NBER WORKING PAPER SERIES

GOODS TRADE, FACTOR MOBILITY AND WELFARE

Stephen J. Redding

Working Paper 18008
http://www.nber.org/papers/w18008

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2012

This paper is a radically changed version of “Goods Trade, Factor Mobility andWelfare,” NBER Working Paper, 18008, 2012. I am grateful to Chang Sun and Tomasz Swiecki for research assistance and Princeton University for research support. I would also like to thank Pablo Fajgelbaum, Gene Grossman, David Nagy, Esteban Rossi-Hansberg, and conference and seminar participants for helpful comments and suggestions. The usual disclaimer applies. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Goods Trade, Factor Mobility and Welfare
Stephen J. Redding
NBER Working Paper No. 18008
April 2012, Revised December 2014
JEL No. F11,F12,F16

ABSTRACT

We develop a quantitative spatial model that incorporates a rich geography of trade and imperfect labor mobility between locations. We provide general results for the existence, uniqueness and comparative statics of the equilibrium. We show how the model can be used to undertake counterfactuals using only data in an initial equilibrium. In these counterfactuals, the welfare gains from trade depend on changes in both domestic trade shares and reallocations of population across locations. We show that factor mobility introduces quantitatively relevant differences in the counterfactual predictions of constant and increasing returns to scale models.

Stephen J. Redding
Department of Economics
and Woodrow Wilson School
Princeton University
Fisher Hall
Princeton, NJ 08544
and NBER
reddings@princeton.edu

1 Introduction

The determinants of the spatial distribution of economic activity is one of the most central issues in economics. Although there is a large literature concerned with this issue, existing theoretical research typically considers stylized settings with a small number of *ex ante* identical locations. Furthermore, existing theoretical research usually makes one of several extreme assumptions about labor mobility: either perfect immobility, perfect mobility or a mechanical relationship between migration flows and relative wages.[1] However, most empirically-observed locations differ substantially from one another in terms of their locational characteristics (e.g. interior versus coast), and existing empirical estimates suggest that labor mobility lies in between the polar extremes of perfect mobility and immobility.[2]

In contrast, we develop a quantitative spatial model that incorporates a large number of potentially asymmetric locations. We allow these locations to differ from one another in terms of their productivity, amenities and transport infrastructure. We incorporate both goods and labor market frictions. Locations can trade with one another subject to bilateral trade costs. Workers are mobile between locations, but have heterogeneous tastes for each location, which provides micro foundations for imperfect labor mobility. Each location faces an upward-sloping supply curve for labor, such that higher real incomes have to paid in order to attract workers with lower idiosyncratic tastes for that location.

Despite the large number of asymmetric locations and the presence of goods and labor market frictions, the model remains highly tractable and amenable to both analytical and quantitative analysis. We provide conditions on the model's parameters for which there exists a unique equilibrium distribution of economic activity. We also provide unambiguous comparative statics for the effect of each location characteristic on economic activity in that location and all other locations. We show that there is one-to-one mapping from the model's parameters and data on wages, population, land area and trade costs to the unobserved characteristics of locations (productivity and amenities). Therefore the model can be inverted to recover exogenous unobserved characteristics from the endogenous variables of the model.

We provide an approach to undertaking model-based counterfactuals for the effects of changes in productivity, amenities and trade costs that does not require observing or making assumptions about the unobserved characteristics of locations. Instead this approach uses only wages, population and trade shares in an initial equilibrium. In contrast to international trade models, in which population is typically exogenous, these counterfactuals yield predictions for the reallocation of population across locations. We show that these population reallocations are consequential for the measurement of the welfare gains from trade for each location. To the extent that some locations experience larger reductions in trade costs than others, population reallocates to these locations and away from other locations, until the price of the immobile factor of production land adjusts such that all locations experience the same welfare gains from

[1] For example, Krugman (1991b) assumes perfectly immobile agricultural workers and perfectly mobile manufacturing workers; Helpman (1998) assumes perfectly mobile workers; Krugman and Venables (1995) assume perfectly immobile workers; Puga (1999) considers both perfectly mobile and perfectly immobile workers; and Fujita, Krugman, and Venables (1999) consider a mechanical relationship between migration and relative wages.

[2] See, for example, Blanchard and Katz (1992) and Bound and Holzer (2000).

the reduction in trade costs. Nonetheless, these population redistributions are not sufficient to equalize real income, because expanding locations have to offer higher real incomes to attract workers with lower idiosyncratic tastes.

To illustrate the role of factor mobility in shaping the impact of reductions in trade costs, we assume central values for the model's parameters from the existing empirical literature. We generate data for a hypothetical economy within the model and undertake counterfactuals for the impact of a transport infrastructure improvement. A large reduced-form empirical literature has estimated the impact of road/railroad construction by comparing locations that are directly treated with the transport infrastructure to locations that are not directly treated. As acknowledged by this literature, such reduced-form regressions cannot capture general equilibrium effects, and face the challenge of distinguishing reallocation from the creation of economic activity. We show that they also mask considerable heterogeneity in treatment effects. Among treated locations, the economic impact of the transport infrastructure depends on the characteristics of the locations that are connected and their centrality within the transport network. Among the untreated locations that are not directly affected by the transport infrastructure, many are indirectly affected because the transport infrastructure reduces transport costs along the least cost route to other locations.

