire. If the income distribution in location k' is $g(y, k')$, then demand for variety j of quality q by consumers

[26] In a slight abuse of notation, I use z to index the homogeneous good. Recall that in the utility function z denotes the quantity of this good.

in location k' is $d_{jk'} = N_{k'} \int \rho_j(y) g(y, k') dy$, which is

$$d_{jk'} = N_{k'} \mathbb{E}_{k'} \left[\frac{\exp(-p_{jk'} q/\theta_q)}{\sum_{j' \in J_q} \exp(-p_{j'k'} q/\theta_q)} \frac{\left[\sum_{j' \in J_q} \exp((y - p_{j'k'}) q/\theta_q)\right]^{\theta_q}}{\sum_{q'} \left[\sum_{j' \in J_{q'}} \exp((y - p_{j'}) q'/\theta_{q'})\right]^{\theta_{q'}}} \right],$$

where $\mathbb{E}_{k'}$ is the expectations operator with respect to the income distribution in location k'. Taking competitors' behavior as given, the optimal prices charged by firm j producing quality q in location k are given by maximizing profits:

$$\max_{\{p_{jk'}\}} \pi_j = \sum_{k'} d_{jk'}(p_{jk'} - c(q, k) - \tau_{qkk'}) - f_q$$
$$\Rightarrow p_{jk'} = c(q, k) + \tau_{qkk'} - \frac{d_{jk'}}{\frac{\partial d_{jk'}}{\partial p_{jk'}}} = c(q, k) + \tau_{qkk'} + \frac{\theta_q}{q}$$

I now define the equilibrium demand level for variety j of quality q in location k'. Denote the number of firms producing varieties of quality q in location k by $n_{q,k}$. Plugging in optimal prices, demand $d_{jk'}$ can be written in terms of (vectors of) the number of firms ($\mathbf{n}$), the unit costs in each location ($\mathbf{c}$), and trade costs (τ).

$$\begin{aligned} d_{jk'} =& N_{k'} \exp(-(c(q, k) + \tau_{qkk'}) q/\theta_q) \mathbb{E}_{k'} \left[\frac{\exp(yq) \left[\sum_\kappa n_{q,\kappa} \exp((-c(q, \kappa) + \tau_{q\kappa k'}) q/\theta_q)\right]^{\theta_q - 1}}{\sum_{q'} \exp(yq') \left[\sum_\kappa n_{q',\kappa} \exp((-c(q, \kappa) + \tau_{q'\kappa k'}) q'/\theta_{q'})\right]^{\theta_{q'}}} \right] \\ =& N_{k'} \exp(-(c(q, k) + \tau_{qkk'}) q/\theta_q) \Gamma_{k'}(q, \mathbf{n}, \mathbf{c}, \tau) \end{aligned}$$

The function $\Gamma_{k'}(q, \mathbf{n}, \mathbf{c}, \tau)$ describes the share of demand in location k' for quality q given the equilibrium prices and locations of all producers.[27] A firm's sales of quality q to k' from k depend on this demand share, population $N_{k'}$, marginal cost $c(q, k)$, and trade cost $\tau_{qkk'}$.

[27] Fajgelbaum, Grossman, and Helpman (2011) introduce a similar demand measure in their equation (22). Their expression subsumes τ by adjusting $\mathbf{n}$ for trade costs to measure "effective varieties" and does not depend on $\mathbf{c}$ because they assume factor-price equalization.

1.3.3 Equilibrium

In equilibrium, labor markets clear and firms earn zero profits. The full-employment condition for each skill ω in each location k is

$$f(\omega, k) = x(z, k)\ell(\omega, z, k) + \int_{q\in Q}\int_{j\in J_q} x(q, k)\ell(\omega, q, k)djdq.$$

Plugging in firms' labor demands and defining $n_{z,k} = 1$, we can write this as

$$f(\omega, k) = w(\omega, k)^{-\sigma}\int_{r\in z\cup Q} n_{r,k}x(r, k)b(\omega, r)^{\sigma}c(r, k)^{\sigma}dr, \tag{1.2}$$

where the variable of integration r includes both the homogeneous good and qualities of the differentiated good. The local income distribution density, which depends on equilibrium wages $w(\omega, k)$, is $g(y, k) = \int_{\omega\in\Omega: w(\omega,k)=y} f(\omega, k)d\omega$.

The free-entry condition says that the profits from producing quality q in location k, $\pi_{q,k}$, are non-positive everywhere and zero where firms are active: $\pi_{q,k} \leq 0\ \forall k$ and $n_{q,k} > 0 \Rightarrow \pi_{q,k} = 0$, where

$$\begin{aligned}\pi_{q,k} =& \sum_{k'} d_{qkk'}(p_{qkk'} - c(q, k) - \tau_{qkk'}) - f_q \\ =& \frac{\theta_q}{q}\exp(-c(q, k)q/\theta_q)\sum_{k'} N_{k'}\exp(-\tau_{qkk'}q/\theta_q)\Gamma_{k'}(q, \mathbf{n}, \mathbf{c}, \tau) - f_q\end{aligned} \tag{1.3}$$

Note that zero-profit condition for the homogeneous, numeraire good is $c(z, k) \geq 1\ \forall k$ and $x(z, k) > 0 \Rightarrow c(z, k) = 1$.

1.3.4 Equilibrium pattern of specialization and trade

Given the distribution of skills across locations, individuals' preferences, and the production technology, the pattern of production and trade in equilibrium is determined by two forces, trade costs and skill intensities. Trade costs, $\tau_{qkk'}$, are important because they shape the pattern of market access and therefore the home-market effect. Skill intensities, governed by $b(\omega, r)$, are important because the relative abundance of skills varies across locations.

I consider two cases of trade costs and two cases of skill intensities. The two potential trade cost matrices are costless trade, in which $\tau_{qkk'} = 0 \ \forall q \ \forall k \ \forall k'$, and trade costs that are small but positive, $\tau_{qkk'} > 0$ for $k \neq k'$ and $\tau_{qkk} = 0$. The two skill-intensity cases are skill intensities that are uniform across products, $b(\omega, r) = b_1(\omega)b_2(r)$, and skill intensities that are increasing in quality, $b(\omega, r)$ weakly log-supermodular.[28]

The resulting four classes of equilibria are summarized in Figure 1.1. I proceed to analyze each in turn. Section 1.3.4.1 describes equilibrium when neither mechanism is active. The resulting pattern of production is indeterminate. Section 1.3.4.2 characterizes equilibrium when only the factor-abundance mechanism is active, while Section 1.3.4.3 does likewise for the home-market effect. In both cases, high-k locations produce high-q varieties. Together, these two sections demonstrate that each mechanism alone is sufficient to cause high-income locations to produce, export, and import high-quality products. Thus, the existing empirical evidence documenting such patterns does not distinguish between the two mechanisms. Section 1.3.4.4 describes equilibrium when both mechanisms are active and shows how to identify the home-market effect for quality after conditioning on plants' skill intensities.

Figure 1.1: Equilibrium pattern of production

	Uniform skill intensities	Quality is skill-intensive
No trade costs	Indeterminate pattern of production	Factor-abundance specialization
Positive trade costs	Home-market-effect specialization	Factor-abundance mechanism + home-market effect

In the following analysis, it is useful to be able to refer to a product's skill intensity. To that end, define a skill-intensity index $i(r)$ with the properties that $i(r) = i(r') \iff b(\omega, r) = h(r, r')b(\omega, r') \ \forall \omega$ for some function $h(r, r')$ and $i(r) > i(r') \Rightarrow r > r'$.[29] It is convenient to choose the labels $i(r)$ such that $i(r)$ is the identity of the lowest r in the set of products with this skill intensity. This allows us to write $b(\omega, r) = h(r, i(r))b(\omega, i(r))$. It also makes $b(\omega, i)$

[28] When $b(\omega, q)$ is log-supermodular, quality is skill-intensive. By allowing z to take any value, I make no assumption on the skill intensity of the homogeneous good, but I assume that there is a value z making $b(\omega, r)$ a log-supermodular function.

[29] Therefore $i(r) = i(r') \Rightarrow c(r, k) = h(r, r')^{\frac{\sigma}{1-\sigma}} c(r', k)$ and $i(r) = i(r') \Rightarrow \ell(r, \omega, k) = h(r, r')^{\frac{\sigma}{1-\sigma}} \ell(r', \omega, k)$.

strictly log-supermodular by definition, whether $b(\omega, r)$ is multiplicatively separable or weakly log-supermodular. In essence, the intensity index $i(r)$ groups together products so that all products in the higher-i group use relatively more skilled labor for any wage schedule.

1.3.4.1 Uniform skill intensities and costless trade

When trade is costless, the zero-profit condition (1.3) reduces to

$$\pi_{q,k} = \exp(-c(q,k)q/\theta_q)\frac{\theta_q}{q}\sum_{k'} N_{k'}\Gamma(q, \mathbf{n}, \mathbf{c}, \mathbf{0}) - f_q \leq 0.$$

In the absence of trade costs, the structure of demand across destinations k' is orthogonal to the location of production. The profits from producing quality q are highest wherever the unit cost $c(q,k)$ is lowest.

When trade is costless and skill intensities are uniform, unit costs are equal across locations, $c(r,k) = c(r)\ \forall k$.[30] Thus, production of any variety is equally profitable across all locations in equilibrium. To reiterate, the labor-market clearing condition (1.2) becomes

$$f(\omega, k) = w(\omega, k)^{-\sigma} b_1(\omega)^{\sigma} \int_{r \in z \cup Q} n_{r,k} x(r,k) b_2(r)^{\frac{\sigma}{1-\sigma}} c(r)^{\sigma} dr.$$

Since nothing inside the integral depends on ω, the composition of local production $n_{r,k}x(r,k)$ is independent of local factor abundance $f(\omega, k)$ in equilibrium.

Result. When trade is costless and skill intensities are uniform, the pattern of production is indeterminate.

1.3.4.2 Skill-intensive quality and costless trade

Now consider the case when trade is costless and $b(\omega, q)$ is (weakly) log-supermodular. As before, costless trade makes the structure of demand across destinations orthogonal to a location's

[30] Note that when $b(\omega, r) = b_1(\omega)b_2(r)$ unit costs are $c(r,k) = b_2(r)^{\frac{\sigma}{1-\sigma}}\left(\int_{\omega\in\Omega} b_1(\omega)^{\sigma} w(\omega,k)^{1-\sigma} d\omega\right)^{\frac{1}{1-\sigma}}$. If $c(r,k) < c(r,k')$, then $c(r',k) < c(r',k')\ \forall r' \in z \cup Q$. Wages will be bid down until factors are employed.

profitability. The most-profitable location is where production costs are lowest.

When $b(\omega, q)$ is log-supermodular, skill abundance imposes structure on the pattern of production through the labor-market clearing condition. In particular, equation (1.2) and the strict log-supermodularity of $f(\omega, k)$ imply, for $k > k'$ and $\omega > \omega'$,

$$\frac{w(\omega,k)^{-\sigma}}{w(\omega',k)^{-\sigma}} \int_{r \in z \cup Q} \frac{b(\omega,r)^{\sigma}}{b(\omega',r)^{\sigma}} \phi(r,\omega',k) dr > \frac{w(\omega,k')^{-\sigma}}{w(\omega',k')^{-\sigma}} \int_{r \in z \cup Q} \frac{b(\omega,r)^{\sigma}}{b(\omega',r)^{\sigma}} \phi(r,\omega',k') dr,$$

where $\phi(r,\omega',k) \equiv \frac{n_{r,k} x(r,k) b(\omega',r)^{\sigma} c(r,k)^{\sigma}}{\int_{r \in z \cup Q} n_{r,k} x(r,k) b(\omega',r)^{\sigma} c(r,k)^{\sigma} dr}$ is a density.[31] $\phi(r,\omega',k)$ describes the output share of quality (product) r in location k when shares are weighted by their production costs and use of skill ω'. Similarly, define the density $\phi_i(\omega',k) \equiv \int_{r:i=i(r)} \phi(r,\omega',k) dr$, which describes the output share of products with skill intensity i in location k, and the expectation operator $\mathbb{E}_{\omega',k}$ with respect to this density, which is $\mathbb{E}_{\omega',k}[\alpha(i)] \equiv \int_i \alpha(i) \phi_i(\omega',k) di$. Using the fact that $\frac{b(\omega,r)}{b(\omega',r)}$ is the same for all r with $i(r) = i$, the inequality can then be written as

$$\frac{w(\omega,k)^{-\sigma}}{w(\omega',k)^{-\sigma}} \mathbb{E}_{\omega',k} \left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}} \right) > \frac{w(\omega,k')^{-\sigma}}{w(\omega',k')^{-\sigma}} \mathbb{E}_{\omega',k'} \left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}} \right). \tag{1.4}$$

Since $b(\omega, i)$ is strictly log-supermodular, $\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}$ is strictly increasing in i and $\mathbb{E}_{\omega',k}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right)$ is a measure of the average skill intensity of output in k.

This labor-market clearing condition implies that more skill-abundant (higher-k) locations produce more skill-intensive (higher-i) products.[32] Since k is skill-abundant relative to k', products made in k are more skill-intensive ($\mathbb{E}_{\omega',k}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right)$ is greater) or skilled labor in k is relatively cheaper ($\frac{w(\omega,k)^{-\sigma}}{w(\omega',k)^{-\sigma}}$ is greater). However, if skilled labor is relatively cheaper in k, then skill-intensive products' unit costs are relatively lower in k, and thus products made in k must be more skill-intensive in equilibrium. Higher-quality varieties are more skill-intensive when $b(\omega, q)$ is (weakly) log-supermodular, so we interpret inequality (1.4) as saying that k absorbs its greater

[31]For expositional convenience, I assume $f(\omega, k) > 0 \ \forall \omega \in \Omega \ \forall k$, so that $k > k, \omega > \omega' \Rightarrow \frac{f(\omega,k)}{f(\omega',k)} > \frac{f(\omega,k')}{f(\omega',k')}$.

[32]See appendix section A.1 for a more formal derivation of this paragraph's explanation.

relative supply of higher skills by producing higher-quality varieties.[33] This result describes the factor-abundance mechanism for quality specialization.

It is important to note that specialization across qualities of the same skill intensity is indeterminate in this case. Inequality (1.4) depends only on i and imposes no restrictions on $\phi(r, \omega, k)$ conditional on $\phi_i(\omega, k)$. Skill-abundant locations produce higher-quality varieties only because such products are more skill-intensive.[34]

What about the equilibrium pattern of demand? Since trade is costless, varieties' prices do not vary across locations, and therefore the fraction of consumers of a given income level purchasing a variety, $\rho_j(y)$ does not vary across locations. Denoting the equilibrium varieties and factor prices by the vectors $\bar{\mathbf{n}}$ and $\bar{\mathbf{c}}$, demand levels $\Gamma_k(q, \bar{\mathbf{n}}, \bar{\mathbf{c}}, \mathbf{0})$ vary with location k solely due to differences in the composition of income. Demand for higher-quality varieties is relatively greater in higher-income locations.

Result. When trade is costless, there is no home-market effect. When quality is skill-intensive and skill-abundant locations are higher-income locations, higher-income locations both produce more of and have greater demand for higher-quality varieties in equilibrium.

The demand levels $\Gamma_k(q, \bar{\mathbf{n}}, \bar{\mathbf{c}}, \mathbf{0})$ will also be important when we consider the interaction of the skill-abundance mechanism and the home-market effect. When trade is costly, we will consider the limiting equilibrium that approaches the skill-intensive-quality, costless-trade equilibrium. But first I describe the case in which the skill-abundance mechanism is absent.

---

[33] This interpretation neglects the skill intensity of the homogeneous good. If the homogeneous good is more skill-intensive, some skill-abundant locations may produce more skill-intensive output by producing more of the homogeneous good rather than higher-quality varieties. When factor intensities vary both across and within goods, the factor-abundance mechanism may operate along both margins. Empirically, Schott (2004) documents that there is little correlation between countries' factor supplies and across-good specialization. Assuming that the homogeneous good is the least skill-intensive product is sufficient to guarantee that high-k locations specialize in high-q varieties.

[34] This result has been derived without any reference to the demand system beyond the fact that costless trade makes consumers' locations irrelevant to the optimal production location. Thus, the empirical investigation of whether the factor-abundance mechanism alone can explain the pattern of specialization does not depend upon the functional form of the preferences in equation (1.1).

1.3.4.3 Uniform skill intensities and costly trade

When skill intensities are uniform, unit costs are multiplicatively separable in (r, k) and can be written as $c(r, k) = b_2(r)^{\frac{\sigma}{1-\sigma}} c(k)$.[35] The labor-market clearing condition (1.2) becomes

$$f(\omega, k) = b_1(\omega)^{\sigma} w(\omega, k)^{-\sigma} c(k) \int_{r \in z \cup Q} n_{r,k} x(r, k) b_2(r)^{\frac{\sigma}{1-\sigma}} dr.$$

Since nothing inside the integral depends on skill ω, the factor-abundance mechanism imposes no restrictions on the equilibrium composition of local production $n_{r,k}x(r, k)$. Any observed relationship between $f(\omega, k)$ and the pattern of specialization results from the demand channel and reflects the connection between $g(y, k)$ and $f(\omega, k)$.

To characterize how specialization is determined by demand, I follow the approach taken by Fajgelbaum, Grossman, and Helpman (2011) to determining the pattern of specialization when trade costs are small. With uniform skill intensities, the zero-profit condition is

$$\pi_{q,k} = \frac{\theta_q}{q} \exp(-b_2(q)^{\frac{\sigma}{1-\sigma}} c(k) q/\theta_q) \sum_{k'} N_{k'} \exp(-\tau_{qkk'} q/\theta_q) \Gamma_{k'}(q, \mathbf{n}, \mathbf{c}, \tau) - f_q \leq 0.$$

Through this condition, demand governs the location of production in equilibrium. Consider two cases, depending on whether wages vary across locations.

When factor prices equalize, $c(k) = 1 \ \forall k$ and $\pi_{q,k}$ varies only with demand. If trade costs are uniform ($\tau_{qkk'} = \tau_q \ \forall k' \neq k$), profits vary with home demand, $\pi_{q,k} > \pi_{q,k''} \iff N_k \Gamma_k(q, \mathbf{n}, \mathbf{c}, \tau) > N_{k'} \Gamma_{k''}(q, \mathbf{n}, \mathbf{c}, \tau)$. When trade costs are sufficiently low, demands approach their costless-trade equilibrium levels $N_k \Gamma_k(q, \mathbf{n}, \mathbf{c}, \mathbf{0})$.[36] Provided that wages are increasing in skill, demand share $\Gamma_k(q, \mathbf{n}, \mathbf{c}, \mathbf{0})$ is strictly log-supermodular in (q, k), as shown in Lemma A.1 in appendix section A.1. High-k locations are high-income locations because they are skill-abundant, and this causes high-k locations to have relatively greater demand for high-q varieties. This makes producing high-q varieties more profitable in high-k locations. When population sizes are equal and locations

[35] $c(r, k) = b_2(r)^{\frac{\sigma}{1-\sigma}} \left(\int_{\omega \in \Omega} b_1(\omega)^{\sigma} w(\omega, k)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}} \equiv b_2(r)^{\frac{\sigma}{1-\sigma}} c(k)$

[36] This equilibrium is distinct from that of the previous subsection.

specialize, Proposition 6 of Fajgelbaum, Grossman, and Helpman (2011) describes the resulting pattern: if location k produces quality q and location $k' < k$ produces quality q', then $q' < q$. Similarly, since higher-k locations have higher relative demand for higher-q varieties, their imports are higher-quality (see Proposition 7 of Fajgelbaum, Grossman, and Helpman 2011). In the case of $\sigma = \infty$ and $N_k = 1$ $\forall k$, the model under consideration reduces exactly to the model described in section VII of Fajgelbaum, Grossman, and Helpman (2011).

When factor prices do not equalize, the location with the lowest $c(k)$ is the most attractive cost-wise for all producers. Producers are willing to locate in higher-cost locations to the extent that these locations have greater demand for their output so that they save on transport costs. In other words, when trade costs are uniform, if $n_{q,k} > 0$ and $c(k) > c(k')$, it must be that $N_k\Gamma_k(q, \mathbf{n}, \mathbf{c}, \tau) > N_{k'}\Gamma_{k'}(q, \mathbf{n}, \mathbf{c}, \tau)$. Qualities are produced where they are in greater demand. When population sizes are equal and trade costs are sufficiently low, this difference in demand is due solely to the income composition of the two locations. As a result, higher-income locations specialize in higher-quality varieties.

Result. When population sizes are equal, skill intensities are uniform, and trade costs are uniform and small, higher-income locations produce, export, and import higher-quality varieties because demand for such qualities is greater in such locations.

Thus, the home-market effect yields equilibrium patterns of production and trade that match the empirical evidence summarized in section 1.2. Since we obtained the same result in the previous section via the factor-abundance mechanism, quality specialization is overdetermined. Each mechanism alone is sufficient to generate the observed patterns.

Result. Higher-income locations disproportionately producing, exporting, and importing higher-quality varieties on average is consistent with the factor-abundance mechanism operating alone or the home-market effect operating alone.

1.3.4.4 Skill-intensive quality and costly trade

This section analyzes what happens when both mechanisms are active. When quality is skill-intensive and trade is costly, the labor-market clearing condition (1.2) and the zero-profit condition (1.3) jointly govern the pattern of quality specialization. The critical result is that demand alone determines specialization across varieties of the same skill intensity. This result underlies the empirical investigation.

First, consider the labor-market-clearing inequality, which is governed by the factor-abundance mechanism. Following section 1.3.4.2, the inequality is

$$\frac{w(\omega,k)^{-\sigma}}{w(\omega',k)^{-\sigma}}\mathbb{E}_{\omega',k}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right) > \frac{w(\omega,k')^{-\sigma}}{w(\omega',k')^{-\sigma}}\mathbb{E}_{\omega',k'}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right).$$

As shown previously, this requires that $\phi_i(\omega',k)$ place more weight on higher-i varieties in higher-k locations. Output of higher-i varieties is relatively greater in higher-k locations. Thus, skill intensities govern the broad pattern of production.

Second, consider the zero-profit condition, which depends on potential customers' incomes through demand levels. To summarize demand, define a market-access term

$$M_{q,k}(\tau) \equiv \sum_{k'} N_{k'} \exp(-\tau_{qkk'}q/\theta_q)\Gamma_{k'}(q,\bar{\mathbf{n}},\bar{\mathbf{c}},\mathbf{0}),$$

where the costless-trade-equilibrium demand levels $\Gamma_k(q,\bar{\mathbf{n}},\bar{\mathbf{c}},\mathbf{0})$ were found in section 1.3.4.2. When trade costs are small, the profits from producing a variety of quality q in location k are approximately

$$\pi_{q,k} \approx \frac{\theta_q}{q}\exp(-c(q,k)q/\theta_q)M_{q,k}(\tau) - f_q. \tag{1.5}$$

When trade costs are small, profits are not sensitive to the locational decisions of other firms. This means that all varieties of a given quality are produced in a single location, and we can identify production locations using the profits expression.

Within skill intensities, demand determines where varieties are produced. When two qualities have the same skill intensity, $i(q) = i(q')$, the location that minimizes the cost of producing a variety of quality q also minimizes the cost of a variety of quality q', $c(q,k) < c(q,k') \iff c(q',k) < c(q',k')$. Thus, if varieties of the same skill intensity are produced in different locations, these differences must be due to differences in market access, $M_{q,k}(\tau)$. In particular, if $c(q,k) \neq c(q,k')$, then firms produce in the higher-cost location because its market-access advantage outweighs its cost disadvantage.

Proposition 1.1 (Within-intensity market access)**.** *When trade costs are small, if* $n_{q,k} > 0$*,* $n_{q',k'} > 0$*, and* $i(q) = i(q')$*, then* $M_{q,k} \geq M_{q,k'}$ *or* $M_{q',k} \leq M_{q',k'}$*.*

Proof. Suppose not. That is, suppose $M_{q,k} < M_{q,k'}$ and $M_{q',k} > M_{q',k'}$. If $c(q,k) \geq c(q,k')$, then by approximation (1.5) $\pi_{q,k'} > \pi_{q,k}$, which contradicts $n_{q,k} > 0$ by the free-entry condition. Similarly, if $c(q,k) \leq c(q,k')$, then by $\pi_{q',k} > \pi_{q',k'}$, which contradicts $n_{q',k'} > 0$. Hence $M_{q,k} \geq M_{q,k'}$ or $M_{q',k} \leq M_{q',k'}$. □

Proposition 1.1 establishes that market access alone governs specialization within qualities of the same skill intensity. An important component of $M_{q,k}(\tau)$ is demand in the location of production, $N_k\Gamma_k(q, \bar{\mathbf{n}}, \bar{\mathbf{c}}, \mathbf{0})$.[37] This is the home-market effect explanation for why high-income locations specialize in high-quality products.

Thus, we have an empirical strategy for distinguishing the two mechanisms. When quality is skill-intensive and trade is costly, both the factor-abundance mechanism and the home-market effect cause high-k locations to specialize in producing high-q varieties. Variation across skill intensities is overdetermined with respect to the two mechanisms. Variation within skill intensities is driven by market access alone. We can therefore identify a lower bound on the home-market effect by examining the pattern of specialization conditional on skill intensities.[38] My empirical strategy is to relate the pattern of specialization across locations to variation in market access after controlling for plants' factor usage.

1.3.5 Taking the theory to plant-level data

The predictions above describe relationships between product quality (q), location (k), skill intensity (i), and market access ($M_{q,k}(\tau)$). These objects can be inferred from observables in the data using

[37] In a many-country world, the "home-market effect" involves an appropriately defined market area, not merely the "home country," as noted at least since Linder (1961, p. 87). When trade costs are uniform, as in Fajgelbaum, Grossman, and Helpman (2011), differences in $M_{q,k}(\tau)$ are due solely to differences in demand in the location of production, $M_{q,k}(\tau) > M_{q,k'}(\tau) \iff N_k\Gamma_k(q, \bar{\mathbf{n}}, \bar{\mathbf{c}}, \mathbf{0}) > N_{k'}\Gamma_{k'}(q, \bar{\mathbf{n}}, \bar{\mathbf{c}}, \mathbf{0})$.

[38] The strategy of using variation in demand within a set of goods of the same factor intensity is similar to the approach used by Davis and Weinstein (2003) to integrate factor-abundance and home-market-effect models. We differ when we go to the data. Whereas Davis and Weinstein (2003) assume that factor intensities are fixed within 3-digit ISIC industries, I use plant-level information to infer factor intensities.

the model and some auxiliary assumptions. The results summarized here are derived in appendix section A.1. Denote the plant index j, so that, for example, plant j's skill intensity is $i(j)$.

I infer product quality from shipments' prices. Assume that the schedule of technological coefficients $b(\omega, q)$ is strictly decreasing in q, so that higher-quality varieties are more costly to produce. Since free-on-board prices are $p(q,k) = c(q,k) + \frac{\theta_q}{q}$, if $c(q,k)$ increases in quality faster than θ_q/q declines in quality, the price of a variety is informative about its quality. I validate this approach in appendix section A.5.1 by using estimated demand shifters to infer product quality. Prices and shifters are strongly positively correlated in my data.

I infer locations' rankings from their per capita incomes, denoted $\bar{y}_k$. Under the assumption that $g(y,k)$ is log-supermodular, average income is a sufficient statistic for k.

I infer skill intensities from the composition and wages of plants' workers. The composition measure assumes that non-production workers are more skilled than production workers. Denote the share of non-production workers employed in a plant with skill intensity i in location k by $share_N(i,k)$. When factor prices equalize, the share of non-production workers reveals a plant's skill intensity, $share_N(i,k) > share_N(i',k') \iff i > i'$, so $share_N(j)$ is a sufficient establishment-level control for skill intensity. When labor is cheaper where it is abundant, plants of all intensities use more skilled workers in skill-abundant locations, and $share_N(i,k)$ is increasing in k. I therefore also use $share_N(j) \times \ln \bar{y}_k$ to control for skill intensity.

The wage measures assume that wages are increasing in skill.[39] If wages are increasing in skill, we can infer the skill intensity i of a producer in location k from its average wage, $\bar{w}(i,k)$, average non-production wage, $\bar{w}_N(i,k)$, or average production wage, $\bar{w}_P(i,k)$. These average wages are all increasing in i. When factor prices equalize, ranking plants by their average wages is equivalent to ranking them by their factor intensities, $\bar{w}_j > \bar{w}_{j'} \iff i(j) > i(j')$. When labor is cheaper where it is abundant, $\bar{w}(i,k)$ is increasing in i, increasing in k, and log-supermodular. I therefore use $\ln \bar{w}_j$ and $\ln \bar{w}_j \times \ln \bar{y}_k$ as controls for skill intensity.

To assess the role of the geography of demand in specialization across US cities, I construct empirical counterparts to the model's market-access term $M_{q,k}(\tau)$ in profit expression (1.5). My

[39] A sufficient condition is $\frac{\partial \ln w(\omega,k)}{\partial \omega} = \frac{\partial \ln b(\omega,q)}{\partial \omega} - \frac{1}{\sigma}\frac{\partial \ln f(\omega,k)}{\partial \omega} > 0 \; \forall q \forall k$. Informally, more skilled individuals have greater absolute advantage than local abundance.

market-access measures are weighted averages of potential customers' per capita incomes, in which the weights reflect potential customers' population size and distance from the location of production.[40] Describing the composition of demand using per capita incomes exploits the fact that this is a sufficient statistic for relative demand for qualities under the model's assumptions. Weighting these incomes by population size and distance reflects the fact that it is more profitable to produce in locations that are more proximate to a larger number of consumers due to distance-related trade costs.

I construct two such market-access measures. Denote log income per capita in destination city d in year t by $\ln \bar{y}_{dt}$, population size by N_{dt}, and the mileage distance between origin o and destination d by $miles_{od}$. The first measure describes the composition of potential customers not residing in the location of production, so it cannot be contaminated by any supply-side mechanism linked to per capita income in the production location. This measure for a plant producing in origin city o is $M^1_{ot} = \sum_{d \neq o} \frac{N_{dt} miles_{od}^{-\eta}}{\sum_{d' \neq o} N_{d't} miles_{od'}^{-\eta}} \ln \bar{y}_{dt}$. I use it to qualitatively establish the relationship between market access and specialization. The second market-access measure includes all potential customers, consistent with the theoretical model of the home-market effect. It is $M^2_{ot} = \sum_{d} \frac{N_{dt} miles_{od}^{-\eta}}{\sum_{d'} N_{d't} miles_{od'}^{-\eta}} \ln \bar{y}_{dt}$. I use this measure to quantify the role of market access in the pattern of within-product specialization. In constructing each measure, I use $\eta = 1$, consistent with the international trade literature and the values estimated in appendix section A.3, and the values of N_{dt}, $miles_{od}$, and $\bar{y}_{dt}$ in the data.

I now turn to the data to characterize the empirical relationships linking product qualities, skill intensities, and market access following the model's guidance.

1.4 Data and empirical setting

This section introduces the empirical setting in which I conduct my investigation. First, I describe the data that I use to characterize the pattern of specialization and exchange between US cities. Additional details are in appendix section A.2. Second, I document that both outgoing shipments

[40] This market-access measure is the "home-market effect for quality" analogue to the measure constructed by Redding and Venables (2004) to test the traditional Krugman (1980) home-market effect.

and incoming shipments within fine product categories exhibit higher prices in higher-income cities. Thus, this empirical setting is suitable for testing theories of quality specialization.

1.4.1 Data

I combine microdata on US manufacturing plants' production and shipments with data describing the characteristics of cities and sectors.

1.4.1.1 Manufacturing microdata

The two plant-level data sources used in this study are the Commodity Flow Survey and the Census of Manufactures. These sources are components of the quinquennial Economic Census; I use confidential microdata from the 1997, 2002, and 2007 editions.[41]

The Commodity Flow Survey (CFS) describes commodity shipments by business establishments in terms of their value, weight, destination ZIP code, transportation mode, and other characteristics.[42] Products are described using the Standard Classification of Transport Goods, a distinct scheme that at its highest level of detail, five digits, defines 512 product categories. Each quarter of the survey year, plants report a randomly selected sample of 20-40 of their shipments in one week.

The Census of Manufactures (CMF) describes a plant's location, industry, inputs and revenues. This census covers the universe of manufacturing plants, which are classified into 473 6-digit NAICS manufacturing industries.[43] The CMF describes establishments' employment of production and non-production workers, production worker hours, production and non-production wages and salaries, book values of equipment and structures, and cost of materials.

In most of the analysis, I define a product as the pairing of a 5-digit SCTG commodity code and a 6-digit NAICS industry code. This results in more narrowly defined products in cases

[41]The Commodity Flow Survey began in 1993. Atalay, Hortaçsu, and Syverson (2014) use 1993 and 1997 CFS microdata to study vertical integration, Hillberry and Hummels (2003, 2008) use 1997 CFS microdata to study how geographic frictions affect trade volumes, and Holmes and Stevens (2010, 2012) use 1997 CFS microdata to study plant size, geography, and trade. The Census of Manufactures has been used in numerous studies.

[42]The US Census Bureau defines an establishment as "a single physical location where business transactions take place or services are performed." The CFS covers manufacturing, mining, wholesale, and select retail and services establishments. This paper only analyzes shipments by manufacturing establishments, which I refer to as "plants."

[43]Information on small establishments is estimated from administrative records rather than reported by the establishment. I exclude these administrative records and imputed observations. See the data appendix A.2 for details.

in which the NAICS industry scheme is more detailed than the SCTG commodity scheme. For example, one product is "footwear" (SCTG 30400) produced by an establishment in "men's footwear (except athletic) manufacturing" (NAICS 316213), which is distinct from "footwear" produced by an establishment in "women's footwear (except athletic) manufacturing" (NAICS 316214). My results are robust to ignoring the NAICS information and using only the SCTG commodity codes to define products.

1.4.1.2 Geographic data

The empirical analysis describes core-based statistical areas (CBSAs), which are 366 metropolitan and 576 micropolitan statistical areas defined by the Office of Management and Budget.[44] I refer to these geographic units as cities. Appendix section A.2 describes how data using other geographies were assigned to CBSAs.

I calculate cities' per capita incomes using data on CBSAs' total populations and personal incomes from the Bureau of Economic Analysis's regional economic profiles for 1997, 2002, and 2007. In my baseline specification, I exclude the employees and income of all establishments in the same 6-digit NAICS industry as the shipping plant when calculating the population and per capita income of its CBSA.[45] Since most manufacturing sectors' workforces and payrolls are small relative to the total populations and incomes of the cities in which they are located, the results obtained without making this adjustment to the per capita income measure are very similar.[46]

1.4.2 Pattern of specialization and trade

This section describes variation in manufacturing shipment prices across US cities.[47] The patterns mirror those found in international trade data. First, outgoing shipments exhibit higher prices in

[44] More than 93% of the US population lived within a CBSA in 2007.

[45] There is therefore variation across plants within a CBSA in the regressors I call "log origin CBSA population" and "log origin CBSA per capita income."

[46] I also obtain similar results when excluding only a plant's own employees and payroll from the per-capita-income calculation.

[47] Recall that all observed "prices" are in fact unit values, the ratio of a shipment's value to its weight in pounds. See data appendix A.2 for details of the sample selection and variable construction.

higher-income cities. This pattern is consistent with quality specialization in which higher-income cities produce higher-price, higher-quality varieties. Second, incoming shipments exhibit higher prices in higher-income cities. This pattern is consistent with non-homothetic preferences in which higher-income consumers demand higher-price, higher-quality varieties.

One concern with inferring qualities from prices is that products may be horizontally differentiated. With horizontal differentiation, two varieties of the same quality can sell at different prices in the same destination, with the high-price variety simply obtaining a smaller market share (Khandelwal, 2010). This raises the concern that high-income locations' specialization in high-price products may only reflect higher costs. However, this objection is unlikely to be problematic for the empirical investigation here.

Unit values are likely to be informative about product quality in this context for three reasons. First, investigations of international trade data distinguishing between raw unit values and quality-adjusted prices have shown unit values to be a meaningful, though imperfect, proxy for quality (Khandelwal, 2010; Feenstra and Romalis, 2012). I obtain similar results in section A.5.1, where I find that estimated demand shifters are positively correlated with unit values. Moreover, these estimated demand shifters exhibit patterns of specialization and factor usage consistent with those found for unit values. Second, my empirical setting allows me to check whether differences in prices across locations only reflect higher costs. Using plant-level data on wages and workers, I can test whether plants shipping from high-income locations charge higher prices only because they have higher labor costs. They don't. Third, consistent with the international evidence presented by Hallak (2006), I find a positive relationship between shipment prices and destinations' per capita income, suggesting that higher-price products are those preferred by higher-income consumers.[48]

The first feature of the US data matching international findings is that outgoing shipments' prices are systematically higher when originating from higher-income cities. To characterize how shipment prices vary with origin characteristics, I estimate linear regressions describing a shipment s of product k by plant j from origin city o to destination city d by transport mode m in year t of

[48] A potential concern is that higher-income consumers pay higher prices for identical products because higher-income consumers are less responsive to price changes (Simonovska, 2013). This would be a concern if the observed price variation were primarily within-plant. Table 1.2 below shows that this is not the case.

the form

$$\ln p_{skjodmt} = \beta_1 \ln \bar{y}_{ot} + \beta_2 \ln N_{ot} + \alpha_0 \ln miles_{skjodmt} + \gamma_{mt} + \gamma_{kdt} + \epsilon_{skjodt},$$

where $p_{skjodmt}$ is the shipment's unit value, $\bar{y}_{ot}$ and N_{ot} are per capita income and total population in the origin CBSA, $miles_{skjodt}$ is the ZIP-to-ZIP mode-specific mileage distance of the shipment, γ_{mt} are mode-year fixed effects, and γ_{kdt} are product-destination-year fixed effects.

Table 1.1: Outgoing shipment prices

Dep var: Log unit value, $\ln p_{skjodmt}$	(1)	(2)	(3)
Log origin CBSA income per capita, $\ln \bar{y}_{ot}$	0.458**	0.448**	0.486**
	(0.0427)	(0.0609)	(0.0607)
Log origin CBSA population, $\ln N_{ot}$	-0.0100*	-0.00519	-0.00737
	(0.00479)	(0.00659)	(0.00657)
Log mileage, $\ln miles_{skjodt}$	0.0400**	0.0521**	0.0496**
	(0.00316)	(0.00406)	(0.00418)
Log orig inc × *differentiation*		0.158**	0.216*
		(0.0555)	(0.0892)
Log orig pop × *differentiation*		-0.00770	-0.0223*
		(0.00575)	(0.0104)
Log mileage × *differentiation*		-0.0109*	-0.0144*
		(0.00446)	(0.00696)
Differentiation measure		Sutton	Khandelwal
R-squared	0.878	0.879	0.879
Observations (rounded)	1,400,000	600,000	600,000
Estab-year (rounded)	30,000	15,000	15,000
Ind-prod-year (rounded)	2,000	1,000	1,000

Standard errors, clustered by CBSA × year, in parentheses

** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include mode × year fixed effects and SCTG5×NAICS6×destination CBSA×year fixed effects.

Table 1.1 characterizes how variation in outgoing shipments' unit values relates to origin characteristics. The first column reports a large, positive origin-income elasticity of shipment prices of 46%. Higher-income cities specialize in the production of higher-price varieties of products, and this pattern of specialization is quite strong. Conditional on the level of per capita income, there is no economically meaningful correlation between origin population size and outgoing shipments' prices.

I proceed to interact the regressors with two measures of the scope for product differentiation. The Sutton (1998) measure, industrial R&D and advertising intensity, infers the scope for quality differentiation from the cost shares of differentiation-related activities. The Khandelwal (2010) measure infers the scope for quality differentiation from the range of estimated demand shifters in US imports. The second and third columns of Table 1.1 show that the positive relationship between origin income per capita and outgoing shipment prices is stronger in products with greater scope for quality differentiation, as classified by both measures. These patterns are consistent with local quality specialization in which higher-income cities specialize in higher-quality products. In products in which there is greater scope for quality differentiation, income differences generate greater differences in the composition of output.

The second feature of the US data matching international findings is that incoming shipments' prices are systematically higher when destined for higher-income cities. To characterize how shipment prices vary with destination characteristics, I estimate linear regressions describing a shipment s of product k by plant j from origin city o to destination city d by transport mode m in year t of the form

$$\ln p_{skjodmt} = \alpha_1 \ln \bar{y}_{dt} + \alpha_2 \ln N_{dt} + \alpha_0 \ln miles_{skjodt} + \gamma_{kt} + \gamma_{mt} + \theta_{ot} + \theta_{kjt} + \epsilon_{skjodt},$$

where $p_{skjodmt}$ is the shipment's unit value, $miles_{skjodt}$ is the ZIP-to-ZIP mileage distance of the shipment, and $\bar{y}_{dt}$ and N_{dt} are per capita income and total population in the destination CBSA. γ_{kt} and γ_{mt} are product-year and mode-year fixed effects that are included in all specifications. The θ fixed effects, which are mutually exclusive and omitted from some specifications, are origin-year and product-plant-year fixed effects.

Table 1.2 reports regressions characterizing how variation in shipment unit values within products relates to destination characteristics. The first column shows that the per-capita-income elasticity of incoming shipment prices is 26%. Higher-income cities import higher-price varieties, which suggests that preferences are non-homothetic. This pattern is attributable to city income

Table 1.2: Incoming shipment prices

Dep var: Log unit value, $\ln p_{skjodmt}$	(1)	(2)	(3)
Log destination CBSA income per capita, $\ln \bar{y}_{dt}$	0.256**	0.165**	0.0481**
	(0.0247)	(0.0191)	(0.00715)
Log destination CBSA population, $\ln N_{dt}$	-0.00450	-0.00336	0.00123
	(0.00266)	(0.00201)	(0.000831)
Log mileage, $\ln miles_{skjodt}$	0.0437**	0.0466**	0.0141**
	(0.00344)	(0.00221)	(0.00103)
R-squared	0.818	0.830	0.916
SCTG5 × NAICS6 × Year FE	Yes	Yes	
Origin CBSA × Year FE		Yes	
Establishment × SCTG5 × Year FE			Yes
Observations (rounded)	1,400,000		
Estab-year (rounded)	30,000		
Ind-prod-year (rounded)	2,000		

Standard errors, clustered by destination CBSA×year, in parentheses
** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include mode × year fixed effects.

composition, not city size per se, as the coefficient on log population reveals. The distance elasticity of incoming shipment prices is about 4%; longer shipments exhibit higher prices.[49]

The second and third columns show that the large majority of the correlation between income per capita and incoming shipment prices is attributable to cities of different income levels purchasing goods from different cities and plants. The second column introduces fixed effects for cities of origin, θ_{ot}. The destination per capita income elasticity falls by about 10 percentage points, indicating that about 40% of this variation is attributable to the composition of cities trading with each other.[50] The coefficients on the other regressors are similar to those in the first column. The third column introduces fixed effects for each plant-product, θ_{kjt}. The within-plant destination-income elasticity of shipment prices is considerably lower, 4.8%. Selling the same product at a higher price

[49]There are at least four possible explanations for the positive correlation between shipment distances and free-on-board prices. First, shipping costs may shift relative demand toward higher-price varieties through the Alchian-Allen effect. Second, non-shipping costs of distance may be included in the fob price. Third, survey respondents may fail to exclude shipping costs from their reported prices. Fourth, plants may charge higher mark-ups when serving more distant/remote locations. The third column of Table 1.2 suggests that this last hypothesis could explain at most one-third of such variation, since the within-establishment mileage elasticity is 1.4%.

[50]Appendix section A.3 shows that cities with more similar income levels trade more intensely with each other.

therefore accounts for at most one-fifth of price variation across destinations of different income levels. This decomposition suggests that changes in markups are not responsible for the majority of the observed correlation between shipment prices and destination incomes. Similarly, the small within-plant-product distance elasticity is evidence against the mark-up explanation for the positive correlation between bilateral distances and free-on-board prices.

These findings demonstrate that the composition of cities' manufactures demand is strongly linked to their income levels. This is consistent with numerous previous empirical studies of both households and countries. Such non-homothetic preferences are necessary for the "home-market effect for quality" hypothesis.[51]

Together, Tables 1.1 and 1.2 demonstrate patterns of specialization that are strongly linked to cities' income levels. Within narrowly defined product categories, higher-income locations both export and import higher-price products than lower-income locations. In addition, Appendix A.3 shows that cities with more similar incomes trade more intensely with each other. These findings mirror those found in international trade data and could be generated by the factor-abundance mechanism or the home-market effect. I now use data on plants' factor inputs to empirically distinguish between these potential explanations.

1.5 Empirical results

This section reports two bodies of empirical evidence. First, observed factor-usage differences explain only about one quarter of the relationship between cities' incomes and the prices of outgoing shipments. This bounds the explanatory power of the factor-abundance mechanism. Second, the market-access measures describing the income composition of proximate potential customers are strongly linked to outgoing shipment prices. The estimated home-market effect explains more than half of the price-income relationship.

My regression results can be understood as decomposing the covariance between plant j's outgoing shipment price and origin o's per capita income, $cov(\ln p_{jo}, \ln \bar{y}_o)$. Omitting notation indicating

[51] Non-homothetic preferences alone are not sufficient to produce the home-market effect, as discussed in section 1.2 and shown in section 1.3.4.2. The home-market effect for quality stems from non-homothetic preferences, economies of scale, and trade costs.

that all these moments are conditional on shipment mileage, origin population size, and destination-product-year fixed effects, we can decompose this covariance into two terms using the law of total covariance:

$$cov(\ln p_{jo}, \ln \bar{y}_o) = \underbrace{cov\left[\mathbb{E}(\ln p_{jo}|i(j)), \mathbb{E}(\ln \bar{y}_o|i(j))\right]}_{\text{between-intensity variation}} + \underbrace{\mathbb{E}_i\left[cov(\ln p_{jo}, \ln \bar{y}_o|i(j))\right]}_{\text{within-intensity variation}}$$

Higher-income origins have higher outgoing shipment prices to the extent that (1) skill-intensive products have higher prices and skill-intensive products are produced in higher-income locations and (2) products of the same skill intensity have higher prices in higher-income locations. The factor-abundance mechanism operates exclusively through the first component, variation across skill intensities. The home-market effect may appear in both components, since higher-income locations have greater demand for higher-quality varieties, regardless of qualities' skill intensities.

Section 1.5.1 shows that the across-skill-intensities component is modest, constituting 27% of the total variation. To introduce a market-access measure M_o, we can further decompose the within-skill-intensity variation into two terms, yielding

$$\begin{aligned} cov(\ln p_{jo}, \ln \bar{y}_o) = & \underbrace{cov\left[\mathbb{E}(\ln p_{jo}|i(j)), \mathbb{E}(\ln \bar{y}_o|i(j))\right]}_{\text{factor usage } = 27\%} \\ & + \underbrace{\mathbb{E}_i\left[cov(\mathbb{E}(\ln p_{jo}|i(j), M_o), \mathbb{E}(\ln \bar{y}_o|i(j), M_o))\right]}_{\text{within-skill-intensity market access } = 54\%} + \underbrace{\mathbb{E}_i\left[\mathbb{E}(cov(\ln p_{jo}, \ln \bar{y}_o|i(j), M_o))\right]}_{\text{residual covariance } = 19\%}. \end{aligned}$$

The first new term is the share of within-skill-intensity variation attributable to higher-price products being produced where proximate potential customers' per capita incomes are higher and higher-income locations being proximate to higher-income potential customers. The second new term is the share of the covariance explained by neither differences in skill intensity nor the market-access measure. Note that this decomposition is conservative with respect to the home-market effect because it restricts attention to within-skill-intensity variation. In section 1.5.2, I find that within-skill-intensity variation in market access accounts for 54% of the observed price-income relationship, while the residual constitutes 19%. Hence, home-market demand explains much of the within-product price-income covariance, and its influence is at least twice as large as observed factor-usage differences.

1.5.1 The factor-abundance hypothesis

This section identifies the share of within-product specialization attributable to differences in observable plant-level factor usage. The canonical factor-abundance theory posits that differences in locations' outputs are explained by differences in the factors employed by their producers. Within groups of products of the same factor intensity, the location of production is indeterminate. That is, under the null hypothesis that differences in factor supplies are the only source of comparative advantage, there should be no correlation between locational characteristics and plants' outputs after controlling for plant-level factor usage. In fact, there is a very strong relationship between income per capita and outgoing shipments prices after controlling for factor inputs. Observed factor usage explains only 27% of the observed covariance between cities' per capita incomes and outgoing shipment prices.

To characterize how shipment prices vary with origin characteristics, I estimate linear regressions describing a shipment s of product k by plant j from origin city o to destination city d by transport mode m in year t of the form

$$\begin{aligned}\ln p_{skjodmt} =& \beta_1 \ln \bar{y}_{ot} + \beta_2 \ln N_{ot} + \alpha_0 \ln miles_{skjodt} + \gamma_{kt} + \gamma_{mt} + \gamma_{kdt} \\ &+ \alpha_1 \ln share_{Njt} + \alpha_2 \ln \frac{K_{jt}}{L_{jt}} + \delta_1 \ln share_{Njt} \ln \bar{y}_{ot} + \delta_2 \ln \frac{K_{jt}}{L_{jt}} \ln \bar{y}_{ot} \\ &+ \alpha_3 \ln \bar{w}_{jt} + \alpha_4 \ln \bar{w}_{Njt} + \alpha_5 \ln \bar{w}_{Pjt} \\ &+ \delta_3 \ln \bar{w}_{jt} \ln \bar{y}_{ot} + \delta_4 \ln \bar{w}_{Njt} \ln \bar{y}_{ot} + \delta_5 \ln \bar{w}_{Pjt} \ln \bar{y}_{ot} + \epsilon_{skjodt}\end{aligned} \tag{1.6}$$

where $share_{Njt}$ is the ratio of the plant's non-production workers to total employees, $\frac{K_{jt}}{L_{jt}}$ is gross fixed assets per worker, $\bar{w}_{jt}$ is average pay per employee, $\bar{w}_{Njt}$ is average pay per non-production worker, and $\bar{w}_{Pjt}$ is average pay per production worker.[52] The interactions of plant-level factor-

[52]The theoretical model emphasized differences in the composition of skill across locations. I also include gross fixed assets per worker as a measure of capital intensity, since this variable has been emphasized in prior empirical work both across countries (Schott, 2004) and across plants (Verhoogen, 2008). Since I cannot construct capital stocks using the perpetual-inventory method with quinquennial data, I use the book value of assets as my measure of plant capital.

usage measures with origin income per capita address the case in which factor prices do not equalize, as described in section 1.3.5.[53] In theory, either $\ln share_{Njt}$ and its interaction with $\ln \bar{y}_{ot}$ or $\ln \bar{w}_{jt}$ and its interaction with $\ln \bar{y}_{ot}$ would be sufficient to characterize plants' skill intensities. In practice, I include a battery of plant-level factor-usage controls to maximize the potential explanatory power of observed factor-usage differences.

Table 1.3 characterizes how variation in shipment unit values relates to origin characteristics and plant-level observables. The first column relates outgoing shipment unit values to origin characteristics controlling for destination fixed effects, as in Table 1.1. The next two columns incorporate the plant-level measures of factor usage and their interactions with income per capita. The second column introduces quantity measures of capital intensity ($\frac{K_{jt}}{L_{jt}}$) and labor usage, the non-production employment share ($share_{Njt}$). The third column adds the three wage measures and therefore corresponds to the regression specified in equation (1.6).

These measures of factor usage are informative predictors of a plant's shipment prices, but they explain only a modest share of the observed origin-income elasticity of outgoing shipment prices. Consistent with the premise that higher-price, higher-quality varieties are more skill-intensive, the coefficients on log non-production worker share, log pay per worker, and log pay per non-production worker are positive and economically large. The negative coefficient on log assets per worker is inconsistent with a model in which higher-price, higher-quality varieties are more capital-intensive.[54] The positive coefficient on the interaction of log pay per production worker and log origin per capita income is consistent with the model-predicted behavior when factor prices are not fully equalized. However, the observed variation in factor usage explains only a small share of the cross-city variation in outgoing shipment prices. Introducing the quantity measures in the second column reduces the origin-income elasticity from 46% to 42%. Incorporating the wage measures in the third column reduces this elasticity to 37%. Thus, these observed factor-usage differences can explain about one-fifth of the origin-income elasticity of shipment prices.

The fourth through sixth columns of Table 1.3 incorporate the control variables in more flexible

[53] Bernard, Redding, and Schott (2013) find that relative factor prices do not equalize within the US when considering two factors, production and non-production workers.

[54] Using very aggregate data, Torstensson (1996) obtains a negative partial correlation between prices and capital per worker when distinguishing between human and physical capital.

Table 1.3: Outgoing shipment prices w/ plant-level factor usage

Dep var: Log unit value, $\ln p_{skjodmt}$	(1)	(2)	(3)	(4)	(5)	(6)
Log origin CBSA income per capita	0.458**	0.421**	0.370**	0.430**	0.381**	0.323**
	(0.0427)	(0.0415)	(0.0397)	(0.0411)	(0.0396)	(0.0381)
Log origin CBSA population	-0.0100*	-0.0139**	-0.0172**	-0.00881	-0.0132**	-0.0161**
	(0.00479)	(0.00464)	(0.00448)	(0.00451)	(0.00437)	(0.00414)
Log mileage	0.0400**	0.0416**	0.0410**	$\surd$	$\surd$	$\surd$
	(0.00316)	(0.00310)	(0.00302)			
Log non-production worker share		0.139**	0.128**		$\surd$	$\surd$
		(0.00797)	(0.00937)			
Log assets per worker		-0.0385**	-0.0552**		$\surd$	$\surd$
		(0.00446)	(0.00470)			
Log non-production worker share		0.0672	0.0398		0.00833	-0.00793
× log per capita income		(0.0355)	(0.0397)		(0.0341)	(0.0390)
Log assets per worker		-0.0194	-0.0579**		0.00127	-0.0246
× log per capita income		(0.0180)	(0.0197)		(0.0195)	(0.0212)
Log pay per worker			0.202**			$\surd$
			(0.0420)			
Log pay per production worker			-0.0258			$\surd$
			(0.0330)			
Log pay per non-production worker			0.0489**			$\surd$
			(0.0146)			
Log pay per worker			0.0106			-0.0887
× log per capita income			(0.120)			(0.122)
Log pay per production worker			0.351**			0.332**
× log per capita income			(0.0930)			(0.0877)
Log pay per non-production worker			0.0915			0.103
× log per capita income			(0.0495)			(0.0555)
R-squared	0.878	0.879	0.879	0.878	0.881	0.882
Observations (rounded)			1,400,000			
Estab-year (rounded)			30,000			
Ind-prod-year (rounded)			2,000			

Standard errors, clustered by CBSA × year, in parentheses
** p<0.01, * p<0.05

Notes: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include SCTG5×NAICS6×destination×year fixed effects and mode × year fixed effects. The fourth through sixth columns include 3-digit-NAICS-specific cubic polynomials in log mileage (4,5,6), log non-production worker share (5,6), log assets per worker (5,6), log pay per worker (6), log pay per production worker (6), and log pay per non-production worker (6).

functional forms. Each variable with an α coefficient in estimating equation (1.6) is entered as a cubic polynomial that varies by 3-digit NAICS industry. Since there are 21 3-digit industries, this introduces 63 regressors for each control variable, yielding a total of 378 regressors.[55] I refrain from reporting the coefficients on these controls.

The results obtained using these more flexible functional forms are quite similar to those in the first three columns of Table 1.3. The origin-income elasticity of 43% is reduced to 38% by the introduction of the quantity controls and further to 32% by the full battery of plant-level factor-usage measures. Thus, differences in plants' observed factor usage explain about one-quarter of the correlation between cities' incomes per capita and outgoing shipment prices. This suggests that the factor-abundance hypothesis has modest explanatory power for the pattern of within-product specialization across US cities.

These results are robust to introducing further information on the skills employed in these plants. I construct city-industry-level measures of employees' schooling from public-use microdata from the Census of Population and American Community Survey. These measures are available for a subset of the observations in the main estimation sample. The results are reported in Appendix Table A.2. The partial-correlation origin-income elasticity of 34% is quite similar to the 32% obtained in Table 1.3.

A potential alternative interpretation of these results would be that the plant-level variables are imperfect controls for plant-level factor usage and that city-level income per capita is informative about plant-level factor usage. Suppose that plants' factor inputs exhibit unobserved differences in quality that are correlated with city-level average income, conditional on plant-level observables. It is plausible that plants with observationally equivalent workforces in terms of non-production-to-production-worker ratios may differ in worker quality. In particular, prior research has documented weak but systematic sorting of workers across cities on unobservable characteristics correlated with higher wages (Davis and Dingel, 2012; De la Roca and Puga, 2012). However, these differences between workers should appear in the plant-level wage data, and the third and sixth columns of Table 1.3 includes plant-level wage measures. The posited unobserved differences in input factor

[55] Using a 3-digit-NAICS-specific translog approximation with the five input measures and a 3-digit-NAICS-specific quadratic in log mileage, for a total of 462 regressors, yields very similar results.

quality would therefore have to be characteristics of worker that raise output quality, are not priced into their wages, and are systematically correlated with city-level incomes, which seems an unlikely explanation for the findings.

Another potential concern is aggregation bias. Though my data describe hundreds of manufacturing product categories, these are less detailed than the most disaggregated product categories in international trade data. I address this concern using data from the Census of Manufactures product trailer, which describes comparable number of product categories and reports quantities for a subset of them. Appendix Table A.3 describes establishments' average unit values from Census of Manufactures data on products for which quantities are reported and reports results that are consistent with those reported in Table 1.3.[56] Though the origin-income elasticity is lower than that found in the CFS data, observed plant-level factor usage explains only about 12% of the total variation.

Finally, one may worry that some other dimension of plant heterogeneity has been omitted. Appendix Table A.4 reports results from CFS data while controlling for plant size. This yields a partial-correlation origin-income elasticity of 30%, similar to that in Table 1.3.

This section has shown that only a modest share of the observed within-product variation in outgoing shipment prices across cities of different income levels is attributable to observable differences in plants' factor usage. Under the null hypothesis that differences in factor abundance alone explain within-product specialization, the partial correlation between origin income per capita and outgoing shipment prices after controlling for plant-level factor usage should be zero. In the presence of a rich set of plant-level controls, the estimated coefficient $\hat{\beta}_1$ in column six of Table 1.3 is 32%, roughly 3/4 of its value in the absence of any plant-level controls. If we were to attribute the full decrease in the value of the coefficient on $\ln \bar{y}_{ot}$ to the factor-abundance mechanism, it would explain about one quarter of the observed variation.[57]

[56]These plant-level average unit values necessarily include shipments destined for the origin CBSA.

[57]To the degree that differences in skill intensities are causally induced by differences in demand, this overstates the explanatory power of the factor-abundance hypothesis.

1.5.2 The home-market effect for quality

This section identifies the share of the covariance between incomes and prices not explained by factor-usage differences that is attributable to home-market demand. I find that cities with greater market access to higher-income households produce higher-price manufactures. This within-intensity home-market effect alone explains twice as much of the covariance between incomes and prices as differences in plants' factor inputs.

The "home-market" effect in fact depends on the composition of demand in all locations potentially served from a location of production, as described in the model by market access $M_{q,k}(\tau)$. A city that is more proximate to another city with many high-income residents has higher relative demand for higher-quality manufactures, ceteris paribus.[58] Section 1.3.5 described two market-access measures. The first, $M^1_{ot} = \sum_{d \neq o} \frac{N_{dt} miles_{od}^{-\eta}}{\sum_{d' \neq o} N_{d't} miles_{od'}^{-\eta}} \ln \bar{y}_{dt}$, omits potential customers residing in the location of production. The identifying assumption when using this measure is that variation across locations in neighboring cities' incomes per capita, after conditioning on plants' inputs and income per capita in the city of production, is related to plants' outputs only through variation in the composition of demand. The second market-access measure, $M^2_{ot} = \sum_{d} \frac{N_{dt} miles_{od}^{-\eta}}{\sum_{d'} N_{d't} miles_{od'}^{-\eta}} \ln \bar{y}_{dt}$, includes all potential customers, consistent with the model. The accompanying identifying assumption is that, after conditioning on plants' inputs, variation across locations in potential consumers' incomes, including residents in the city of production, is related to plants' outputs only through variation in the composition of demand.[59]

Table 1.4 demonstrates that market access plays a significant role in explaining the origin-income elasticity of shipment prices. The first column reports a regression that flexibly controls for the quantity measures of factor usage, the non-production worker share and assets per worker, like the fifth column of Table 1.3. The second column adds the first market-access measure. This

[58] Fajgelbaum, Grossman, and Helpman (2011) derive their results in a setting in which the cost of exporting to another location is the same across all locations. Thus, in their model the home-market effect depends only on the difference in income composition between the location of production and the rest of the world. When trade costs are not uniform, the home-market effect depends on a production location's access to every other market, i.e. the matrix of bilateral trade costs. Measuring multilateral market access has received considerable attention in empirical assessments of the new economic geography, e.g. Redding and Venables (2004). See Lugovskyy and Skiba (2012) for a discussion of market access in the context of quality specialization.

[59] This identifying assumption would be violated by unobserved quality-improving inputs or technologies that were correlated with city-level income per capita but not correlated with my plant-level measures of inputs.

reduces the origin-income elasticity from 38% to 24%, which is a large change. Controlling for factor quantities only reduced the elasticity from 43% to 38%. This implies that the income composition of proximate potential customers *other than those in the city of production* is more than twice as quantitatively important as the capital intensity and the non-production employee share in explaining the pattern of shipment prices. In locations with better access to high-income customers, plants produce higher-price products. This evidence suggests that the geography of demand influences the pattern of within-product specialization. The third column uses the second market-access measure, which includes the income of potential customers in the city of production in the weighted average. This reduces the origin-income elasticity to 13.6%, a reduction of nearly 25 percentage points compared to column 1.

Table 1.4 reports similar results when controlling for factor usage using the wage measures. The fourth column is a regression that flexibly controls for the quantity and wage measures of factor usage, like the sixth column of Table 1.3. The fifth column introduces the first market-access measure, reducing the origin-income elasticity from 32.3% to 19.5%. The factor-usage measures reduced this elasticity by 10 percentage points, so the income composition of proximate potential customers other than those in the city of production has considerable explanatory power. The sixth column uses the second market-access measure that includes residents in the city of production. This reduces the origin-income elasticity to 12%, a reduction of more than 20 percentage points compared to column 4. Thus, market access explains substantially more of the observed relationship between income per capita and outgoing shipment prices than differences in plants' factor usage.[60]

These results can be succinctly summarized as a decomposition of the covariance between incomes and prices.[61] After controlling for population size and shipment mileage, differences in observed factor usage are responsible for 27% of the covariance between outgoing shipment prices and origin income per capita. Conditional on factor usage, the first market-access measure, which

[60] If the demand-side mechanism induces differences in the factor intensities used to produce different qualities, then market access *causally* explains an even greater fraction of the observed correlation. I am conservative in controlling for observed factor usage prior to attributing the change in coefficients to the introduction of the market-access measure.

[61] Comparing the estimated elasticities yields shares very similar to this decomposition. However, that back-of-the-envelope calculation does not account for the changing set of regressors. The law of total covariance is an identity, so the numbers reported in this paragraph are an exact decomposition.

Table 1.4: Outgoing shipments and market access

Dep var: Log unit value	(1)	(2)	(3)	(4)	(5)	(6)
Log origin CBSA income per capita	0.381**	0.240**	0.136**	0.323**	0.195**	0.120*
	(0.0396)	(0.0412)	(0.0495)	(0.0381)	(0.0396)	(0.0481)
Log origin CBSA population	-0.0132**	-0.00233	-0.00814	-0.0161**	-0.00591	-0.0118**
	(0.00437)	(0.00435)	(0.00425)	(0.00414)	(0.00412)	(0.00410)
Market access (excl orig) M^1_{ot}		1.144**			1.075**	
		(0.129)			(0.125)	
Market access M^2_{ot}			1.040**			0.899**
			(0.131)			(0.126)
R-squared	0.881	0.881	0.881	0.882	0.882	0.882
Observations (rounded)	1,400,000					
Estab-year (rounded)	30,000					
Ind-prod-year (rounded)	2,000					

Standard errors, clustered by CBSA × year, in parentheses
** p<0.01, * p<0.05

Notes: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include SCTG5×NAICS6×destination×year fixed effects and mode × year fixed effects. Unreported controls include 3-digit-NAICS-specific cubic polynomials in log mileage (1-6), log non-production worker share (1-6), log assets per worker (1-6), log pay per worker (4-6), log pay per production worker (4-6), and log pay per non-production worker (4-6). Also unreported are the interactions of log origin income per capita with the five input variables.

omits residents in the city of production, accounts for 34% of the total covariance, leaving 39% as residual variation. The second market-access measure, which follows the model by including residents in the city of production, accounts for 54% of the total covariance, leaving 19% as residual variation.

This section has established the role of home-market demand in explaining the pattern of outgoing shipment prices. The income composition of proximate potential customers is strongly associated with outgoing shipment prices. Consistent with the model, plants located near higher-income potential customers sell products at higher average prices. The income composition of potential customers other than those in the location of production is quantitatively more important for explaining the origin-income elasticity of outgoing shipment prices than observed plant-level factor usage. When including individuals residing in the city of production, the income composition of potential customers explains at least half of the observed origin-income elasticity of shipment prices. This is consistent with a model in which the home-market effect plays a large role in determining the pattern of quality specialization.

1.5.3 Additional evidence

This section briefly summarizes a series of results in Appendix A.5 that are consistent with the results reported above.

Appendix section A.5.1 characterizes the pattern of quality specialization using estimated demand shifters instead of outgoing shipments' unit values as the dependent variable. Due to data constraints, these are only available in 2007. The empirical results are consistent with the unit-value findings for the influence of market access, though factor usage exhibits greater explanatory power. The origin-income elasticity of the plant-product estimated demand shifter is 41%, remarkably similar to the 43% origin-income elasticity of outgoing shipment prices. This covariance between income per capita and estimated demand shifter decomposes into factor-intensity differences (46%), within-intensity market-access differences (48%), and residual variation (7%). The greater explanatory power of plants' factor inputs primarily reflects less residual variation, not a dramatically weakened role for the income composition of proximate potential customers. Home-market demand plays a substantial role in quality specialization, at least as large as that explained by the factor-abundance mechanism.

Appendix section A.5.2 uses another moment of the income distribution to identify the role of demand in quality specialization. Conditional on average income, cities with higher variance in household income have higher incoming shipment prices. I then show that cities with greater income dispersion have higher outgoing shipment prices, and this is not due to greater dispersion in the wages or skills of workers employed at the plants shipping these products. This is consistent with the home-market effect under the Fajgelbaum, Grossman, and Helpman (2011) demand system in an equilibrium in which most individuals purchase low-quality varieties.

Appendix section A.5.3 shows that the patterns founds in domestic shipments are also found in export shipments destined for foreign markets. The origin-income elasticity of export prices is 42%. After controlling for plants' factor inputs, this elasticity is 30%. After controlling for both factor inputs and market access, this elasticity becomes negative and statistically indistinguishable from zero. Home-market demand explains a greater share of within-product variation in export prices than differences in factor usage.

1.6 Conclusions

Two prominent theories predict that high-income locations specialize in producing and exporting high-quality products. The Linder (1961) conjecture, formalized by Fajgelbaum, Grossman, and Helpman (2011), emphasizes the role of high-income customers' demand for high-quality products. The canonical factor-proportions theory focuses on the abundant supply of capital and skills in high-income locations. Prior empirical evidence does not separate the contributions of these mechanisms because each makes the same predictions about country-level trade flows.

In this paper, I combine microdata on manufacturing plants' shipments and inputs with data on locations' populations and incomes to quantify each mechanism's role in quality specialization across US cities. I develop a model that nests both mechanisms to guide my empirical investigation. The theory's basic insight is that the factor-abundance mechanism operates exclusively through plants' input usage. Conditional on plant-level factor intensity, demand determines quality specialization. I implement my empirical strategy using US microdata because the Commodity Flow Survey and Census of Manufactures describe plants located in many cities of varying income levels. In doing so, I document that US cities exhibit the same patterns found in international trade data that have been interpreted as evidence of quality specialization. My empirical investigation finds that home-market demand explains at least as much of the specialization across US cities as differences in plants' factor inputs.

This finding is significant because the two mechanisms have distinct implications for welfare, inequality, and trade policy. The large share of quality specialization attributable to market access suggests that a location's capacity to profitably produce high-quality products depends significantly on the income composition of neighboring locations. As a result, geography influences specialization in part because economic developments in neighboring locations may shift local demand for quality. To the degree that demand shapes entry and product availability, individuals may gain by living in locations where other residents' incomes are similar to theirs. Finally, since market access is affected by trade policy, governments may have scope to influence quality specialization.

Chapter 2

The Comparative Advantage of Cities

Donald R. Davis and Jonathan I. Dingel[1]

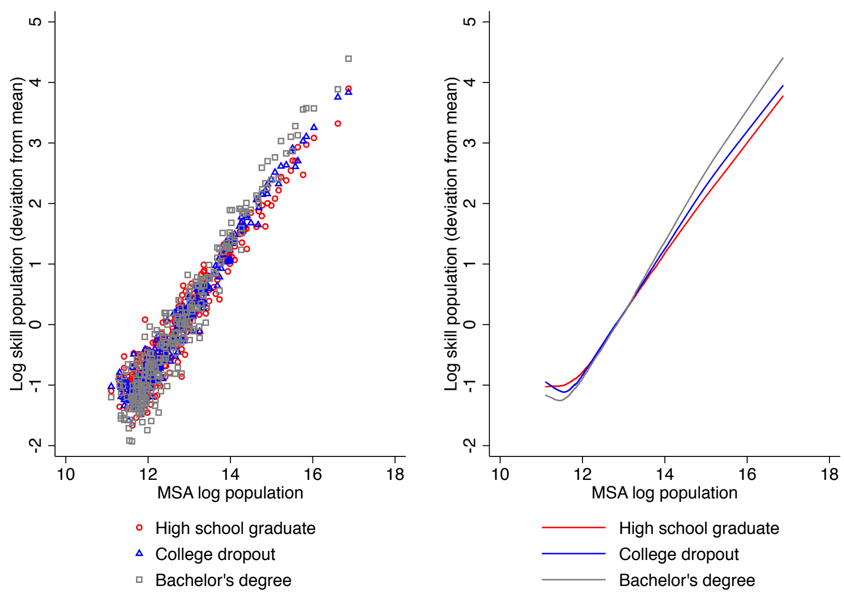
[1]We thank numerous people and seminar audiences for helpful comments and suggestions, especially Bernard Salanié, Bruno Strulovici, Daniel Sturm, and Jonathan Vogel. We thank Yuxiao Huang and especially Antonio Miscio for research assistance.

2.1 Introduction

The distributions of skills, occupations, and industries vary substantially and systematically across US cities. Figures 2.1 through 2.3 illustrate this with three selected examples for each.

- Figure 2.1 plots the population of three educational attainment categories against total metropolitan area population.[2] The left panel plots the data; the right panel plots a locally weighted regression for each category. While each educational category's population rises with metropolitan population, the relative levels also exhibit a systematic relationship with city size. Comparing elasticities, the population with a bachelor's degree rises with city size faster than the population of college dropouts, which in turn rises faster than the population of high-school graduates.

Figure 2.1: Populations of three educational groups across US metropolitan areas



Data source: 2000 Census of Population microdata via IPUMS-USA

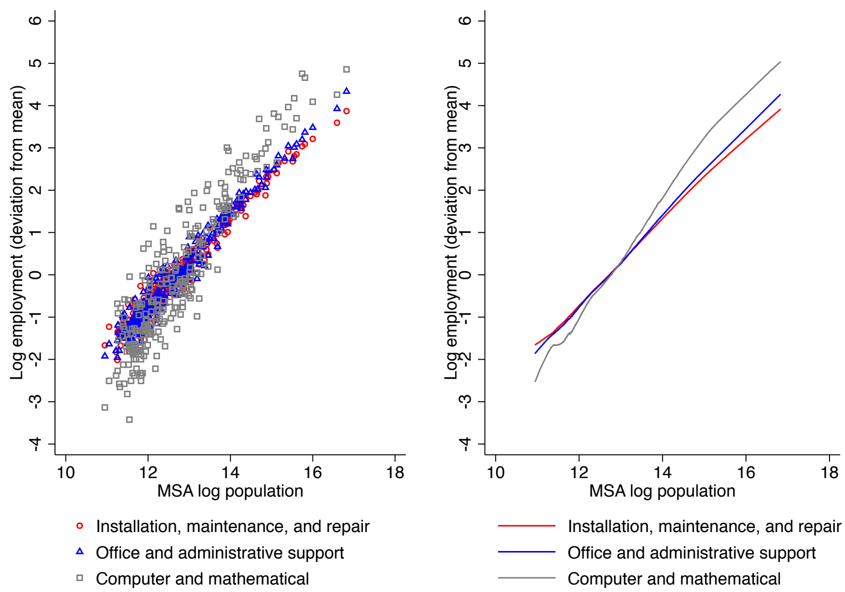
- Figure 2.2 plots metropolitan area employment in three occupational categories.[3] Computer

[2]We use the terms cities and metropolitan areas interchangeably, as is customary in the literature. These three educational groups comprise about 70 percent of the employed metropolitan population (see Table 2.1).

[3]The occupations are SOC 49, 43, and 15 in the 2000 Occupational Employment Statistics data.

and mathematical employment rises with city size faster than office and administrative employment, which in turn rises faster than installation, maintenance and repair employment. These sectors also differ in their employee characteristics. Nationally, the average individual in computer and mathematical occupations has about two more years of schooling than the average individual in office and administrative support and three more years than those in installation, maintenance, and repair.

Figure 2.2: Employment in three occupations across US metropolitan areas

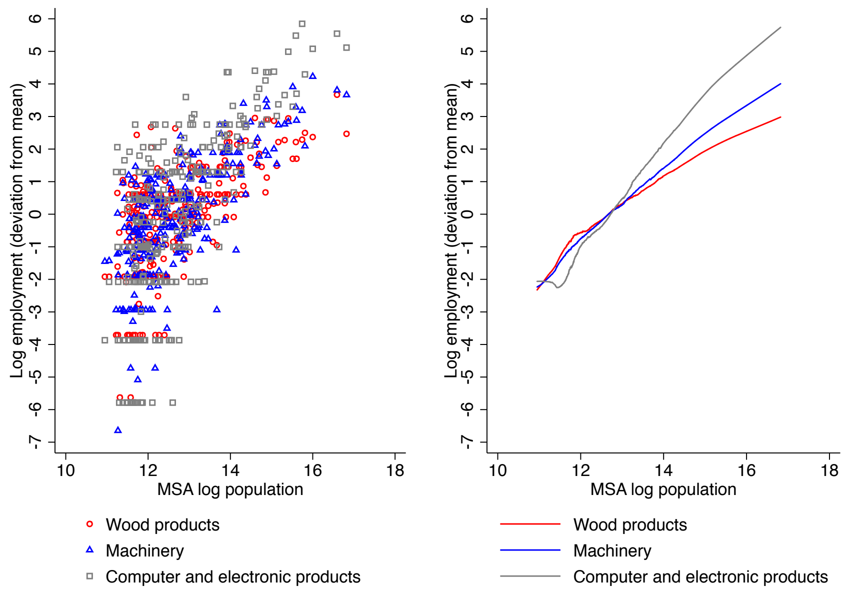


Data source: Occupational Employment Statistics 2000

- Figure 2.3 plots employment in three manufacturing industries.[4] Employment in computer and electronic products rises with city size faster than machinery, which in turn rises faster than wood products. On average, computer and electronic employees have about two more years of education than wood products employees.

[4]The industries are NAICS 321, 333, and 334 in the 2000 County Business Patterns data. Employment levels cluster at particular values due to censored observations. See appendix B.3 describing the data.

Figure 2.3: Employment in three manufacturing industries across US metropolitan areas



Data source: County Business Patterns 2000

Together, these three figures suggest that larger cities are skill-abundant and specialize in skill-intensive activities. Explaining these patterns involves fundamental questions about the spatial organization of economic activity. What determines the distribution of skills across cities? What determines the distribution of occupations and industries across cities? How are these two phenomena interrelated? In this paper, we develop a theory describing the comparative advantage of cities that predicts such a pattern of skills and sectors in a manner amenable to empirical investigation.

As we describe in section 2.2, prior theories describing cities' sectoral composition have overwhelmingly focused on the polar cases in which cities are either completely specialized "industry towns" or perfectly diversified hosts of all economic activities (Helsley and Strange, 2012). Yet Figures 2.2 and 2.3 make clear both that reality falls between these poles and that sectoral employment shares are systematically related to cities' sizes. In this paper, we integrate modern trade theory with urban economics by introducing a spatial-equilibrium model in which the comparative advantage of cities is jointly governed by the comparative advantage of individuals and their locational choices. Our theory both describes the intermediate case in which cities are incompletely specialized and relates the pattern of specialization to cities' observable characteristics. It makes strong, testable predictions about the distributions of skills and sectors across cities that we take to the data.

Section 2.3 introduces our model of a system of cities with heterogeneous internal geographies. Cities are *ex ante* homogeneous, so cross-city heterogeneity is an emergent outcome of the choices made by freely mobile individuals. Agglomeration economies make cities with larger, more skilled populations exhibit higher total factor productivity (TFP). Locations within cities are heterogeneous and more desirable locations are relatively scarce, as is customary in land-use models (Fujita and Thisse, 2002, Ch 3). These cities are populated by heterogeneous individuals with a continuum of skill types, and these individuals may be employed in a continuum of sectors. Comparative advantage causes more skilled individuals to work in more skill-intensive sectors, as in Sattinger (1975), Costinot (2009), and Costinot and Vogel (2010). There is a complementarity between individual income and locational attractiveness, so more skilled individuals are more willing to pay for more attractive locations and occupy these locations in equilibrium, as in the differential rents model of Sattinger (1979).

In equilibrium, agglomeration, individuals' comparative advantage, and heterogeneity across internal locations within cities combine to deliver a rich set of novel predictions. Agglomeration causes larger cities to have higher TFP, which makes a location within a larger city more attractive than a location of the same innate desirability within a smaller city. For example, the best location within a larger city is more attractive than the best location within a smaller city due to the difference in TFP. Since more skilled individuals occupy more attractive locations, larger cities are skill-abundant. The most skilled individuals in the population live only in the largest city and more skilled individuals are more prevalent in larger cities, consistent with the pattern shown in Figure 2.1. By individuals' comparative advantage, the most skill-intensive sectors are located exclusively in the largest cities and larger cities specialize in the production of skill-intensive output. More skill-intensive sectors exhibit higher population elasticities of sectoral employment, as suggested in Figures 2.2 and 2.3. Our model therefore predicts an urban hierarchy of skills and sectors.

We examine the model's predictions about the spatial distribution of skills and sectors across US cities using data from the 2000 Census of Population, County Business Patterns, and Occupational Employment Statistics described in section 2.4. We use two empirical approaches. The first involves regression estimates of the population elasticities of educational, occupational, and industrial populations akin to those shown in Figures 2.1 through 2.3. The second involves pairwise comparisons governed by the monotone likelihood ratio property, as per Costinot (2009).[5]

Section 2.5 reports the results, which provide support for our model's predictions about the spatial pattern of skills and sectors. Characterizing skills in terms of three or nine educational groups, we find that larger cities are skill-abundant. Among US-born individuals, cities' skill distributions typically exhibit the monotone likelihood ratio property.[6] Characterizing sectors in terms of 21 manufacturing industries or 22 occupational categories, we find that larger cities specialize in skill-intensive sectors. While sectors do not exhibit the monotone likelihood ratio property as reliably as skills, there is systematic variation in cities' sectoral distributions that is consistent with the novel predictions of our theory.

In short, when mobile individuals optimally choose locations and sectors, larger cities will have more skilled populations and thereby comparative advantage in skilled activities. These features are consistent with US data.

2.2 Related literature

Our contributions are related to a diverse body of prior work. Our focus on high-dimensional labor heterogeneity is related to recent developments in labor and urban economics. Our theoretical approach integrates elements from the systems-of-cities literature, land-use theory, and international trade. Our model yields estimating equations and pairwise inequalities describing the comparative advantage of cities that are related to prior reduced-form empirical work in urban economics, despite a contrast in theoretical underpinnings.

Our theory describes a continuum of heterogeneous individuals. A large share of systems-of-

[5]The distributions $f_c(\sigma)$ and $f_{c'}(\sigma)$ exhibit the monotone likelihood ratio property if, for any $\sigma > \sigma'$, $\frac{f_c(\sigma)}{f_{c'}(\sigma)} \geq \frac{f_c(\sigma')}{f_{c'}(\sigma')}$.