We show that the average treatment effect of the transport infrastructure depends in a quantitatively relevant way on the degree of both goods and labor market frictions. In general, lower levels of factor mobility imply larger average treatment effects for wages, but smaller average treatment effects for population and land prices. For example, as we vary the Fréchet shape determining the degree of heterogeneity in worker tastes (and hence the degree of labor mobility) from 3 to 5, we find that the average treatment effects for population and land rents can vary from around 50 to 70 percent. Across this range of values for labor mobility, we find that the reallocation effects of the transport improvement are large relative to its effect on welfare. Given data on population and wages before and after the transport improvement, and assuming constant unobserved characteristics of locations, we show how the structure of the model can be used to estimate the parameters that determine the size of goods and labor market frictions.

While we first develop these results in a model with constant returns to scale, we later extend the analysis to incorporate agglomeration forces from consumer love of variety, increasing returns to scale and transport costs. A key implication of the introduction of these agglomeration forces is that the measure of goods produced by a location is endogenous to its population. Nevertheless, both the constant and increasing returns to scale models have a one-to-one mapping from location characteristics (productivity, amenities, land supplies and trade costs) to populations and wages. Therefore, assuming the same elasticity of trade with respect to trade costs and the same elasticity of population with respect to real income, both models can calibrated to the same initial equilibrium populations and wages through the appropriate choice of the unobserved productivities and amenities for each location.

In an international trade context, where labor is perfectly immobile across countries, the two models have the same counterfactual predictions for the impact of reductions in trade costs when calibrated in this way. In contrast, when labor is imperfectly mobile across locations, the two models necessarily have different counterfactual predictions even when calibrated in this way. As trade costs fall, population reallocates across locations, which leads to endogenous changes in the measure of goods produced by each location in the increasing returns model. These endogenous changes in the measure of goods produced in turn affect trade shares, and hence lead to different counterfactual predictions for wages, trade shares and populations from the constant returns model. We show that these differences in predictions for the welfare gains from trade can be quantitatively relevant for plausible reductions in trade costs (around 10 percent of the overall welfare gains).

Finally, we explore another form of imperfect labor mobility based on the distinction between regions and countries. We assume that labor is imperfectly mobile across regions within countries but perfectly immobile between countries. We show that the general equilibrium of the model can be characterized using a directly analogous approach to before. Counterfactuals again can be undertaken using only the values of endogenous variables in an initial equilibrium. Imperfect labor mobility within countries implies that expected utility is the same across all regions within a given country. This common level of expected utility is equal to that of a hypothetical region with a geometric mean of the regional domestic trade shares and regional populations for that country. At the regional level, measuring each region's welfare gains from trade using its domestic trade share without controlling for its change in population can lead to substantial discrepancies from the true welfare gains from trade (ranging up to 30 percent for plausible reductions in trade costs). At the national level, measuring the common change in expected utility using the domestic trade share for the country as a whole provides a much better approximation to the true welfare gains from trade. The success of this approximation depends on the extent to which the change in the country's aggregate domestic trade share approximates a weighted average of the change in the geometric means of regional domestic trade shares and populations.

Our paper is related to the literature on economic geography including Krugman (1991a,b), Helpman (1998), Hanson (2005), Redding and Sturm (2008), Ramondo, Rodríguez-Clare, and Saborio (2012), Coşar and Fajgelbaum (2013), and Caliendo, Parro, Rossi-Hansberg, and Sarte (2014).[3] Within this line of research, Allen and Arkolakis (2014) develop an Armington model with perfect labor mobility and trade costs, and provide general conditions for the existence, uniqueness and stability of equilibrium. In this Armington setting, the differentiation of goods by location of origin provides a dispersion force. In contrast, in our framework, the immobility of land provides the dispersion force. This difference proves to be consequential for the measurement of the welfare gains from trade, which depend on endogenous reallocations of the mobile factor in our setting (through the demand for land), whereas they do not in an Armington framework, without augmenting that framework with externalities.

This economic geography literature typically assumes either perfect labor immobility, perfect labor mobility or a mechanical migration process. In contrast, we develop a model of imperfect labor mobility based on heterogenous worker tastes for each location, where perfect labor mobility corresponds to the

[3] See also Davis and Weinstein (2002), Desmet and Rossi-Hansberg (2014), Fujita, Krugman, and Venables (1999), Hanson (1996, 1997), Head and Ries (2001), Redding and Venables (2004), and Rossi-Hansberg (2005).

limiting case in which this distribution is degenerate. This approach to modeling imperfect labor mobility follows a line of research dating back to McFadden (1974), including Artuc, Chaudhuri, and McLaren (2010), Kennan and Walker (2011), Grogger and Hanson (2011), Moretti (2011) and Busso, Gregory, and Kline (2013).[4] We incorporate such imperfect labor mobility into a general equilibrium trade model with a rich geography of trade costs. We build on the tools introduced by Allen and Arkolakis (2014) to provide general results for the existence, uniqueness and comparative statics of the equilibrium in this setting with both imperfect labor mobility and goods market frictions.