[6]Relative to our theory, foreign-born individuals with less than a high-school education tend to disproportionately locate in large US cities. Data from 1980, when foreign-born individuals were a substantially smaller share of the US population, suggest this reflects particular advantages that large cities offer foreign-born individuals rather than a general tendency for the unskilled to locate in large cities. See section 2.5.1.2.

cities theories describe a homogeneous population (Abdel-Rahman and Anas, 2004). Most previous examinations of heterogeneous labor have only described two skill levels, typically labeled skilled and unskilled.[7] To describe greater heterogeneity, we assume a continuum of skills, like Behrens, Duranton, and Robert-Nicoud (2012) and Davis and Dingel (2012).[8] Understanding the distribution of skills across cities with more than two types is valuable for at least three reasons. First, a large literature in labor economics has described important empirical developments such as wage polarization, job polarization, and simultaneous changes in between- and within-group inequality that cannot be explained by a model with two homogeneous skill groups (Acemoglu and Autor, 2011). Second, these developments have counterparts in cross-city variation in inequality and skill premia (Baum-Snow and Pavan, 2011; Davis and Dingel, 2012). Third, we document systematic patterns in the cross-city distribution of skills at high levels of disaggregation, which suggests that individuals within broad skill categories are imperfect substitutes.[9]

Our model is a novel integration of systems-of-cities theory with land-use theory. The Alonso-Muth-Mills model of a single city describes a homogeneous population of residents commuting to a central business district (Brueckner, 1987). In that model, higher rents for locations with shorter commutes equalize utility across locations in equilibrium. When individuals are heterogeneous and value the rent-distance tradeoff differently, the single city's equilibrium rent schedule is the upper envelope of individuals' bid-rent functions (von Thünen, 1826; LeRoy and Sonstelie, 1983; Fujita and Thisse, 2002, Ch 3). Models of a system of cities have incorporated the Alonso-Muth-Mills urban structure in which all individuals are indifferent across all locations within a city as a city-level congestion mechanism (Abdel-Rahman and Anas, 2004; Behrens, Duranton, and Robert-Nicoud, 2012). Our novel contribution is to describe multiple cities with internal geographies when

[7]We focus on theories in which labor is heterogeneous in some asymmetric sense (e.g. more skilled individuals have absolute advantage in tasks or more skilled individuals generate greater human-capital spillovers). There are also models describing matching problems, such as Helsley and Strange (1990) and Duranton and Puga (2001), in which labor is heterogeneous in a horizontal characteristic.

[8]Eeckhout, Pinheiro, and Schmidheiny (2011) describe a model with three skill types.

[9]A long line of empirical work describes cross-city variation in skill distributions in terms of the share of residents who have a college degree (Glaeser, 2008). Most closely related to our work is Hendricks (2011), who finds a weak relationship between cities' industries and college shares.

individuals are not spatially indifferent across all locations.[10] The essential idea is that individuals choosing between living in Chicago or Des Moines simultaneously consider in what parts of Chicago and what parts of Des Moines they might locate. Though these tradeoffs appear obvious, we are not aware of a prior formal analysis. Considering both dimensions simultaneously is more realistic in both the description of the economic problem and the resulting predicted cross-city skill distributions. Since we have a continuum of heterogeneous individuals, we obtain equilibrium rent schedules that are integrals rather than upper envelopes of a discrete number of bid-rent functions.[11]

Our model belongs to a long theoretical tradition describing factor-supply-driven comparative advantage, dating from the Heckscher-Ohlin theory formalized by Samuelson (1948). In international contexts, theorists have typically taken locations' factor supplies as exogenously endowed. Since individuals are mobile across cities, our theory endogenizes cities' factor supplies while describing how the composition of output is governed by comparative advantage. Our approach to comparative advantage with a continuum of factors and a continuum of sectors follows a large assignment literature and is most closely related to the recent work of Costinot (2009) and Costinot and Vogel (2010).[12] While these recent papers assume that countries' factor endowments exhibit the monotone likelihood ratio property, we obtain the result that cities' skill distributions exhibit this property as an equilibrium outcome. Thus, from a theoretical perspective, cities within a country constitute a natural setting to examine these theories of comparative advantage. Moreover, the assumption of a common production technology is likely more appropriate within than between economies, and data from a single economy are likely more consistent and comparable than data combined across countries.

The Heckscher-Ohlin model has been the subject of extensive empirical investigation in international economics. A pair of papers describe regional outputs using this framework. Davis and Weinstein (1999) run regressions of regional outputs on regional endowments, employing the framework of Leamer (1984), but they abstract from the issue of labor mobility across regions. Bernstein

[10] In order to tractably characterize multiple cities with internal geographies and heterogeneous agents, we neglect the business-vs-residential land-use problem studied by Lucas and Rossi-Hansberg (2002)

[11] Our continuum-by-continuum approach to a differential rents model is in the spirit of Sattinger (1979).

[12] Sattinger (1993) surveys the assignment literature.

and Weinstein (2002) consider the two-way links between endowments and outputs, concluding that if we know regions' outputs, we know with considerable precision the inputs used, but not vice versa. For these reasons, traditional Heckscher-Ohlin models did not appear a promising way to explain regional differences in sectoral composition.

Our theory predicts systematic variation in sectoral composition in the form of an urban hierarchy of sectors. Prior systems-of-cities theories have overwhelmingly described polarized sectoral composition: *specialized cities* that have only one tradable sector and *perfectly diversified cities* that have all the tradable sectors (Abdel-Rahman and Anas, 2004; Helsley and Strange, 2012). A recent exception is Helsley and Strange (2012), who examine whether the equilibrium level of coagglomeration is efficient. While Helsley and Strange (2012) make minimal assumptions in order to demonstrate that Nash equilibria are generically inefficient when there are interindustry spillovers, we make strong assumptions that yield testable implications about the distribution of sectoral activity across cities.

Our model's equilibrium exhibits a hierarchy of cities and sectors, as larger cities produce a superset of the goods produced in smaller cities. Models in central place theory, dating from Christaller (1933) through Hsu, Holmes, and Morgan (2013), have attributed this hierarchy property to the interaction of industry-specific scale economies and geographic market access based on the distance between firms located in distinct city centers. It is interesting that our model yields the hierarchy property in the absence of both. Our theory links the hierarchy of sectors to a hierarchy of skills shaped by the internal geography of cities, neither of which have been considered in central place theory.

A recent empirical literature has demonstrated significant agglomeration and coagglomeration of industries relative to the null hypothesis of locations being (uniformly) randomly assigned in proportion to local population (Ellison and Glaeser, 1997; Duranton and Overman, 2005; Ellison, Glaeser, and Kerr, 2010). Our model's predictions are consistent with these findings. Since our theory says that sectors are ranked in terms of their relative employment levels, at most one sector could exhibit employment proportionate to total population. All other sectors will exhibit geographic concentration. Similarly, since sectors more similar in skill intensity will exhibit more similar relative employment levels, the cross-city distribution of sectoral employment will be consistent with skill-related coagglomeration. We obtain these results in the absence of industry-specific scale economies and industry-pair-specific interactions or spillovers.

Our empirical work follows directly from our model's predictions about the cross-city distribution of sectoral activity relating cities' and sectors' characteristics. There is a small empirical literature describing variation in cities' sectoral composition, but this work has not been tightly tied to theory. This is likely because theories describing specialized or perfectly diversified cities provide limited guidance to empirical investigations of data that fall between the extremes. Holmes and Stevens (2004) survey the spatial distribution of economic activities in North America. In examining the empirical pattern of specialization, they show that agriculture, mining, and manufacturing are disproportionately in smaller cities, while finance, insurance, real estate, professional, and management activities are disproportionately in larger cities. However, they do not reference a model or theoretical mechanism that predicts this pattern to be the equilibrium outcome. Seminal work by Vernon Henderson explores theoretically and empirically the relationship between city size and industrial composition (Henderson, 1991). Henderson (1974) theoretically describes the polar cases of specialized and perfectly diversified cities (Helsley and Strange, 2012), while our model predicts incomplete industrial specialization. Henderson has argued that localization economies link cities' and industries' sizes, while our theory relies on urbanization economies and individuals' comparative advantage.[13] Despite these contrasts, our theory yields estimating equations for the population elasticities of sectoral employment that are closely related to the reduced-form regressions of employment shares on population that Henderson (1983) estimated for a few select industries. Our theory provides an explicit microfoundation for these regressions for an arbitrary number of sectors. Moreover, it predicts that we can order these elasticities by skill intensity. It also describes how to compare the sectoral composition of groups of cities ordered by size, nesting the comparison of large and medium-size cities made by Henderson (1997). While our urbanization-based theory abstracts from the localization economies emphasized by Henderson, we believe future work should

[13]The literature traditionally distinguishes two types of external economies of scale (Henderson, 1987, p.929). Localization economies are within-industry, reflecting the scale of activity in that industry in that location. Urbanization economies are general, reflecting the scale of all economic activity in a location. Beyond scale, Lucas (1988) has stressed the composition of a location's human capital. The agglomeration process generating city-level productivities in our theory incorporates both scale and composition effects.

seek to integrate these distinct approaches.

2.3 Model

We develop a general-equilibrium model in which L heterogeneous individuals choose a city, a location within that city, and a sector in which to produce. There are C discrete cities ($c \in \mathbb{C} = \{1, \ldots, C\}$), a continuum of skills, and a continuum of sectors. We study the consequences for city total factor productivity and the cross-city distributions of skills and sectors.

2.3.1 Preferences, production, and places

Individuals consume a freely traded final good. This final good is the numeraire and produced by combining a continuum of freely traded, labor-produced intermediate goods indexed by $\sigma \in \Sigma$. These have prices $p(\sigma)$ that are independent of location because trade costs are zero. Locations are characterized by their city c and their (inverse) desirability $\tau \in \mathcal{T}$, so they have rental prices $r(c, \tau)$.

Final-goods producers have a CES production function

$$Q = \left\{ \int_{\sigma \in \Sigma} B(\sigma)[Q(\sigma)]^{\frac{\epsilon-1}{\epsilon}} d\sigma \right\}^{\frac{\epsilon}{\epsilon-1}}, \tag{2.1}$$

where the quantity of intermediate good σ is $Q(\sigma)$, $\epsilon > 0$ is the elasticity of substitution between intermediates, and $B(\sigma)$ is an exogenous technological parameter. The profits of final-goods producers are given by

$$\Pi = Q - \int_{\sigma \in \Sigma} p(\sigma)Q(\sigma)d\sigma. \tag{2.2}$$

Heterogeneous individuals use their labor to produce intermediate goods. There is a mass of L heterogeneous individuals with skills ω that have the cumulative distribution function $F(\omega)$ and density $f(\omega)$ on support $\Omega \equiv [\underline{\omega}, \bar{\omega}]$. The productivity of an individual of skill ω in sector σ at location τ in city c is

$$q(c, \tau, \sigma; \omega) = A(c)T(\tau)H(\omega, \sigma). \tag{2.3}$$

$A(c)$ denotes city-level total factor productivity, which results from agglomeration and is taken as given by individuals. $T(\tau)$ reflects the productivity effects of location within the city, which in a canonical case is the cost of commuting to the central business district.[14] We assume that $T(\tau)$ is continuously differentiable and $T'(\tau) < 0$, which is just a normalization that higher-τ locations are less desirable. We assume that the twice-differentiable function $H(\omega, \sigma)$ is strictly log-supermodular in ω and σ and strictly increasing in ω.[15] The former governs comparative advantage, so that higher-ω individuals are relatively more productive in higher-σ sectors.[16] The latter says that absolute advantage is indexed by ω, so that higher-ω individuals are more productive than lower-ω individuals in all sectors. Each individual inelastically supplies one unit of labor, so her income is her productivity times the price of the output produced, $q(c, \tau, \sigma; \omega)p(\sigma)$.

Locations within each city are heterogeneous, with the desirability of a location indexed by $\tau \geq 0$. The most desirable location is denoted $\tau = 0$, so higher values of τ denote greater distance from the ideal location. The supply of locations with desirability greater than τ is $S(\tau)$.[17] This is a strictly increasing function, since the supply of available locations increases as one lowers one's minimum standard of desirability. $S(0) = 0$, since there are no locations better than the ideal. We assume $S(\tau)$ is twice continuously differentiable. Locations are owned by absentee landlords who spend their rental income on the final good. The city has sufficient land capacity that everyone can reside in the city and the least desirable locations are unoccupied. We normalize the reservation value of unoccupied locations to zero, so $r(c, \tau) \geq 0$.

Individuals choose their city c, location τ, and sector σ to maximize utility. An individual's utility depends on their consumption of the numeraire final good, which is their income after paying

[14] As written, $T(\tau)$ indexes the desirability of the location for its productive advantages, but a closely related specification makes $T(\tau)$ describe a location's desirability for its consumption value. The production and consumption interpretations yield very similar results but differ slightly in functional form. For expositional clarity, we use the production interpretation given by equations (2.3) and (2.4) in describing the model in the main text and present the consumption interpretation in appendix B.1.

[15] In $\mathbb{R}^2$, a function $H(\omega, \sigma)$ is strictly log-supermodular if $\omega > \omega'$, $\sigma > \sigma' \Rightarrow H(\omega, \sigma)H(\omega', \sigma') > H(\omega, \sigma')H(\omega', \sigma)$.

[16] We refer to higher-ω individuals as more skilled and higher-σ sectors as more skill-intensive.

[17] In the special case of the classical von Thünen model, τ describes physical distance from the central business district and the supply is $S(\tau) = \pi\tau^2$.

their locational cost:

$$U(\omega, c, \tau, \sigma) = A(c)T(\tau)H(\omega, \sigma)p(\sigma) - r(c, \tau). \tag{2.4}$$

Denote the endogenous quantity of individuals of skill ω residing in city c at location τ and working in sector σ by $L \times f(\omega, c, \tau, \sigma)$.

City-level TFP, $A(c)$, reflects agglomeration gains derived from both population size and composition. $A(c)$ is higher when a city contains a larger and more skilled population. Denote the endogenous quantity of individuals of skill ω residing in city c by $L \times f(\omega, c) = L \times \int_{\sigma \in \Sigma} \int_{\tau \in \mathcal{T}} f(\omega, c, \tau, \sigma) d\tau d\sigma$. Total factor productivity is

$$A(c) = J\left(L, \int_{\omega \in \Omega} j(\omega) f(\omega, c) d\omega\right), \tag{2.5}$$

where $J(\cdot, \cdot)$ is a positive function increasing in each of its arguments and $j(\omega)$ is a positive, non-decreasing function.

2.3.2 Equilibrium

In a competitive equilibrium, individuals maximize utility, final-good producers and landowners maximize profits, and markets clear. Individual maximize their utility by their choices of city, location, and sector such that

$$f(\omega, c, \tau, \sigma) > 0 \iff \{c, \tau, \sigma\} \in \arg\max U(\omega, c, \tau, \sigma). \tag{2.6}$$

Profit maximization by final-good producers yields demands for intermediates

$$Q(\sigma) = I\Big(\frac{p(\sigma)}{B(\sigma)}\Big)^{-\epsilon}, \tag{2.7}$$

where $I \equiv L \sum_c \int_\sigma \int_\omega \int_\tau q(\omega, c, \tau, \sigma) p(\sigma) f(\omega, c, \tau, \sigma) d\tau d\omega d\sigma$ denotes total income and these producers' profits are zero. Profit maximization by absentee landlords engaged in Bertrand competition causes unoccupied locations to have rental prices of zero,

$$r(c,\tau) \times \left(S'(\tau) - L \int_{\sigma\in\Sigma} \int_{\omega\in\Omega} f(\omega,c,\tau,\sigma) d\omega d\sigma \right) = 0 \ \forall c \ \forall \tau. \tag{2.8}$$

Market clearing requires the endogenous quantity of individuals of skill ω residing in city c at location τ and working in sector σ, $L \times f(\omega,c,\tau,\sigma)$, to be such that the supply of a location type is greater than or equal to its demand, the demand and supply of intermediate goods are equal, and every individual lives somewhere.

$$S'(\tau) \geq L \int_{\omega\in\Omega} \int_{\sigma\in\Sigma} f(\omega,c,\tau,\sigma) d\sigma d\omega \quad \forall c \ \forall \tau \tag{2.9}$$

$$Q(\sigma) = \sum_{c\in\mathbb{C}} Q(\sigma,c) = L \sum_{c\in\mathbb{C}} \int_{\omega\in\Omega} \int_{\tau\in\mathcal{T}} q(c,\tau,\sigma;\omega) f(\omega,c,\tau,\sigma) d\omega d\tau \quad \forall \sigma \tag{2.10}$$

$$f(\omega) = \sum_{c\in\mathbb{C}} f(\omega,c) = \sum_{c\in\mathbb{C}} \int_{\sigma\in\Sigma} \int_{\tau\in\mathcal{T}} f(\omega,c,\tau,\sigma) d\tau d\sigma \quad \forall \omega \tag{2.11}$$

A competitive equilibrium is a set of functions $Q : \Sigma \to \mathbb{R}^+$, $f : \Sigma \times \mathbb{C} \times \mathcal{T} \times \Omega \to \mathbb{R}^+$, $A : \mathbb{C} \to \mathbb{R}^+$, $r : \mathbb{C} \times \mathcal{T} \to \mathbb{R}^+$, and $p : \Sigma \to \mathbb{R}^+$ such that conditions (2.6) through (2.11) hold.

2.3.3 An autarkic city

We begin by considering a single city, denoted c, with exogenous population $L(c)$ and skill distribution $F(\omega)$. With fixed population, autarky TFP is fixed by equation (2.5). We describe individuals' choices of sectors and locations to solve for the autarkic equilibrium.

To solve, we exploit the fact that locational and sectoral argument enters individuals' utility functions separably. Individuals' choices of their sectors are independent of their locational decisions:

$$\arg\max_{\sigma} A(c)T(\tau)H(\omega,\sigma)p(\sigma) - r(c,\tau) = \arg\max_{\sigma} H(\omega,\sigma)p(\sigma)$$

Define the assignment function $M(\omega) = \arg\max_{\sigma} H(\omega,\sigma)p(\sigma)$ so that we can write $G(\omega) \equiv$

$H(\omega, M(\omega))p(M(\omega))$. By comparative advantage, $M(\omega)$ is increasing.[18] By absolute advantage, more skilled individuals earn higher nominal incomes and $G(\omega)$ is a strictly increasing function.[19]

Individuals' choices of their locations are related to their sectoral decisions in the sense that willingness to pay for more desirable locations depends on the skill component of income $G(\omega)$. Within the city, individual choose their optimal location:

$$\max_{\tau} A(c)T(\tau)G(\omega) - r(c, \tau)$$

Competition among landlords ensures that the most desirable locations are those occupied, so the least desirable occupied site $\bar{\tau}(c) \equiv \max_{\tau}\{\tau : \; f(\omega, c, \tau, \sigma) > 0\}$ in a city of population $L(c)$ is defined by $L(c) = S(\bar{\tau}(c))$. More desirable locations have higher rental prices.

Lemma 2.1 (Populated locations)**.** *In equilibrium,* $S(\tau) = L \int_0^{\tau} \int_{\sigma \in \Sigma} \int_{\omega \in \Omega} f(\omega, c, x, \sigma) d\omega d\sigma dx \; \forall \tau \leq \bar{\tau}(c)$, $r(c, \tau)$ *is strictly decreasing in* τ $\forall \tau < \bar{\tau}(c)$, *and* $r(c, \bar{\tau}(c)) = 0$.

Individuals of higher skill have greater willingness to pay for more desirable locations. That is, $\frac{\partial^2}{\partial \tau \partial \omega} A(c)T(\tau)G(\omega) < 0$ because locational advantages complement individual productivity. As a result, in equilibrium higher-ω individuals occupy lower-τ locations.

Lemma 2.2 (Autarky locational assignments)**.** *In autarkic equilibrium, there exists a continuous and strictly decreasing locational assignment function* $N : \mathcal{T} \to \Omega$ *such that* $f(\omega, c, \tau, \sigma) > 0 \iff N(\tau) = \omega$, $N(0) = \bar{\omega}$ *and* $N(\bar{\tau}(c)) = \underline{\omega}$.

This assignment function is obtained by equating supply and demand of locations:

$$S(\tau) = L \int_0^{\tau} \int_{\sigma \in \Sigma} \int_{\omega \in \Omega} f(\omega, c, x, \sigma) d\omega d\sigma dx$$
$$\Rightarrow N(\tau) = F^{-1}\left(\frac{L(c) - S(\tau)}{L(c)}\right)$$

Given individuals' equilibrium locations within the city, the schedule of locational rental prices

[18] Lemma 1 of Costinot and Vogel (2010) shows that $M(\omega)$ is continuous and strictly increasing in equilibrium.

[19] Absolute advantage across all sectors is far from necessary. The weaker condition that productivity is increasing in skill at the equilibrium assignments, $\frac{d}{d\omega} H(\omega, M(\omega)) > 0$, is sufficient.

supporting these assignments comes from combining individuals' utility-maximizing decisions and the boundary condition $r(c, \bar{\tau}(c)) = 0$.

Lemma 2.3 (Autarky locational prices)**.** *In autarkic equilibrium, $r(c, \tau)$ is continuously differentiable on $\tau \geq 0$ and given by $r(c, \tau) = -A(c) \int_{\tau}^{\bar{\tau}(c)} T'(t) G(N(t)) dt$* for $\tau \leq \bar{\tau}(c)$.

The properties of interest in a competitive equilibrium are characterized by the assignment functions M and N. In the autarkic equilibrium, more skilled individuals work in more skill-intensive sectors and live in more desirable locations.

2.3.4 A system of cities

The previous section described a single city with an exogenous population. We now describe a system of cities in which these populations are endogenously determined in spatial equilibrium. Take cities' TFPs, which will be endogenously determined in equilibrium, as given for now and order the cities so that $A(C) \geq A(C-1) \geq \cdots \geq A(2) \geq A(1)$.[20] In autarky, τ was a sufficient statistic for the attractiveness of a location. Now a location's attractiveness, which we denote by γ, depends both on city-level TFP and its innate desirability within the city.

Definition 2.1. The *attractiveness* of a location in city c of desirability τ is $\gamma = A(c)T(\tau)$.

Cities with higher TFP have larger populations. Consider two cities, c and c', that differ in productivity, with $A(c) > A(c')$. The city with greater TFP will have greater population, $L(c) > L(c')$. If it did not, the least desirable occupied location in city c would be more desirable than the least desirable occupied location in city c', $\bar{\tau}(c) \leq \bar{\tau}(c')$, since the supply of locations, $S(\tau)$, is common across cities. Since TFP is also higher in c, this would make the least attractive occupied location in city c more attractive than the least attractive occupied location in city c', $A(c)T(\bar{\tau}(c)) > A(c')T(\bar{\tau}(c'))$. In equilibrium, the least desirable occupied location in each city has a price of zero, $r(c, \bar{\tau}(c)) = r(c', \bar{\tau}(c')) = 0$, by lemma 2.1. In that case, every individual would agree that living in c at $\bar{\tau}(c)$ is strictly better than living in c' at $\bar{\tau}(c')$ (because $A(c)T(\bar{\tau}(c))G(\omega) >$

[20] Individuals take these TFPs as given. For now, we can assume these differences in total factor productivity are exogenously given. We describe their endogenous determination in section 2.3.6.

$A(c')T(\bar{\tau}(c'))G(\omega))$, which contradicts the definition of $\bar{\tau}(c')$ as an occupied location. So the city with higher TFP must have a larger population.

A smaller city's locations are a subset of those in a larger city in terms of attractiveness. For every location in the less populous city, there is a location in the more populous city that is equally attractive. The location in city c' of desirability τ' is equivalent to a location τ in city c, given by $A(c)T(\tau) = A(c')T(\tau')$. The equally attractive location in the larger city has higher TFP but lower innate desirability. That is, an individual who is indifferent between Chicago and Des Moines lives closer to the most desirable location in Des Moines than the most desirable location in Chicago, $\tau' = T^{-1}\left(\frac{A(c)}{A(c')}T(\tau)\right) < \tau$. The more populous city also has locations that are strictly more attractive than the best location in the less populous city; locations of attractiveness $\gamma > A(c')T(0)$ are found in c and not in c'. In equilibrium, two locations of equal attractiveness must have the same price, so we can describe the rental price of a location of attractiveness γ as $r_\Gamma(\gamma)$.

To characterize locational assignments and prices in the system of cities, we first characterize assignments and prices in terms of γ. The solution is analogous to that derived in the autarkic case. We then translate these assignments and prices into functions of c and τ.

Individuals of higher skill have greater willingness to pay for more attractive locations, so in equilibrium higher-ω individuals live in higher-γ locations.

Lemma 2.4 (Locational assignments)**.** *In equilibrium, there exists a continuous and strictly increasing locational assignment function $K : \Gamma \to \Omega$ such that $f(\omega, c, \tau, \sigma) > 0 \iff A(c)T(\tau) = \gamma$ and $K(\gamma) = \omega$, where $K(\underline{\gamma}) = \underline{\omega}$, and $K(\bar{\gamma}) = \bar{\omega}$.*

To obtain an explicit expression for $K : \Gamma \to \Omega$, we can denote the supply of locations offering benefits γ or greater as $S_\Gamma(\gamma)$. The supply function is

$$S_\Gamma(\gamma) = \sum_{c:A(c)T(0)\geq\gamma} S\left(T^{-1}\left(\frac{\gamma}{A(c)}\right)\right)$$

By definition $S_\Gamma(\bar{\gamma}) = 0$ and by the fact that the best locations are populated $S_\Gamma(\underline{\gamma}) = L$. Lemmas 2.1 and 2.4 allow us to say that $S_\Gamma(\gamma) = L\int_\gamma^{\bar{\gamma}} f(K(x))K'(x)dx$, so $K(\gamma) = F^{-1}\left(\frac{L - S_\Gamma(\gamma)}{L}\right)$. These locational assignments yield an expression for equilibrium locational prices.

Lemma 2.5 (Locational prices)**.** *In equilibrium, $r_\Gamma(\gamma)$ is continuously differentiable on $[\underline{\gamma}, \bar{\gamma}]$ and given by $r_\Gamma(\gamma) = \int_{\underline{\gamma}}^{\gamma} G(K(x))dx$.*

Therefore, the determination of locational assignments and prices within the system of cities is analogous to determining these locational assignments and prices for an autarkic city with a supply of locations that is the sum of locations across the system of cities. The task that remains is to translate these assignments and prices from γ back to c, τ.

The city limit is $\bar{\tau}(c) = T^{-1}\left(\frac{\gamma}{A(c)}\right)$. Since we are not focused on rental prices, a sufficient characterization is $r(c, \tau) = r_\Gamma(A(c)T(\tau))$. For locational assignments, we obtain an explicit expression that characterizes cities' skill distributions in terms of ω and c.

Lemma 2.6 (A city's skill distribution)**.** *The population of individuals of skill ω in city c is*

$$f(\omega, c) = \begin{cases} \frac{-K^{-1\prime}(\omega)}{A(c)L} T^{-1\prime}\left(\frac{K^{-1}(\omega)}{A(c)}\right) S'\left(T^{-1}\left(\frac{K^{-1}(\omega)}{A(c)}\right)\right) & \text{if } A(c)T(0) \geq K^{-1}(\omega) \\ 0 & \text{otherwise} \end{cases}.$$

The relative populations of individuals of skill ω depends on the relative supply of locations of attractiveness $K^{-1}(\omega)$. Since higher-ω individuals live in more attractive locations and the most attractive locations are found exclusively in the larger city, there is an interval of high-ω individuals who reside exclusively in the larger city. Individuals of abilities below this interval are found in both cities, and individuals of equal skill reside in equally attractive locations that have equal rental prices.

2.3.5 Cities' populations and sectors

When more desirable locations within cities are scarcer, our model implies that larger cities will exhibit relatively more skilled populations and produce relatively more in skill-intensive sectors. The first result comes from large cities offering more attractive locations, which are occupied by more skilled individuals in equilibrium. The second result comes from larger cities' TFP advantages being sector-neutral, so that sectoral composition is governed by skill composition.

Larger cities have relatively more skilled populations. We have already shown that the most skilled individuals live exclusively in the largest city because it offers the most attractive locations.

What about individuals of skill levels found in multiple cities? The distribution of skills is governed by the interaction of city TFPs, $A(c)$, and the supply of locations within cities yielding at least $T(\tau) = z$, which is $S(T^{-1}(z))$. We say that the latter is *well-behaved* when more desirable locations are scarcer and their distribution satisfies a common regularity condition.[21]

Definition 2.2. The supply of locations within cities is *well-behaved* if the density $-\frac{\partial}{\partial z} S(T^{-1}(z))$ is decreasing and log-concave.

When the supply of locations within cities is well-behaved, more skilled individuals are relatively more prevalent in larger cities throughout the skill distribution. Consider a location in Des Moines, which has attractiveness γ'. Higher-TFP Chicago has more locations of attractiveness $\gamma \geq \gamma'$, since $S(\tau)$ is increasing. If $S(T^{-1}(z))$ is well-behaved, then Chicago's supply of γ' locations relative to γ'' locations, for any $\gamma' > \gamma''$, is greater than that of Des Moines. Since higher-ω individuals live in higher-γ locations, this means the population of Chicago is relatively more skilled when comparing any two skill types. Proposition 2.1 states this result for a system of cities; its proof is in appendix B.2.

Proposition 2.1 (Skill abundance)**.** *If the supply of locations within cities is well-behaved, then $f(\omega, c)$ is log-supermodular.*

Our next set of results concern the sectoral output of cities. Since individuals' choices of sectors are independent of their cities and locations, the cross-city skill distribution governs the cross-city output distribution. Larger cities are relatively skill-abundant and more skilled individuals work in more skill-intensive sectors, so larger cities produce relatively more in skill-intensive sectors.

These patterns of specialization and trade are closely related to the high-dimensional model of endowment-driven comparative advantage introduced by Costinot (2009), but in our setting cities' populations are endogenously determined.[22] Since at equilibrium larger cities' productivity advantages are sector-neutral differences in total factor productivity, $f(\omega, c)$ is log-supermodular,

[21] Log-concavity is most commonly used in economics in the context of probability distributions (Bagnoli and Bergstrom, 2005).

[22] Assumption 2 in Costinot (2009)'s factor endowment model is that countries' exogenous endowments are such that countries can be ranked according to the monotone likelihood ratio property. Proposition 2.1 identifies sufficient conditions for cities' *equilibrium* skill distributions to exhibit this property.

and $H(\omega, \sigma)$ is log-supermodular, our economy's equilibrium satisfies Definition 4, Assumption 2, and Assumption 3 of Costinot (2009).[23] The result is that $Q(\sigma, c)$ is log-supermodular. As shown by Corollaries 2 and 3 in Costinot (2009), we can therefore rank the relative output ($Q(\sigma, c)$), employment ($f(\sigma, c) \equiv \int_{\tau \in \mathcal{T}} \int_{\omega \in \Omega} f(\omega, c, \tau, \sigma) d\omega d\tau$), and revenue ($R(\sigma, c) \equiv p(\sigma)Q(\sigma, c)$) of any two sectors in any two cities.

Proposition 2.2 (Comparative Advantage)**.** *If $f(\omega, c)$ is log-supermodular, then $Q(\sigma, c)$, $f(\sigma, c)$, and $R(\sigma, c)$ are log-supermodular, so that for $c \geq c'$ and $\sigma \geq \sigma'$, the following inequalities hold true*

$$
\begin{aligned}
Q(\sigma, c)Q(\sigma', c') &\geq Q(\sigma, c')Q(\sigma', c) \\
f(\sigma, c)f(\sigma', c') &\geq f(\sigma, c')f(\sigma', c) \\
R(\sigma, c)R(\sigma', c') &\geq R(\sigma, c')R(\sigma', c)
\end{aligned}
\tag{2.12}
$$

These inequalities characterize the pattern of comparative advantage across cities.[24]

2.3.6 Endogenizing cities' total factor productivities

Our exposition of equilibrium in sections 2.3.4 and 2.3.5 took cities' total factor productivities as exogenously given. When the conditions of Proposition 2.1 are satisfied, a city that has higher total factor productivity $A(c)$ is larger and has a skill distribution $f(\omega, c)$ that likelihood ratio dominates those of cities with lower TFPs. Thus, this spatial pattern can be supported by endogenous productivity processes that make the city-level characteristic $A(c)$ higher when the city contains a larger and more skilled population, such as the class of agglomeration functions described by

[23] Definition 4 of Costinot (2009) requires that factor productivity vary across countries (cities) in a Hicks-neutral fashion. Since productivity $A(c)T(\tau)$ varies both across and within cities, our production function $q(c, \tau, \sigma; \omega)$ does not satisfy this requirement for arbitrary locational assignments. However, in equilibrium, our economy does exhibit this property. In the production interpretation of $T(\tau)$, equilibrium productivity $q(c, \tau, \sigma; \omega) = K^{-1}(\omega)H(\omega, \sigma)$ does not vary across ω-occupied locations and is log-supermodular in ω and σ. In the notation of equation (6) in Costinot (2009), $a(\gamma) = 1$ and $h(\omega, \sigma) = K^{-1}(\omega)H(\omega, \sigma)$, satisfying Definition 4 and Assumption 3.

[24] A traditional definition of comparative advantage refers to locations' autarkic prices. In our setting, autarky means prohibiting both trade of intermediate goods and migration between cities. Since individuals are spatially mobile, cities do not have "factor endowments", and we must specify the autarkic skill distributions. If we consider an autarkic equilibrium with the skill distributions from the system-of-cities equilibrium, then larger cities have lower relative autarkic prices for higher-σ goods because they are skill-abundant in the MLRP sense, as shown by Costinot and Vogel (2010, p. 782).

equation (2.5). Numerous agglomeration processes may generate such productivity benefits, and we do not attempt to distinguish between them here.

2.4 Empirical approach and data description

We examine the predictions of Propositions 2.1 and 2.2 using two approaches. The first involves regression estimates of the population elasticities of educational, occupational, and industrial populations. The second involves pairwise comparisons governed by the monotone likelihood ratio property.

Empirically testing our model requires data on cities' skill distributions, sectors' skill intensities, and cities' sectoral employment. We use public-use microdata from the US Census of Population to identify the first two. The latter is described by data from County Business Patterns and Occupational Employment Statistics. The Census of Population describes individuals' educational attainments, geographic locations, places of birth, occupations, and industries. County Business Patterns describes cities' industrial employment. Occupational Employment Statistics describes cities' occupational employment. We combine these various data at the level of (consolidated) metropolitan statistical areas (MSAs); see appendix B.3 for details.

2.4.1 Empirical tests

Propositions 2.1 and 2.2 say that the distribution of skills across cities, $f(\omega, c)$, and the distribution of sectoral employment across cities, $f(\sigma, c)$, are log-supermodular functions. Log-supermodularity has many implications; we focus on two that are amenable to empirical testing. If the function $f(\nu, c)$ is log-supermodular, then

- a linear regression $\ln f(\nu, c) = \alpha_\nu + \beta_\nu \ln L(c) + \epsilon_{\nu,c}$ in which α_ν are fixed effects and $L(c)$ is city population yields $\beta_\nu \geq \beta_{\nu'} \iff \nu \geq \nu'$;
- if $\mathcal{C}$ and $\mathcal{C}'$ are distinct sets and $\mathcal{C}$ is greater than $\mathcal{C}'$ ($\inf_{c \in \mathcal{C}} L(c) > \sup_{c' \in \mathcal{C}'} L(c')$), then $\sum_{c \in \mathcal{C}} f(\nu, c) \sum_{c' \in \mathcal{C}'} f(\nu', c') \geq \sum_{c \in \mathcal{C}} f(\nu', c) \sum_{c' \in \mathcal{C}'} f(\nu, c') \; \forall \nu > \nu'$.

The first implication, which we will refer to as the "elasticity test," says that the city-population elasticity of the population of a skill type in a city $f(\omega, c)$ is increasing in skill ω.[25] Similarly, the population elasticity of sectoral employment $f(\sigma, c)$ is increasing in skill intensity σ. The elasticity test examines the patterns suggested by Figures 2.1 through 2.3, where steeper slopes correspond to higher elasticities. Our theory thus provides a structure to interpret previous work describing the population elasticities of sectoral employment, such as Henderson (1983) and Holmes and Stevens (2004).[26] The second implication, which we will refer to as the "pairwise comparisons test", says that if cities are divided into bins ordered by population sizes, then in any pairwise comparison of two bins and two skills/sectors, the bin containing more populous cities will have relatively more of the more skilled type.[27]

2.4.2 Skills

Following a large literature, we use observed educational attainment as a proxy for individuals' skills.[28] Educational attainment is a coarse measure, but it is the best measure available in data describing a large number of people across detailed geographic locations. To describe cities' skill distributions, we aggregate individual-level microdata to the level of metropolitan statistical areas. A large literature in urban economics describes variation in terms of two skill groups, typically college and non-college workers. Following Acemoglu and Autor (2011), we use a minimum of three skill groups. The Census 2000 microdata identify 16 levels of educational attainment, from "no schooling completed" to "doctoral degree". We define three skill groups of approximately equal size among the working population: high-school degree or less; some college or associate's degree;

[25] The linear regression may understood as a first-order Taylor approximation: $\ln f(\nu, c) \approx \ln f(\nu, c^*) + \frac{\partial \ln f(\nu,c^*)}{\partial \ln L(c)}(\ln L(c) - \ln L(c^*)) + \epsilon = \alpha_\nu + \beta_\nu \ln L(c) + \epsilon_{\nu,c}$, where $\beta_\nu = \frac{\partial \ln f(\nu,c^*)}{\partial \ln L(c)}$ is increasing in ν by log-supermodularity of $f(\nu, c)$.

[26] Henderson (1983) regresses employment shares on population levels, but reports "percent $\mathcal{T}$ share / percent $\mathcal{T}$ population", which is equal to $\beta_\sigma - 1$ in our notation. Similarly, Holmes and Stevens (2004) describe how location quotients, a city's share of industry employment divided by its share of total employment, vary with city size. In our notation, a location quotient is $LQ(\sigma, c) = \frac{f(\sigma,c)/\sum_{c'} f(\sigma,c')}{L(c)/L}$, so the $L(c)$-elasticity of $LQ(\sigma, c)$ is $\beta_\sigma - 1$.

[27] The pairwise comparisons test follows from taking sums twice of each side of $f(\nu, c)f(\nu', c') \geq f(\nu', c)f(\nu, c')$ given $c > c' \; \forall c \in C \; \forall c' \in C' \; \forall \nu > \nu'$.

[28] Costinot and Vogel (2010) show that log-supermodularity of factor supplies in an observed characteristic and unobserved skill ω is sufficient for mapping a theory with a continuum of skills to data with discrete characteristics.

and bachelor's degree or more. In a more ambitious approach, we also consider nine skill groups, ranging from individuals who never reached high school (3 percent of the population) to those with doctoral degrees (1 percent).[29] Table 2.1 shows the population shares of each of these skill groups in 2000.

Table 2.1: Skill groups by educational attainment

Skill (3 groups)	Population share	Skill (9 groups)	Population share
High school or less	0.35	Less than high school	0.03
		High school dropout	0.07
		High school graduate	0.24
Some college	0.32	College dropout	0.24
		Associate's degree	0.08
BA or more	0.33	Bachelor's degree	0.21
		Master's degree	0.08
		Professional degree	0.03
		Doctoral degree	0.01

Population shares are percentages of full-time, full-year prime-age workers.
Source: Census 2000 microdata via IPUMS-USA

2.4.3 Sectors

In our model, workers produce freely traded sectoral outputs indexed by σ that are used to produce the final good. In the international trade literature, it is common to interpret sectors in models of comparative advantage as industries. Recent work in both international and labor economics has emphasized a perspective focused on workers completing tasks, which empirical work has frequently operationalized as occupations (Grossman and Rossi-Hansberg, 2008; Acemoglu and Autor, 2011). We will implement empirical tests using each. We define sectors to be the 21 manufacturing industries in the three-digit stratum of the North American Industry Classification System (NAICS) or the 22 occupational categories in the two-digit stratum of the Standard Occupational Classification (SOC). We suspect that the assignment of workers to sectors is better characterized as assignments to occupations than assignments to industries, since virtually all industries employ both skilled and unskilled workers. Our measures of cross-sectoral variation in skill intensities in the following section are consistent with this conjecture.

[29] Individuals with doctorates typically earn less than individuals with professional degrees, so it may be inappropriate to treat PhDs as higher-ω individuals than professionals.