Our analysis is also related to the recent quantitative trade literature, including Eaton and Kortum (2002), Alvarez and Lucas (2007), Arkolakis, Costinot, and Rodriguez-Clare (2012), Caliendo and Parro (2012), Costinot, Donaldson, and Komunjer (2012), Eaton, Kortum, Neiman, and Romalis (2011), Fieler (2011), Hsieh and Ossa (2011) and Ossa (2011).[5] As this literature is concerned with international trade, it makes the standard assumption that labor is perfectly immobile between countries. In contrast, our analysis is specifically concerned with the determinants of the spatial distribution of economic activity within countries, where labor is likely to be imperfectly mobile across locations.

Finally, our work relates to the empirical literature has examined the relationship between economic activity and transport infrastructure, including Donaldson (2014), Baum-Snow (2007), Duranton and Turner (2012), Faber (2014) and Michaels (2008). The main focus of this line of research has been the use of quasi-experimental variation in transport infrastructure to estimate the average impact on treated locations relative to untreated locations. In contrast, we use a structural model of economic geography to highlight general equilibrium effects, heterogeneous treatment effects among treated and untreated locations, and the role of labor mobility in shaping the impact of transport infrastructure improvements.

The remainder of the paper is structured as follows. Section 2 introduces the baseline version of our quantitative spatial model. Section 3 calibrates the model's parameters to central estimates from the existing empirical literature and examines how factor mobility shapes the impact of transport infrastructure improvements. Section 4 introduces agglomeration forces as a result of the combination of consumer love of variety, increasing returns to scale and transport costs. Section 5 explores the implications of different degrees of factor mobility within and between countries. Section 6 concludes.

2 Theoretical Framework

We consider an economy consisting of many (potentially asymmetric) locations indexed by $i, n \in N$. Locations can differ from one another in terms of land supply, productivity, amenities and their geographical location relative to one another. We allow for imperfect geographic mobility of both goods and factors of production. Bilateral trade costs for goods are assumed to take the iceberg form, such that d_{ni} units of a

[4] Desmet, Nagy, and Rossi-Hansberg (2014) develop an alternative approach that uses data on actual and desired migration flows and measures of subjective well-being to quantify migration restrictions between countries.

[5] A longer tradition in international trade that has examined the extent to which goods and factor movements are complements or substitutes (as in Markusen 1983, Mundell 1957 and Jones 1967) and the contribution of lumpiness in the distribution of relative factor endowments across regions in influencing country trade (as in Courant and Deardorff 1992, 1993).

good must be shipped from one location for one unit to arrive in another location, where $d_{ni} > 1$ for $n \neq i$ and $d_{nn} = 1$. Land and labor are the two factors of production. Land is perfectly immobile. In contrast, labor is imperfectly mobile across locations, because of idiosyncratic shocks to worker preferences for each location.[6]

2.1 Consumer Preferences

Preferences for worker ω residing in location n depend on goods consumption (C_n), residential land use (H_{Un}) and an idiosyncratic amenity shock to the utility from residing in that location:[7]

$$U_n(\omega) = b_n(\omega) \left(\frac{C_n(\omega)}{\alpha} \right)^{\alpha} \left(\frac{H_{Un}(\omega)}{1-\alpha} \right)^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{1}$$

The goods consumption index (C_n) is defined over consumption of a fixed continuum of goods $j \in [0,1]$:

$$C_n = \left[\int_0^1 c_{nj}^{\rho} dj \right]^{\frac{1}{\rho}}, \tag{2}$$

where the CES parameter (ρ) determines the elasticity of substitution between goods ($\sigma = 1/(1-\rho)$). The corresponding dual price index for goods consumption (P_n) is:

$$P_n = \left[\int_0^1 p_{nj}^{1-\sigma} dj \right]^{\frac{1}{1-\sigma}}, \qquad \sigma = \frac{1}{1-\rho}. \tag{3}$$

The idiosyncratic amenity shocks ($b_n(\omega)$) capture the idea that workers have heterogeneous preferences for living in each location (e.g. different preferences for climate, proximity to the coast etc). We assume that these amenity shocks are drawn independently across locations and workers from a Fréchet distribution:

$$G_n(b) = e^{-B_n b^{-\epsilon}}, \tag{4}$$

where the scale parameter B_n determines average amenities for location n and the shape parameter ϵ controls the dispersion of amenities across workers for each location. Each worker is endowed with one unit of labor that is supplied inelasticity with zero disutility.

2.2 Production

Each location draws an idiosyncratic productivity $z(j)$ for each good j. Productivity is independently drawn across goods and locations from a Fréchet distribution:

$$F_i(z) = e^{-A_i z^{-\theta}}, \tag{5}$$

where the scale parameter A_i determines average productivity for location i and the shape parameter θ controls the dispersion of productivity across goods.

[6] While we interpret the idiosyncratic shocks in terms of worker preferences for each region, there is an isomorphic interpretation of this specification in terms of shocks to worker productivity for each region.

[7] For empirical evidence using U.S. data in support of the constant housing expenditure share implied by the Cobb-Douglas functional form, see Davis and Ortalo-Magné (2011).

Goods are homogeneous in the sense that one unit of a given good is the same as any other unit of that good. Each good is produced with labor under conditions of perfect competition according to a linear technology.[8] The cost to a consumer in location n of purchasing one unit of good j from location i is therefore:

$$p_{ni}(j) = \frac{d_{ni} w_i}{z_i(j)}, \tag{6}$$

where w_i denotes the wage in location i.

2.3 Expenditure Shares and Price Indices

The representative consumer in location n sources each good from the lowest-cost supplier to that location. Using equilibrium prices (6) and the properties of the Fréchet distribution following Eaton and Kortum (2002), the share of expenditure of location n on goods produced by location i is:

$$\pi_{ni} = \frac{A_i \left(d_{ni} w_i\right)^{-\theta}}{\sum_{s \in N} A_s \left(d_{ns} w_s\right)^{-\theta}}, \tag{7}$$

where the elasticity of trade with respect to trade costs is determined by the Fréchet shape parameter for productivity θ. Using the domestic trade share (π_{nn}) and noting that $d_{nn} = 1$, the consumption goods price index can be written solely in terms of this domestic trade share, wages and parameters:

$$P_n^{-\theta} = \gamma \left[\sum_{i \in N} A_i \left(d_{ni} w_i\right)^{-\theta} \right] = \frac{\gamma^{-\theta} A_n w_n^{-\theta}}{\pi_{nn}}, \tag{8}$$

where $\gamma = \left[\Gamma\left(\frac{\theta - (\sigma - 1)}{\theta} \right) \right]^{\frac{1}{1-\sigma}}$ and $\Gamma(\cdot)$ denotes the Gamma function. To ensure a finite value for the price index, we require $\theta > \sigma - 1$.

2.4 Residential Choices and Income

Given the specification of consumer preferences (1), the corresponding indirect utility function is:

$$U_n(\omega) = \frac{b_n(\omega) v_n(\omega)}{P_n^{\alpha} r_n^{1-\alpha}}, \tag{9}$$

where r_n is the land rent for location n; $v_n(\omega)$ is the income of worker ω in location n; with perfectly competitive labor markets, $v_n(\omega)$ is the same across workers ω within a given region n. Since indirect utility is a monotonic function of the amenity draw, it too has a Fréchet distribution:

$$G_n(U) = e^{-\psi_n U^{-\epsilon}}, \qquad \psi_n = B_n \left(v_n / P_n^{\alpha} r_n^{1-\alpha} \right)^{\epsilon}.$$

Each worker chooses the location that offers her the highest utility after taking into account her idiosyncratic preferences. Using the above distribution of indirect utility, the probability that a worker chooses to live in location $n \in N$ is:

$$\frac{L_n}{\bar{L}} = \frac{B_n \left(v_n / P_n^{\alpha} r_n^{1-\alpha} \right)^{\epsilon}}{\sum_{k \in N} B_k \left(v_k / P_k^{\alpha} r_k^{1-\alpha} \right)^{\epsilon}}, \tag{10}$$

[8] Although, to simplify the exposition, we assume that land is only used residentially, it is straightforward to also allow land to be used commercially, as shown in the web appendix.

where the elasticity of population with respect to real income is determined by the Fréchet shape parameter for consumer tastes ϵ. For finite values of ϵ, each location faces a population supply curve that is upward sloping in real income. This upward-sloping supply curve implies that real income in general differs across locations, because higher real incomes have to be paid to attract workers with lower idiosyncratic tastes for a location. Expected utility for a worker across locations is:

$$\bar{U} = \delta \left[\sum_{k \in N} B_k \left(v_k / P_k^{\alpha} r_k^{1-\alpha} \right)^{\epsilon} \right]^{\frac{1}{\epsilon}}, \tag{11}$$

where $\delta = \Gamma\left((\epsilon - 1)/\epsilon\right)$ and $\Gamma(\cdot)$ denotes the Gamma function. To ensure a finite value for expected utility, we require $\epsilon > 1$.

An implication of the Fréchet distribution for utility is that expected utility conditional on living in location n is the same across all locations n and equal to expected utility for the economy as a whole. On the one hand, more attractive location characteristics directly raise the utility of a worker with a given idiosyncratic utility draw, which increases expected utility. On the other hand, more attractive location characteristics attract workers with lower idiosyncratic utility draws, which reduces expected utility. With a Fréchet distribution of utility, these two effects exactly offset one another. Therefore, although real income in general differs across locations, expected utility (taking into account idiosyncratic shocks) is the same across locations. Hence this common value for expected utility captures the welfare gains from trade for all locations.

Expenditure on land in each location is redistributed lump sum to the workers residing in that location, as in Helpman (1998). Therefore total income in each location (v_n) equals labor income plus expenditure on residential land:

$$v_n L_n = w_n L_n + (1-\alpha) v_n L_n = \frac{w_n L_n}{\alpha}. \tag{12}$$

Labor income in each location equals expenditure on goods produced in that location:

$$w_i L_i = \sum_{n \in N} \pi_{ni} w_n L_n. \tag{13}$$

Land market clearing implies that the equilibrium land rent can be determined from the equality of land income and expenditure:

$$r_n = \frac{(1-\alpha) v_n L_n}{H_n} = \frac{1-\alpha}{\alpha} \frac{w_n L_n}{H_n}. \tag{14}$$