We measure industrial employment in a metropolitan area using data from the 2000 County Business Patterns. We measure occupational employment in a metropolitan area using estimates from the 2000 BLS Occupational Employment Statistics. See appendix B.3 for details.

2.4.4 Skill intensities

Our theory makes the strong assumption that $H(\omega, \sigma)$ is strictly log-supermodular so that sectors are ordered with respect to their skill intensities. In our empirical work, we infer sectors' skill intensities from the data using the observable characteristics of the workers employed in them. We use microdata from the 2000 Census of Population, which contains information about workers' educational attainments, industries, and occupations. We use the average years of schooling of workers employed in a sector as a measure of its skill intensity.[30] In doing so, we control for spatial differences by regressing years of schooling on both sectoral and city fixed effects, but we have found that omitting the city fixed effects has little effect on the estimated skill intensities. Table 2.2 reports the five least skill-intensive and five most skill-intensive sectors among both the 21 manufacturing industries and the 22 occupational categories. There is considerably greater variation in average years of schooling across occupational categories than across industries.[31] This may suggest that the "assignment to occupations" interpretation of our model will be a more apt description of the data than the "assignments to industries" interpretation.

[30] Autor and Dorn (2013) rank occupations by their skill level according to their mean log wage. Our assumption of absolute advantage is consistent with such an approach. Using average log wages as our measure of skill intensity yields empirical success rates comparable to and slightly higher on average than those reported in section 2.5. We use years of schooling rather than wages as our measure of sectoral skill intensities since nominal wages may also reflect compensating differentials or local amenities.

[31] The standard deviations of average years of schooling across occupational categories, industries, and manufacturing industries are 2.2, 1.0, and 0.9, respectively.

Table 2.2: Sectoral skill intensities

SOC	Occupational category	Skill intensity	NAICS	Manufacturing industry	Skill intensity
45	Farming, Fishing, and Forestry	9.3	315	Apparel	10.7
37	Building & Grounds Cleaning	10.9	314	Textile Product Mills	11.4
35	Food Preparation and Serving	11.4	316	Leather and Allied Product	11.7
47	Construction and Extraction	11.5	313	Textile Mills	11.7
51	Production	11.6	337	Furniture and Related Products	11.7
29	Healthcare Practitioners and Technical	15.6	312	Beverage and Tobacco Products	13.1
21	Community and Social Services	15.8	336	Transportation Equipment	13.2
25	Education, Training, and Library	16.5	324	Petroleum and Coal Products	13.5
19	Life, Physical, and Social Science	17.1	334	Computer & Electronic Products	14.1
23	Legal	17.3	325	Chemical	14.1

Source: Census 2000 microdata via IPUMS-USA.

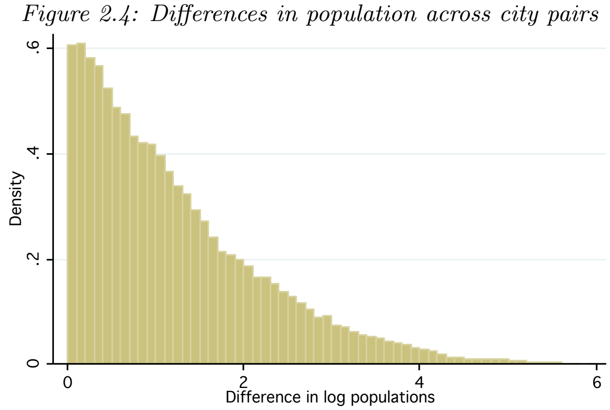
2.4.5 Pairwise weights

The most disaggregate implications of Propositions 2.1 and 2.2 are inequalities describing the number of individuals residing (employed) in two cities and two skill groups (sectors). Empirically testing these pairwise predictions involves evaluating as many as ten million of these inequalities and summarizing the results. An important question is whether each of these comparisons should be considered equally informative.

An unweighted summary statistic assigns equal weight to correctly predicting that Chicago (population 9 million) is relatively more skilled than Des Moines (population 456 thousand) and correctly predicting that Des Moines is relatively more skilled than Kalamazoo (population 453 thousand). Given the numerous idiosyncratic features of the real world omitted from our parsimonious theory, the former comparison seems much more informative about the relevance of our theory than the latter. Similarly, an unweighted summary statistic treats comparisons involving high school graduates (24 percent of the workforce) and comparisons involving PhDs (1 percent of the workforce) equally, while these differ in their economic import.

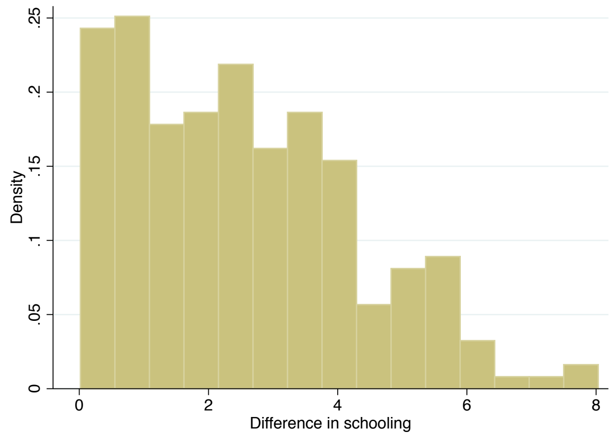
Following Trefler (1995), we report weighted averages of success rates in addition to unweighted statistics. In describing skill distributions, we weight each pairwise comparison by the two cities' difference in log population. When we consider nine skill groups, we also report a case where we weight by the product of the two skill groups' population shares. Figure 2.4 shows the distribution of differences in log population across city pairs. Since the majority of city pairs have quite small differences in log population, the unweighted and weighted statistics may yield substantially different results. In describing sectoral distributions, we weight pairwise comparisons by the two cities' difference in log population, the two sectors' difference in skill intensity, or the product of these. Figure 2.5 shows the distribution of differences in skill intensity across occupational pairs. While not as right-skewed as the distribution of differences in log population, this distribution may cause the unweighted and weighted statistics to differ. Figure 2.6 shows the distribution of differences in skill intensity across industries. The median difference between occupations is 2.3 years while the median difference between manufacturing industries is only 0.9 years. This relative compression in skill differences in the industrial data suggests that it may prove harder to make strong statements about differences across cities in industries than in occupations. Figures 2.5 and 2.6 underscore the importance of looking at weighted comparisons.

Figure 2.4: Differences in population across city pairs



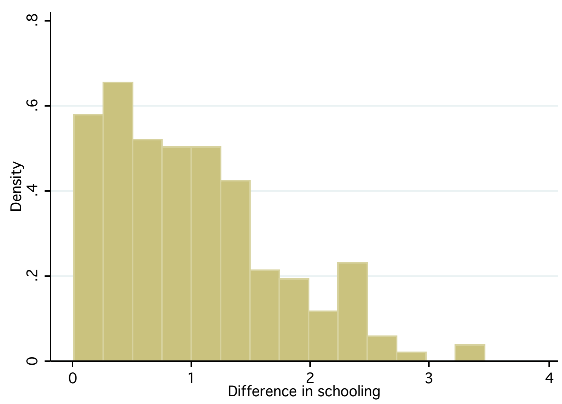
Data source: Census 2000 PHC-T-3

Figure 2.5: Differences in skill intensities across occupational pairs



231 pairwise comparisons between 22 2-digit SOC occupational categories.
Data source: 2000 Census of Population microdata via IPUMS-USA.

Figure 2.6: Differences in skill intensities across industrial pairs



210 pairwise comparisons between 21 3-digit NAICS manufacturing industries.
Data source: 2000 Census of Population microdata via IPUMS-USA.

2.5 Empirical results

In this section, we test our predictions relating cities' sizes to their distributions of skill, occupational employment, and industrial employment. First, we examine whether populations are log-supermodular in educational attainment and city size. This prediction is a much stronger characterization of cities' skill distributions than the well known fact that larger cities typically have a greater share of college graduates. Second, we examine whether the spatial pattern of sectoral employment is governed by this spatial pattern of skills. Our theory's predictions are more realistic than completely specialized or perfectly diversified cities and more specific than theories allowing arbitrary patterns of interindustry spillovers.

The data are broadly consistent with both of our novel predictions. Skill distributions regularly exhibit the monotone likelihood ratio property, though international migration plays an important role in the largest US cities that is omitted from our model. More skill-intensive sectors are relatively larger in more populous cities, on average. However, cities' sectoral distributions do not exhibit the monotone likelihood ratio property as often as cities' skill distributions do. One interpretation of this result is that skill-driven comparative advantage plays an important role in determining the spatial pattern of production, but localization and coagglomeration economies may also play a role.[32] We show that there are not systematic violations of our predicted pattern of comparative advantage.

2.5.1 Larger cities are relatively more skilled

This subsection tests our prediction that larger cities have relatively more skilled populations. We empirically describe skill abundance using the two tests described in section 2.4.1. We first do these exercises using three skill groups defined by educational attainment levels and then repeat them using nine very disaggregated skill groups.

2.5.1.1 Three skill groups

The elasticity test applied to the three skill groups across 270 metropolitan areas is reported in Table 2.3. The results match our theory's prediction that larger cities will have relatively more people from higher skill groups. The population elasticities are monotonically increasing in educational attainment and the elasticities differ significantly from each other. In anticipation of issues related to international immigration that arise when we examine nine skill groups, the second column of the table reports the population elasticities of US-born individuals for these three educational categories. The estimated elasticities are slightly lower, since foreign-born individuals are more concentrated in larger cities, but the differences between the elasticities are very similar.

[32] Since the employment data do not distinguish employees by birthplace, another possibility is that the disproportionate presence of low-skill foreign-born individuals in larger cities influences sectoral composition in a manner not described by our theory.

Table 2.3: Population elasticities of educational groups

Dep var: $\ln f(\omega, c)$	All	US-born
β_{ω_1} HS or less	0.96	0.90
$\times$ log population	(0.011)	(0.016)
β_{ω_2} Some college	1.00	0.97
$\times$ log population	(0.010)	(0.012)
β_{ω_3} BA or more	1.10	1.07
$\times$ log population	(0.015)	(0.017)

Standard errors, clustered by MSA, in parentheses.

Sample is all full-time, full-year employees residing in 270 metropolitan areas.

The pairwise comparison test examines ordered groups of cities to see if the relative population of the more skilled is greater in larger cities. Implementing this test involves defining groups of cities. Ordering cities by population, we partition the 270 metropolitan areas in our data into 2, 3, 5, 10, 30, 90, and 270 sets of cities. Making pairwise comparisons between three skill groups and as many as 270 metropolitan areas involves computing up to 108,945 inequalities.[33] Note that prior work typically describes a contrast between large and small cities for high and low skills, whereas our most aggregated comparison is between large and small cities for three skill groups.

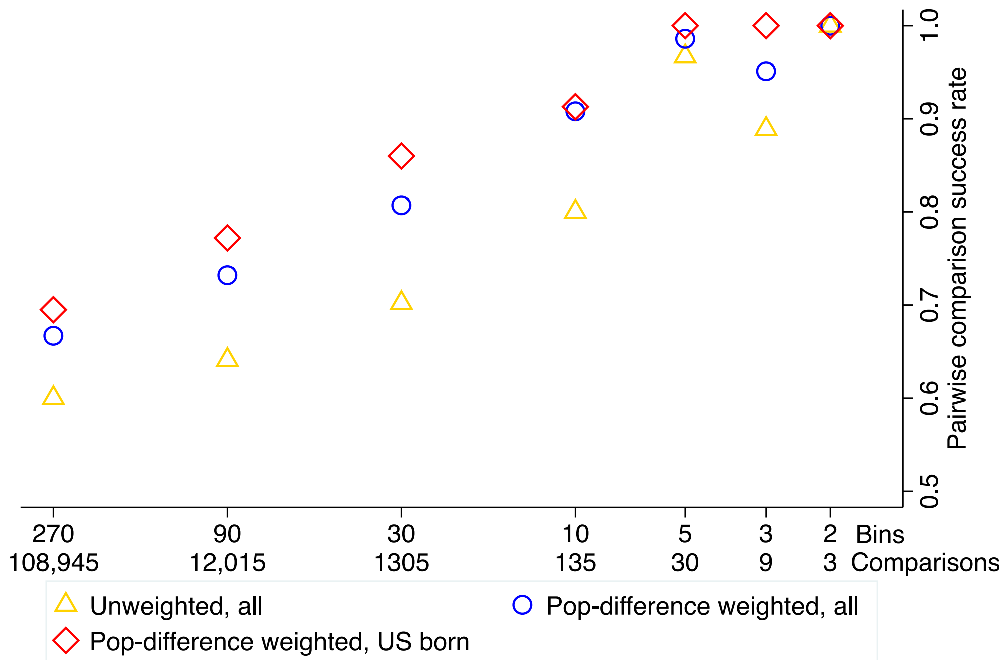
Figure 2.7 and Appendix Table B.1 summarize the results of these tests using various sets of cities, weights, and birthplaces. In the unweighted comparisons, the success rate ranges from 60 percent when comparing individual cities to 97 percent when comparing five groups of cities to 100 percent for the standard case of two groups of cities. Weighting the comparisons by the population difference generally yields a higher success rate.[34] When we weight by population differences, the success rate is 67 percent when comparing individual cities, 98 percent for five groups of cities, and 100 percent for the simple comparison of large versus small cities.[35]

[33]With n city groups and m skill groups, we make $\frac{n(n-1)}{2}\frac{m(m-1)}{2}$ comparisons. For example, $\frac{270\times269}{2}\frac{3\times2}{2} = 108,945$.

[34]Despite the fact that the success rate of the Des-Moines-Kalamazoo comparisons is actually higher than the Chicago-Des-Moines comparisons.

[35]Our comparisons of two or five groups of cities are analogous to the empirical exercises presented in Eeckhout, Pinheiro, and Schmidheiny (2011) and Bacolod, Blum, and Strange (2009).

Figure 2.7: Pairwise comparisons of three skill groups



2.5.1.2 Nine skill groups

We next examine our tests for the case with nine skill groups. Starting with the elasticity test, Table 2.4 shows, contrary to our model's predictions, that those not completing high school are highly prevalent in larger cities. The second column reveals that this result is due to the presence of foreign-born individuals with low educational attainment in larger cities. If we restrict attention to US-born individuals, we can only reject the hypothesis that $\beta_\omega \geq \beta_{\omega'} \iff \omega \geq \omega'$ in one of thirty-six comparisons, the case where $\beta_{\omega_2} = 0.94 > 0.90 = \beta_{\omega_3}$.[36] In short, the elasticity test provides strong support for our theory when we examine the US-born population.[37]

How should we interpret the difference between the spatial distribution of skills among the population as a whole and among US-born individuals? One possibility is that immigrants strongly prefer larger cities for reasons omitted from our model, causing less-skilled foreign-born individuals

[36]The estimated elasticities for the tails of the skill distribution have larger standard errors. This likely reflects greater sampling noise for scarce educational categories; for example, the median (C)MSA had 34 observations of full-time, full-year employees with a PhD in the 5 percent public-use 2000 Census microdata.

[37]Interestingly, among US-born individuals, the nine estimated elasticities naturally break into the three more aggregate educational attainment categories that we used above: $\beta_{\omega_1}, \beta_{\omega_2}, \beta_{\omega_3} \in (0.91, 0.94)$; $\beta_{\omega_4}, \beta_{\omega_5} \in (0.96, 0.98)$; $\beta_{\omega_6}, \beta_{\omega_7}, \beta_{\omega_8}, \beta_{\omega_9} \in (1.06, 1.09)$.

Table 2.4: Population elasticities of educational groups, 2000

Dep var: $\ln f(\omega, c)$	All	US-born
β_{ω_1} Less than HS	1.17	0.91
× log population	(0.039)	(0.028)
β_{ω_2} High school dropout	1.03	0.94
× log population	(0.017)	(0.020)
β_{ω_3} High school graduate	0.93	0.90
× log population	(0.013)	(0.016)
β_{ω_4} College dropout	1.00	0.98
× log population	(0.011)	(0.013)
β_{ω_5} Associate's degree	1.00	0.96
× log population	(0.014)	(0.016)
β_{ω_6} Bachelor's degree	1.10	1.07
× log population	(0.015)	(0.017)
β_{ω_7} Master's degree	1.12	1.09
× log population	(0.018)	(0.019)
β_{ω_8} Professional degree	1.12	1.09
× log population	(0.018)	(0.019)
β_{ω_9} PhD	1.11	1.06
× log population	(0.035)	(0.033)

Standard errors, clustered by MSA, in parentheses.

Sample is all full-time, full-year employees residing in 270 metropolitan areas.

to disproportionately locate in larger cities. This would be consistent with an established literature that describes agglomeration benefits particular to unskilled foreign-born individuals, such as linguistic enclaves (Edin, Fredriksson, and Aslund, 2003; Bauer, Epstein, and Gang, 2005).[38]

Eeckhout, Pinheiro, and Schmidheiny (2011) articulate another possibility, in which an economic mechanism they term "extreme-skill complementarity" causes less skilled individuals, foreign-born or US-born, to disproportionately reside in larger cities. Larger cities' benefits for immigrants serve as a "tie breaker" that causes the foreign-born to choose larger cities in equilibrium. This theory predicts that in the absence of foreign-born low-skilled individuals, US-born low-skilled individuals would disproportionately locate in larger cities.

We attempt to distinguish between these hypotheses by looking at the skill distributions of US cities two decades earlier. In 2000, foreign-born individuals were 11 percent of the US population, while in 1980 they constituted about 6 percent. More importantly, in 2000, foreign-born individuals

[38] Another potential mechanism is that immigrants may find larger cities' combination of higher nominal wages and higher housing prices more attractive than natives (Diamond, 2012), possibly because they remit their nominal incomes abroad or demand less housing than US-born individuals.

Table 2.5: *Population elasticities of educational groups, 1980*

Population share		Population elasticities All	US-born
0.06	β_{ω_1} Less than grade 9	0.99	0.89
	$\times$ log population	(0.028)	(0.030)
0.11	β_{ω_2} Grades 9-11	1.00	0.98
	$\times$ log population	(0.019)	(0.021)
0.33	β_{ω_3} Grade 12	0.97	0.95
	$\times$ log population	(0.013)	(0.015)
0.08	β_{ω_4} 1 year college	1.04	1.03
	$\times$ log population	(0.018)	(0.018)
0.13	β_{ω_5} 2-3 years college	1.09	1.07
	$\times$ log population	(0.018)	(0.018)
0.13	β_{ω_6} 4 years college	1.10	1.08
	$\times$ log population	(0.018)	(0.018)
0.13	β_{ω_7} 5+ years college	1.13	1.11
	$\times$ log population	(0.022)	(0.022)

Standard errors, clustered by MSA, in parentheses

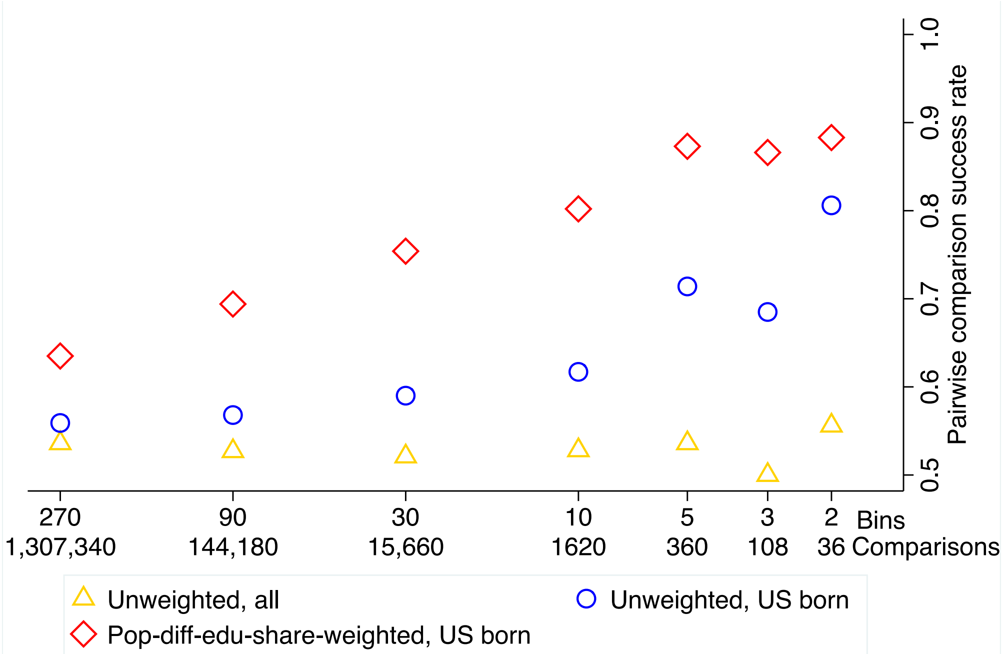
Sample is full-time, full-year employees residing in 253 metropolitan areas.

constituted nearly 80 percent of the lowest skill group, while in 1980 they were less than one third of the lowest skill group. If our hypothesis that foreign-born individuals are particularly attracted to larger cities is correct, then the population elasticity of less-skilled types should be lower when foreign-born shares are lower. Table 2.5 demonstrates that this is the case in 1980. It does not provide any evidence that the least skilled were overrepresented in larger cities in 1980, among either the population as a whole or US-born individuals.[39] Reconciling these results with the model of Eeckhout, Pinheiro, and Schmidheiny (2011) would require that the production function changed from top-skill complementarity in 1980 to extreme-skill complementarity in 2000.

We now turn to the pairwise comparisons for the case with nine skill groups in 2000, summarized in Figure 2.8 and Appendix Table B.2. These test inequalities for 36,315 individual city pairs for each pairing of the nine skill groups. In both the unweighted and weighted comparisons, our theory does best in predicting comparisons of skill groups that have a high school degree or higher attainment. Fewer than 50 percent of the comparisons yield the correct inequality when the "less than high school" skill group is involved in the comparison. As would be expected from the

[39] The educational categories in Table 2.5 differ from prior tables because Census microdata collected prior to 1990 identify coarser levels of educational attainment in terms of years of schooling rather than highest degree attained.

Figure 2.8: Pairwise comparisons of nine skill groups



previous results, these comparisons are considerably more successful when restricted to the US-born population. When these are also weighted by population differences and education shares, the overall success rate in comparing individual cities rises to 64 percent.

Figure 2.8 and Appendix Table B.3 show how the pairwise comparison success varies when we group cities by size. When we restrict attention to the US-born, the unweighted success rate respectively for individual cities, five and two groups of cities are 56 percent, 71 percent, and 81 percent. If, in addition, we weight successes by education shares and population differences, the success rates for individual cities, five and two groups of cities are 64 percent, 87 percent, and 88 percent, respectively. In short, for the case of nine skill groups, the raw comparisons for individual cities including the foreign born show very modest success. As in the elasticities test, restricting attention to the US-born population yields significant improvement. Likewise, there is considerably greater success as we group cities and as we weight them by the overall prevalence of the education group in the labor force. Overall, we consider this solid support for our theory.

2.5.2 Larger cities specialize in skill-intensive sectors

This section examines the spatial pattern of sectoral employment. In our theory, larger cities are relatively more skilled, cities' equilibrium productivity differences are Hicks-neutral, and sectors can be ordered by their skill intensity, so larger cities employ relatively more labor in skill-intensive sectors. We established that larger cities are relatively more skilled in section 2.5.1. We now examine whether larger cities are relatively specialized in skill-intensive sectors. Since employment levels in both industries and occupations are readily available in the data, we test the employment implications of Proposition 2.2.[40]

2.5.2.1 The spatial distribution of occupations

We first implement the elasticities test and the pairwise comparisons test interpreting sectors as occupations. We begin with a visualization of the elasticity results. Figure 2.9 plots the 22 occupational categories' estimated population elasticities of employment against their skill intensities, measured as the average years of schooling of individuals employed in that occupation.[41] There is a clear positive relationship. Outliers in the figure include close-to-unitary elasticities for the relatively skilled occupations in education, healthcare, and social services, which may reflect non-traded status. On the other side, computer and mathematical occupations have an elasticity that is quite high relative to their average schooling.

We can also look at this more formally. With the population elasticities of occupations in hand, the hypothesis that $\beta_\sigma \geq \beta_{\sigma'} \iff \sigma \geq \sigma'$ involves 231 $(= 22 \times 21/2)$ comparisons of the estimated coefficients.[42] This hypothesis is rejected at the five-percent significance level in 46 comparisons, so the success rate is 80 percent.

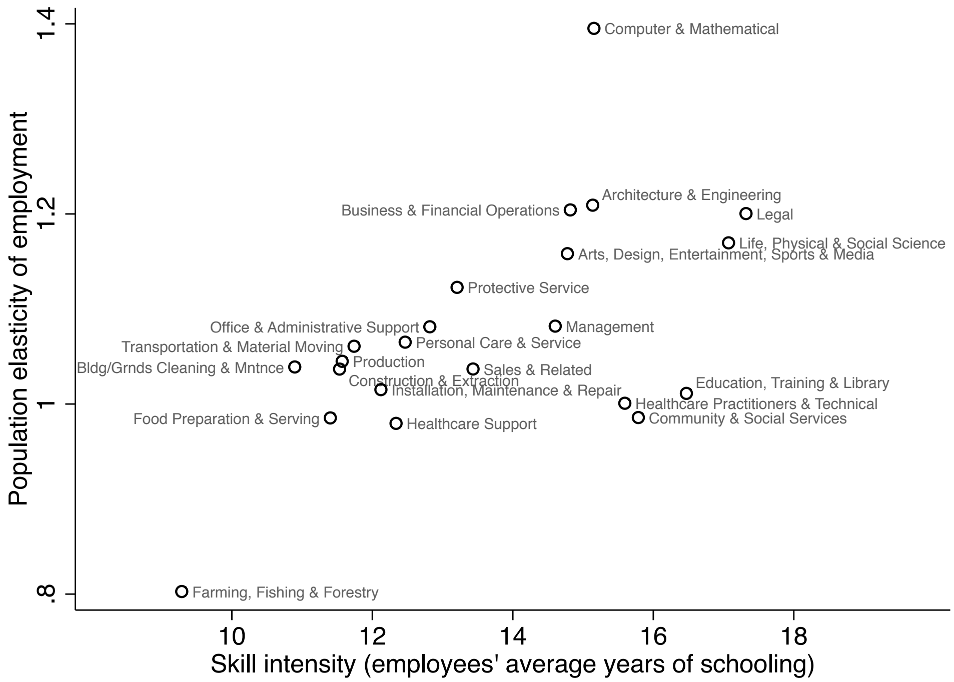
The results for pairwise comparisons for occupations appear in Figure 2.10 and Appendix Table B.5. When we do this for 276 cities and 22 occupations, we have a total of 8,766,450 pairwise

[40] Section 2.5.1 showed that US-born individuals better match our model's predictions about the distribution of skills. Unfortunately, the County Business Patterns and Occupational Employment Statistics data describe employment counts, not individual employees' characteristics, so we cannot address the birthplace issues in this section.

[41] These elasticities are estimated without including zero-employment observations. The results obtained when including those observations are similar.

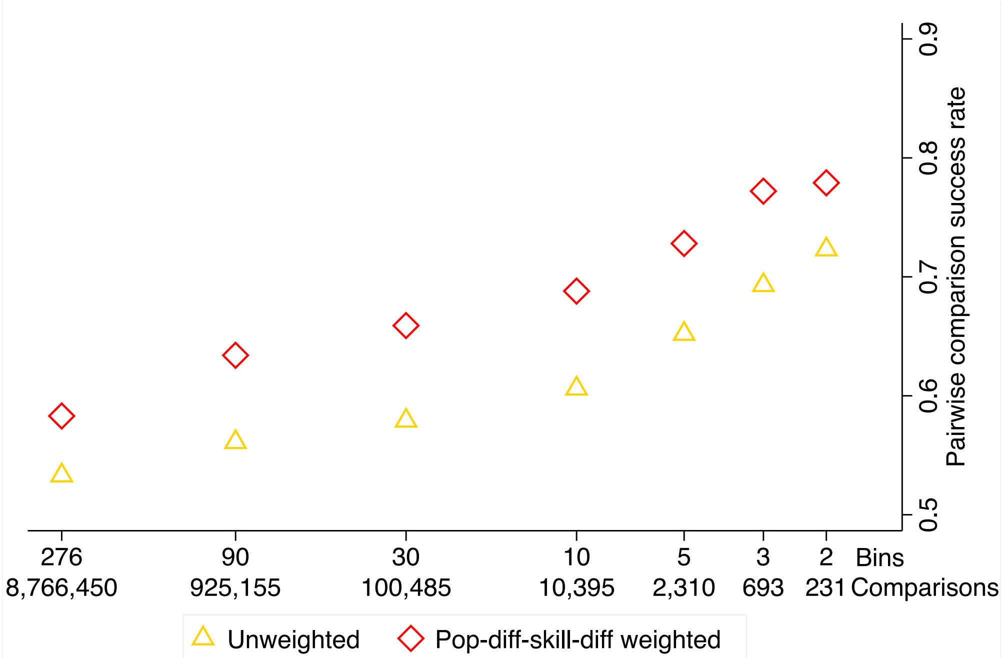
[42] The elasticity estimates appear in Appendix Table B.4.

Figure 2.9: Occupations' population elasticities and skill intensities



comparisons, of which 53 percent are correct. This is low compared to our results for skills. When we stay with individual cities but weight by population and skill differences, this rises above 58 percent. We can maintain the weighting and consider it for cities grouped by size into, for example, 30, 5, or 2 groups. The corresponding proportion of successes rises respectively to 66, 73, and 78 percent. While the results for occupations are not as strong as the results for skills, there are nonetheless quite informative patterns – even when we group cities into five size-based bins, we get nearly three-fourths of the pairwise comparisons correct across the 22 occupational categories.

Figure 2.10: Pairwise comparisons of 22 occupational categories



2.5.2.2 The spatial distribution of industries

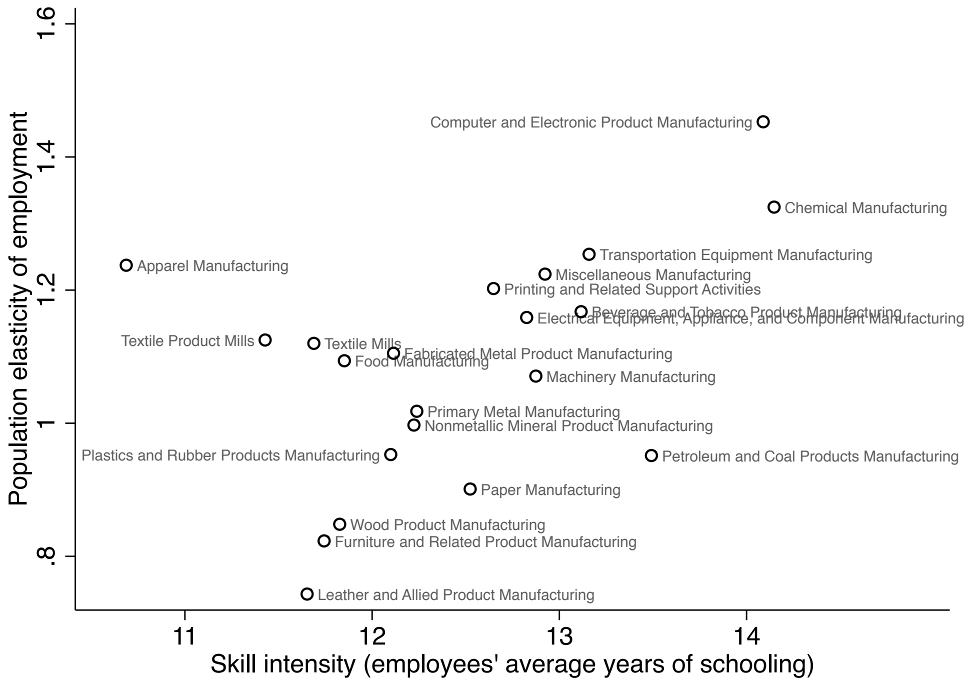
We now implement the elasticities test and the pairwise comparisons test interpreting sectors as manufacturing industries.[43] A visualization of the elasticity test appears in Figure 2.11. Again, as predicted by our theory, there is a clear positive relationship so that the population elasticity of industry employment is rising with the skill intensity of the industry. The apparel industry is an outlier, with low average education and a high population elasticity of employment. This may reflect the share of apparel industry employees who are less-skilled foreign-born individuals, consistent with our previous discussion of skills. Testing the hypothesis that $\beta_\sigma \geq \beta_{\sigma'} \iff \sigma \geq \sigma'$ for the 21 manufacturing industries involves 210 ($= 21 \times 20/2$) comparisons of these estimated elasticities.[44] This hypothesis is rejected in 26 comparisons, so the elasticity implication holds true for manufacturing industries about 87 percent of the time.[45] This success rate is higher than the

[43]We focus on manufacturing industries since we believe they have the lowest trade costs, but we have found broadly similar results when using all industries.

[44]The elasticity estimates appear in Appendix Table B.6.

[45]If we restrict the data to uncensored observations, which reduces the sample considerably, this hypothesis is rejected in 32 comparisons, for an 85 percent success rate. See appendix B.3 for a discussion of censoring in County Business Patterns data.

Figure 2.11: Industries' population elasticities and skill intensities

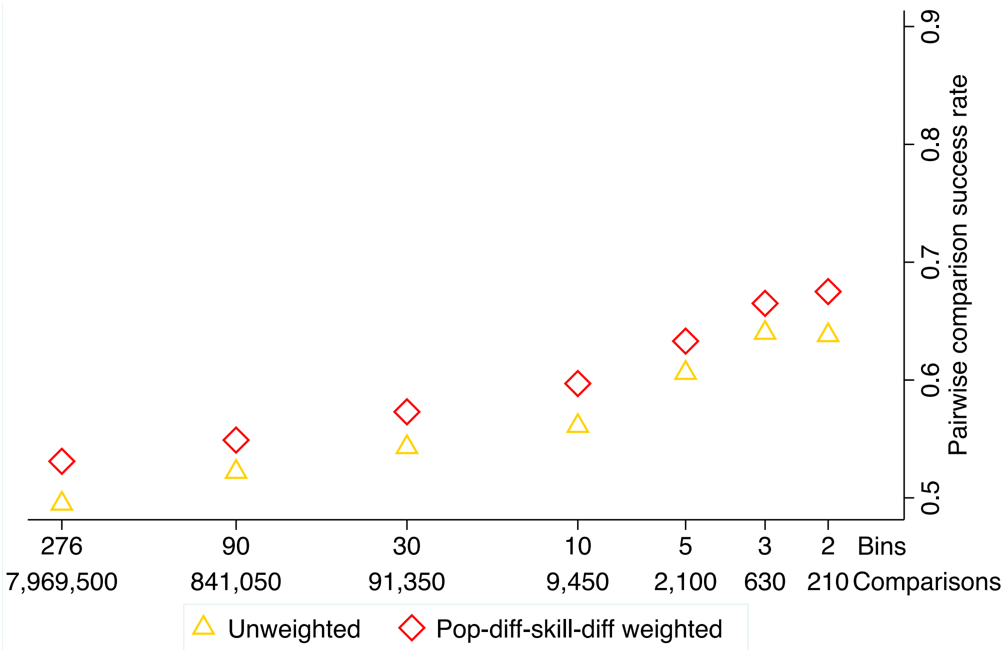


corresponding statistic for occupational elasticities.

The pairwise comparisons results for industries appear in Figure 2.12 and Appendix Table B.7. When we do this for 276 individual cities and 21 industries, we have a total of nearly 8 million pairwise comparisons, of which just half are correct. Weighting this by skill and population differences raises this to 54 percent, again low compared to our results for pairwise comparisons of skills. We can maintain the weighting and consider this for cities grouped by size into 30, 5, or 2 groups. The corresponding proportion of successes rises respectively to 58, 66, and 71 percent. These are low relative to the prior results on occupations and even more so relative to the results on skills. Nonetheless, they do show that there is systematic variation across cities of different sizes in the composition of manufacturing.[46] Note that prior work contrasting large and medium-size cities, Henderson (1997), is analogous to our comparisons of two or three groups of cities ordered by population.

[46]These results are not driven solely by the largest metropolitan areas; excluding the ten largest cities from pairwise comparisons of occupations and industries yields similar success rates.

Figure 2.12: Pairwise comparisons of 21 manufacturing industries



2.5.3 Testing for systematic failures of comparative advantage

Our results for the cross-city distributions of skills, industries, and occupations demonstrate systematic patterns in line with our theory's predictions. While demonstrating predictive power, the pairwise comparisons also fall well short of 100 percent success. This is not surprising, given that our model's parsimony stems from making strong assumptions that omit various features that influence the real world. An important question is whether our theory's unsuccessful pairwise predictions are merely idiosyncratic deviations from the pattern of comparative advantage or are systematic violations of our predicted pattern.

Sattinger (1978) develops an approach to test for such systematic violations in the form of systematic intransitivity in the pattern of comparative advantage. It is possible for the data to exhibit, for $c > c' > c''$ and $\sigma > \sigma' > \sigma''$, $\frac{f(\sigma,c)}{f(\sigma',c)} \geq \frac{f(\sigma,c')}{f(\sigma',c')}$ and $\frac{f(\sigma',c')}{f(\sigma'',c')} \geq \frac{f(\sigma',c'')}{f(\sigma'',c'')}$ without exhibiting $\frac{f(\sigma,c)}{f(\sigma'',c)} \geq \frac{f(\sigma,c'')}{f(\sigma'',c'')}$. With hundreds of metropolitan areas and dozens of sectors, it is easy to find three cities and three sectors in the data exhibiting such intransitivity. But do intransitivities arise systematically? Sattinger (1978) shows that if $\ln f(\sigma, c)$ is a polynomial function of $\hat{\beta}_\sigma$ and $\ln population_c$, then there can be systematic intransitivity only if $\ln f(\sigma, c)$ is a function of higher-order interactions of $\hat{\beta}_\sigma$ and $\ln population_c$. We therefore added quadratic terms and their interactions to our elasticity regressions. These did little to improve the regression's adjusted R^2, and F-tests yielded p-values that did not come close to rejecting the null that these additional terms were uninformative. There is no evidence of systematic intransitivity in comparative advantage. While our theory's predictive successes are systematic, the empirical departures from our theory appear to be idiosyncratic.

2.6 Discussion and conclusions

In this paper, we introduce a model that simultaneously characterizes the distribution of skills and sectors across cities. We describe a high-dimensional economic environment that is a system of cities in which cities' internal geographies exhibit substantive heterogeneity and individuals' comparative advantage governs the distribution of sectoral employment. Our model achieves two aims. First, we obtain "smooth" predictions, in the sense that cities' skill and sectoral distributions will be highly overlapping. These are more realistic than prior theories describing cities that are perfectly sorted along skills or polarized in terms of sectoral composition. Second, we obtain "strong" predictions, in the sense that cities' skill and sectoral distributions will exhibit systematic variation according to the monotone likelihood ratio property. These are more precise than the predictions of many prior theories of the spatial organization of economy activity and guide our empirical investigation.

Examining data on US metropolitan areas' populations, occupations, and industries in the year 2000 reveals systematic variation in the cross-city distribution of skills and sectors that is consistent with our theory. Larger cities are skill-abundant. Our results using three equal-sized categories of educational attainment are quite strong. Even disaggregated to nine educational categories, the cross-city distribution of US-born individuals is well described by our theory.

Empirically, we find that larger cities specialize in skill-intensive activities. More skill-intensive occupations and industries tend to have higher population elasticities of employment. In making pairwise comparisons, our model does better in describing the pattern of occupational employment than industrial employment. This is consistent with a recent emphasis in the literature on workers performing tasks. Our results demonstrate that metropolitan skill distributions shape the comparative advantage of cities, though the fact that our sectoral-employment predictions do not perform as well as our skill-distribution predictions is consistent with the idea that localization or coagglomeration economies omitted from our model may also be important.

We believe that our framework is amenable to both theoretical and empirical applications and extensions. The “smoothness” resulting from the simultaneous consideration of cross- and within-city heterogeneity in a continuum-by-continuum environment would make our model amenable to theoretical analyses of the consequences of commuting costs, globalization, and skill-biased technical change. The “strong” character of our predictions and their demonstrated relevance for describing US cities in 2000 suggest that their examination in other settings, such as economies at different stages of development or in different historical periods, would be interesting.

Chapter 3

Consumer Cities in General Equilibrium

Jonathan I. Dingel[1]

[1]I am grateful to Don Davis for invaluable discussions and guidance. I thank Jessie Handbury, Corinne Low, David Munroe, Eric Verhoogen, Jonathan Vogel, and participants in the Columbia University applied microeconomics and international trade colloquia for helpful comments.