2.5 General Equilibrium

The general equilibrium of the model can be represented by the measure of workers (L_n) in each location $n \in N$, the share of each location's expenditure on goods produced in other locations (π_{ni}) and the wage in each location (w_n). Using labor income (13), the trade share (7), the price index (8), residential choice probabilities (10) and land market clearing (14), this equilibrium triple $\{L_n, \pi_{ni}, w_n\}$ solves the following system of equations for all $i, n \in N$. First, each location's income must equal expenditure on the goods produced in that location:

$$w_i L_i = \sum_{n \in N} \pi_{ni} w_n L_n. \tag{15}$$

Second, location expenditure shares are:

$$\pi_{ni} = \frac{A_i \left(d_{ni} w_i\right)^{-\theta}}{\sum_{k \in N} A_k \left(d_{nk} w_k\right)^{-\theta}}. \tag{16}$$

Third, residential choice probabilities imply:

$$\frac{L_n}{\bar{L}} = \frac{B_n \left(\frac{A_n}{\pi_{nn}}\right)^{\frac{\alpha\epsilon}{\theta}} \left(\frac{L_n}{H_n}\right)^{-\epsilon(1-\alpha)}}{\sum_{k \in N} B_k \left(\frac{A_k}{\pi_{kk}}\right)^{\frac{\alpha\epsilon}{\theta}} \left(\frac{L_k}{H_k}\right)^{-\epsilon(1-\alpha)}}. \tag{17}$$

2.6 Existence and Uniqueness

We now show that there exists a unique general equilibrium that solves the system of equations (15)-(17). Using the the requirement that labor income for each location equals expenditure on goods produced in that location (13), we obtain one system of equations for the wages and populations of locations $n \in N$ as a function of parameters and a transformation of expected utility ($\bar{W}$):

$$\bar{W}^{-\theta} = \frac{w_n^{1+\theta} L_n / A_n}{\gamma^{-\theta} \left[\sum_{k \in N} d_{kn}^{-\theta} B_k^{\frac{\theta}{\alpha\epsilon}} H_k^{\theta\left(\frac{1-\alpha}{\alpha}\right)} w_k^{1+\theta} L_k^{1-\theta\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}\right]}, \tag{18}$$

where $\bar{W} = \left[\alpha^{\epsilon} \left(\frac{1-\alpha}{\alpha}\right)^{\epsilon(1-\alpha)} \left(\bar{U}/\delta\right)^{\epsilon} \left(\bar{L}\right)^{-1}\right]^{1/\alpha\epsilon}$, as shown in the web appendix.

Using the price index (8), we obtain a second system of equations for the wages and populations of locations $n \in N$ as a function of parameters and the transformation of expected utility ($\bar{W}$):

$$\bar{W}^{-\theta} = \frac{w_n^{-\theta} B_n^{-\frac{\theta}{\alpha\epsilon}} H_n^{-\theta\left(\frac{1-\alpha}{\alpha}\right)} L_n^{\theta\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}{\gamma^{-\theta} \left[\sum_{k \in N} A_k \left(d_{nk} w_k\right)^{-\theta}\right]}, \tag{19}$$

as also shown in the web appendix.

Under the assumption that trade costs are symmetric ($d_{nk} = d_{kn}$), these two wage systems imply the following closed-form solution linking the endogenous variables for each location $n \in N$:

$$w_n^{1+2\theta} A_n^{-1} B_n^{\frac{\theta}{\alpha\epsilon}} H_n^{\theta\left(\frac{1-\alpha}{\alpha}\right)} L_n^{1-\theta\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)} = \kappa, \tag{20}$$

where κ is a scalar. If equation (20) holds, then any functions w_n and L_n that satisfy the system of equations (18) will also satisfy the system of equations (19) (and vice versa). In the proposition below, we prove that equation (20) is the unique relationship between w_n and L_n that satisfies both systems. Substituting this relationship (20) into (19), we obtain the following system of equations that uniquely determines the equilibrium population of each location as a function of the parameters of the model:

$$L_n^{\tilde{\theta}\gamma_1} A_n^{-\tilde{\theta}} B_n^{-\frac{\tilde{\theta}(1+\theta)}{\alpha\epsilon}} H_n^{-\frac{\tilde{\theta}(1+\theta)(1-\alpha)}{\alpha}} = \bar{W}^{-\theta}\gamma^{-\theta}\left[\sum_{k\in N} d_{nk}^{-\theta} A_k^{\frac{\tilde{\theta}(1+\theta)}{\theta}} B_k^{\frac{\tilde{\theta}\theta}{\alpha\epsilon}} H_k^{\frac{\tilde{\theta}\theta(1-\alpha)}{\alpha}} \left(L_k^{\tilde{\theta}\gamma_1}\right)^{\frac{\gamma_2}{\gamma_1}}\right], \tag{21}$$

$$\tilde{\theta} = \frac{\theta}{1+2\theta},$$

$$\gamma_1 = 1 + (1+\theta)\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha}\right),$$

$$\gamma_2 = 1 - \theta\left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha}\right) < \gamma_1,$$

where equilibrium expected utility ($\bar{W}$) is implicitly determined by the requirement that the labor market clears across all locations: $\sum_{n\in N} L_n = \bar{L}$.

Proposition 1 *Given the land area, productivity and amenity parameters {H_n, A_n, B_n} and bilateral trade frictions {d_{ni}} for all locations $n, i \in N$, there exist unique equilibrium populations (L_n^*), wages (w_n^*) and trade shares (π_{ni}^*).*

Proof. See the web appendix. ■

Having determined unique equilibrium populations (L_n) from the system of equations (21), we can solve for equilibrium wages (w_n) as a function of populations from the closed-form solutions (20), and we can solve for equilibrium trade shares (π_{ni}) as a function of wages from the expenditure shares (7).

Intuitively, as population concentrates in a location this bids up land prices, so that the inelastic supply of land ensures the existence of a unique equilibrium distribution of population across locations.

2.7 Comparative Statics

Although we allow for both goods and labor market frictions, and consider a large number of locations that can differ from one another in productivity, amenities, land supplies and bilateral trade costs, the model admits closed-form expressions for the comparative statics of the endogenous variables with respect to the relative value of these location characteristics. To characterize these comparative statics, we re-write the system of equations for equilibrium populations (21) as the following implicit function:

$$\begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_N \end{pmatrix} = \begin{pmatrix} \Omega_1^I \\ \vdots \\ \Omega_N^I \end{pmatrix} - \begin{pmatrix} \Omega_1^{II} \\ \vdots \\ \Omega_N^{II} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{22}$$

$$\Omega_n^I = L_n^{\tilde{\theta}\gamma_1} A_n^{-\theta} B_n^{-\frac{\tilde{\theta}(1+\theta)}{\alpha\epsilon}} H_n^{-\frac{\tilde{\theta}(1+\theta)(1-\alpha)}{\alpha}},$$

$$\Omega_n^{II} = \sum\nolimits_{k\in N} \Omega_{nk}^{II},$$

$$\Omega_{nk}^{II} = \bar{W}^{-\theta}\gamma^{-\theta} d_{nk}^{-\theta} A_k^{\frac{\tilde{\theta}(1+\theta)}{\theta}} B_k^{\frac{\tilde{\theta}\theta}{\alpha\epsilon}} H_k^{\frac{\tilde{\theta}\theta(1-\alpha)}{\alpha}} \left(L_k^{\tilde{\theta}\gamma_1}\right)^{\frac{\gamma_2}{\gamma_1}},$$

where Ω_n^{II} has an interpretation as a *market access* term that captures the goods market access of each location (depending on trade costs d_{nk}) to the characteristics of other locations.

The implicit function (Ω_n) is monotonically decreasing in the productivities, amenities and land supplies of all locations and monotonically increasing in the trade costs between all pairs of locations, as shown in the web appendix, where we report the closed-form solutions for each of these derivatives. An implication is that the equilibrium population of each location (L_n) depends solely on the *relative* rather than the absolute levels of these characteristics. The reason is that proportional changes in the absolute level of these characteristics across all locations simply lead to a change in the common level of expected utility $(\bar{W})$, so as to ensure that the labor market clears, while leaving the relative levels of population across the locations unchanged. The implicit function (Ω_n) is also monotonically increasing in own population (L_n) and monotonically decreasing in the population of other locations (L_k for $k \neq n$). Therefore the system of equations for equilibrium populations (21) satisfies gross substitution and yields the following unambiguous comparative static predictions.

Proposition 2 *A location n's equilibrium population (L_n) is increasing in its productivity (A_n), amenities (B_n) and land supply (H_n) relative to the values of these characteristics for all other locations $k \neq n$ and decreasing in its trade costs (d_{nk}) relative to the trade costs for all other locations $k \neq n$.*

Proof. See the web appendix. ■

In these comparative statics, the effects of changes in location characteristics on equilibrium populations are shaped by both goods and labor market frictions $\{\theta, \epsilon\}$. These parameters enter γ_1, γ_2 and Ω_{nk}^{II} in the implicit function (22) and hence shape the sensitivity of equilibrium populations to changes in productivity (A_n), amenities (B_n), land supplies (H_n) and trade costs (d_{ni}). Intuitively, locations with higher productivity, more attractive amenities, larger land supplies and lower trade costs attract larger populations, but the trade elasticity θ and the population supply elasticity ϵ determine the sensitivity of equilibrium populations to differences in these characteristics.

2.8 Recovering Location Fundamentals

Given values for the model's parameters $\{\alpha, \theta, \epsilon\}$, a parameterization of bilateral trade costs $\{d_{ni}\}$ and data on populations, wages and land supplies $\{L_n, w_n, H_n\}$, we now show that the solution to the general equilibrium of the model can be used to recover the unobserved location characteristics of amenities (B_n) and productivities (A_n).