3.1 Introduction

By many accounts, big cities are attractive places to reside. A best-selling popular account of modern cities, Florida (2002), says that metropolitan areas' fortunes depend on the quality of life they offer innovative individuals. News stories describe firms pressured to move to denser metropolitan areas because "a suburban location has become a liability in recruiting the best and brightest young bankers, who want to live in Manhattan or Brooklyn, not in Stamford."[2] Glaeser, Kolko, and Saiz (2001) describe the various attractions of "consumer cities" and argue that "the future of cities increasingly depends on whether cities are attractive places for consumers to live."

But urban economists have concluded that large cities are unattractive to consumers. Their conclusion follows from an indifference condition that is at the heart of spatial-equilibrium theory, originally laid out in Roback (1982). In spatial equilibrium, the advantage of larger cities' higher nominal wages must be offset by a consumption disadvantage so that individuals are indifferent between larger and smaller cities.[3] If larger cities offered both higher nominal wages and lower amenity-adjusted local prices, no one would live in smaller cities. Empirically, individuals with similar demographic and educational characteristics earn higher wages in larger cities. This observation causes Glaeser and Gottlieb (2009, p.1000) to write that "the spatial equilibrium model allows us to easily reject the view that consumer amenities are the primary force driving urban concentration in the United States."

This chapter revisits that conclusion and shows that the consumption motive can play a first-order role in spatial variation in wage distributions when individuals are heterogeneous. I introduce a general-equilibrium model in which larger cities' consumption benefits explain their higher nominal incomes. Cities' characteristics are endogenous outcomes of individuals' locational choices. Larger cities offer a greater variety of local goods and services. They also exhibit higher housing prices due to congestion costs. Heterogeneous consumers have non-homothetic preferences. Higher-income consumers spend a larger fraction of their income on local goods and services and a

[2] Charles V. Bagli, "Regretting Move, Bank May Return to Manhattan," *New York Times*, 8 June 2011. See also Alejandra Cancino, "In search for talent, companies relocating to downtown Chicago," *Chicago Tribune*, 31 July 2011.

[3] Albouy (2012) shows that the implied consumption disadvantages are overstated if one fails to account for federal taxes and non-labor income.

smaller fraction of their incomes on housing. This causes higher-income individuals to find larger cities relatively more desirable places to locate. This consumption motive induces spatial sorting: higher-income individuals tend to live in larger cities.

The model demonstrates that the inference that agglomeration is driven by production stems from assumptions about the homogeneity of individuals, not spatial equilibrium. Differences in consumption opportunities are the model's only force for urban concentration, yet an econometrician analyzing model-generated data in which individual skills are imperfectly observed would find a positive city-size wage premium due to differences in the composition of cities' populations. The more populous city has higher nominal income per capita and higher housing prices. The larger city's skill distribution stochastically dominates that of the smaller city, and its nominal wage distribution inherits this property. Consumption motives alone could explain a large body of facts about the spatial pattern of wages and prices.

I develop several new facts about the spatial choices of retirees to demonstrate that the consumer-cities hypothesis is more than a theoretical curiosity. Since retirees do not work, their locational choices must reflect consumption concerns. Yet the data do not reveal an exodus of the recently retired from larger cities. Moreover, more educated retirees are more likely to live in larger cities. Theories positing a skill-size complementarity in production cannot explain such a pattern amongst those not producing. This evidence suggests that an account of cities omitting consumption motives is incomplete.

In isolating the role of non-homothetic preferences in explaining the spatial allocation of heterogeneous talent, my consumer-cities theory complements the theories of Behrens, Duranton, and Robert-Nicoud (2012) and Davis and Dingel (2012) in describing systems of cities with heterogeneous individuals in general equilibrium.[4] A number of authors (Glaeser, Kolko, and Saiz, 2001; Combes, Duranton, and Gobillon, 2008; Lee, 2010) have suggested the relevance of consumption motives, but their role has thus far not been formalized in a general-equilibrium setting. The model

[4]The first draft of this chapter was completed in April 2011, when it shared Columbia economics' Vickrey Prize. Contemporaneous with my research, Behrens, Duranton, and Robert-Nicoud (2012) developed an extension of their theory that is related to my model. See appendix H in the June 2012 version of their working paper. Behrens, Duranton, and Robert-Nicoud (2012) do not note the implications of the consumer-cities hypothesis for interpreting empirical wage patterns nor present empirical evidence distinguishing between consumption- and production-motivated agglomeration.

presented here starkly assumes *no* production-related agglomeration mechanisms in order to examine the extent to which empirical patterns in systems of cities could be explained by consumption motives alone. In reality, both production and consumption forces likely explain larger cities' higher average wages and skills. Silicon Valley is more than a case of correlated, geeky consumer tastes. The value of offering a competing, consumption-driven explanation for the stylized facts established by empirical work is that it forces us to revisit our interpretation of empirical evidence and consider the mechanisms' differing welfare and policy implications. This chapter is a first step towards identifying and understanding the relative importance of consumption and production motives in systems of cities.

Section 3.2 describes how my theory unifies emerging literatures on spatial sorting and consumption motives. Section 3.3 describes empirical patterns that are puzzling for theories neglecting the benefits of agglomeration for consumption purposes. Section 3.4 formalizes a parsimonious, general-equilibrium model in which consumption motives alone produce a system of cities in which larger cities exhibit higher nominal wages and more skilled populations. Section 3.5 concludes.

3.2 Related literature

This chapter unifies two recent lines of thought in the systems-of-cities literature. An empirical line of work examines the role of labor heterogeneity in determining the spatial pattern of economic outcomes, particularly the degree to which spatial sorting causes differences across individuals to result in differences across places. The second line of literature considers the benefits that cities provide to individuals as consumers rather than producers. The model presented here uses the consumption motive, in the form of differences in willingness to pay for access to cities' local goods and services, to explain the sorting of persons with different characteristics across different cities.

3.2.1 Cities' incomes and populations

Nominal wages are higher in larger cities. This section describes recent empirical work showing that a substantial share of this spatial variation in incomes is attributable to spatial variation in individuals' characteristics.

Individuals living in larger cities exhibit meaningful differences in observable characteristics (Bacolod, Blum, and Strange, 2009; Glaeser and Resseger, 2010). Perhaps the most obvious is spatial variation in educational attainment. A greater fraction of the population has a university degree in larger cities, and a greater fraction of these university graduates have masters or professional degree in larger cities.[5] Since individuals with greater educational attainment have higher average incomes, these population differences generate income differences across cities.

Recent empirical work shows that individuals living in larger cities tend to also have unobservable characteristics that are correlated with higher nominal income. These studies identify spatial sorting on unobserved characteristics by using longitudinal data and wage regressions with individual fixed effects.[6] They suggest that a meaningful share of larger cities' higher nominal incomes are attributable to the composition of their populations. Gibbons, Overman, and Pelkonen (2013), using longitudinal data on UK individuals, suggest that the spatial distribution of individual characteristics accounts for fourth-fifths or more of spatial wage disparities. They also show that there is a positive correlation between area effects and individual characteristics associated with higher wages. Combes, Duranton, and Gobillon (2008), using population variables from more than a century earlier to instrument for employment area density and land area, estimate a coefficient on density that is half that usually found in the literature. They show that aggregating their French data to the area level without controlling for sorting nearly doubles the coefficient on density and therefore stress the "failure of previous literature to control properly for unobserved individual heterogeneity." In discussing the relative importance of the problems that both the quantities and qualities of workers are endogenous, Combes, Duranton, Gobillon, and Roux (2010) conclude that "the sorting of workers across places is a quantitatively more important issue than their indiscriminate agglomeration in highly productive locations."

Why do more skilled, higher-income individuals tend to locate in larger cities? There is no standard theory predicting the positive skill-size relationship. In concluding their empirical work,

[5] In the Census 2000 microdata, amongst the full-time, full-year employed population ages 25-60, a doubling of metropolitan population is associated with a 2.4 percentage-point increase in the share with a bachelor's degree or higher and a 0.4 percentage-point increase in the share of bachelor's degree holders with a master's or professional degree.

[6] The identifying assumption underlying the use of individual fixed effects is that movement is random conditional on observable characteristics.

Combes, Duranton, and Gobillon (2008) suggest three possible explanations:

1. The most talented go to the largest towns to find scope for their abilities.

2. Workers learn more in larger cities.

3. Larger cities offer amenities that appeal more to workers earning higher wages.

The first explanation has been formalized by Nocke (2006) in partial equilibrium and Behrens, Duranton, and Robert-Nicoud (2012) in general equilibrium. The second explanation has recently been explored by Davis and Dingel (2012). This chapter introduces a model of the third mechanism, in which larger cities' local consumption opportunities appeal more to higher-income individuals.

3.2.2 Consumer cities and non-homothetic preferences

Recent work in urban economics has drawn attention to the role of cities as centers of consumption. Glaeser, Kolko, and Saiz (2001) argue that urban economics has often assumed "that cities are good for production and bad for consumption", thereby neglecting an important dimension of city life. The authors discuss four types of urban consumption attractions – variety of goods and services, aesthetics and physical setting, good public services, and transport speed – and provide evidence suggesting that these city elements influence individuals' locational choices.

Two prior theories describe the spatial allocation of heterogeneous workers when locational choices are motivated by consumption concerns.

Gyourko, Mayer, and Sinai (2013) present a two-city model in which individuals' productivities and nominal incomes are location-invariant in order to study spatial skewness in housing prices and incomes. The two cities differ exogenously in their elasticities of housing supply. Consumers at every income level have idiosyncratic tastes for one city or the other. Higher-income consumers exhibit greater willingness to pay for their preferred location. In the location with a less elastic housing supply, high rents cause lower-income consumers to move away. This "superstar city" has a nominal income distribution that stochastically dominates that of the city with a more elastic housing supply.

The model presented here differs in at least two important respects. First, its asymmetric outcomes are emergent, since its locational fundamentals are symmetric. The model shows how consumer spending acts as an agglomeration force by increasing the variety of local goods and services.[7] Heterogeneous consumers' differing willingness to pay for access to this variety of local goods and services sustains an equilibrium with cities of different sizes. Second, it exhibits the positive correlations between cities' populations, housing prices, and nominal income distributions found in the data. In the Gyourko, Mayer, and Sinai (2013) model, the "superstar city" with high housing prices and high nominal incomes may be more or less populous than the other city.

Lee (2010) introduces a partial-equilibrium model with exogenous city characteristics to motivate empirical work that examines whether more skilled individuals' nominal wages rise less quickly with city size. Like the model presented here, Lee (2010) posits that higher-income individuals spend a larger fraction of their budget on differentiated local goods and services relative to housing. His Proposition 1 states that if all individuals are indifferent across cities, more skilled individuals' nominal incomes will increase with city size less quickly. Lee (2010) claims that this prediction is unique to the consumer-cities hypothesis. It is not. As Black, Kolesnikova, and Taylor (2009) show, skill premia will be negatively correlated with housing prices whenever the income elasticity of housing demand is less than one and all individuals are indifferent across cities. This prediction follows from the spatial-indifference condition, regardless of the agglomeration force.

Empirically, cities' skill premia and population sizes tend to be positively related. Lee (2010) finds a negative correlation amongst individuals employed in medical occupations, but this pattern does not hold more broadly. Davis and Dingel (2012) show that college graduates' wages increase with metropolitan size faster than high school graduates' wages, so that there is a positive correlation between college premia and city sizes. Bacolod, Blum, and Strange (2009) show that high school graduates' wages increase with metropolitan size faster than high school dropouts' wages. Thus, theories predicting negative premia-size correlations, regardless of the agglomeration mechanism, are generally not empirically relevant.

The model put forth in this chapter differs from the model in Lee (2010) in three important respects. First, it is general equilibrium, so housing prices, goods prices, and the measure of varieties are determined by market-clearing conditions. In Lee (2010), healthcare workers produce

[7] This love-of-variety mechanism has a long history in economic geography. See Krugman (1991b) and Handbury and Weinstein (2011).

medical services that aren't consumed and consume local varieties that aren't produced. Second, city characteristics are endogenous, so the system-of-cities phenomena of interest result from the posited economic mechanisms. Third, some individuals are inframarginal residents, so the Black, Kolesnikova, and Taylor (2009) result does not apply and skill premia are not necessarily negatively related to city size. I provide examples of parameter values such that the model's equilibrium exhibits a positive size-premium relationship.

In sum, while urban economists have drawn attention to consumption motives in recent work, their potential importance has been understated. The following section presents patterns in the spatial distribution of retirees suggesting that consumption motives are empirically relevant. The model presented in the subsequent section demonstrates how, in a general-equilibrium setting, consumption motives alone could explain the cross-sectional spatial pattern of skills and wages.

3.3 Empirical evidence from retiree populations

In this section, I examine the spatial distribution of retiree populations.[8] I document three empirical patterns that are puzzling for the traditional production-motive conclusion but consistent with the consumer-cities hypothesis. Both the canonical spatial-equilibrium model and the consumer-city model presented here are static, so they do not formally describe life-cycle dynamics. But their contrasting economic mechanisms have contrasting implications for the behavior of retirees, who finance their consumption from location-independent wealth.

3.3.1 Contrasting predictions

If the net benefits of agglomeration stem from production advantages, retirees should find large cities unattractive. The canonical spatial-equilibrium model with homothetic preferences says that the higher wages of workers in larger cities compensate them for the combined effect of higher local prices or lower consumer amenities. When workers retire, the wage-earning motivation for

[8] An older literature studied aggregate net retiree migration flows, neglecting households' characteristics (Graves and Waldman, 1991; Clark and Hunter, 1992; Duncombe, Robbins, and Wolf, 2001). In a recent article, Chen and Rosenthal (2008) use Census microdata to study how households' migration decisions vary with their educational attainments. I discuss their findings below.

locating in large cities ceases.[9] The theory then predicts that all retirees should move to the city with the *lowest* nominal wages, since the spatial indifference condition reveals this location to have the best combination of consumer amenities and local prices.[10] This prediction also holds in production-driven agglomeration theories with heterogeneous labor, such as Behrens, Duranton, and Robert-Nicoud (2012) and Davis and Dingel (2012). If larger cities are attractive because they improve individuals' skills or allow individuals to apply their skills to a larger market, these benefits are irrelevant to retirees. Retirees of all skill levels will move to the location where workers' wages are lowest.

If cities' income and skills reflect consumption-driven locational decisions, some retirees should find large cities attractive. In the consumer-city model presented in this chapter, all individuals' incomes are location-independent. High-wage individuals tend to live in large cities because they find large cities' combination of high housing prices and low variety-adjusted non-tradables prices relatively attractive. Low-wage individuals spend a larger share of their income on housing, so they tend to prefer small cities. Upon retirement, individuals have no reason to change location because they consume out of permanent income and locations' consumption opportunities do not depend upon labor-force participation. Retirees with higher expenditure levels find larger cities relatively more attractive.

Empirical reality falls between these two extreme predictions, since some retirees do change locations upon retirement but not all move to the city with the lowest nominal wage. A more appropriate inquiry is to ask which theory better describes the data, since each necessarily omits many aspects. To do so, consider introducing idiosyncratic moving costs and idiosyncratic permanent-income shocks that are both orthogonal to individuals' and cities' characteristics. These weaken the theories' sharp predictions while allowing the posited economic mechanisms to drive the central tendencies in the data.

The weakened predictions of the production-driven agglomeration theory are that individuals in large cities tend to relocate upon retirement, they move to locations with lower nominal wages,

[9] Albouy (2012) analyzes the importance of location-independent income in Rosen-Roback exercises. He identifies the location-dependent income share as the 75% of average household income that is labor income.

[10] This logic dates at least to Henderson (1974, p.643), who says that if capital owners do not work as laborers, they will "avoid the high cost of living or housing in cities... by living in the countryside."

and individuals of all skill levels have equal reason to relocate. In the frictionless theory, all retirees would move to the location with the lowest nominal wages. Idiosyncratic bilateral moving costs mean that not all retirees will move and not all movers will go to the same location. Idiosyncratic income shocks do not alter the relative desirability of destinations. All retirees' moves will be shifts to locations with lower nominal wages than their city of origin, since these destinations have lower amenity-adjusted local prices. Since moving costs and income shocks are orthogonal to individuals' characteristics, the profile of destinations should be common across skill groups.

The weakened predictions of the consumer-cities theory are that retirees will sometimes move, they will move to both more and less populous destinations, and higher-income retirees should find larger cities relatively more attractive for the same reason that higher-income working individuals find larger cities attractive. In the frictionless case, individuals have no reason to relocate upon retirement because they consume out of permanent income, just as they did while working. The introduction of idiosyncratic moving costs does not change this. But income shocks will cause some retirees to move when the benefit of moving to a better location for consumption exceeds their bilateral moving cost. Since unexpected income shocks are mean zero, the destination profile of retirees inherits the characteristics of the income shocks. For example, if the income-shock distribution is symmetric with a median of zero, then moves to locations with a higher average nominal wage than the individual's current location are as likely as moves to locations with a lower nominal wage. Finally, since higher-income individuals find larger cities more attractive, this means, conditional on changing metropolitan areas, more educated retirees are more likely to select larger cities as their destination.

Thus, the production-driven-agglomeration and consumer-cities hypotheses make contrasting predictions about the frequency of retiree relocations, the contrast between the characteristics of a mover's origin and destination cities, and the destination profiles of different skill groups.

Comparing employed and retired individuals' behavior to distinguish between consumption- and production-driven agglomeration assumes that the most prominent difference between these two groups is the absence of the wage-earning motivation to locate in large cities. Is this a reasonable assumption? Workers and retirees are similar in terms of their levels of consumption. Hurst (2008) shows that consumption expenditure changes little for most individuals upon retirement, except for a decline in work-related expenses and a shift towards home-produced meals. Are workers and retirees similar in terms of their consumer tastes? Consumers' concerns likely do change with age. For example, older individuals may be more sensitive to unpleasant weather conditions or less concerned with varieties of local nightlife than younger counterparts. To the extent that tastes differ, this empirical comparison will be less informative. But there seems little reason to believe that older individuals find larger cities' consumption opportunities much more attractive than younger individuals do, the rationalization necessary to accept the retiree evidence below while denying the relevance of the consumer-cities hypothesis for the working-age population.

3.3.2 Empirical evidence

Three pieces of evidence show that the behavior of retirees is closer to the prediction of the consumer-cities theory than the prediction of the production agglomeration theory. First, the majority of individuals do not change residences upon retiring. Second, retirees who do move across metropolitan areas do not show a strong tendency to relocate to metropolitan areas with smaller populations or lower nominal wages. Third, more educated retirees, including those who move across metropolitan areas, more frequently reside in metropolitan areas with larger populations and higher nominal wages.

I use microdata from the 2000 US Census of Population (Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek, 2010) describing individuals age 60 or older who are not in the labor force, who I henceforth refer to as "retirees".[11] The Census asks individuals about their current residence and their residence five years prior. This allows me to study both the location of retirees and retirees' residential moves. The Census also reports non-employed individuals' previous occupations if they worked sometime during the five years prior, allowing me to attempt to identify those retirees who recently exited the labor force.

First, most retirees continue to reside in the same metropolitan area as when they were working. 95% of all retirees live in the same metropolitan area as five years prior. Amongst those who exited

[11] I use 60 as the age cutoff because there is a notable acceleration of the "not in labor force" rate at this age. The results of analyzing individuals age 65 or older are very similar. The not in labor force rates are 42% at 60 and 70% at 65.

the labor force in the last five years, 92.5% reside in the same metropolitan area as five years prior. Amongst retirees who recently exited the labor force and changed their residential domicile in the previous five years, only 30% changed their metropolitan area of residence.[12]

The fact that the majority of individuals do not change residences upon retiring is compatible with two starkly different explanations. The first is the consumer-cities hypothesis put forward in this chapter. The second is that the costs of relocation are quite high, so individuals have limited mobility. If the latter were true, the low rate of retiree relocations would not be evidence in favor of the consumer-cities hypothesis. While most urban economists believe that the US population is relatively mobile, in the rest of the empirical work that follows, I condition on individuals changing the metropolitan area in which they reside and use the behavior of these movers to distinguish between the consumption-driven and production-driven agglomeration hypotheses.

The second piece of evidence is that retirees who move across metropolitan areas do not tend to relocate to small, low-wage cities. Their locations generally mirror those of the population as a whole. Considering only retirees who changed metropolitan areas in the prior five years, a regression of log retiree population on log working population yields an elasticity of .83 and an R^2 of .68. While this elasticity is lower than that obtained for the retiree population as a whole (.93), it is still much closer to the consumer-cities hypothesis of unity than the negative correlation predicted by the canonical theory.

The data do not reveal a strong tendency for retirees to move from large cities to small cities. Of retirees who changed metropolitan areas between 1995 and 2000, 62.7% moved from a more populous metropolitan area to a less populous location. If individuals chose their destinations randomly, 50.7% of retiree moves would have been from larger to smaller cities.[13] Therefore, the data reveal a tendency for retirees to move from more populous to less populous cities. But the claim that the net benefits of agglomeration favor production over consumption implies that everyone will move to smaller cities upon retirement, and retiree locational changes fall far short of that pattern.

[12]77% of all retirees and 74% of recent retirees live in the same residence as five years prior.

[13]I assume that individuals chose their destinations randomly, with each potential destination assigned a probability equal to its share of the US national population. If the probability of choosing each destination were uniform, the benchmark would be 83.1%.

Individuals' relocations do not show a strong tendency for retirees to move from high-wage to low-wage cities. Since larger cities exhibit higher observed nominal wages, the wage differences associated with retirees' relocations are quite similar to the population differences.[14] The random-destination benchmark says that retirees will move to a metropolitan area with a lower wage than their origin 50.2% of the time; the data show that retirees move to a lower-wage location 64% of the time. This is weak evidence for the canonical hypothesis.

These findings echo the results reported in Chen and Rosenthal (2008). Those authors find that there is a tendency for married retiree households who migrate to move to locations with lower nominal wages. However, these moves are relatively muted in comparison to the hypothesis that households will retire to the location with the lowest nominal wage. The greatest changes found by Chen and Rosenthal (2008) are moves by married households in which both spouses are college educated and the head of household is between ages 60 and 65. Conditional on changing metropolitan areas, these households typically move to a city with an annual nominal wage effect that is $933 lower (in 2000 US dollars).[15] In the 2000 data, the standard deviation of cities' wage effects exceeds $1300 and the difference between the cities with the highest and lowest wage effects exceeds $9000.[16] Thus, while these individuals do relocate to lower-wage cities, their moves do not exhibit changes of the magnitude implied by theories in which agglomeration only benefits productivity and retirees move to the location with the lowest nominal wages. The wage differences associated with moves by retiree households of other ages, educational attainments, and marital status are even smaller. In fact, Chen and Rosenthal (2008) find no statistically significant wage downgrades associated with metropolitan area changes for unmarried college-educated individuals over age 65.

The third piece of evidence for the consumer-cities hypothesis is the spatial distribution of

[14] I estimate city's nominal wage effects using data on full-time, full-year employees with high school degrees and bachelor's degrees. I regress individuals' log weekly wages on city fixed effects and individual controls (demographic characteristics interacted with educational attainment). The population elasticity of this nominal wage effect is 6.2%, in line with estimates typically obtained when failing to control for unobserved individual heterogeneity (Combes, Duranton, and Gobillon, 2008). The R^2 from regressing cities' wage effects on their log population is 0.53, so characterizing retirees' relocations in terms of wage differences or population differences yields quite similar results.

[15] In the notation of Chen and Rosenthal (2008), $Q_H - Q_B = -2\bar{w}_{jt}$, so their coefficient of 1865.16 in their Table A4 shows a wage effect that is $932.58 lower.

[16] These numbers describe the wage effects whose estimation is described in footnote 14.

retiree skills. Since educational attainment is strongly correlated with lifetime income, differences between retirees in terms of educational attainment likely reflect difference between retirees in terms of consumption expenditure. I split the population into four educational groups: those with less than a high school degree, high school graduates, college dropouts, and those with an associate's degree or higher educational attainment. The first two groups each constitute approximately one third of the retiree population; the latter two are approximately one sixth each.[17]

Table 3.1: The working-population elasticity of retiree populations

	All retirees		Retirees who changed MSAs			
	(1)	(2)	(3)	(4)	(5)	(6)
Log working population	0.919		0.828		0.755	
	(0.015)		(0.031)		(0.032)	
Log working population × HSD		0.912		0.856		0.716
		(0.021)		(0.030)		(0.029)
Log working population × HSG		0.908		0.812		0.765
		(0.020)		(0.033)		(0.034)
Log working population × CD		0.939		0.819		0.785
		(0.017)		(0.043)		(0.044)
Log working population × CG		0.978		0.917		0.856
		(0.017)		(0.039)		(0.041)
Only US-born					Yes	Yes
N	270	1080	270	1080	270	1080
R^2	0.918	0.884	0.683	0.634	0.640	0.591

Standard errors (clustered by MSA) in parentheses
Dependent variable is log retiree population in odd-numbered columns and log retiree population by educational group in even-numbered columns

Table 3.1 shows that more educated retirees are more likely to reside in larger metropolitan areas. This is true both for all retirees and for those changing metropolitan areas in the previous five years. The first column reports the total-population elasticity of the retiree population. The second column reports separate elasticities for each educational group and shows that more educated retirees are more likely to live in larger cities. This is partly attributable to the fact that more educated workers are more likely to live in larger cities and few workers change metropolitan areas

[17] I obtain similar results when I divide the population into three educational groups of roughly equal size.

upon retirement. However, restricting attention to the population of retirees who changed their metropolitan area of residence in the previous five years, the fourth column of Table 3.1 shows that more educated retirees are more likely to live in larger cities. An exception is that high-school dropouts are more likely to live in larger cities than high school graduates and college dropouts, but this tendency is driven by the locational choices of foreign-born retirees, as shown in column six. Amongst US-born retirees, the total-population elasticity of retiree population is strictly increasing in educational attainment. An established literature describes agglomeration benefits particular to less educated foreign-born individuals, such as linguistic enclaves, that are omitted from my model (Edin, Fredriksson, and Aslund, 2003; Bauer, Epstein, and Gang, 2005). The overrepresentation of foreign-born high-school-dropout retirees in large cities mirrors their overrepresentation in the working population.

These three pieces of empirical evidence demonstrate that the locational behavior of retirees favors the consumer-cities hypothesis over the canonical interpretation of spatial-equilibrium theory. First, the vast majority of retirees do not change residences or metropolitan areas upon exiting the labor force, consistent with the hypothesis that locations reflect consumption decisions financed out of permanent income. Second, retirees do not exhibit a strong tendency to move from high-wage cities to low-wage cities, contrary to the behavior implied by the standard production-driven spatial-equilibrium account. Third, retired residents in larger cities are more educated, consistent with the hypothesis that large cities complement skills through the consumption channel.

This evidence motivates a formalization of the consumer-cities hypothesis. The following section introduces a simple model with heterogenous labor and non-homothetic preferences in which larger cities exhibit higher nominal incomes and more skilled populations.

3.4 A model of consumer cities in general equilibrium

I describe an economy made up of two cities, denoted c and c', whose characteristics are endogenously determined by individuals' locational choices. The outcomes are driven by non-homothetic preferences exhibiting love of variety in consumption of non-tradables.

3.4.1 Consumer preferences

This section describes consumers and their preferences. Housing is a necessity, so the income elasticity of housing demand is less than one. Individuals with higher nominal incomes value greater variety of local goods and services relative to housing costs more than individuals with lower nominal incomes, since lower-income individual spend a larger share of their income on housing.[18] This is the agglomeration force that attracts higher-income individuals to larger cities.

There are N individuals in the economy. Individuals are heterogeneous because they vary in their effective units of labor. Individual i is endowed with and inelastically supplies l_i units of labor. Labor endowments are distributed according to the cumulative distribution function $G(l)$ with support on $[l_{\min}, l_{\max}]$.

Consumers must pay for their housing and commuting costs, r_c or $r_{c'}$, before purchasing goods. There are two types of goods – tradables and non-tradables. The tradable is a homogeneous good that I use as the numeraire. It is produced with labor at constant returns to scale, and I choose units so that the wage per effective unit of labor is also the numeraire. The non-tradable good has differentiated Dixit-Stiglitz varieties, whose elasticity of substitution is σ. These are produced by firms with increasing returns to scale, and the local Dixit-Stiglitz price indices are P_c and $P_{c'}$. The non-tradables and tradable have budget shares of α and $1 - \alpha$, respectively. The indirect utility of individual i in city c is

$$U_{ic} = \frac{l_i - r_c}{P_c^{\alpha}} + l_i \epsilon_{ic} \tag{3.1}$$

The idiosyncratic values of cities are such that $\epsilon_{ic} - \epsilon_{ic'} \sim F(\cdot)$, with $F(0) = \frac{1}{2}$. Following Tabuchi and Thisse (2002) and Moretti (2011), these idiosyncratic valuations cause local labor supply curves to be upward sloping in the real wage.[19] The fundamentals are symmetric so that if prices are equal

[18]This is similar to the logic presented by Black, Gates, Sanders, and Taylor (2002), who attribute gays' greater concentration in high-amenity cities to their smaller housing expenditure shares, which are due to their lower likelihood of child rearing.

[19]Canonical spatial-equilibrium models such as Roback (1982) have perfectly elastic local labor supplies, which is the case when $F(\cdot)$ is degenerate and $\epsilon_{ic} - \epsilon_{ic'} = 0$. In Moretti (2011), $\epsilon_{ic} - \epsilon_{ic'}$ is uniformly distributed. Davis and Dingel (2013) provide for microfoundations for upward sloping local labor supply curves by introducing heterogeneity in the internal geography of cities. In the interest of brevity, I follow Moretti (2011) by using the idiosyncratic-preferences specification.

across locations, the median individual of every skill level is indifferent between the two cities.

This specification of preferences is particularly tractable because after paying the necessary housing and commuting cost, every consumer allocates $1-\alpha$ of her remaining expenditure to the tradable numeraire and α to the differentiated good. This feature means that preferences are "homothetic at the margin" and permits aggregation of individual demands to the city level, provided that every individual in city c has income greater than r_c. Denote the total earnings in a city by L_c and population by N_c. Total expenditure on the non-tradables is $X_c = \alpha(L_c - r_c N_c)$.

These preferences are non-homothetic because housing and commuting costs are necessities that have income elasticities of demand of zero.[20] As a result, individuals with higher incomes find cities with lower Dixit-Stiglitz price indices relatively more attractive. The probability that an individual of skill l_i finds city c preferable to c' is

$$Pr(U_{ic} > U_{ic'}) = 1 - F\left(\left(P_{c'}^{-\alpha} - P_c^{-\alpha}\right) + \frac{1}{l_i}\left(r_c P_c^{-\alpha} - r_{c'} P_{c'}^{-\alpha}\right)\right)$$

This probability rises with skill when the relative price of non-tradables is lower in city c:

$$\begin{aligned}
\frac{\partial Pr(U_{ic} > U_{ic'})}{\partial l_i} &= F'\left(\left(P_{c'}^{-\alpha} - P_c^{-\alpha}\right) + \frac{1}{l_i}\left(r_c P_c^{-\alpha} - r_{c'} P_{c'}^{-\alpha}\right)\right)\frac{1}{l_i^2}\left(\frac{r_c}{P_c^{\alpha}} - \frac{r_{c'}}{P_{c'}^{\alpha}}\right) \\
\frac{\partial Pr(U_{ic} > U_{ic'})}{\partial l_i} &> 0 \iff \frac{P_c^{\alpha}}{r_c} < \frac{P_{c'}^{\alpha}}{r_{c'}}
\end{aligned}$$

In contrast, suppose that individuals had homothetic preferences over the tradable numeraire, non-tradable Dixit-Stiglitz varieties, and housing taking the form $U_{ic} = \frac{l_i}{P(1,P_c,r_c)} + l_i \epsilon_{ic}$, where $P(\cdot)$ is the homothetic price index. Then $\frac{\partial Pr(U_{ic} > U_{ic'})}{\partial l_i} = 0$ and the fraction of individuals finding city c preferable to city c' is independent of skill. The model's key mechanism is the fact that the preferences in equation (3.1) exhibit love of variety for non-tradables and are non-homothetic so that individuals' willingness to pay higher housing prices to access greater local variety varies with their incomes.[21]

---

[20] This stark assumption is one of convenience. The necessary assumption is only that the income elasticity of demand for housing is less than one. Notowidigdo (2011) and Ganong and Shoag (2012) provide evidence from the Consumer Expenditure Survey that the income elasticity of housing demand is below one. Glaeser, Kahn, and Rappaport (2008) show that the income elasticity of housing space demand is far below one.

[21] As written, individuals consume every local variety available in their location. The model would produce the

3.4.2 Production

Housing and commuting costs are determined by the usual Alonso-Mills-Muth internal urban structure, in which these costs are increasing in the local population N_c due to congestion. Every resident of city c pays the net urban cost $r_c = \theta N_c^\gamma$, where $\theta, \gamma > 0$.[22]

The tradable good is produced with labor at constant returns to scale. This makes the nominal wage per effective unit of labor the same across cities, so an individual's labor income is simply l.

Dixit-Stiglitz varieties are produced by firms with increasing returns to scale (Dixit and Stiglitz, 1977; Dingel, 2009). To produce a quantity x, a firm incurs a fixed cost f and constant marginal cost m. An individual firm has labor demand $l = mx + f$. Each firm chooses to produce a distinct variety, and there is free entry so that each firm's profits are zero. Since each firm hires $f\sigma$ units of labor, the number of firms in city c is $n_c = \frac{X_c}{f\sigma}$, where X_c is total local expenditure on the non-tradable. Therefore, the Dixit-Stiglitz price index is

$$P_c = \left(\frac{X_c}{f\sigma}\right)^{\frac{1}{1-\sigma}} \frac{m\sigma}{\sigma - 1} \tag{3.2}$$

3.4.3 Equilibrium

Since individuals earn the same nominal income in both cities, their locational decisions depend on local prices, r_c and P_c, and their idiosyncratic valuations of locations. Denote $\pi \equiv P_{c'}^{-\alpha} - P_c^{-\alpha}$ and $\Gamma \equiv r_c P_c^{-\alpha} - r_{c'} P_{c'}^{-\alpha}$, so that the fraction of individuals with labor endowment l preferring city c' to c is $F\left(\pi + \frac{\Gamma}{l}\right)$. These individual decisions aggregate to determine the cities' populations and local expenditure.

The population and total expenditure in each city are

$$N_c = N \int \left[1 - F\left(\pi + \frac{\Gamma}{l}\right)\right] dG(l) \qquad N_{c'} = N \int F\left(\pi + \frac{\Gamma}{l}\right) dG(l)$$
$$L_c = N \int l \cdot \left[1 - F\left(\pi + \frac{\Gamma}{l}\right)\right] dG(l) \qquad L_{c'} = N \int l \cdot F\left(\pi + \frac{\Gamma}{l}\right) dG(l)$$

---

same results if consumers valued greater local variety because they wished to buy an ideal variety in a number of goods categories, in the spirit of Lancaster (1979). Shamus Khan ("The New Elitists", *New York Times*, 7 Jul 2012) provides anecdotal accounts of higher-income individuals' greater concern with variety and product differentiation.

[22]See Behrens, Duranton, and Robert-Nicoud (2012) for a derivation of this urban cost structure.

The existence of a symmetric equilibrium is straightforward. When $L_c = L_{c'}$ and $N_c = N_{c'}$, local prices in the two locations are equal, $r_c = r_{c'}$ and $P_c = P_{c'}$, so $\pi = \Gamma = 0$ and the median individual at every skill level is indifferent between the two locations. $N_c = N \int [1 - F(0)]\, dG(l) = N \int \left[\frac{1}{2}\right] dG(l) = \frac{1}{2}N = N_{c'}$ and similarly $L_c = L_{c'} = \frac{1}{2}L$.

Of greater interest are asymmetric equilibria. These occur when there is a solution to the set of equations above in which $N_c \neq N_{c'}$ and $L_c \neq L_{c'}$.

Love of variety in local goods and services is the agglomeration force that is necessary for an asymmetric equilibrium. If there were no love of variety ($\sigma = +\infty$), then the relevant price index would simply be the price of housing ($\Gamma = r_c - r_{c'}$ and $\pi = 0$). Due to congestion costs, the larger city would be less attractive to all individuals, so this cannot be an equilibrium. Suppose that $N_c > N_{c'}$ so that $\Gamma > 0$. Then $F\left(\frac{\Gamma}{l}\right) > \frac{1}{2}\ \forall\ l$, so $N_c < N_{c'}$, contradicting the premise. In the absence of an agglomeration force, individuals disperse, causing the two locations to have equal population sizes.

Sufficient conditions for the existence of such equilibria depend on the distributions $F(\cdot)$ and $G(\cdot)$. However, we can establish a number of properties of these asymmetric equilibria without specifying functional forms for $F(\cdot)$ and $G(\cdot)$.

3.4.4 Cross-city patterns in asymmetric equilibria

In asymmetric equilibria, the two cities differ in both their aggregate and individual-level outcomes. At the city level, the two locations differ in their populations and aggregate income. These generate differences in the cities' prices. At the micro level, the larger city has skill and wage distributions that stochastically dominate those of the smaller city. These patterns match the empirical facts usually ascribed to production motives.

The two cities' differences in population and aggregate income imply that the more populous city has higher housing prices and a lower Dixit-Stiglitz price index. The higher housing prices follow directly from congestion raising local urban costs, r_c. In equilibrium, these higher local urban costs must be accompanied by a lower local Dixit-Stiglitz price index P_c. If both prices were higher, the city would be relatively less attractive to consumers of all income levels and therefore the less populous location.[23] The Dixit-Stiglitz price index is lower when total local expenditures on these varieties is higher, so aggregate income is higher in the more populous city.

These differences in cities' prices cause the larger city to have a more skilled population and higher individual incomes. Since the larger city has higher housing costs and a lower Dixit-Stiglitz price index, it is more attractive to higher-income individuals who spend a larger fraction of their income on Dixit-Stiglitz varieties. Thus, the fraction of individuals of skill level l living in the larger city is increasing in their labor endowment l.[24] This means that the larger city's skill distribution stochastically dominates that of the smaller city. Since the nominal wage per effective unit of labor is location-invariant, this causes the larger city's nominal income distribution to stochastically dominate that of the smaller city.

These skill and wage distributions can explain the empirical finding that observationally similar individuals earn higher nominal wages in larger cities, which is usually interpreted as suggesting that cities are not driven by consumption motives. Since the larger city's income distribution stochastically dominates that of the smaller city, an econometrician imperfectly observing individuals' skills would conclude that similar individuals earn higher nominal wages in the larger city. For example, if the econometrician compared the average nominal wages of individuals in a skill interval, $l_i \in [\underline{l}, \bar{l}]$, first-order stochastic dominance would cause the average in the larger city to be higher.[25]

Thus, this simple model in which individuals' nominal incomes are location-invariant yields asymmetric equilibria replicating a number of established empirical regularities in systems of cities. The larger city has higher average nominal wages, higher housing prices, a nominal wage distribution that stochastically dominates the wage distribution of the smaller city, and a greater fraction of higher-skilled individuals. When individuals are heterogeneous, consumption motives can play a first-order role in determining price and wage patterns in the cross section of cities.

---

[23]Formally, if $P_c > P_{c'}$ and $r_c > r_{c'}$, then $\pi + \frac{\Gamma}{l_i} > 0 \; \forall l_i$ so $N_c < N_{c'}$.

[24]Recall that $\frac{\partial Pr(U_{ic} > U_{ic'})}{\partial l_i} > 0 \iff \frac{P_c^\alpha}{r_c} < \frac{P_{c'}^\alpha}{r_{c'}}$.

[25]These "skill intervals" might correspond to various levels of educational attainment, for example. There is considerable wage variation within the class of individuals holding a given degree.

3.4.5 Existence of asymmetric equilibria

In this section, I describe the conditions for existence of asymmetric equilibria under various distributional assumptions.

3.4.5.1 Perfectly elastic labor supplies

Suppose that $F(\cdot)$ is a degenerate distribution at $\epsilon_{ic} - \epsilon_{ic'} = 0$ so that there is no heterogeneity amongst individuals who have the same labor endowment. In this case, all individuals who have the same labor endowment live in the same location unless $\frac{l_i - r_c}{P_c^\alpha} = \frac{l_i - r_{c'}}{P_{c'}^\alpha}$ so that they are indifferent. In an asymmetric equilibrium in which $L_c \neq L_{c'}$ and $N_c \neq N_{c'}$, denote the value of l_i satisfying this equality as l^*.