Proposition 3 *Given the model parameters $\{\alpha, \theta, \epsilon\}$, a parameterization of bilateral trade costs $\{d_{ni}\}$ and data on populations, wages and land supplies $\{L_n, w_n, H_n\}$, there exist unique values of amenities (B_n) and productivities (A_n) that are consistent with the data up to a normalization that corresponds to a choice of units in which to measure amenities and productivities.*

Proof. See the web appendix. ■

To solve for unobserved productivities and amenities, we use the recursive structure of the model. First, given data on population and wages $\{L_n, w_n\}$, we can use the equality of income and expenditures (15) and trade shares (16) to solve for the unobserved productivities $\{A_n\}$ for which the endogenous variables are an equilibrium of the model. From these solutions for unobserved productivities $\{A_n\}$ and population and wages $\{L_n, w_n\}$, we immediately obtain trade shares $\{\pi_{ni}\}$. Second, given data on population and wages $\{L_n, w_n\}$, we can use land market clearing (14) to solve for land rents $\{r_n\}$. Third, given data on wages $\{w_n\}$ and the solutions for productivity and trade shares $\{A_n, \pi_{ni}\}$, we can use the relationship between price indices and trade shares (8) to solve for price indices $\{P_n\}$. Finally, using data on population and wages $\{L_n, w_n\}$ and the solutions for land rents and price indices $\{r_n, P_n\}$, we can use the residential choice probabilities (10) to solve for the unobserved amenities $\{B_n\}$ for which the endogenous variables are an equilibrium of the model.

2.9 Counterfactuals

The system of equations for general equilibrium (15)-(17) provides an approach for undertaking model-based counterfactuals that uses only parameters and the values of endogenous variables in the initial equilibrium (as in Dekle, Eaton, and Kortum 2007). In contrast to standard trade models, these model-based counterfactuals yield predictions for the reallocation of labor across locations.

The system of equations for general equilibrium (15)-(17) must hold both before and after a change in trade frictions, productivity or amenities. We denote the value of variables in the counterfactual equilibrium with a prime (x') and the relative value of variables in the counterfactual and initial equilibria by a hat ($\hat{x} = x'/x$). Using this notation, the system of equations for the counterfactual equilibrium (15)-(17) can be re-written as follows:

$$\hat{w}_i \hat{L}_i Y_i = \sum_{n \in N} \pi'_{ni} \hat{w}_n \hat{L}_n Y_n, \tag{23}$$

$$\pi'_{ni} = \frac{\pi_{ni} \hat{A}_i \left(\hat{d}_{ni} \hat{w}_i \right)^{-\theta}}{\sum_{k \in N} \pi_{nk} \hat{A}_k \left(\hat{d}_{nk} \hat{w}_k \right)^{-\theta}}, \tag{24}$$

$$\lambda'_n = \frac{\hat{B}_n \hat{A}_n^{\frac{\alpha\epsilon}{\theta}} \hat{\pi}_{nn}^{-\frac{\alpha\epsilon}{\theta}} \hat{\lambda}_n^{-\epsilon(1-\alpha)} \lambda_n}{\sum_{k \in N} \hat{B}_k \hat{A}_k^{\frac{\alpha\epsilon}{\theta}} \hat{\pi}_{kk}^{-\frac{\alpha\epsilon}{\theta}} \hat{\lambda}_k^{-\epsilon(1-\alpha)} \lambda_k}, \tag{25}$$

where $Y_i = w_i L_i$ denotes labor income and $\lambda_n = L_n/\bar{L}$ denotes the population share of each location in the initial equilibrium. For example, a reduction in trade costs holding productivity and amenities constant corresponds to $\hat{d}_{ni} < 1$, $\hat{A}_n = 1$ and $\hat{B}_n = 1$, while an increase in productivity corresponds to $\hat{A}_n > 1$.

2.10 Welfare Gains from Trade

We now examine the implications of imperfect factor mobility for the welfare gains from trade. Using the price index (8), income equals expenditure (12) and land market clearing (14), expected utility for the economy as a whole (11) can be written in terms of the population and domestic trade share for all locations:

$$\bar{U} = \delta \left[\sum_{n \in N} B_n \left(\frac{\left(\frac{A_n}{\pi_{nn}}\right)^{\frac{\alpha}{\theta}} H_n^{1-\alpha} L_n^{-(1-\alpha)}}{\alpha \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \gamma^{\alpha}} \right)^{\epsilon} \right]^{\frac{1}{\epsilon}}. \tag{26}$$

Intuitively, expected utility depends on productivity (A_n) and amenities (B_n) for each location, the domestic trade share for each location (π_{nn}) (since this affects the consumption price index), and the population for each location (L_n) (since this affects the demand for land and hence the price of land).

As discussed above, population mobility and the Fréchet distribution for amenities imply that the expected utility conditional on living in each location is equal to the above expected utility for the economy as a whole. Noting that the denominator in the residential choice probabilities (10) is a power function of expected utility (11), we can use these residential choice probabilities together with income equals expenditure (12), land market clearing (14) and the goods price index (8) to write the common level of expected utility solely in terms of the domestic trade share and population of an individual location:

$$\bar{U}_n = \bar{U} = \frac{\delta B_n^{\frac{1}{\epsilon}} \left(\frac{A_n}{\pi_{nn}}\right)^{\frac{\alpha}{\theta}} H_n^{1-\alpha} L_n^{-\left(\frac{1}{\epsilon}+(1-\alpha)\right)}}{\alpha \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \gamma^{\alpha} \left(\bar{L}\right)^{-\frac{1}{\epsilon}}}, \qquad \forall\, n. \tag{27}$$

Population mobility implies that this relationship must hold for each location. Locations with higher productivity (A_n), better amenities (B_n), better goods market access to other locations (lower π_{nn}) and higher supplies of land (H_n) have higher populations, which bids up the price of land until expected utility conditional on living in each location is the same for all locations.