Suppose that $L_c > L_{c'}$. The population and total expenditure in each city are

$$N_c = N \int_{l^*}^{\infty} dG(l) \qquad\qquad N_{c'} = N \int_{-\infty}^{l^*} dG(l)$$
$$L_c = N \int_{l^*}^{\infty} l \; dG(l) \qquad\qquad L_{c'} = N \int_{-\infty}^{l^*} l \; dG(l)$$

A sufficient condition for the existence of an asymmetric equilibrium in this case is $\theta \left(\frac{N}{2}\right)^\gamma < l_{\min} < \theta N^\gamma$. The first inequality ensures that there exists a spatial allocation of the population such that all individuals have enough income to afford their housing and commuting costs. The second inequality ensures that congestion costs are sufficiently high that not all individuals can live in a single location in equilibrium. This condition means that there exists a value l^* such that the four equations above are satisfied and $\frac{l^* - r_c}{P_c^\alpha} = \frac{l^* - r_{c'}}{P_{c'}^\alpha}$.

While such an asymmetric equilibrium exhibits all the properties described in section 3.4.4, these outcomes are very stark. The skill and wage distributions of the two cities do not overlap because the least skilled individual in the larger city has income equal to that of the most skilled individual in the smaller city. Nonetheless, this result is theoretically interesting because it demonstrates that consumption motives alone can generate a system of cities with a positive city-size wage premium. The essential elements are heterogenous labor and non-homothetic preferences. To generate a more empirically relevant cross-city pattern of outcomes, I relax the assumption that local labor supplies are perfectly elastic.

3.4.5.2 Uniform distributions

Suppose that, as in Moretti (2011), individuals' idiosyncratic valuations are uniformly distributed such that $\epsilon_{ic} - \epsilon_{ic'} \sim U(-s, s)$. And suppose that individuals' labor endowments are uniformly distributed such that $l \sim U(a, a+b)$.[26] Aggregate expenditure in the economy is $L = N\left(a + \frac{b}{2}\right)$. If every type of laborer is present in each city ($0 < Pr(U_{ic} > U_{ic'}) < 1$), then the populations and expenditures are

$$N_c = N\frac{1}{2}\left(1 - \frac{\pi}{s}\right) - N\frac{\Gamma}{2s}\ln\left(\frac{a+b}{a}\right) \qquad N_{c'} = N - N_c$$
$$L_c = L\frac{1}{2}\left(1 - \frac{\pi}{s}\right) - L\frac{\Gamma}{2s}\frac{1}{a+\frac{b}{2}} \qquad L_{c'} = L - L_c$$

where Γ and π are defined as before.

Defining $\lambda_N = \frac{N_c}{N}, \lambda_L = \frac{L_c}{L},\ k_1 \equiv \left(\frac{\alpha}{f\sigma}\right)^{\frac{1}{1-\sigma}}\frac{m}{\rho}, k_2 \equiv \theta N^{\gamma}, k_3 \equiv k_2 \ln\left(\frac{a+b}{a}\right), k_4 \equiv k_2\frac{1}{a+\frac{b}{2}}$ and plugging in the values of π and Γ using equation (3.2), the equilibrium can be expressed as two equations in two unknowns:

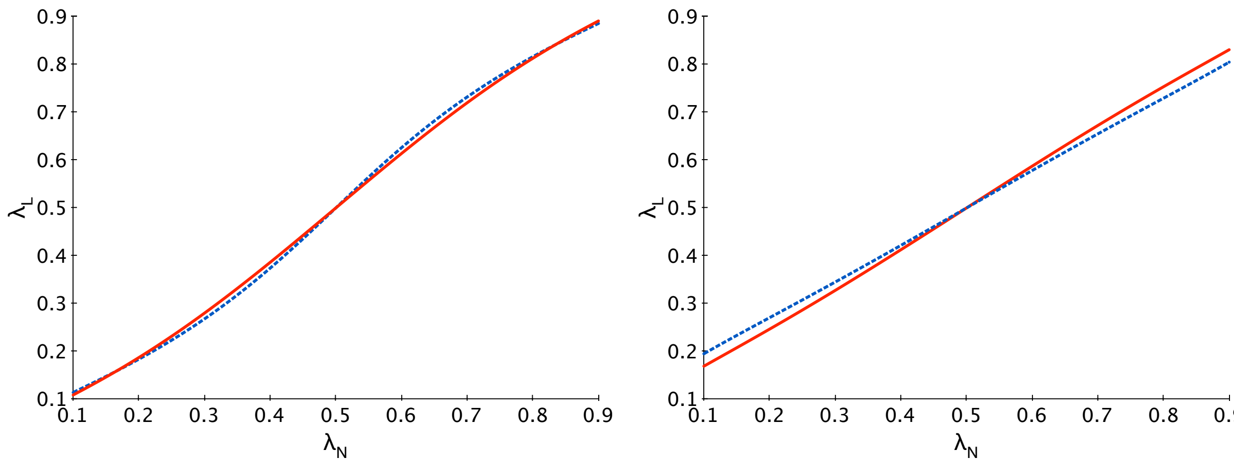
$$\begin{aligned}
\lambda_N = &\frac{1}{2} + \frac{1}{2s}\left(k_1 - k_3\lambda_N^{\gamma}\right)\left(\lambda_L L - k_2 N\lambda_N^{\gamma+1}\right)^{\frac{\alpha}{\sigma-1}} \\
&- \frac{1}{2s}\left(k_1 - k_3(1-\lambda_N)^{\gamma}\right)\left((1-\lambda_L)L - k_2N(1-\lambda_N)^{\gamma+1}\right)^{\frac{\alpha}{\sigma-1}} \\
\lambda_L = &\frac{1}{2} + \frac{1}{2s}\left(k_1 - k_4\lambda_N^{\gamma}\right)\left(\lambda_L L - k_2 N\lambda_N^{\gamma+1}\right)^{\frac{\alpha}{\sigma-1}} \\
&- \frac{1}{2s}\left(k_1 - k_4(1-\lambda_N)^{\gamma}\right)\left((1-\lambda_L)L - k_2N(1-\lambda_N)^{\gamma+1}\right)^{\frac{\alpha}{\sigma-1}}
\end{aligned}$$

To aid interpretation, think of the two equations as defining λ_L as an implicit function of λ_N. Note that λ_N determines relative housing prices and λ_L determines relative Dixit-Stiglitz prices. The first equation describes the spatial allocation of aggregate expenditure, λ_L, necessary to realize a given spatial allocation of aggregate population, λ_N. That is, given relative housing prices, what relative Dixit-Stiglitz prices would induce the population to choose their locations in such a way as

[26] Therefore, $l_{\min} = a$ and $l_{\max} = a + b$.

to yield those relative housing prices? The second equation describes the expenditure pattern that is induced by a given spatial allocation of aggregate population. That is, fixing relative housing prices, what Dixit-Stiglitz prices are consistent with the population choosing their locations in such a way as to yield those Dixit-Stiglitz prices? An equilibrium satisfies both equations simultaneously.

Figure 3.1: Equilibrium conditions



NOTE: The parameters for the left figure are $\alpha = .3, \sigma = 5, a = 10, s = .1, \gamma = .2, k_1 = .5, N = 1, b = 5, \theta = 2.5$. In the right figure, $\sigma = 4$. The solid red line satisfies the first equation. The dashed blue line satisfies the second equation.

As is common in the new economic geography, this system of equations does not have a tractable analytical solution for asymmetric equilibria. I therefore characterize the properties of the system numerically. Figure 3.1 provides two examples. The left figure shows a system in which there are both symmetric and asymmetric equilibria. The right figure, with a lower value of σ, shows a system in which the only equilibrium is the symmetric one.

The properties of the left figure are straightforward. Since the fundamentals are symmetric, one asymmetric equilibrium is just a relabeling of the other. In the right figure, the agglomeration force is stronger because individuals place greater value on the variety of local goods and services. From the left figure to the right figure, both lines rotate clockwise around the $\lambda_N = \lambda_L = \frac{1}{2}$ origin, but the dashed blue line moves farther, so that it now falls below the solid red line for all values $\lambda_N > \frac{1}{2}$. Both lines rotate clockwise because the stronger agglomeration force makes relative Dixit-Stiglitz prices more responses to differences in total local expenditure. The dashed blue line rotates farther than the solid red line because higher-income individuals are relatively more responsive to differences in relative Dixit-Stiglitz prices when σ is lower.[27]

In sum, the existence of an asymmetric equilibrium depends on the strength of the agglomeration force relative to congestion costs and individuals' idiosyncratic tastes for cities. If the agglomeration force is too strong, differences in the variety of local goods and services across cities make the larger city more attractive to all individuals and the second city cannot exist in an asymmetric equilibrium. If the agglomeration force is too weak, differences in variety are insufficient to attract enough individuals to generate that variety. In between, asymmetric equilibria exist and exhibit the properties described in section 3.4.4.

3.4.6 Skill premia and city size

When labor is heterogeneous and preferences are non-homothetic, consumer cities can match the empirical cross-city pattern of skills and wages. The model demonstrates that larger cities can exhibit both higher nominal incomes and more skilled populations without an agglomeration mechanism that increases productivity. Lee (2010) presents a partial-equilibrium model in which the pattern of skill premia reflects consumption motives. In his model, individuals have reservation utilities and local labor supplies are perfectly elastic. Since higher-income individuals are more willing to pay for greater local variety, skilled individuals' nominal incomes should not increase with city size as quickly as less skilled individuals' incomes.[28] Empirically, skill premia, measured as college wage premia, increase with city size (Davis and Dingel, 2012).[29] While this fact rejects the particular hypothesis described by the model of Lee (2010), it does not imply that consumption motives are unimportant in determining cities' skills and wages.

When local labor supplies are less than perfectly elastic and skill is imperfectly observed, a system of consumer cities may exhibit a positive correlation between skill premia and city sizes. I define the observed skill premium in a city as the average nominal wage of skilled individuals in that city divided by the average nominal wage of unskilled individuals in that city. The observed

[27] When $\epsilon_{ic} - \epsilon_{ic'} \sim U(-s, s)$, $\frac{\partial^2 Pr(U_{ic} > U_{ic'})}{\partial l_i \partial \sigma} = \frac{1}{2s} \frac{1}{l_i} \frac{\alpha k_1}{(\sigma-1)^2} \left(r_{c'} X_{c'}^{\frac{\alpha}{\sigma-1}} \ln X_{c'} - r_c X_c^{\frac{\alpha}{\sigma-1}} \ln X_c \right) < 0$.

[28] Since Lee (2010) does not model multiple cities, it would be more accurate to describe this as a comparative static than a prediction about the cross-city pattern of premia.

[29] Lee (2010) restricts his study to medical occupations, which are an exception to the general pattern.

cross-city pattern of skill premia is jointly determined by the economy-wide skill distribution, how the continuum of skills is divided into two skill groups, and within-group variation in incomes.[30]

I now provide a numerical example. Consider the asymmetric equilibrium in the left panel of Figure 3.1. Recall that $l \sim U(10, 15)$. If the population is divided into two groups in which those with $l \geq 12.5$ are "skilled" and those with $l \leq 12.5$ are "unskilled", the observed skill premia of the two cities are equal in the asymmetric equilibrium. When the skilled-unskilled cutoff is less than 12.5, the larger city's skill premium is greater than that of the smaller city.

This demonstrates that a consumption-driven system of cities can explain the empirical finding that city sizes and skill premia are positively correlated. Of course, a numerical example does not demonstrate that a positive size-premium correlation is a general property of the model. But it demonstrates that the indifference-condition test of the consumer-cities hypothesis proposed by Lee (2010) depends on the assumption of perfectly elastic labor supplies and precisely observed skill levels. Relaxing those assumptions, the model presented here is compatible with the empirical pattern of skill premia.

3.5 Conclusion

Popular and academic discussions of cities emphasize their values as places to consume. But consumption motives have hitherto been relegated to a second-order force shaping the system of cities. Since nominal wages are higher in larger cities, urban economists have concluded that agglomeration is primarily about larger cities' advantages for production. For example, Rappaport (2008, p.549) says that "the observed positive correlation between wages and density places an upper bound on the importance of quality of life as a source of local crowdedness." This chapter shows that larger cities' higher nominal incomes and more skilled populations could be explained by consumption motives alone. If higher-income individuals value greater variety of local goods and services more relative to housing prices, they will find larger cities relatively more attractive. These differing valuations of heterogeneous individuals support heterogeneous cities as an equilibrium outcome.

[30] Davis and Dingel (2012) show how within-group heterogeneity and spatial sorting determine the pattern of skill premia in an economy with production motives for agglomeration.

Consumption-driven agglomeration may be as important as production-driven agglomeration.

The logic of the consumer-cities hypothesis presented in this chapter is more general than the particular functional forms used to demonstrate my results. For example, larger cities may exhibit consumption benefits valued by high-income individuals in dimensions other than variety. Suppose that product quality is endogenous and firms upgrade their product quality by incurring fixed costs, as in Sutton (1991). In larger markets, firms would produce higher-quality goods, so cities with higher total expenditure would have a lower quality-adjusted price index rather than a lower variety-adjusted price index.[31] If high-income consumers spent a larger share of their income on these goods or were more willing to pay for quality, they would move to larger cities in order to access higher-quality products. This variant of the model would yield the same predicted skill and wage distributions, but its predictions about observable outcomes in the goods market would be in terms of quality rather than variety.

It seems unlikely that individuals' productivities are spatially invariant. But it seems equally implausible that cities' consumption opportunities are similarly valued by heterogeneous individuals. The purpose of this chapter is to introduce a parsimonious, general-equilibrium model of a system of cities in which consumption motives are the agglomeration mechanism. It shows that consumer cities can account for a range of empirical facts often identified as evidence that larger cities' advantages must stem from production-driven agglomeration. While Glaeser, Kolko, and Saiz (2001) argue that the role of cities as centers of consumption is increasingly important relative to their role as centers of production, this theory shows that consumer cities could already account for fundamental facts about systems of cities.

[31] Berry and Waldfogel (2010) provide evidence that product quality increases with market size in daily newspapers, where quality is produced with fixed costs.

Bibliography

Abdel-Rahman, H. M., and A. Anas (2004): "Theories of systems of cities," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. F. Thisse, vol. 4, chap. 52, pp. 2293–2339. Elsevier.

Acemoglu, D., and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4, pp. 1043–1171. Elsevier.

Aghion, P., R. Blundell, R. Griffith, P. Howitt, and S. Prantl (2009): "The Effects of Entry on Incumbent Innovation and Productivity," *The Review of Economics and Statistics*, 91(1), 20–32.

Albouy, D. (2012): "Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas," mimeo.

Anderson, J. E. (2011): "The Gravity Model," *Annual Review of Economics*, 3(1), 133–160.

Atalay, E., A. Hortaçsu, and C. Syverson (2014): "Vertical Integration and Input Flows," *The American Economic Review*, 104(4), 1120–1148.

Autor, D. H., and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103(5), 1553–97.

Bacolod, M., B. S. Blum, and W. C. Strange (2009): "Skills in the city," *Journal of Urban Economics*, 65(2), 136–153.

Bagnoli, M., and T. Bergstrom (2005): "Log-concave probability and its applications," *Economic Theory*, 26(2), 445–469.

Baldwin, R., and J. Harrigan (2011): "Zeros, Quality, and Space: Trade Theory and Trade Evidence," *American Economic Journal: Microeconomics*, 3(2), 60–88.

Bauer, T., G. Epstein, and I. Gang (2005): "Enclaves, language, and the location choice of migrants," *Journal of Population Economics*, 18(4), 649–662.

Baum-Snow, N., and R. Pavan (2011): "Inequality and City Size," mimeo.

Beath, J., and Y. Katsoulacos (1991): *The Economic Theory of Product Differentiation*. Cambridge University Press.

Behrens, K., G. Duranton, and F. Robert-Nicoud (2012): "Productive cities: Sorting, selection and agglomeration," mimeo.

Bergstrand, J. H. (1990): "The Heckscher-Ohlin-Samuelson Model, the Linder Hypothesis and the Determinants of Bilateral Intra-industry Trade," *Economic Journal*, 100(403), 1216–29.

Bergstrand, J. H., and P. Egger (2011): "Gravity Equations and Economic Frictions in the World Economy," in *Palgrave Handbook of International Trade*, ed. by D. Bernhofen, R. Falvey, D. Greenaway, and U. Kreickemeier. Palgrave Macmillan.

Bernard, A. B., S. J. Redding, and P. K. Schott (2013): "Testing for Factor Price Equality with Unobserved Differences in Factor Quality or Productivity," *American Economic Journal: Microeconomics*, 5(2), 135–63.

Bernasconi, C. (2013): "Similarity of income distributions and the extensive and intensive margin of bilateral trade flows," Economics Working Papers 115, University of Zurich.

Bernstein, J. R., and D. E. Weinstein (2002): "Do endowments predict the location of production?: Evidence from national and international data," *Journal of International Economics*, 56(1), 55–76.

Berry, S., and J. Waldfogel (2010): "Product Quality and Market Size," *Journal of Industrial Economics*, 58(1), 1–31.

Berry, S. T. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242–262.

Black, D., G. Gates, S. Sanders, and L. Taylor (2002): "Why Do Gay Men Live in San Francisco?," *Journal of Urban Economics*, 51(1), 54–76.

Black, D., N. Kolesnikova, and L. Taylor (2009): "Earnings Functions When Wages and Prices Vary by Location," *Journal of Labor Economics*, 27(1), 21–47.

Bombardini, M., G. Gallipoli, and G. Pupato (2012): "Skill Dispersion and Trade Flows," *American Economic Review*, 102(5), 2327–48.

Brueckner, J. K. (1987): "The structure of urban equilibria: A unified treatment of the muth-mills model," in *Handbook of Regional and Urban Economics*, ed. by E. S. Mills, vol. 2, chap. 20, pp. 821–845. Elsevier.

Bureau of Transportation Statistics, and US Census Bureau (2004): "2002 Economic Census: 2002 Commodity Flow Survey," EC02TCF-US, US Department of Transportation and US Department of Commerce, https://1bts.rita.dot.gov/pdc/user/products/src/products.xml?p=1836.

——— (2010): "2007 Economic Census: 2007 Commodity Flow Survey," EC07TCF-US, US Department of Transportation and US Department of Commerce, https://1bts.rita.dot.gov/pdc/user/products/src/products.xml?p=3193.

Caron, J., T. Fally, and J. R. Markusen (2012): "Skill Premium and Trade Puzzles: a Solution Linking Production and Preferences," NBER Working Paper 18131.

Chen, Y., and S. S. Rosenthal (2008): "Local amenities and life-cycle migration: Do people move for jobs or fun?," *Journal of Urban Economics*, 64(3), 519–537.

Choi, Y. C., D. Hummels, and C. Xiang (2009): "Explaining import quality: The role of the income distribution," *Journal of International Economics*, 78(2), 293–303.

Christaller, W. (1933): *Die zentralen Orte in Suddeutschland*. Gustav Fischer, Jena.

Clark, D. E., and W. J. Hunter (1992): "The Impact of Economic Opportunity, Amenities and Fiscal Factors on Age-Specific Migration Rates," *Journal of Regional Science*, 32(3), 349–365.

Combes, P.-P., G. Duranton, and L. Gobillon (2008): "Spatial wage disparities: Sorting matters!," *Journal of Urban Economics*, 63(2), 723–742.

Combes, P.-P., G. Duranton, L. Gobillon, and S. Roux (2010): "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in *Agglomeration Economics*, ed. by E. L. Glaeser, pp. 15–65. NBER.

Costinot, A. (2009): "An Elementary Theory of Comparative Advantage," *Econometrica*, 77(4), 1165–1192.

Costinot, A., and J. Vogel (2010): "Matching and Inequality in the World Economy," *Journal of Political Economy*, 118(4), 747–786.

Davis, D. R., and J. I. Dingel (2012): "A Spatial Knowledge Economy," NBER Working Paper 18188, National Bureau of Economic Research, Inc.

——— (2013): "The Comparative Advantage of Cities," Columbia University mimeo.

Davis, D. R., and D. E. Weinstein (1999): "Economic geography and regional production structure: An empirical investigation," *European Economic Review*, 43(2), 379–407.

——— (2003): "Market access, economic geography and comparative advantage: an empirical test," *Journal of International Economics*, 59(1), 1–23.

De la Roca, J., and D. Puga (2012): "Learning by working in big cities," mimeo.

Deardorff, A. V. (1984): "Testing trade theories and predicting trade flows," in *Handbook of International Economics*, ed. by R. W. Jones, and P. B. Kenen, vol. 1 of *Handbook of International Economics*, chap. 10, pp. 467–517. Elsevier.

Deaton, A., and J. Muellbauer (1980): *Economics and Consumer Behavior*, no. 9780521296762 in Cambridge Books. Cambridge University Press.

Diamond, R. (2012): "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," mimeo.

Dingel, J. I. (2009): "The basics of 'Dixit-Stiglitz lite'," teaching mimeo.

Disdier, A.-C., and K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade," *The Review of Economics and Statistics*, 90(1), 37–48.

Dixit, A. K., and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67(3), 297–308.

Dollar, D., E. N. Wolff, and W. J. Baumol (1988): "The Factor-Price Equalization Model and Industry Labor Productivity: An Empirical Test across Countries," in *Empirical Methods for International Trade*, ed. by R. C. Feenstra, chap. 2, pp. 23–48. MIT Press.

Duncombe, W., M. Robbins, and D. A. Wolf (2001): "Retire to where? A discrete choice model of residential location," *International Journal of Population Geography*, 7(4), 281–293.

Duranton, G., and H. G. Overman (2005): "Testing for Localization Using Micro-Geographic Data," *Review of Economic Studies*, 72(4), 1077–1106.

Duranton, G., and D. Puga (2001): "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," *American Economic Review*, 91(5), 1454–1477.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Edin, P.-A., P. Fredriksson, and O. Aslund (2003): "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment," *The Quarterly Journal of Economics*, 118(1), 329–357.

Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2011): "Spatial Sorting," mimeo.

Ellison, G., and E. L. Glaeser (1997): "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105(5), 889–927.

Ellison, G., E. L. Glaeser, and W. R. Kerr (2010): "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns," *American Economic Review*, 100(3), 1195–1213.

Fajgelbaum, P., G. M. Grossman, and E. Helpman (2011): "Income Distribution, Product Quality, and International Trade," *Journal of Political Economy*, 119(4), 721 – 765.

Falvey, R. E. (1981): "Commercial policy and intra-industry trade," *Journal of International Economics*, 11(4), 495–511.

Federal Trade Commission (1981): *Statistical Report: Annual Line of Business Report 1975.*

Feenstra, R. C., and J. Romalis (2012): "International Prices and Endogenous Quality," NBER Working Paper 18314.

Fieler, A. C. (2011): "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 79(4), 1069–1101.

Flam, H., and E. Helpman (1987): "Vertical Product Differentiation and North-South Trade," *American Economic Review*, 77(5), 810–22.

Florida, R. (2002): *The Rise of the Creative Class.* Basic Books.

Fujita, M., and J.-F. Thisse (2002): *Economics of Agglomeration.* Cambridge University Press.

Gabszewicz, J. J., A. Shaked, J. Sutton, and J. Thisse (1981): "International Trade in Differentiated Products," *International Economic Review*, 22(3), pp. 527–534.

GANONG, P., AND D. SHOAG (2012): "Why Has Regional Convergence in the U.S. Stopped?," mimeo.

GIBBONS, S., H. G. OVERMAN, AND P. PELKONEN (2013): "Area Disparities in Britain: Understanding the Contribution of People vs. Place Through Variance Decompositions," *Oxford Bulletin of Economics and Statistics*, p. http://dx.doi.org/10.1111/obes.12043.

GLAESER, E., AND M. RESSEGER (2010): "The complementarity between cities and skills," *Journal of Regional Science*, 50(1), 221–244.

GLAESER, E. L. (2008): *Cities, Agglomeration, and Spatial Equilibrium*, The Lindahl Lectures. Oxford University Press.

GLAESER, E. L., AND J. D. GOTTLIEB (2009): "The Wealth of Cities: Agglomeration Economies and Spatial Equilibrium in the United States," *Journal of Economic Literature*, 47(4), 983–1028.

GLAESER, E. L., M. E. KAHN, AND J. RAPPAPORT (2008): "Why do the poor live in cities? The role of public transportation," *Journal of Urban Economics*, 63(1), 1–24.

GLAESER, E. L., J. KOLKO, AND A. SAIZ (2001): "Consumer city," *Journal of Economic Geography*, 1(1), 27–50.

GRAVES, P. E., AND D. M. WALDMAN (1991): "Multimarket Amenity Compensation and the Behavior of the Elderly," *American Economic Review*, 81(5), 1374–81.

GROSSMAN, G. M. (2004): "The Distribution of Talent and the Pattern and Consequences of International Trade," *Journal of Political Economy*, 112(1), pp. 209–239.

GROSSMAN, G. M., AND G. MAGGI (2000): "Diversity and Trade," *American Economic Review*, 90(5), 1255–1275.

GROSSMAN, G. M., AND E. ROSSI-HANSBERG (2008): "Trading Tasks: A Simple Theory of Offshoring," *American Economic Review*, 98(5), 1978–97.

GYOURKO, J., C. MAYER, AND T. SINAI (2013): "Superstar Cities," *American Economic Journal: Economic Policy*, 5(4), 167–99.

HALLAK, J. C. (2006): "Product quality and the direction of trade," *Journal of International Economics*, 68(1), 238–265.

——— (2010): "A Product-Quality View of the Linder Hypothesis," *The Review of Economics and Statistics*, 92(3), 453–466.

HALLAK, J. C., AND P. K. SCHOTT (2011): "Estimating Cross-Country Differences in Product Quality," *The Quarterly Journal of Economics*, 126(1), 417–474.

HALLAK, J. C., AND J. SIVADASAN (2013): "Product and process productivity: Implications for quality choice and conditional exporter premia," *Journal of International Economics*, 91(1), 53 – 67.

HANDBURY, J. (2012): "Are Poor Cities Cheap for Everyone? Non-Homotheticity and the Cost of Living Across U.S. Cities," mimeo.

Handbury, J., and D. E. Weinstein (2011): "Is New Economic Geography Right? Evidence from Price Data," NBER Working Paper 17067, National Bureau of Economic Research, Inc.

Hanson, G. H., and C. Xiang (2004): "The Home-Market Effect and Bilateral Trade Patterns," *American Economic Review*, 94(4), 1108–1129.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–61.

Helpman, E., and P. R. Krugman (1989): *Trade Policy and Market Structure*. MIT Press.

Helsley, R. W., and W. C. Strange (1990): "Matching and agglomeration economies in a system of cities," *Regional Science and Urban Economics*, 20(2), 189–212.

——— (2012): "Coagglomeration and the Scale and Composition of Clusters," UC Berkeley mimeo.

Henderson, J. V. (1974): "The Sizes and Types of Cities," *American Economic Review*, 64(4), 640–56.

——— (1983): "Industrial Bases and City Sizes," *American Economic Review*, 73(2), 164–68.

——— (1987): "General equilibrium modeling of systems of cities," in *Handbook of Regional and Urban Economics*, ed. by E. S. Mills, vol. 2, chap. 23, pp. 927–956. Elsevier.

——— (1991): *Urban Development: Theory, Fact, and Illusion*. Oxford University Press.

Henderson, V. (1997): "Medium size cities," *Regional Science and Urban Economics*, 27(6), 583–612.

Hendricks, L. (2011): "The Skill Composition Of U.S. Cities," *International Economic Review*, 52(1), 1–32.

Hillberry, R., and D. Hummels (2003): "Intranational Home Bias: Some Explanations," *The Review of Economics and Statistics*, 85(4), 1089–1092.

——— (2008): "Trade responses to geographic frictions: A decomposition using micro-data," *European Economic Review*, 52(3), 527–550.

Holmes, T. J., and J. J. Stevens (2004): "Spatial distribution of economic activities in North America," in *Handbook of Regional and Urban Economics*, ed. by J. V. Henderson, and J. F. Thisse, vol. 4, chap. 63, pp. 2797–2843. Elsevier.

——— (2010): "An Alternative Theory of the Plant Size Distribution with an Application to Trade," NBER Working Paper 15957, National Bureau of Economic Research, Inc.

——— (2012): "Exports, borders, distance, and plant size," *Journal of International Economics*, 88(1), 91–103.

Hsu, W.-T., T. J. Holmes, and F. Morgan (2013): "Optimal City Hierarchy: A Dynamic Programming Approach to Central Place Theory," mimeo.

Hummels, D., and P. J. Klenow (2005): "The Variety and Quality of a Nation's Exports," *American Economic Review*, 95(3), 704–723.

Hummels, D., and A. Skiba (2004): "Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture," *Journal of Political Economy*, 112(6), 1384–1402.

Hurst, E. (2008): "Understanding Consumption in Retirement: Recent Developments," in *Recalibrating Retirement Spending and Saving*, ed. by J. Ameriks, and O. S. Mitchell. Oxford University Press.

Khandelwal, A. (2010): "The Long and Short (of) Quality Ladders," *Review of Economic Studies*, 77, 1450–1476.

Krugman, P. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70(5), 950–59.

——— (1991a): *Geography and trade*. MIT Press.

——— (1991b): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–99.

Lancaster, K. (1979): *Variety, Equity and Efficiency: Product Variety in an Industrial Society*. Blackwell.

Leamer, E. E. (1984): *Sources of International Comparative Advantage: Theory and Evidence*. MIT Press.

Lederman, D., and W. F. Maloney (eds.) (2012): *Does What You Export Matter? In Search of Empirical Guidance for Industrial Policies*. World Bank.

Lee, S. (2010): "Ability sorting and consumer city," *Journal of Urban Economics*, 68(1), 20–33.

Lehmann, E. (1955): "Ordered Families of Distributions," *The Annals of Mathematical Statistics*, 26(3), 399–419.

LeRoy, S. F., and J. Sonstelie (1983): "Paradise lost and regained: Transportation innovation, income, and residential location," *Journal of Urban Economics*, 13(1), 67–89.

Linder, S. B. (1961): *An Essay on Trade and Transformation*. John Wiley & Sons, Ltd.

Lucas, R. E. (1988): "On the mechanics of economic development," *Journal of Monetary Economics*, 22(1), 3–42.

Lucas, R. E., and E. Rossi-Hansberg (2002): "On the Internal Structure of Cities," *Econometrica*, 70(4), 1445–1476.

Lugovskyy, V., and A. Skiba (2012): "How Geography Affects Quality," mimeo.

Markusen, J. R. (1986): "Explaining the Volume of Trade: An Eclectic Approach," *American Economic Review*, 76(5), 1002–11.

Mayer, T., and S. Zignago (2011): "Notes on CEPII's distances measures: The GeoDist database," Working Papers 2011-25, CEPII.

McFadden, D. (1978): “Modelling the Choice of Residential Location,” in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlvist, L. Lundqvist, F. Snickars, and J. Weibull, pp. 75–96. North-Holland.

Moretti, E. (2011): “Local Labor Markets,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4, chap. 14, pp. 1237–1313. Elsevier.

Murphy, K. M., and A. Shleifer (1997): “Quality and trade,” *Journal of Development Economics*, 53(1), 1–15.

Nocke, V. (2006): “A Gap for Me: Entrepreneurs and Entry,” *Journal of the European Economic Association*, 4(5), 929–956.

Notowidigdo, M. J. (2011): “The Incidence of Local Labor Demand Shocks,” Working Paper 17167, National Bureau of Economic Research.

Pierce, J., and P. Schott (2012): “A Concordance between Ten-Digit U.S. Harmonized System Codes and SIC/NAICS Product Classes and Industries,” *Journal of Economic and Social Measurement*, 37(1-2), 61–96.

Rappaport, J. (2008): “Consumption amenities and city population density,” *Regional Science and Urban Economics*, 38(6), 533–552.

Redding, S. (1996): “The Low-Skill, Low-Quality Trap: Strategic Complementarities between Human Capital and R&D,” *Economic Journal*, 106(435), 458–70.

Redding, S., and A. J. Venables (2004): “Economic geography and international inequality,” *Journal of International Economics*, 62(1), 53–82.

Roback, J. (1982): “Wages, Rents, and the Quality of Life,” *Journal of Political Economy*, 90(6), 1257–78.

Ruggles, S., J. T. Alexander, K. Genadek, R. Goeken, M. B. Schroeder, and M. Sobek (2010): “Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database],” Minneapolis, MN: Minnesota Population Center.

Samuelson, P. A. (1948): “International trade and the equalisation of factor prices,” *The Economic Journal*, 58(230), 163–184.

Sattinger, M. (1975): “Comparative Advantage and the Distributions of Earnings and Abilities,” *Econometrica*, 43(3), 455–68.

——— (1978): “Comparative Advantage in Individuals,” *The Review of Economics and Statistics*, 60(2), 259–67.

——— (1979): “Differential Rents and the Distribution of Earnings,” *Oxford Economic Papers*, 31(1), 60–71.

——— (1993): “Assignment Models of the Distribution of Earnings,” *Journal of Economic Literature*, 31(2), 831–80.

Schott, P. K. (2003): “One Size Fits All? Heckscher-Ohlin Specialization in Global Production,” *American Economic Review*, 93(3), 686–708.

——— (2004): "Across-product Versus Within-product Specialization in International Trade," *The Quarterly Journal of Economics*, 119(2), 646–677.

Silva, J. M. C. S., and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88(4), 641–658.

Simonovska, I. (2013): "Income Differences and Prices of Tradables," mimeo.

Sutton, J. (1991): *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. MIT Press.

——— (1998): *Technology and Market Structure: Theory and History*. MIT Press.

——— (2012): *Competing in Capabilities: The Globalization Process*. Oxford University Press.

Tabuchi, T., and J.-F. Thisse (2002): "Taste heterogeneity, labor mobility and economic geography," *Journal of Development Economics*, 69(1), 155–177.

Torstensson, J. (1996): "Can factor proportions explain vertical intra-industry trade?," *Applied Economics Letters*, 3(5), 307–309.

Trefler, D. (1995): "The Case of the Missing Trade and Other Mysteries," *American Economic Review*, 85(5), 1029–46.

Verhoogen, E. A. (2008): "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *The Quarterly Journal of Economics*, 123(2), 489–530.

von Thünen, J. H. (1826): *Der Isolierte Staat in Beziehung auf Landschaft und Nationalökonomie.*

Appendix A

Appendix for Chapter 1

A.1 Theory appendix

Equilibrium pattern of specialization and trade

Skill-intensive quality and costless trade

Inequality (1.4) says that skill-abundant locations specialize in skill-intensive products. Recall:

$$k > k', \omega > \omega' \Rightarrow \frac{w(\omega,k)^{-\sigma}}{w(\omega',k)^{-\sigma}}\mathbb{E}_{\omega',k}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right) > \frac{w(\omega,k')^{-\sigma}}{w(\omega',k')^{-\sigma}}\mathbb{E}_{\omega',k'}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right)$$

To arrive at the result, we'll first consider the case of factor-price equalization and then describe the more general case.

When wages are equal across locations, $w(\omega,k) = w(\omega)\ \forall k$, this inequality simplifies to

$$\mathbb{E}_{\omega',k}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right) > \mathbb{E}_{\omega',k'}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right).$$

In this case, $\phi(r,\omega',k) = \frac{n_{r,k}x(r,k)b(\omega',r)^{\sigma}c(r)^{\sigma}}{\int_{r\in z\cup Q} n_{r,k}x(r,k)b(\omega',r)^{\sigma}c(r)^{\sigma}dr} = \frac{n_{r,k}x(r,k)\ell(\omega',r)}{\int_{r\in z\cup Q} n_{r,k}x(r,k)\ell(\omega',r)dr}$ is a density that describes ω'-use-weighted shares of total output of r in location k. The inequality says that the average skill intensity of output is higher in k than k'. This requires that $\phi_i(\omega',k)$ put more weight on some higher values of i than $\phi_i(\omega',k')$ does so that the average skill intensity of output produced in k is higher than in k'.[1] This is the factor-abundance mechanism for quality specialization.

In the absence of factor-price equalization, a slightly longer chain of reasoning delivers the same conclusion. Suppose that inequality (1.4) were true but $\phi_i(\omega', k)$ did not place greater weight on higher values of q in higher-k locations (i.e. $\phi_i(\omega', k) = \phi_i(\omega', k')$). If so, $w(\omega, k)^{-\sigma}$ must be strictly log-supermodular to satisfy this inequality. If $w(\omega, k)^{-\sigma}$ is log-supermodular, then $c(r, k)^{-\sigma}$ is log-supermodular, since $c(r, k)^{1-\sigma} = \int_{\omega \in \Omega} b(\omega, r)^{\sigma} w(\omega, k)^{1-\sigma} d\omega$ and $b(\omega, r)$ is log-supermodular (Lehmann, 1955). In the absence of trade costs, varieties of skill intensity i are only produced in location k when $c(i, k) = \min_{k'} c(i, k')$. Thus, log-supermodularity of $c(i, k)^{-\sigma}$ and the zero-profit condition imply that $\phi_i(\omega', k)$ is log-supermodular in (i, k). As a result, $\mathbb{E}_{\omega', k}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right) > \mathbb{E}_{\omega', k'}\left(\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\right)$. Therefore skill-abundant locations specialize in skill-intensive, high-quality varieties.

Uniform skill intensities and costly trade

Lemma A.1. *When factor prices equalize and wages are increasing in skill, $\Gamma_k(q, \mathbf{n}, \mathbf{c}, \mathbf{0})$ is strictly log-supermodular in (q, k).*

Proof. The proof of this lemma is quite similar to the proof of Lemma 1 in Fajgelbaum, Grossman, and Helpman (2011). The demand level is

$$\Gamma_k(q, \mathbf{n}, \mathbf{c}, \mathbf{0}) = \int \underbrace{\frac{\exp(yq)\left[\sum_{k''} n_{q,k''} \exp(-c(q, k'')q/\theta_q)\right]^{\theta_q - 1}}{\sum_{q'} \exp(yq')\left[\sum_{k''} n_{q',k''} \exp(-c(q, k'')q'/\theta_{q'})\right]^{\theta_{q'}}}}_{\equiv \Psi(y, q, \mathbf{n}, \mathbf{c}, \mathbf{0})} g(y, k) dy$$

$\exp(yq)$ is strictly log-supermodular (SLSM), so $\Psi(y, q, \mathbf{n}, \mathbf{c}, \mathbf{0})$ is SLSM in (y, q). Since $w(\omega)$ is increasing and $f(\omega, k)$ is SLSM, $g(y, k)$ is SLSM. Since $\Psi(y, q, \mathbf{n}, \mathbf{c}, \mathbf{0})$ is SLSM in (y, q) and $g(y, k)$ is SLSM in (y, k), $\Gamma_k(q, \mathbf{n}, \mathbf{c}, \mathbf{0})$ is strictly log-supermodular in (q, k) (Lehmann, 1955). $\square$

[1] Note that the inequality of expectations holds for an arbitrary $\omega > \omega'$. Therefore, $\{\frac{b(\omega,i)^{\sigma}}{b(\omega',i)^{\sigma}}\}_{\omega, \omega' \in \Omega, \omega > \omega'}$ is a class of strictly increasing functions. If this class were all increasing functions, we would conclude that $\phi_i(\omega', k)$ stochastically dominates $\phi_i(\omega', k')$, since $\mathbb{E}_{\tilde{x}}(u(x)) \geq \mathbb{E}_{\tilde{x}'}(u(x)) \; \forall u'(x) > 0 \iff F_{\tilde{x}}(x) \leq F_{\tilde{x}'}(x) \; \forall x$, where F is the cumulative distribution function.

Taking the theory to plant-level data

If $b(\omega, q)$ is strictly decreasing in q, higher-quality varieties are more costly to produce, $c(q,k) > c(q',k) \iff q > q'$.

$$\frac{\partial c(q,k)}{\partial q} = \frac{-\sigma}{\sigma-1} c(q,k)^{\sigma} \int b(\omega,q)^{\sigma-1} w(\omega,k)^{1-\sigma} \frac{\partial b(\omega,q)}{\partial q} d\omega$$
$$\frac{\partial b(\omega,q)}{\partial q} < 0 \ \forall \omega \Rightarrow \frac{\partial c(q,k)}{\partial q} > 0$$

If $g(y,k)$ is log-supermodular, average income is a sufficient statistic for k. $k > k'$ if and only if $g(y,k)$ likelihood-ratio dominates $g(y,k')$, so $\mathbb{E}_k(y) > \mathbb{E}_{k'}(y) \iff k > k'$.

The composition measure assumes that non-production workers are more skilled than production workers. Denote the fraction of workers of skill ω labeled as non-production by $l(N,\omega)$ and the fraction labeled production as $l(P,\omega) = 1 - l(N,\omega)$. Denote the share of non-production workers employed in a plant with skill intensity i in location k by $share_N(i,k) \equiv \frac{\int_\omega \ell(\omega,i(q),k) l(N,\omega) d\omega}{\int_\omega \ell(\omega,i(q),k) d\omega}$. If $l(N,\omega)$ is strictly increasing in ω, then $share_N(i,k)$ is strictly increasing in i.[2] Inside the factor-price equalization (FPE) set, $\ell(\omega,q,k) = \ell(\omega,q) \ \forall k$ and therefore $share_N(i,k) = share_N(i) \ \forall k$. Outside the FPE set, if $w(\omega,k)^{-\sigma}$ is log-supermodular, $share_N(i,k)$ is strictly increasing in k.[3] I therefore use $share_N(j) \times \ln \bar{y}_k$ as an additional control for skill intensity.

The wage measures assume that wages are increasing in skill, ω.[4] If wages are increasing in skill, we can infer the skill intensity of a plant's variety from its average wage. The average wage at a plant producing quality q with skill intensity $i(q)$ in location k is

$$\bar{w}(i,k) = \bar{w}(q,k) = \int_\omega \frac{w(\omega,k)\ell(\omega,i(q),k)}{\int_{\omega'} \ell(\omega',i(q),k) d\omega'} d\omega = \int_\omega w(\omega,k)\varphi(\omega,i,k) d\omega,$$

[2]This assumption is analogous to Property (28) in Costinot and Vogel (2010), which connects observable and unobservable skills. If $l(N,\omega)$ is strictly increasing in ω, then choosing the labeling scheme of worker type $t = N$ or $t = P$ with $N > P$ makes $l(t,\omega)$ a strictly log-supermodular function. Since $\ell(\omega,i,k)$ is strictly log-supermodular in (ω,i) and strict log-supermodularity is preserved by integration, the integral $\int_\omega \ell(\omega,i,k) l(t,\omega) d\omega$ is strictly log-supermodular in (t,i). As a result, the ratio $\int_\omega \ell(\omega,i,k) l(N,\omega) d\omega / \int_\omega \ell(\omega,i,k) l(P,\omega) d\omega$ is strictly increasing in i. Therefore $share_N(i,k)$ is strictly increasing in i.

[3]This follows from $\ell(\omega,i,k)$ log-supermodular in (ω,k) and $l(t,\omega)$ log-supermodular in (t,ω).

[4]A sufficient condition is $\frac{\partial \ln w(\omega,k)}{\partial \omega} = \frac{\partial \ln b(\omega,q)}{\partial \omega} - \frac{1}{\sigma} \frac{\partial \ln f(\omega,k)}{\partial \omega} > 0 \ \forall q \forall k$. Informally, more skilled individuals have greater absolute advantage than local abundance.

where $\varphi(\omega, i(q), k) \equiv \frac{\ell(\omega,i(q),k)}{\int_{\omega'} \ell(\omega',i(q),k)d\omega'}$ is a density that is strictly log-supermodular in (ω, i), which means that $\varphi(\omega, i, k)$ likelihood-ratio dominates $\varphi(\omega, i', k)$ if and only if $i > i'$.[5] The average wage $\bar{w}(i, k)$ is therefore strictly increasing in skill intensity i. We can similarly define the average wage of production workers by

$$\bar{w}_P(i, k) = \int \frac{w(\omega, k)\ell(\omega, q, k)l(P, \omega)}{\int \ell(\omega', q, k)l(P, \omega')d\omega'} d\omega = \int w(\omega, k)\varphi_P(\omega, i, k)d\omega,$$

and an analogous average wage $\bar{w}_N(i, k)$ for non-production workers with density $\varphi_N(\omega, i, k)$. $\varphi_P(\omega, i, k)$ and $\varphi_N(\omega, i, k)$ are strictly log-supermodular in (ω, i), so these average wages are strictly increasing in skill intensity i.

Inside the FPE set, wages equalize across locations, $\bar{w}(q, k) = \bar{w}(q)\ \forall k$, and ranking plants by their averages wages is equivalent to ranking them by their factor intensities, $\bar{w}_j > \bar{w}_{j'} \iff i(j) > i(j')$. $\bar{w}_P(q, k)$ and $\bar{w}_N(q, k)$ also have these properties. This motivates using these average wages as establishment-level controls for skill intensity. Outside the FPE set, if $w(\omega, k)^{1-\sigma}$ is log-supermodular, $w(\omega, k)\varphi(\omega, i, k)$ is strictly log-supermodular in (ω, i) and in (ω, k). As a result, $\bar{w}(i, k)$ is increasing in i, increasing in k, and log-supermodular. I therefore use $\bar{w}_j \times \ln \bar{y}_k$ as an additional establishment-level control for skill intensity.

Describing the composition of demand using per capita incomes exploits the fact that this is a sufficient statistic for relative demand for qualities under the model's assumptions. Lemma 1 in Fajgelbaum, Grossman, and Helpman (2011) shows that, when $g(y, k)$ is log-supermodular and trade costs are small, relative demand for higher-q varieties is greater in the higher-k location. That is, when $g(y, k)$ is log-supermodular, $\Gamma_{k'}(q, \bar{\mathbf{n}}, \bar{\mathbf{c}}, \mathbf{0})$ is log-supermodular in (q, k). As mentioned previously, when $g(y, k)$ is log-supermodular, income per capita is a sufficient statistic for k.

[5] $\bar{w}(i, k)$ and $\varphi(\omega, i, k)$ can be written in terms of the skill intensity $i(q)$ because $i(q) = i(q') \Rightarrow \frac{\ell(\omega,q,k)}{\int_{\omega'} \ell(\omega',q,k)d\omega'} = \frac{\ell(\omega,q',k)}{\int_{\omega'} \ell(\omega',q',k)d\omega'}$.

A.2 Data appendix

A.2.1 Public data

A.2.1.1 Geography

All the reported results describe core-based statistical areas (CBSAs). ZIP-code tabulation areas (ZCTAs), counties, and public-use microdata areas (PUMAs) were assigned to OMB-defined CBSAs using the MABLE Geocorr2K geographic correspondence engine from the Missouri Census Data Center. I use the CBSA geographies as defined in November 2008. These consist of 366 metropolitan areas and 574 micropolitan statistical areas.

In the gravity regressions and in constructing the market-access measures, I define the mileage distance between two CBSAs as the geodetic distance between their population centers. These population centers are the population-weighted average of the latitude and longitude coordinates of all ZCTAs within the CBSA, using population counts from the 2000 Census. I define the mileage distance from a CBSA to itself as the population-weighted average of the pairwise geodetic distances between all the ZCTAs within the CBSA. For five CBSAs containing only one ZCTA, I generated this mileage distance to self using the predicted values obtained by projecting mileage distance to self for the other 935 CBSAs onto their land areas.

A.2.1.2 Locational characteristics

Data on CBSAs' aggregate populations and personal incomes come from the BEA's regional economic profiles for 1997, 2002, and 2007, data series CA30.

Data on the distribution of household incomes for the 1997 and 2002 samples were constructed from county-level estimates reported in U.S. Census Bureau, 2000 Census Summary File 3, Series P052. Data on the distribution of household incomes at the CBSA level for the 2007 sample were obtained from U.S. Census Bureau, 2005-2009 American Community Survey 5-Year Estimates, Series B19001.

City-industry college shares were constructed from the 2000 Census and 2005-2009 American Community Surveys microdata made available via IPUMS-USA (Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek, 2010). City-industry means and standard deviations of years of schooling and wages from the 2000 Census and 2005-2009 American Community Surveys microdata made available via IPUMS-USA (Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek, 2010).

Locations' latitudes and longitudes were compiled from various sources. Latitude and longitude coordinates for US ZIP-code Tabulation Areas (ZCTAs) were obtained from the 2000 Census.[6] The geodetic distances for export shipments in appendix section A.5.3 were calculated using latitude and longitude coordinates for major Canadian cities, which were constructed by aggregating Canadian dissemination areas' populations and coordinates in the 2006 Census Geographic Attribute File from Statistics Canada, and coordinates of each nation's capital or main city from Mayer and Zignago (2011).

A.2.1.3 Industrial and product characteristics

The Sutton (1998) R&D and advertising intensity measure of scope for vertical differentiation is provided at the SIC72 level in Federal Trade Commission (1981). These were mapped to 1987 SIC codes using the Bartlesman, Becker, and Gray concordance from Jon Haveman's website and to 1997, 2002, and 2007 (via 2002) NAICS codes using concordances from the US Census. For industries to which multiple SIC72 industries were mapped, I calculated the weighted average of intensities, using 1975 sales as weights.

The Khandelwal (2010) 'ladder' measure of scope for vertical differentiation was mapped from HS10 product codes to 6-digit NAICS codes using the Pierce and Schott (2012) concordance. For industries to which multiple commodities were mapped, I calculated the weighted average of ladder lengths, using the initial period import values reported by Khandelwal as weights.

In estimating demand shifters, I mapped the Feenstra and Romalis (2012) estimates of $\hat{\sigma}$ and $\hat{\lambda}$ from SITC revision 2 commodity codes to 5-digit SCTG product codes using a United Nations SITC-HS concordance and a Statistics Canada HS-SCTG concordance. When multiple parameter estimates mapped to a single 5-digit SCTG product code, I used the median values of $\hat{\sigma}$ and $\hat{\lambda}$. The results are robust to using the arithmetic mean.

[6] Downloaded from http://www.census.gov/tiger/tms/gazetteer/zcta5.txt in December 2012.

A.2.2 Confidential Census data

I use establishment-level microdata from the 1997, 2002, and 2007 editions of the Commodity Flow Survey (CFS) and Census of Manufactures (CMF).

A.2.2.1 Commodity Flow Survey

I use data describing shipments by manufacturing plants from the Commodity Flow Survey, a component of the quinquennial Economic Census. Each quarter of the survey year, establishments report a randomly selected sample of 20-40 of their shipments from a given week and describe them in terms of commodity content, value, weight, destination, transportation mode, and other characteristics. The approximately 100,000 establishments sampled by the CFS were selected using a stratified sampling design reflecting the Commodity Flow Survey's objectives (Bureau of Transportation Statistics and US Census Bureau, 2010); of the approximately 350,000 manufacturing establishments in the United States, about 10,000 per year appear in my estimation sample.[7]

These data are analogous to firm-level customs data with four important distinctions. First, the data describe shipments at the establishment level rather than at the firm level. Second, the geographic detail of ZIP-to-ZIP shipments is orders of magnitude more precise than the distance measures used to describe international transactions. Each shipment's mileage was estimated by BTS/Census using routing algorithms and an integrated, intermodal transportation network developed for that purpose. Third, establishments report a sample of their shipments in the survey, not a complete record of all transactions. Each quarter of the survey year, establishments report a randomly selected sample of 20-40 of their shipments in one week. The CFS data include statistical weights that can be used to estimate aggregate shipment flows. Fourth, the CFS uses a distinct product classification scheme, the Standard Classification of Transport Goods, that is related to the Harmonized System used in international trade data. At its highest level of detail, five digits, the SCTG defines 512 product categories.[8]

[7] The 2002 CFS sample is roughly half that of the 1997 and 2007 surveys, sampling about 50,000 establishments in total and a proportionate number that appear in my estimation sample (Bureau of Transportation Statistics and US Census Bureau, 2004).

[8] By comparison, the HS scheme has 97 2-digit and about 1400 4-digit commodity categories.

The Commodity Flow Survey microdata include statistical weights so that observations can be summed to obtain estimated totals that are representative. Each shipments' associated "tabulation weight" is the product of seven component weights (Bureau of Transportation Statistics and US Census Bureau, 2010, Appendix C). The products of four of these weights (shipment weight, shipment nonresponse weight, quarter weight, and quarter nonresponse weight) scale up an establishment's shipments to estimate the establishments' total annual shipments. The other three component weights (establishment-level adjustment weight, establishment weight, industry-level adjustment weight) scale up establishments' total shipments to estimate national shipments. The 1997 and 2002 microdata report only the tabulation weights, while the 2007 microdata report all seven component weights.

The demand estimation performed in section A.5.1 requires measures of establishments' market shares, which are calculated from estimates of their total sales of that product in a destination market. These shares are estimated using the first four component weights. These establishment-level measures should not be scaled up by the latter three component weights, such as the probability of the establishment being selected into the CFS sample. As a result, it is only possible to estimate the plant-product demand shifters using the 2007 microdata, which include the component weights required to estimate market shares.

The CFS classifies shipments' commodity contents using the Standard Classification of Transported Goods (SCTG), a coding system based on the Harmonized System (HS) classification that was introduced in the 1997 CFS. At its highest level of detail, five digits, the SCTG defines 512 product categories. By comparison, the HS scheme has 97 2-digit and about 1400 4-digit commodity categories. I exclude from my analysis all SCTG product categories whose 5-digit product codes end in 99, since these are catch-all categories such as 24399 "Other articles of rubber."

I calculate shipment unit values by dividing shipment value by shipment weight. All my analyses of these unit values are within-product comparisons or regressions incorporating product fixed effects. These unit values are proxies for producer prices, because they do not include shipping costs or shelving costs that may appear in the retail consumer price.

Each shipment is reported to have been sent by any combination of eight transportation modes. In much of the analysis, I restrict attention to unimodal shipments, which account for more than

80% of shipments by value and 90% by weight (Bureau of Transportation Statistics and US Census Bureau, 2010, Table 1b).

A.2.2.2 Census of Manufactures

Each manufacturing plant appearing in the Commodity Flow Survey also appears in the Census of Manufactures, which describes plant-level characteristics such as wage bills, production and non-production employees, and capital stocks. Very small establishments do not report detailed production data to the CMF. Instead, the Census Bureau uses data from administrative records from other agencies, such as tax records, to obtain information on revenues and employment. It then imputes other variables.

The product trailer of the CMF describes the products produced at each establishment. For all products, establishments report the total value of their annual output. For a subset of products, establishments report both values and quantities.[9] I calculate unit values by dividing product value shipped by product quantity shipped; these unit values are used in appendix Table A.3.

A.2.2.3 Longitudinal Business Database

I also use information from the Longitudinal Business Database (LBD), which is a census of US business establishments and firms with paid employees. Microdata from the 1997, 2002, and 2007 editions are a combination of survey and administrative records. I use the LBD for two purposes. First, I use the records in linking establishments across the 2002 CFS and CMF data sets. Second, I use an establishment's first year appearing in the LBD, which is a comprehensive census, to calculate plants' ages.

A.2.2.4 Combining the CFS and CMF

I matched shipment-level observations in CFS data to establishment-level characteristics in CMF data using unique establishment identifiers called Census File Numbers (1997) and Survey Unit Identifiers (2002, 2007). I also used information in the LBD to address the switch from CFNs to SUIs in 2002.

[9] The set of products for which product quantity shipped data are collected has shrunk over time.

Each data source contains information on the geographic location of an establishment. The CMF reports the county in which an establishment is located. The CFS reports the ZIP code and state in all survey years and some information about core-based statistical areas in 2002 and 2007.

A.2.2.5 Sample selection

Though the CFS and CMF data are very rich descriptions of establishments and their shipments, some of the observations exhibit limitations that warrant their exclusion from the estimation sample.

CFS & CMF: I exclude establishments not belonging to a OMB-defined CBSA. I exclude a small number of establishments for which the CFS and CMF do not report the same CBSA. I exclude SCTG5-NAICS6 pairs in which fewer than five establishments report shipping a commodity.

CFS: I restrict the sample to unimodal shipments, which constitute more than 80% of shipments by value and 90% by weight (Bureau of Transportation Statistics and US Census Bureau, 2010). I exclude shipments with unit values more than two standard deviations from the product mean. I exclude destination-product pairs for which only one establishment ships that product to that destination. I exclude shipments that are the unique instance of that commodity being shipped by that establishment.

CMF: I exclude establishments whose information in the Census of Manufactures are derived from administrative records rather than directly reported. I exclude establishments whose employment levels or wage bills are imputed in the 2002 and 2007 CMF.[10] I exclude establishments with wages that lie below the 1st percentile or above the 99th percentile of the wage distribution for manufacturing establishments. I exclude CBSA-NAICS6 pairs in which total employment is reported to be more than 10% of residential population.

A.3 Gravity appendix

This appendix characterizes the pattern of manufactures shipments between US cities using a gravity model of shipment volumes. Gravity regressions relate the volume of trade to the origin's economic size in terms of output produced, the destination's economic size in terms of consumer

[10]The 1997 CMF does not identify variables that have been imputed, so I am unable to exclude such observations.

expenditure, and trade frictions between the origin and destination.[11] Hallak (2010) and Bernasconi (2013) use gravity models of trade flows between countries to assess whether locations with more similar income levels trade more with each other, controlling for origin characteristics, destination characteristics, and bilateral trade frictions. They find that countries with more similar income distributions trade more with each other, as predicted by Linder (1961). I find a similar pattern of trade flows between US cities.

The baseline gravity specification is

$$\ln X_{odst} = \eta \cdot \ln \text{miles}_{od} + \beta \cdot |\ln \bar{y}_{ot} - \ln \bar{y}_{dt}| + \gamma_{ost} + \gamma_{dst} + \epsilon_{odst}$$

where X_{odst} is the volume of shipments in sector s sent from origin o to destination d in year t, miles_{odt} is the distance between the two locations, $|\ln \bar{y}_{ot} - \ln \bar{y}_{dt}|$ is the difference in their log per capita incomes, γ_{ost} and γ_{dst} origin-sector-year and destination-sector-year fixed effects, and ϵ_{odst} is a residual reflecting both random sampling and potential measurement error.[12]

This baseline specification is estimated using observations with strictly positive shipment volumes. In fact, there are many zeros in the trade matrix, and these non-positive shipment volumes reflect economic mechanisms, like trade costs. For example, every city ships a positive amount to itself in every sector $X_{oost} > 0 \ \forall o \forall s \forall t$.

I use two approaches to correct for the non-random nature of zeros. First, I implement the Heckman (1979) two-step selection correction. The first-step probit regression, which has the same regressors on the right-hand side, yields an estimated probability of a strictly positive shipment volume for each origin-destination-sector-year. The second-step estimating equation is

$$\ln X_{odst} = -\eta \cdot \ln \text{miles}_{od} + \beta \cdot |\ln \bar{y}_{ot} - \ln \bar{y}_{dt}| + \delta \frac{\phi(\Phi^{-1}(\hat{\rho}_{odst}))}{\hat{\rho}_{odst}} + \gamma_{ost} + \gamma_{dst} + \epsilon_{odst},$$

[11] See Anderson (2011) and Bergstrand and Egger (2011) for surveys of the gravity literature. Importantly, the correct notions of economic size account for "multilateral resistance" terms that depend on the locations' bilateral trade frictions with all trading partners. These are captured by fixed effects in my regressions.

[12] Observations of X_{odst}include sampling error since a representative sample of shipments is used to estimate the total shipment volume. See the data appendix A.2 for details. Noise in the dependent variable will not bias the estimated coefficients of interest.

where $\hat{\rho}_{odst}$ is the predicted probability from the probit regression, ϕ and Φ are the probability and cumulative density functions of the normal distribution, and $\frac{\phi(\Phi^{-1}(\hat{\rho}_{odst}))}{\hat{\rho}_{odst}}$ is the inverse Mills ratio. The second approach I use to address zeros is the Poisson pseudo-maximum likelihood estimator introduced by Silva and Tenreyro (2006) to estimate the gravity equation in levels.

$$\mathbb{E}(X_{odst}) = \text{miles}_{od}^{-\eta} |\ln \bar{y}_{ot} - \ln \bar{y}_{dt}|^{\beta} \exp(\gamma_{ost} + \gamma_{dst})$$

This estimation approach allows me to include observations for which $X_{odst} = 0$.[13]

Table A.1 reports the result of estimating this gravity regression for aggregate manufactures shipment volumes. The first column reports the result of estimating the baseline gravity specification via OLS. The second column uses the squared difference in incomes, $(\ln \bar{y}_{ot} - \ln \bar{y}_{dt})^2$, as a regressor instead of the absolute difference. The third column adds the two-step Heckman (1979) correction for selection into strictly positive trade flows. The fourth and fifth columns use the Silva and Tenreyro (2006) Poisson pseudo-maximum likelihood estimator for all observations and strictly positive trade flows, respectively.

In all the relevant specifications, the difference between two cities' average income levels is negatively correlated with the level of trade between them. The Linder pattern of trade, in which locations with more similar incomes trade more intensely with each other, holds true for manufactures shipments between US cities. The finding that locations disproportionately demand products that are produced in locations of similar income levels suggests two elements that any model explaining within-product specialization must incorporate. First, preferences are non-homothetic, so the composition of demand varies with locations' incomes. Second, high-income locations have comparative advantage in producing products that are particularly attractive to high-income consumers. As described in section 1.2, these patterns are compatible with the factor-abundance mechanism or home-market effect determining the pattern of within-product specialization.

The distance elasticity of trade, η, is estimated to be near one. Most of the estimates are about 0.9, while the two-step Heckman specification implies that shipment volumes are notably

[13] Silva and Tenreyro (2006) emphasize that their estimation procedure addresses a concern that higher moments of $\exp(\epsilon_{odst})$ are correlated with the regressors, which would cause the estimated coefficients in the log-linear specification to be inconsistent.

Table A.1: Shipment volumes (2007)

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	Heckman	PPML w/ zeros	PPML w/o zeros
Dep var:	$\ln X_{od}$	$\ln X_{od}$	$\ln X_{od}$	X_{od}	X_{od}
$\ln \text{miles}_{od}$	-0.916**	-0.917**	-1.773**	-0.952**	-0.831**
	(0.00745)	(0.00715)	(0.0104)	(0.0170)	(0.0170)
$\lvert \ln \bar{y}_o - \ln \bar{y}_d \rvert$	-1.696**		-1.074**	-0.430**	-1.241**
	(0.0413)		(0.0454)	(0.0840)	(0.0934)
$(\ln \bar{y}_o - \ln \bar{y}_d)^2$		-2.657**			
		(0.0854)			
Inverse Mills $\frac{\phi(\Phi^{-1}(\hat{\rho}_{od}))}{\hat{\rho}_{od}}$			2.434**		
			(0.0222)		
R-squared	0.347	0.347	0.378		
Observations (rounded)	175,000	175,000	175,000	850,000	175,000
Origin CBSAs (rounded)			900		
Destination CBSAs (rounded)			950		

Standard errors in parentheses

** p<0.01, * p<0.05

NOTES: Aggregate shipment volume by establishments in the CFS and CMF between distinct CBSAs in 2007. All regressions include origin and destination fixed effects. Standard errors are bootstrapped with 50 repetitions in columns 1-3 and heteroskedastic-robust in columns 4-5.

more sensitive to the distance between origin and destination. These results are consistent with the central tendency of the vast international literature summarized by Disdier and Head (2008) and the elasticity of domestic shipments reported in Table 1 of Hillberry and Hummels (2008) for geographically aggregate shipment volumes.

A.4 Tables appendix

Table A.2: Outgoing shipment prices with city-industry schooling measures

Dep var: Log unit value	(1)	(2)	(3)	(4)	(5)
Log origin CBSA income per capita	0.529**	0.472**	0.392**	0.404**	0.340**
	(0.0546)	(0.0529)	(0.0517)	(0.0587)	(0.0570)
Log origin CBSA population	-0.0217**	-0.0241**	-0.0253**	-0.0209**	-0.0232**
	(0.00672)	(0.00635)	(0.00607)	(0.00731)	(0.00688)
Log non-production worker share		-0.0149	-0.0536	-0.0171	-0.0580
× log per capita income		(0.0450)	(0.0527)	(0.0454)	(0.0529)
Log assets per worker		0.00456	-0.0250	0.00485	-0.0251
× log per capita income		(0.0251)	(0.0295)	(0.0245)	(0.0290)
Log pay per worker			0.0680		0.0933
× log per capita income			(0.146)		(0.146)
Log pay per production worker			0.265*		0.249*
× log per capita income			(0.110)		(0.113)
Log pay per non-production worker			0.0207		0.0184
× log per capita income			(0.0661)		(0.0665)
City-industry mean years schooling				0.0191	-0.000797
× log per capita income				(0.0325)	(0.0332)
R-squared	0.883	0.885	0.886	0.885	0.887
Note	flex miles	ctrl qnt flex	ctrl all flex	ctrl qnt flex + school flex	ctrl all flex + school flex
Observations (rounded)			1,000,000		
Estab-year (rounded)			22,500		
Ind-prod-year (rounded)			2,000		

Standard errors, clustered by CBSA × year, in parentheses

** $p<0.01$, * $p<0.05$

NOTES: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include SCTG5×NAICS6×destination×year fixed effects and mode × year fixed effects. Unreported controls are 3-digit-NAICS-specific third-order polynomials in log mileage (columns 1-5), log non-production worker share (2-5), log assets per worker (2-5), log pay per worker (3,5), log pay per producer worker (3,5), log pay per non-production worker (3,5), and city-industry means years of schooling (4-5).

Table A.3: Establishments' prices and origin characteristics

Dep var: CMF Log unit value	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log origin CBSA income per capita	0.218**	0.204**	0.193**	0.192**	0.155**	0.142**	0.0591
	(0.0384)	(0.0384)	(0.0392)	(0.0409)	(0.0373)	(0.0392)	(0.0459)
Log origin CBSA population	-0.000772	-0.00120	-0.00151	-0.000374	0.00380	0.00345	-0.00274
	(0.00484)	(0.00481)	(0.00473)	(0.00501)	(0.00483)	(0.00492)	(0.00490)
Log non-production worker share		0.0564**	0.0471**	√		√	√
		(0.00539)	(0.00652)				
Log assets per worker		-0.0125**	-0.0159**	√		√	√
		(0.00395)	(0.00404)				
Log non-production share		0.0261	0.0320	0.0147		0.0125	0.00937
× log per capita income		(0.0268)	(0.0343)	(0.0311)		(0.0310)	(0.0310)
Log assets per worker		-0.0229	-0.0183	-0.0229		-0.0226	-0.0181
× log per capita income		(0.0226)	(0.0213)	(0.0198)		(0.0197)	(0.0198)
Log pay per worker			0.109**	√		√	√
			(0.0281)				
Log pay per production worker			-0.0411*	√		√	√
			(0.0187)				
Log pay per non-production worker			-0.0257*	√		√	√
			(0.0124)				
Log pay per worker			-0.0729	0.00495		0.00157	-0.0108
× log per capita income			(0.170)	(0.147)		(0.146)	(0.147)
Log pay per production worker			-0.0408	-0.0751		-0.0752	-0.0922
× log per capita income			(0.120)	(0.101)		(0.100)	(0.101)
Log pay per non-production worker			-0.00298	0.00839		0.00891	0.00244
× log per capita income			(0.0575)	(0.0531)		(0.0527)	(0.0531)
Market access (excl origin) M^1_{ot}					0.564**	0.469**	
					(0.136)	(0.136)	
Log std dev household income							0.0804
							(0.0460)
Market access M^2_{ot}							0.403**
							(0.113)
R-squared	0.914	0.914	0.914	0.917	0.914	0.917	0.917
Obs (rounded)				100000			
Estab-year (rounded)				27500			
Ind-prod-year (rounded)				8000			

Standard errors, clustered by CBSA × year, in parentheses
** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CMF reporting a product with quantity shipped. All regressions include NAICS6×year and NAICS product code×year fixed effects. The fourth, sixth, and seventh columns include 3-digit-NAICS-specific third-order polynomials in log non-production worker share, log assets per worker, log pay per worker, log pay per producer worker, and log pay per non-production worker.

Table A.4: Outgoing shipment prices with plant-size controls

Dep var: Log unit value	(1)	(2)	(3)
Log origin CBSA income per capita	0.412**	0.366**	0.298**
	(0.0415)	(0.0402)	(0.0384)
Log origin CBSA population	-0.00823	-0.0112*	-0.0150**
	(0.00454)	(0.00441)	(0.00413)
Log non-production worker share		0.0253	0.0194
× log per capita income		(0.0345)	(0.0398)
Log assets per worker		0.000542	-0.0246
× log per capita income		(0.0191)	(0.0210)
Log pay per worker			-0.0819
× log per capita income			(0.121)
Log pay per production worker			0.363**
× log per capita income			(0.0876)
Log pay per non-production worker			0.133*
× log per capita income			(0.0555)
R-squared	0.880	0.882	0.883
Observations (rounded)		1,400,000	
Estab-year (rounded)		30,000	
Ind-prod-year (rounded)		2,000	

Standard errors, clustered by CBSA × year, in parentheses

** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include SCTG5×NAICS6×destination×year fixed effects and mode × year fixed effects. Unreported controls are 3-digit-NAICS-specific third-order polynomials in log mileage (columns 1-3), log establishment size (1-3), log non-production worker share (2-3), log assets per worker (2-3), log pay per worker (3), log pay per producer worker (3), and log pay per non-production worker (3).

A.5 Supplementary appendix

This appendix reports further empirical evidence consistent with the results presented in section 1.5. First, estimated demand shifters exhibit the same patterns as outgoing shipment prices. Second, cities with greater income dispersion have higher outgoing shipment prices, consistent with the model's demand system in an equilibrium in which most individuals purchase low-quality varieties. Third, export shipments exhibit patterns consistent with those found in domestic transactions.

A.5.1 Estimated demand shifters

This section characterizes the pattern of within-product specialization across US cities and its determinants using estimated demand shifters. As previously described, consumer love of variety in the presence of horizontal differentiation breaks the price-quality mapping by allowing high-cost varieties to sell alongside low-cost varieties of the same quality. Section 1.5 addresses this concern by including a variety of plant-level cost measures, which were not available to researchers analyzing aggregate trade flows between countries. This section addresses the concern a second time by estimating demand shifters for each plant-product pair. The empirical results are consistent with the unit-value findings.

The demand-shifter approach assigns higher quality valuations to products that have higher market shares, conditional on price (Berry, 1994; Khandelwal, 2010; Sutton, 2012). As described in the data appendix, it is only possible to calculate plants' market shares in the 2007 edition of the Commodity Flow Survey. This considerably reduces the number of observations compared to the number underlying the previously presented results.

To estimate demand shifters, I use the "non-homothetic CES preferences" of Feenstra and Romalis (2012).[14] In this specification, the sales volume s_{jd} of product j in destination market d in 2007 is described by

[14] One merit of this demand system is its computational simplicity. Because the nested-logit demand system used in the Fajgelbaum, Grossman, and Helpman (2011) model uses quality levels as nests, its estimation would require a computationally intensive iterative approach. Products must be assigned to quality nests in order to estimate the demand system, and product qualities must be inferred by estimating demand.

$$\ln s_{jd} = (\sigma - 1)(\ln q_j + \lambda \ln \bar{y}_d \ln q_j - \ln p_{jd}) + \gamma_d + \epsilon_{jd}$$

where q_j is the product-quality shifter, p_{jd} is price, $\bar{y}_d$ is per capita income in market d, γ_d captures both aggregate expenditure and the price index in the destination market, and ϵ_{jd} captures both idiosyncratic demand shocks and measurement error. The parameter λ governs how consumer valuation of quality varies with income; the parameter σ is the elasticity of substitution and price elasticity of demand. In the demand system, p_{jd} is the price paid by the consumer, while in my data the observed price $\check{p}_{jd}$ excludes shipping costs.[15] I therefore include the shipment mileage from establishment to destination and the shipment mileage interacted with price, $\ln miles_{jd}$ and $\frac{1}{\check{p}_{jdt}} \ln miles_{jd}$, as additional regressors to control for shipping costs.[16] I use sectoral estimates of $\hat{\lambda}$ and $\hat{\sigma}$ from Feenstra and Romalis (2012) in order to estimate q_j in the linear regression

$$\frac{\ln s_{jd} + (\hat{\sigma} - 1)\ln \check{p}_{jd}}{(1 + \hat{\lambda} \ln \bar{y}_d)(\hat{\sigma} - 1)} = \ln q_j + \eta_1 \ln miles_{jd} + \eta_2 \frac{1}{\check{p}_{jd}} \ln miles_{jd} + \tilde{\gamma}_d + \tilde{\epsilon}_{jd}$$

where $\tilde{\gamma}_d$ and $\tilde{\epsilon}_{jd}$ are rescaled versions of γ_d and ϵ_{jd}. These regressions are estimated product-by-product, for 1000 products defined by SCTG5-NAICS6 codes, for the 2007 sample.[17]

Table A.5 describes how these estimated demand shifters relate to the observable characteristics of products, plants, and cities. The first column reports that the estimated demand shifters are strongly positively correlated with plants' prices. This validates the use of prices in inferring the pattern of quality specialization earlier in the paper. The second column shows that plants with higher estimated demand shifters are located in cities with higher per capita incomes. The 41% origin-income elasticity of the estimated demand shifter is remarkably similar to the 43% origin-

[15] In international trade parlance, demand depends on the "cost-insurance-and-freight" price while my data reports the "free-on-board" price. If the consumer price reflects both multiplicative and additive trade costs, τ^m_{jd} and τ^a_{jd}, then $p_{jd} = \check{p}_{jd}\tau^m_{jd} + \tau^a_{jd}$ and $\ln p_{jd} \approx \ln \check{p}_{jd} + (\tau^m_{jd} - 1) + \frac{1}{\check{p}_{jd}}\tau^a_{jd}$. Assuming that τ^m_{jd} and τ^a_{jd} are functions of the shipment distance motivates the inclusion of $\ln miles_{jd}$ and $\frac{1}{\check{p}_{jd}} \ln miles_{jd}$as regressors.

[16] Omitting these regressors has very little impact on the estimated demand shifters and the subsequent results relating these shifters to city and plant characteristics.

[17] I obtain very similar results if I define products using only 5-digit SCTG codes. See the appendix section A.2 for details of how I mapped the Feenstra and Romalis (2012) parameter estimates to these product codes.

Table A.5: Estimated demand shifters

Dep var: $\ln \hat{q}_j$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log unit value	0.746**										
	(0.0102)										
Log CBSA income per capita		0.412**	0.270**	0.157	0.0856	0.265**	0.164	0.237**	0.124	0.0451	0.0401
		(0.0840)	(0.0819)	(0.0831)	(0.105)	(0.0857)	(0.112)	(0.0793)	(0.0814)	(0.101)	(0.114)
Log CBSA population		0.00577	-0.00259	0.00629	0.00197	0.0170	0.0119	-0.000985	0.00793	0.00375	-0.0154
		(0.0120)	(0.0114)	(0.0109)	(0.0113)	(0.0115)	(0.0119)	(0.0109)	(0.0106)	(0.0109)	(0.0107)
Market access (excl origin) M^1_{ot}				0.807**		1.034**			0.822**		
				(0.241)		(0.250)			(0.241)		
Market access M^2_{ot}					0.681*		0.904**			0.715*	
					(0.277)		(0.319)			(0.278)	
Log std dev orig household income											0.317**
											(0.119)
Log non-production worker share			0.0907**	0.0892**	0.0907**			√	√	√	√
			(0.0135)	(0.0136)	(0.0135)						
Log assets per worker			-0.0496**	-0.0485**	-0.0484**			√	√	√	√
			(0.0109)	(0.0109)	(0.0109)						
Log non-production worker share			0.0106	0.00320	-0.00394			0.0212	0.0116	0.00263	0.0152
× log per capita income			(0.0727)	(0.0735)	(0.0728)			(0.0784)	(0.0793)	(0.0786)	(0.0780)
Log assets per worker			0.0119	0.0136	0.0230			0.0277	0.0300	0.0417	0.0294
× log per capita income			(0.0592)	(0.0594)	(0.0590)			(0.0623)	(0.0624)	(0.0620)	(0.0622)
Log pay per worker			0.409**	0.406**	0.407**			√	√	√	√
			(0.0485)	(0.0486)	(0.0493)						
Log pay per production worker			0.0881**	0.0844*	0.0844*			√	√	√	√
			(0.0330)	(0.0333)	(0.0331)						
Log pay per non-production worker			0.0376	0.0363	0.0358			√	√	√	√
			(0.0232)	(0.0232)	(0.0232)						
Log pay per worker			0.198	0.201	0.170			0.0201	0.0224	-0.00358	0.0309
× log per capita income			(0.224)	(0.224)	(0.231)			(0.220)	(0.221)	(0.226)	(0.220)
Log pay per production worker			0.196	0.180	0.170			0.0874	0.0687	0.0571	0.0819
× log per capita income			(0.148)	(0.149)	(0.148)			(0.166)	(0.167)	(0.167)	(0.166)
Log pay per non-production worker			-0.0965	-0.104	-0.115			-0.127	-0.132	-0.145	-0.145
× log per capita income			(0.120)	(0.119)	(0.121)			(0.113)	(0.112)	(0.114)	(0.110)
R-squared	0.857	0.614	0.626	0.627	0.626	0.615	0.615	0.646	0.647	0.647	0.646
Standard errors	cl estab	cl cbsa	cl cbsa	cl cbsa	cl cbsa	cl cbsa	cl cbsa	cl cbsa	cl cbsa	cl cbsa	cl cbsa
Observations (rounded)						13,000					
Establishments (rounded)						10,000					
Ind-prod (rounded)						1,000					

Clustered standard errors in parentheses

** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CFS and CMF. All regressions include SCTG5 × NAICS6 × year fixed effects. Unreported controls in columns 8-11 are 3-digit-NAICS-specific cubic polynomials in log non-production worker share, log assets per worker, log pay per worker, log pay per production worker, and log pay per non-production worker.

income elasticity of outgoing shipment prices. The third and eighth columns demonstrate that this positive relationship persists after controlling for plants' input usage. Qualitatively consistent with the result in section 1.5.1, observed differences in plant-level factor usage explain less than half of the observed correlation between plants' output characteristics and per capita incomes. The fourth and ninth columns replicate the finding that the income composition of proximate potential customers, excluding those in the city of production, is strongly positively associated with a plant's output profile. The 11-percentage-point decline in the origin-income elasticity caused by introducing the first-market access measure after controlling for factor-usage differences suggests that proximity to these customers explains at least one-quarter of the observed variation. The sixth column demonstrates that introducing the first market-access measure prior to controlling for factor usage would result in a change in the origin-income elasticity of essentially the same magnitude as controlling for factor usage. The fifth, seventh, and tenth columns demonstrate that the second market-access measure, which includes residents in the city of production, has greater explanatory power. The eleventh column replicates the findings of section 1.5.2 by demonstrating a positive relationship with origin CBSA income dispersion and the market-access measure.

These results can be succinctly summarized as a decomposition of the covariance between incomes and shifters. After controlling for population size, differences in observed factor usage are responsible for 46% of the covariance between per capita incomes and estimated demand shifters. Conditional on factor usage, the conservative market-access measure that omits residents in the city of production accounts for 25% of the total covariance, leaving 30% as residual variation. The model-consistent market-access measure that includes residents in the city of production accounts for 48% of the total covariance, leaving 7% as residual variation.[18] Thus, the observed pattern of specialization and inferences about its determinants obtained using estimated demand shifters are similar to those obtained by examining unit values.

[18] Numbers sum to 101% due to rounding.

A.5.2 Income dispersion

This section documents the relationship between the second moment of the income distribution and shipment prices. Cities with greater dispersion in household income exhibit higher prices for both incoming and outgoing shipments. The latter is not explained by dispersion in workers' wages or skills. These findings are consistent with a demand-side mechanism linking local income distributions to the pattern of quality specialization.

In the Fajgelbaum, Grossman, and Helpman (2011) model, income inequality is linked to quality specialization because the income distribution determines the composition of local demand for quality. In general, the effect of greater income dispersion on relative demand for quality is ambiguous. The authors' Proposition 2(iii) shows that, when there are two qualities and the majority of individuals at all income levels consume the low-quality variety, a mean-preserving spread of the income distribution raises local relative demand for the high-quality variety in their demand system. Since the converse would hold if a majority of individuals consumed high-quality varieties, there is no general theoretical result for the correlation between income dispersion and relative demand for quality.

A few theories link income dispersion to specialization through supply-side mechanisms that are absent from the model in section 1.3. In Grossman and Maggi (2000) and Bombardini, Gallipoli, and Pupato (2012), locations with more diverse skill distributions have comparative advantage in sectors in which skills are more substitutable.[19] In Grossman (2004), imperfect labor contracting causes locations with more diverse skill distributions to have comparative advantage in sectors in which the most talented individuals' contributions are more easily identified. Applying these models to question at hand involves reinterpreting them as theories of intrasectoral specialization. For example, if different skills were less substitutable in the production of higher-quality products, these models would predict that locations with greater skill dispersion would specialize in lower-

[19] Though both papers describe locations with greater skill dispersion specializing in sectors with greater substitutability of skills, these two papers differ considerably. Grossman and Maggi (2000) compare two countries and two sectors, one in which output is supermodular in the two workers' talents and another in which output is submodular in talents. They assume that talent is perfectly observed. Bombardini, Gallipoli, and Pupato (2012) describe imperfectly observed skills and CES production functions that vary in their elasticities of substitution between skills.

quality varieties.[20] If they were more substitutable, the reverse prediction would result.

In light of these theoretical ambiguities, I rely on the distinction between income dispersion among local consumers and skill dispersion among local workers to empirically distinguish between the demand-side and supply-side mechanisms. Income dispersion among all potential customers influences the demand channel. In the supply-side theories (appropriately reinterpreted to describe specialization within sectors), only skill dispersion among those working in the industry in question is relevant. Thus, I construct two types of empirical measures: the standard deviation of household income within each city and the standard deviations of years of schooling and weekly wages within each city-industry pair. The former proxies for the demand-side mechanism; the latter for the supply-side. I proceed to include these measures in linear regressions describing shipment prices.

Table A.6 documents how shipment prices are related to income and skill dispersion in the destination and origin cities. The first two columns report the result of adding the standard deviation of household income in the shipment destination to multivariate regressions like those appearing in Table 1.2. The first column omits any origin-city characteristics; the second column includes origin-city fixed effects. In each case, the standard deviation of household income is strongly positively related to the price of incoming shipments. This is consistent with an equilibrium in which a more dispersed income distribution has more households in the right tail of the distribution who purchase higher-price, higher-quality varieties.

The next five columns of Table A.6 relate outgoing shipment prices to income and skill dispersion in the shipment origin. These regressions all include destination-product-year fixed effects and control variables with industry-specific third-order polynomials, like those regressions appearing in the last three columns of Tables 1.3. The third column introduces the standard deviation of household income as a regressor alongside origin characteristics and shipment mileage. The fourth column adds controls for plant-level factor usage in quantities and wages. The standard deviation of household income in the shipment origin is strongly, positively related to the outgoing shipment price. This is consistent with models in which income dispersion generates demand for high-price, high-quality varieties or skill dispersion generates comparative advantage in high-price,

[20] Grossman and Maggi (2000, p.1255,1271) cite quality control as an example of supermodular production in which less dispersion yields comparative advantage.

high-quality varieties. The fifth and sixth columns repeat the third and fourth columns for the subset of observations for which city-industry-level measures of skill dispersion are available so that we can contrast income dispersion amongst potential customers with skill dispersion amongst workers employed in production. The sixth column introduces an additional control, city-industry mean years of schooling, that is available for these observations. Among this subsample, the coefficients on the log standard deviation of household income are more than double their values for the full sample. The key result, appearing in the seventh column, is that controlling for the logs of the standard deviations of years of schooling and weekly wages at the city-industry level leaves the coefficients on origin characteristics virtually unaltered. Skill dispersion on the supply side appears unrelated to outgoing shipment prices. These findings suggest that a demand-side mechanism links the local income distribution to specialization.

Table A.6: Shipment prices and income dispersion

Dep var: Log unit value, $\ln p_{skjodmt}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log destination CBSA income per capita	0.140**	0.0988**					
	(0.0267)	(0.0232)					
Log destination CBSA population	-0.00720**	-0.00806**					
	(0.00248)	(0.00205)					
Log origin CBSA income per capita	0.455**		0.363**	0.259**	0.387**	0.194**	0.195**
	(0.0184)		(0.0492)	(0.0464)	(0.0691)	(0.0724)	(0.0725)
Log origin CBSA population	-0.00936**		-0.0137**	-0.0212**	-0.0321**	-0.0332**	-0.0345**
	(0.00169)		(0.00502)	(0.00468)	(0.00744)	(0.00735)	(0.00774)
Log mileage	0.0421**	0.0461**	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
	(0.00288)	(0.00219)					
Log std dev dest household income	0.118**	0.109**					
	(0.0256)	(0.0224)					
Log std dev orig household income			0.112*	0.110*	0.230**	0.231**	0.219**
			(0.0516)	(0.0478)	(0.0812)	(0.0792)	(0.0818)
Log std dev orig-ind weekly wage							0.00521
							(0.0143)
Std dev orig-ind years schooling							0.00561
							(0.00998)
R-squared	0.819	0.830	0.878	0.882	0.883	0.887	0.887
Standard errors			cl cbsa × year	cl cbsa × year	cl cbsa × year	cl cbsa × year	cl cbsa × year
Origin CBSA x Year FE		Yes					
SCTG5 x NAICS6 x Year FE	Yes	Yes					
SCTG5 x NAICS6 x dest CBSA x Year FE			Yes	Yes	Yes	Yes	Yes
Observations (rounded)	1,400,000	1,400,000	1,400,000	1,400,000	1,000,000	1,000,000	1,000,000
Estab-year (rounded)	30,000	30,000	30,000	30,000	22,500	22,500	22,500
Ind-prod-year (rounded)	2,000	2,000	2,000	2,000	2,000	2,000	2,000
Note			flex miles	flex ctrl	flex miles	flex ctrl + school flex	flex ctrl + school flex

Robust standard errors in parentheses
** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CFS and CMF. All shipments are to a domestic destination CBSA distinct from the origin CBSA. All regressions include mode × year fixed effects. The third through seventh columns include 3-digit-NAICS-specific third-order polynomials in log mileage (3-7), log non-production worker share (4,6,7), log assets per worker (4,6,7), log pay per worker (4,6,7), log pay per production worker (4,6,7), log pay per non-production worker (4,6,7), and city-industry mean years of schooling (6,7).

A.5.3 Export shipments

This section examines export shipments. My empirical investigation was motivated in part by a growing international trade literature on quality specialization. I implemented my empirical strategy using plant-level data from US cities of varying income levels. The analysis above described shipments destined for US cities, which account for the vast majority of US manufactures output, to characterize how shipments' characteristics are related to the characteristics of their production locations. This section shows that the patterns found in domestic shipments are also found in export shipments destined for foreign markets.

Export shipments by US manufacturing plants exhibit price patterns consistent with those observed in domestic shipments. Table A.7 presents results for regressions analogous to those presented in Tables 1.3 and 1.4 using shipments sent to foreign destinations. The sample size is considerably smaller, since exports represent less than 8% of shipments by value and 4% by weight (Bureau of Transportation Statistics and US Census Bureau, 2010).[21] I calculate the mileage distance from origin CBSA to foreign destination using latitude and longitude coordinates.[22]

The estimated coefficients are consistent with those reported for shipments to domestic destinations. The origin-income elasticity of export prices is 42%. After controlling for plant-level factor usage, the origin-income elasticity is 30%. Upon introduction of the market-access measures, the origin-income elasticity becomes negative and statistically indistinguishable from zero. The dispersion of household income in the origin CBSA is positively related to export shipment prices, though this relationship is statistically insignificant, presumably due to the small sample size. Thus, the empirical relationships between export shipment prices, factor usage, and the demand measures are in line with those found for shipments to domestic destinations.

[21] This small sample size prevents me from estimating demand shifters using export shipments.

[22] For Canadian destinations, I use the coordinates of major Canadian cities. For other countries, I use the coordinates of the capital or main city. See data appendix A.2 for details.

Table A.7: Export shipments

Dep var: Log unit value, $\ln p_{skjodmt}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Log origin CBSA income per capita	0.434**	0.401**	0.334**	0.419**	0.369**	0.304**	0.305*	0.218	0.312**	-0.0724	-0.215
	(0.117)	(0.115)	(0.114)	(0.117)	(0.107)	(0.105)	(0.151)	(0.135)	(0.106)	(0.114)	(0.138)
Log origin CBSA population	0.00876	0.00335	0.00333	0.00920	-0.000623	-0.00218	0.00104	-0.00998	-0.00337	0.0288*	0.0112
	(0.0129)	(0.0125)	(0.0124)	(0.0127)	(0.0119)	(0.0115)	(0.0134)	(0.0129)	(0.0116)	(0.0118)	(0.0116)
Log mileage	0.0787**	0.0812**	0.0815**								
	(0.0205)	(0.0202)	(0.0198)								
Log non-production worker share		0.169**	0.141**								
		(0.0222)	(0.0271)								
Log assets per worker		-0.0451**	-0.0601**								
		(0.0128)	(0.0134)								
Log non-production worker share		-0.00972	-0.0296		-0.105	-0.0258					
× log per capita income		(0.0825)	(0.0969)		(0.0839)	(0.0978)					
Log assets per worker		-0.0432	-0.0787		-0.0811	-0.119*					
× log per capita income		(0.0418)	(0.0455)		(0.0503)	(0.0541)					
Log pay per worker			0.283**								
			(0.0961)								
Log pay per production worker			-0.0633								
			(0.0623)								
Log pay per non-production worker			0.00207								
			(0.0395)								
Log pay per worker			-0.0302			-0.485					
× log per capita income			(0.379)			(0.423)					
Log pay per production worker			0.430			0.590*					
× log per capita income			(0.238)			(0.279)					
Log pay per non-production worker			0.153			0.400*					
× log per capita income			(0.145)			(0.159)					
Log std dev orig household income							0.187	0.153			
							(0.146)	(0.137)			
Market access (excl origin) M^1_{ot}										2.477**	
										(0.327)	
Market access M^2_{ot}											1.961**
											(0.375)
R-squared	0.821	0.823	0.824	0.823	0.828	0.831	0.823	0.831	0.831	0.832	0.831
Note				miles flex	qnt flex	all flex	miles flex	all flex	all flex	all flex	all flex
Observations (rounded)						64,000					
Estab-year (rounded)						10,000					
Ind-prod-year (rounded)						2,000					

Standard errors, clustered by CBSA × year, in parentheses
** p<0.01, * p<0.05

NOTES: Manufacturing establishments in the CFS and CMF. All shipments are exports to a foreign destination. All regressions include SCTG5 × NAICS6 × year fixed effects and destination × year fixed effects. The fourth through eleventh columns include 3-digit-NAICS-specific third-order polynomials in log mileage (4-11), log non-production worker share (5, 6, 8-11), log assets per worker (5, 6, 8-11), log pay per worker (6, 8-11), log pay per production worker (6, 8-11), and log pay per non-production worker (6, 8-11).

Appendix B

Appendix for Chapter 2

B.1 Consumption interpretation

The production and consumption interpretations yield very similar results but differ slightly in notation. In the consumption interpretation, an individual's productivity and utility are

$$q(c, \tau, \sigma; \omega) = A(c)H(\omega, \sigma) \tag{B.1}$$

$$U(\omega, c, \tau, \sigma) = T(\tau)\left[A(c)H(\omega, \sigma)p(\sigma) - r(c, \tau)\right] \tag{B.2}$$

where $T(\tau)$ determines the value of the individual's disposable income after paying his or her locational price.[1] In this interpretation, preferences are non-homothetic in a manner akin to that of Gabszewicz, Shaked, Sutton, and Thisse (1981). Higher-income individuals are more willing to pay for higher-quality locations because a more desirable location complements their higher consumption of tradables.

In this case, instead of $\gamma = A(c)T(\tau) = A(c')T(\tau') \iff r(c, \tau) = r(c', \tau') = r_\Gamma(\gamma)$, the appropriate equivalence between two locations is their "amenity-amplified price", which is $T(\tau)r(c, \tau)$. So the equivalence statement is now $\gamma = A(c)T(\tau) = A(c')T(\tau') \iff T(\tau)r(c, \tau) = T(\tau')r(c', \tau') = r_\Gamma(\gamma)$. The results in lemma 2.1 are unaltered, though the proof is modified to use the relevant $U(\omega, c, \tau, \sigma)$. The expressions for $K : \Gamma \to \Omega$, $\bar{\gamma}$, and $\underline{\gamma}$ are unaltered. This

[1] Recall that the final good is the numeraire.

leaves the conclusions of lemmas 2.4, 2.5, and 2.6 intact. The locational price schedule is given by $r(c,\tau) = \frac{r_\Gamma(A(c)T(\tau))}{T(\tau)} = A(c)\frac{r_\Gamma(\gamma)}{\gamma}$.

These locational prices do not appear in the proof of Proposition 2.1 nor the endogenous definition of $A(c)$. When evaluated at equilibrium, occupied locations' productivities $q(c,\tau,\sigma;\omega) = A(c)H(\omega,\sigma)$ differ across cities in a Hicks-neutral fashion that satisfies Costinot's Definition 4 (see footnote 23), so Proposition 2.2 applies. As a result, the predictions about cities' population, sectors, and productivities described in sections 2.3.5 and 2.3.6 are unaltered by interpreting $T(\tau)$ as describing consumption benefits rather than production benefits.

B.2 Proofs

Proof of Lemma 2.1:

Proof. Suppose that $\exists \tau' < \bar{\tau}(c) : \; S(\tau') > L\int_0^{\tau'}\int_{\sigma\in\Sigma}\int_{\omega\in\Omega} f(\omega,c,x)d\omega d\sigma dx$. Then $\exists \tau \leq \tau'$: $S'(\tau) > L\int_{\sigma\in\Sigma}\int_{\omega\in\Omega} f(\omega,c,\tau)d\omega d\sigma$. Then $r(c,\tau)) = 0 \leq r(c,\bar{\tau}(c))$, so $U(\omega,c,\tau,\sigma) > U(\omega,c,\bar{\tau}(c),\sigma)\ \forall\omega\forall\sigma$ since $T(\tau)$ is strictly decreasing. This contradicts the definition of $\bar{\tau}(c)$, since $\bar{\tau}(c)$ is a location that maximizes utility for some individual. Therefore $S(\tau) = L\int_0^{\tau}\int_{\sigma\in\Sigma}\int_{\omega\in\Omega} f(\omega,c,x)d\omega d\sigma dx\ \forall\tau \leq \bar{\tau}(c)$.

Suppose that $\exists \tau',\tau'' : \tau' < \tau'' \leq \bar{\tau}(c)$ and $r(c,\tau') \leq r(c,\tau'')$. Then $U(\omega,c,\tau',\sigma) > U(\omega,c,\tau'',\sigma)\ \forall\omega\forall\sigma$ since $T(\tau)$ is strictly decreasing. This contradicts the result that τ'' maximizes utility for some individual. Therefore $r(c,\tau)$ is strictly decreasing in $\tau\ \forall\tau \leq \bar{\tau}(c)$.

Suppose $r(c,\bar{\tau}(c)) > 0$. Then by its definition as a populated location, $\exists\omega :\ A(c)T(\bar{\tau}(c))G(\omega) - r(c,\bar{\tau}(c)) \geq A(c)T(\bar{\tau}(c)+\epsilon)G(\omega)\ \forall\epsilon > 0$. This inequality is false for all ω for sufficiently small ϵ, by the continuity of $T(\tau)$. Therefore $r(c,\bar{\tau}(c)) = 0$. □

Proof of Lemma 2.2:

Proof. Much of our argument follows the proof of Lemma 1 in Costinot and Vogel (2010). Define $f(\omega,c,\tau) \equiv \int_{\sigma\in\Sigma} f(\omega,c,\tau,\sigma)d\sigma$. Define $\Omega(\tau) \equiv \{\omega \in \Omega | f(\omega,c,\tau) > 0\}$ and $\mathcal{T}(\omega) \equiv \{\tau \in [0,\bar{\tau}(c)] | f(\omega,c,\tau) > 0\}$.

1. $\mathcal{T}(\omega) \neq \emptyset$ by equation (2.11) and $f(\omega) > 0$. $\Omega(\tau) \neq \emptyset\ \forall\tau \leq \bar{\tau}(c)$ by lemma 2.1.

2. $\Omega(\tau)$ is a non-empty interval for $\tau \in [0, \bar{\tau}(c)]$. Suppose not, such that $\omega < \omega' < \omega''$ with $\omega, \omega'' \in \Omega(\tau)$ and $\omega' \notin \Omega(\tau)$. $\exists \tau' : \omega' \in \Omega(\tau')$. Suppose $\tau' > \tau$. By utility maximization

$$A(c)T(\tau')G(\omega') - r(c, \tau') \geq A(c)T(\tau)G(\omega') - r(c, \tau)$$
$$A(c)T(\tau)G(\omega) - r(c, \tau) \geq A(c)T(\tau')G(\omega) - r(c, \tau')$$

These jointly imply $(T(\tau') - T(\tau))(G(\omega') - G(\omega)) \geq 0$, contrary to $\tau' > \tau$ and $\omega' > \omega$. The $\tau' < \tau$ case is analogous, using ω' and ω''. Therefore $\Omega(\tau)$ is a non-empty interval. The same pair of inequalities proves that for $\tau < \tau' \leq \bar{\tau}(c)$, if $\omega \in \Omega(\tau)$ and $\omega' \in \Omega(\tau')$, then $\omega \geq \omega'$.

3. $\Omega(\tau)$ is singleton for all but a countable subset of $[0, \bar{\tau}(c)]$. Follow Costinot and Vogel (2010).

4. $\mathcal{T}(\omega)$ is singleton for all but a countable subset of Ω. Follow Costinot and Vogel (2010).

5. $\Omega(\tau)$ is singleton for $\tau \in [0, \bar{\tau}(c)]$. Suppose not, such that there exists $\tau \in [0, \bar{\tau}(c)]$ for which $\Omega(\tau)$ is not singleton. By step two, $\Omega(\tau)$ is an interval, so $\mu[\Omega(\tau)] > 0$, where μ is the Lebesgue measure over $\mathbb{R}$. By step four, we know that $\mathcal{T}(\omega) = \{\tau\}$ for μ-almost all $\omega \in \Omega(\tau)$. Hence condition (2.11) implies

$$f(\omega, c, \tau) = f(\omega)\delta^{Dirac}[1 - \mathbf{1}_{\Omega(\tau)}] \quad \text{for } \mu\text{-almost all } \omega \in \Omega(\tau), \tag{B.3}$$

where δ^{Dirac} is a Dirac delta function. Combining equations (2.9) and (B.3) with $\mu[\Omega(\tau)] > 0$ yields $S'(\tau) = +\infty$, which contradicts our assumptions about $S(\tau)$.

Step 5 means there is a function $N : \mathcal{T} \to \Omega$ such that $f(\omega, c, \tau) > 0 \iff N(\tau) = \omega$. Step 2 says N is weakly decreasing. Since $\Omega(\tau) \neq \emptyset \ \forall \tau \leq \bar{\tau}(c)$, N is continuous and satisfies $N(0) = \bar{\omega}$ and $N(\bar{\tau}(c)) = \underline{\omega}$. Step 4 means that N is strictly decreasing on $(0, \bar{\tau}(c))$. □

Proof of the explicit expression of $N(\tau)$ that follows Lemma 2.2:

$$
\begin{aligned}
S(\tau) &= L\int_0^\tau \int_{\sigma\in\Sigma}\int_{\omega\in\Omega} f(\omega,c,x,\sigma)d\omega d\sigma dx \\
&= L\int_0^\tau \int_{\omega\in\Omega} f(\omega)\delta^{Dirac}[x-N^{-1}(\omega)]d\omega dx \\
&= L\int_0^\tau \int_{\tau'} f(N(\tau'))\delta^{Dirac}[x-\tau']N'(\tau')d\tau' dx \\
&= -L\int_0^\tau f(N(x))N'(x)dx = L(1-F(N(\tau))) \\
\Rightarrow N(\tau) &= F^{-1}\left(\frac{L-S(\tau)}{L}\right)
\end{aligned}
$$

Proof of Lemma 2.3:

Proof. By utility maximization

$$
\begin{aligned}
A(c)T(\tau)G(N(\tau)) - r(c,\tau) &\geq A(c)T(\tau+d\tau)G(N(\tau)) - r(c,\tau+d\tau) \\
A(c)T(\tau+d\tau)G(N(\tau+d\tau)) - r(c,\tau+d\tau) &\geq A(c)T(\tau)G(N(\tau+d\tau)) - r(c,\tau)
\end{aligned}
$$

Together, these inequalities imply

$$
\begin{aligned}
\frac{A(c)T(\tau+d\tau)G(N(\tau)) - A(c)T(\tau)G(N(\tau))}{d\tau} &\leq \frac{r(c,\tau+d\tau)-r(c,\tau)}{d\tau} \\
&\leq \frac{A(c)T(\tau+d\tau)G(N(\tau+d\tau)) - A(c)T(\tau)G(N(\tau+d\tau))}{d\tau}
\end{aligned}
$$

Taking the limit as $d\tau \to 0$, we obtain $\frac{\partial r(c,\tau)}{\partial \tau} = A(c)T'(\tau)G(N(\tau))$. Integrating from τ to $\bar{\tau}(c)$ and using the boundary condition $r(c,\bar{\tau}(c)) = 0$ yields $r(c,\tau) = -A(c)\int_\tau^{\bar{\tau}(c)} T'(t)G(N(t))dt$. □

Proof of Lemma 2.4:

This proof is analogous to the proof of lemma 2.2.

Proof of Lemma 2.5:

Proof. By utility maximization

$$\begin{aligned}\gamma G(K(\gamma)) - r_\Gamma(\gamma) \geq& (\gamma + d\gamma) G(K(\gamma)) - r_\Gamma(\gamma + d\gamma)\\(\gamma + d\gamma) G(K(\gamma + d\gamma)) - r_\Gamma(\gamma + d\gamma) \geq& \gamma G(K(\gamma + d\gamma)) - r_\Gamma(\gamma)\end{aligned}$$

Together, these inequalities imply

$$\frac{(\gamma + d\gamma)G(K(\gamma + d\gamma)) - \gamma G(K(\gamma + d\gamma))}{d\gamma} \geq \frac{r_\Gamma(\gamma + d\gamma) - r_\Gamma(\gamma)}{d\gamma} \geq \frac{(\gamma + d\gamma)G(K(\gamma)) - \gamma G(K(\gamma))}{d\gamma}$$

Taking the limit as $d\gamma \to 0$, we obtain $\frac{\partial r_\Gamma(\gamma)}{\partial \gamma} = G(K(\gamma))$. Integrating from $\underline{\gamma}$ to γ and using the boundary condition $r_\Gamma(\underline{\gamma}) = 0$ yields $r_\Gamma(\gamma) = \int_{\underline{\gamma}}^{\gamma} G(K(x))dx$. □

Proof of Lemma 2.6:

Proof. In city c, the population of individuals with skills between ω and $\omega + d\omega$ is

$$L \int_{\omega}^{\omega + d\omega} f(x, c)dx = S\left(T^{-1}\left(\frac{K^{-1}(\omega)}{A(c)}\right)\right) - S\left(T^{-1}\left(\frac{K^{-1}(\omega + d\omega)}{A(c)}\right)\right)$$

Taking the derivative with respect to $d\omega$ and then taking the limit as $d\omega \to 0$ yields the population of ω in c. □

In the course of proving Proposition 2.1, we use the following lemma.

Lemma B.1. *Let $f(z) : \mathbb{R} \to \mathbb{R}^{++}$ and $g(x, y) : \mathbb{R}^2 \to \mathbb{R}^{++}$ be C^2 functions. If f is decreasing and log-concave and $g(x, y)$ is increasing in x, decreasing in y, and submodular, then $f(g(x, y))$ is log-supermodular in (x, y).*

Proof. $f(g(x, y))$ is log-supermodular in x and y if and only if

$$\frac{\partial^2 \ln f}{\partial x \partial y} = \frac{\partial \ln f}{\partial z} g_{xy} + \frac{\partial^2 \ln f}{\partial z^2} g_x g_y > 0$$

$f_z < 0$, $\frac{\partial^2 \ln f}{\partial z^2}$, $g_x > 0$, $g_y < 0$, and $g_{xy} < 0$ are sufficient for this inequality to be true. □

More concisely, an increasing, log-convex transformation of an increasing, supermodular function is log-supermodular.

Proof of Proposition 2.1:

Proof. The population of skill ω in city c can be written as

$$f(\omega, c) = -\frac{1}{L}\frac{\partial}{\partial \omega} S\left(T^{-1}\left(\frac{K^{-1}(\omega)}{A(c)}\right)\right) = \frac{{K^{-1}}'(\omega)}{A(c)L}\left[\left(-\frac{\partial}{\partial z} S\left(T^{-1}\left(z\right)\right)\right)\Big|_{z=\frac{K^{-1}(\omega)}{A(c)}}\right]$$

Since $\frac{{K^{-1}}'(\omega)}{A(c)L}$ is multiplicatively separable in ω and c, $f(\omega, c)$ is log-supermodular if and only if the term in brackets is log-supermodular. Since $K^{-1}(\omega)$ and $A(c)$ are increasing functions, if $-\frac{\partial}{\partial z} S(T^{-1}(z))$ is a decreasing and log-concave function, then $f(\omega, c)$ is log-supermodular by lemma B.1. □

B.3 Data

Data sources: Our metropolitan population data are from the US Census website (2000). Our data on individuals' demographics, educational attainments, geographic locations, and sectors of employment come from the 5 percent sample of the 2000 US Census and the 1 percent metro sample of the 1980 US Census made available by IPUMS-USA (Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek, 2010). Our data on industrial employment come from the 2000 County Business Patterns, available from the US Census Bureau website. Our data on occupational employment come from the 2000 Occupational Employment Statistics, available from the Burea of Labor Statistics website.

Geography: We use (consolidated) metropolitan statistical areas as defined by the OMB as our unit of analysis.

The smallest geographic unit in the IPUMS-USA microdata is the public-use microdata area (PUMA), which has a minimum of 100,000 residents. We map the PUMAs to metropolitan statistical areas (MSAs) using the MABLE Geocorr2K geographic correspondence engine from the Missouri Census Data Center. In some sparsely populated areas, a PUMA is larger than a metropolitan area. We drop six MSAs in which fewer than half of the residents of the only relevant PUMA live within the metropolitan area. As a result, there are 270 MSAs when we use Census of Population data.

The 1980 Census of Population IPUMS-USA microdata do not identify PUMAs, so we use the "metarea" variable describing 253 consolidated MSAs for the regressions in Table 2.5.

The County Business Patterns data describe 318 metropolitan statistical areas. These correspond to a mix of OMB-defined primary and consolidated metropolitan statistical areas outside New England and New England county metropolitan areas (NECMAs). We aggregate these into OMB-defined (consolidated) metropolitan statistical areas to obtain 276 MSAs.

The Occupational Employment Statistics data describe 331 (primary) metropolitan statistical areas. We aggregate these into OMB-defined (consolidated) metropolitan statistical areas to obtain observations for 276 MSAs.

Skill distribution: Our sample of individuals includes all full-time, full-year prime-age workers, defined as individuals 25 to 55 years of age who reported working at least 35 hour per week and 40 weeks in the previous year. Using the "educd" variable from IPUMS, we construct nine levels of educational attainment: less than high school, high school dropout, high school graduate, some college, associate's degree, bachelor's degree, master's degree, professional degree, and doctorate. There is at least one observation in every educational category in every metropolitan area.

Sectoral skill intensity: Using the same sample of full-time, full-year prime-age workers, we measure a sector's skill intensity by calculating the average years of schooling of its employees after controlling for spatial differences in average schooling. We calculate years of schooling using the educational attainment "educd" variable from IPUMS at its finest level of disaggregation. For instance, this means that we distinguish between those whose highest educational attainment is sixth grade or eighth grade. We use the "indnaics" and "occsoc" variables to assign individuals to their 3-digit NAICS and 2-digit SOC sectors of employment. Aggregating observations to the MSA-sector level, weighted by the IPUMS-provided person weights, we regress the average years of schooling on MSA and sectoral dummies. The sectoral dummy coefficients are our measure of skill intensities.

Industrial employment: There 96 3-digit NAICS industries, of which 21 are manufacturing industries. 75 of these industries, including all 21 manufacturing industries, appear in both the

Census of Population microdata and the County Business Patterns data. The County Business Patterns data are an almost exhaustive account of US employer establishments. When necessary to protect the confidentiality of individual establishments, employment in an industry in a location is reported as falling within an interval rather than its exact number. In our empirical work, we use the midpoints of these intervals as the level of employment. There are 390 (C)MSA-manufacturing-industry pairs, out of $5796 = 21 \times 276$, in which there are zero establishments. The County Business Patterns data omit self-employed individuals and employees of private households, railroads, agriculture production, the postal service, and public administrations. See the CBP methodology webpage for details.

Occupational employment: There are 22 2-digit SOC occupations. Across 331 (P)MSAs, there should be 7282 metropolitan-occupation observations. The 2000 BLS Occupational Employment Statistics contain employment estimates for 7129 metropolitan-occupation observations, none of which are zero. The 153 omitted observations "may be withheld from publication for a number of reasons, including failure to meet BLS quality standards or the need to protect the confidentiality of [BLS] survey respondents." 31 of these observations are for "farming occupations", and the vast majority of them are for less populous metropolitan areas. We assign these observations values of zero when they are included in our calculations.

B.4 Tables

Table B.1: *Pairwise comparisons of three skill groups*

Bins	Weights	Birthplace	College vs some college	College vs HS or less	Some college vs HS or less	Total comparisons	Average
2	Unweighted	All	1.00	1.00	1.00	3	1.00
2	Pop diff	All	1.00	1.00	1.00	3	1.00
2	Unweighted	US-born	1.00	1.00	1.00	3	1.00
2	Pop diff	US-born	1.00	1.00	1.00	3	1.00
3	Unweighted	All	1.00	1.00	.667	9	.889
3	Pop diff	All	1.00	1.00	.853	9	.951
3	Unweighted	US-born	1.00	1.00	1.00	9	1.00
3	Pop diff	US-born	1.00	1.00	1.00	9	1.00
5	Unweighted	All	1.00	1.00	.900	30	.967
5	Pop diff	All	1.00	1.00	.959	30	.986
5	Unweighted	US-born	1.00	1.00	1.00	30	1.00
5	Pop diff	US-born	1.00	1.00	1.00	30	1.00
10	Unweighted	All	.800	.844	.756	135	.8
10	Pop diff	All	.914	.935	.875	135	.908
10	Unweighted	US-born	.800	.867	.756	135	.808
10	Pop diff	US-born	.914	.951	.875	135	.913
30	Unweighted	All	.763	.717	.625	1,305	.702
30	Pop diff	All	.875	.845	.699	1,305	.806
30	Unweighted	US-born	.775	.759	.699	1,305	.744
30	Pop diff	US-born	.886	.878	.817	1,305	.86
90	Unweighted	All	.671	.666	.587	12,015	.641
90	Pop diff	All	.788	.774	.634	12,015	.732
90	Unweighted	US-born	.672	.686	.639	12,015	.666
90	Pop diff	US-born	.790	.800	.727	12,015	.772
270	Unweighted	All	.629	.616	.556	108,945	.6
270	Pop diff	All	.717	.695	.588	108,945	.667
270	Unweighted	US-born	.624	.635	.589	108,945	.616
270	Pop diff	US-born	.712	.726	.647	108,945	.695

Note: The number of cities per "bin" may differ by one, due to the integer constraint.

Table B.2: Pairwise comparisons of nine skill groups with one city per bin

Unweighted comparisons									
		LHS	HSD	HS	CD	AA	BA	MA	Pro
2	HSD	.423							
3	HS	.399	.413						
4	CD	.428	.486	.587					
5	AA	.43	.483	.571	.483				
6	BA	.476	.555	.644	.619	.602			
7	MA	.484	.558	.643	.614	.615	.528		
8	Pro	.484	.57	.645	.617	.604	.524	.499	
9	PhD	.49	.548	.598	.576	.577	.521	.501	.511
Population-difference weighted comparisons of US-born population									
		LHS	HSD	HS	CD	AA	BA	MA	Pro
2	HSD	.568							
3	HS	.488	.435						
4	CD	.583	.569	.649					
5	AA	.552	.53	.616	.453				
6	BA	.644	.65	.738	.695	.682			
7	MA	.648	.651	.738	.686	.695	.544		
8	Pro	.654	.654	.73	.676	.676	.533	.493	
9	PhD	.611	.605	.651	.605	.617	.502	.476	.497

Table B.3: Pairwise comparisons of nine skill groups

Bins	Weights	Total comparisons	All individuals success rate	US-born success rate
2	Unweighted	36	.556	.806
2	Population differences	36	.556	.806
2	Educational shares	36	.754	.883
2	Pop diff × edu shares	36	.754	.883
3	Unweighted	108	.5	.685
3	Population differences	108	.54	.744
3	Educational shares	108	.707	.836
3	Pop diff × edu shares	108	.736	.866
5	Unweighted	360	.536	.714
5	Population differences	360	.55	.767
5	Educational shares	360	.747	.844
5	Pop diff × edu shares	360	.756	.873
10	Unweighted	1620	.528	.617
10	Population differences	1620	.551	.696
10	Educational shares	1620	.67	.715
10	Pop diff × edu shares	1620	.725	.802
30	Unweighted	15,660	.521	.59
30	Population differences	15,660	.545	.656
30	Educational shares	15,660	.632	.666
30	Pop diff × edu shares	15,660	.699	.754
90	Unweighted	144,180	.527	.568
90	Population differences	144,180	.555	.626
90	Educational shares	144,180	.594	.614
90	Pop diff × edu shares	144,180	.657	.694
270	Unweighted	1,307,340	.536	.559
270	Population differences	1,307,340	.564	.605
270	Educational shares	1,307,340	.567	.579
270	Pop diff × edu shares	1,307,340	.612	.635

Note: The number of cities per “bin” may differ by one, due to the integer constraint.

Table B.4: Occupational employment population elasticities

β_{σ_1} Farming, Fishing, and Forestry Occupations	0.803	$\beta_{\sigma_{12}}$ Sales and Related Occupations	1.037
× log population	(0.048)	× log population	(0.010)
β_{σ_2} Building and Grounds Cleaning and Maintenance	1.039	$\beta_{\sigma_{13}}$ Management occupations	1.082
× log population	(0.011)	× log population	(0.015)
β_{σ_3} Food Preparation and Serving Occupations	0.985	$\beta_{\sigma_{14}}$ Arts, Design, Entertainment, Sports, and Media	1.158
× log population	(0.011)	× log population	(0.019)
β_{σ_4} Construction and Extraction Occupations	1.037	$\beta_{\sigma_{15}}$ Business and Financial Operations Occupations	1.204
× log population	(0.014)	× log population	(0.018)
β_{σ_5} Production Occupations	1.045	$\beta_{\sigma_{16}}$ Architecture and Engineering Occupations	1.209
× log population	(0.025)	× log population	(0.026)
β_{σ_6} Transportation and Material Moving Occupations	1.061	$\beta_{\sigma_{17}}$ Computer and Mathematical Occupations	1.395
× log population	(0.014)	× log population	(0.034)
β_{σ_7} Installation, Maintenance, and Repair Workers	1.015	$\beta_{\sigma_{18}}$ Healthcare Practitioners and Technical Occupations	1.001
× log population	(0.011)	× log population	(0.014)
β_{σ_8} Healthcare Support Occupations	0.980	$\beta_{\sigma_{19}}$ Community and Social Services Occupations	0.986
× log population	(0.013)	× log population	(0.020)
β_{σ_9} Personal Care and Service Occupations	1.065	$\beta_{\sigma_{20}}$ Education, Training, and Library Occupations	1.011
× log population	(0.017)	× log population	(0.017)
$\beta_{\sigma_{10}}$ Office and Administrative Support Occupations	1.081	$\beta_{\sigma_{21}}$ Life, Physical, and Social Science Occupations	1.170
× log population	(0.010)	× log population	(0.030)
$\beta_{\sigma_{11}}$ Protective Service Occupations	1.123	$\beta_{\sigma_{22}}$ Legal Occupations	1.200
× log population	(0.014)	× log population	(0.022)
Observations	5943	Observations	5943
R-squared	0.931	R-squared	0.931
Occupation FE	Yes	Occupation FE	Yes

Standard errors, clustered by MSA, in parentheses

Table B.5: Pairwise comparisons of occupations

Bins	Weights	Comparisons	Success rate
2	Unweighted	231	0.723
2	Popuation difference × skill difference	231	0.779
3	Unweighted	693	0.693
3	Popuation difference × skill difference	693	0.772
5	Unweighted	2,310	0.652
5	Popuation difference × skill difference	2,310	0.728
10	Unweighted	10,395	0.606
10	Popuation difference × skill difference	10,395	0.688
30	Unweighted	100,485	0.579
30	Popuation difference × skill difference	100,485	0.659
90	Unweighted	925,155	0.561
90	Popuation difference × skill difference	925,155	0.634
276	Unweighted	8,766,450	0.533
276	Popuation difference × skill difference	8,766,450	0.583

Note: The number of cities per “bin” may differ by one, due to the integer constraint.

Table B.6: Industrial employment population elasticities

β_{σ_1} Apparel Manufacturing	1.237	1.024	$\beta_{\sigma_{11}}$ Nonmetallic Mineral Product Manufacturing	1.018	0.955
× log population	(0.070)	(0.148)	× log population	(0.036)	(0.042)
β_{σ_2} Textile Product Mills	1.125	0.905	$\beta_{\sigma_{12}}$ Paper Manufacturing	0.901	0.539
× log population	(0.056)	(0.135)	× log population	(0.063)	(0.104)
β_{σ_3} Leather and Allied Product Manufacturing	0.743	0.147	$\beta_{\sigma_{13}}$ Printing and Related Support Activities	1.202	1.122
× log population	(0.099)	(0.284)	× log population	(0.036)	(0.047)
β_{σ_4} Furniture and Related Product Manufacturing	1.120	1.000	$\beta_{\sigma_{14}}$ Electrical Equipment, Appliance & Component	1.159	0.813
× log population	(0.050)	(0.076)	× log population	(0.074)	(0.111)
β_{σ_5} Textile Mills	0.823	0.352	$\beta_{\sigma_{15}}$ Machinery Manufacturing	1.071	0.960
× log population	(0.105)	(0.208)	× log population	(0.055)	(0.069)
β_{σ_6} Wood Product Manufacturing	0.848	0.608	$\beta_{\sigma_{16}}$ Miscellaneous Manufacturing	1.224	1.208
× log population	(0.055)	(0.085)	× log population	(0.044)	(0.059)
β_{σ_7} Fabricated Metal Product Manufacturing	1.094	1.036	$\beta_{\sigma_{17}}$ Beverage and Tobacco Product Manufacturing	1.168	1.010
× log population	(0.048)	(0.050)	× log population	(0.065)	(0.147)
β_{σ_8} Food Manufacturing	0.953	0.864	$\beta_{\sigma_{18}}$ Transportation Equipment Manufacturing	1.254	0.940
× log population	(0.050)	(0.067)	× log population	(0.075)	(0.101)
β_{σ_9} Plastics and Rubber Products Manufacturing	1.105	0.975	$\beta_{\sigma_{19}}$ Petroleum and Coal Products Manufacturing	0.951	0.393
× log population	(0.056)	(0.070)	× log population	(0.074)	(0.308)
$\beta_{\sigma_{10}}$ Primary Metal Manufacturing	0.997	0.449	$\beta_{\sigma_{20}}$ Computer and Electronic Product Manufacturing	1.453	1.254
× log population	(0.078)	(0.107)	× log population	(0.075)	(0.108)
			$\beta_{\sigma_{21}}$ Chemical Manufacturing	1.325	0.992
			× log population	(0.065)	(0.098)
Observations	5406	2130	Observations	5406	2130
R-squared	0.564	0.541	R-squared	0.564	0.541
Industry FE	Yes	Yes	Industry FE	Yes	Yes
Only uncensored obs		Yes	Only uncensored obs		Yes

Standard errors, clustered by MSA, in parentheses

Table B.7: Pairwise comparisons of manufacturing industries

Bins	Weights	Comparisons	Success rate
2	Unweighted	210	.638
2	Popuation difference × skill difference	210	.675
3	Unweighted	630	.640
3	Popuation difference × skill difference	630	.665
5	Unweighted	2,100	.606
5	Popuation difference × skill difference	2,100	.633
10	Unweighted	9,450	.561
10	Popuation difference × skill difference	9,450	.597
30	Unweighted	91,350	.543
30	Popuation difference × skill difference	91,350	.573
90	Unweighted	841,050	.522
90	Popuation difference × skill difference	841,050	.549
276	Unweighted	7,969,500	.495
276	Popuation difference × skill difference	7,969,500	.531

Note: The number of cities per "bin" may differ by one, due to the integer constraint.