An implication of this result is that the domestic trade share in the open economy equilibrium (π_{nn}^T), populations in the closed and open economies (L_n^A and L_n^T), the trade elasticity (θ), the elasticity of population supply with respect to real income (ϵ) and the consumption goods share (α) are sufficient statistics for the welfare gains from trade:

$$\frac{\bar{U}_n^T}{\bar{U}_n^A} = \frac{\bar{U}^T}{\bar{U}^A} = \left(\frac{1}{\pi_{nn}^T}\right)^{\frac{\alpha}{\theta}} \left(\frac{L_n^A}{L_n^T}\right)^{\frac{1}{\epsilon}+(1-\alpha)}, \qquad \forall\, n, \tag{28}$$

where we use the superscript T to denote the trade equilibrium and the superscript A to denote the autarky equilibrium; we have used $\pi_{nn}^A = 1$; and in general $L_n^A \neq L_n^T$.

Intuitively, if some locations have better market access than others in the open economy (as reflected in a lower open economy domestic trade share π_{nn}), the opening of goods trade will lead to a larger reduction in the consumption price index in these locations. This larger reduction in the consumption price index in turn creates an incentive for migration from locations with worse market access to those with better market access. This labor mobility provides the mechanism that restores equilibrium, as the price of land is bid up in locations with better market access and bid down in those with worse market access, until expected utility is equalized across all locations. Therefore, computing the common value for the welfare gains from trade across all locations involves taking into account not only domestic trade shares (which affect consumption price indices) but also population redistributions (which affect the price of the immobile factor land).

Although factor mobility ensures the equalization of expected utility across all locations, real income is not equalized, because of the heterogeneity in workers' idiosyncratic tastes for locations. Each location faces an upward sloping supply curve for workers, as higher real income has to be paid to attract workers with lower realizations for idiosyncratic tastes for that location. Only in the special case of no idiosyncratic heterogeneity in worker tastes ($\epsilon \to \infty$) is there perfect labor mobility and real income equalization across locations. In this special case, the common level of expected utility is given by:

$$\bar{U}_n = \bar{U} = \frac{\left(\frac{A_n}{\pi_{nn}}\right)^{\frac{\alpha}{\theta}} H_n^{1-\alpha} L_n^{-(1-\alpha)}}{\alpha \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \gamma^{\alpha}}, \qquad \forall\, n, \tag{29}$$

and the welfare gains from trade are:

$$\frac{\bar{U}_n^T}{\bar{U}_n^A} = \frac{\bar{U}^T}{\bar{U}^A} = \left(\frac{1}{\pi_{nn}^T}\right)^{\frac{\alpha}{\theta}} \left(\frac{L_n^A}{L_n^T}\right)^{1-\alpha}, \qquad \forall\, n, \tag{30}$$

which corresponds to the limiting case of (27) and (28) in which $\epsilon \to \infty$.

In the opposite polar extreme of perfect labor immobility, expected utility takes the same form as in (29), except that expected utility in general differs across locations in the absence of labor mobility:

$$\bar{U}_n = \frac{\left(\frac{A_n}{\pi_{nn}}\right)^{\frac{\alpha}{\theta}} H_n^{1-\alpha} L_n^{-(1-\alpha)}}{\alpha \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \gamma^{\alpha}} \neq U_k, \qquad n \neq k. \tag{31}$$

Similarly, the welfare gains from trade in general differ across locations under perfect labor immobility:

$$\frac{\bar{U}_n^T}{\bar{U}_n^A} = \left(\frac{1}{\pi_{nn}^T}\right)^{\frac{\alpha}{\theta}} \neq \frac{\bar{U}_k^T}{\bar{U}_k^A}, \qquad n \neq k. \tag{32}$$

Intuitively, when labor is perfectly immobile, locations with better access to markets in the open economy experience larger welfare gains from trade, because population reallocations no longer provide a mechanism for utility equalization through changes in the price of land.

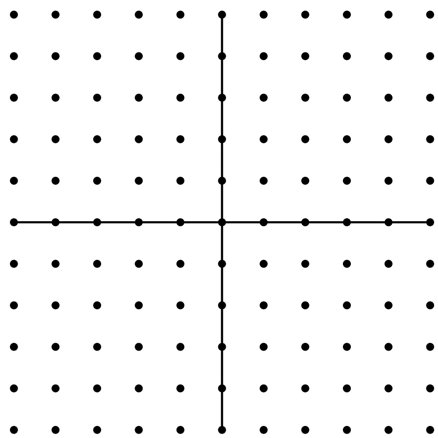
3 Quantitative Analysis

To examine the quantitative properties of the model, we first assume parameter values and generate data for a hypothetical economy from the model. Second, we examine the implications of a reduction in trade costs on the organization of economic activity within this economy. Third, we examine how the model's quantitative predictions for the impact of this reduction in trade costs depend on its parameter values. Fourth, we suppose that a researcher only observes data on employment and wages before and after the reduction in trade costs. Under the assumption that productivity and amenities are unchanged over time, we show that data on the endogenous variables of the model before and after the reduction in trade costs can be used to estimate its parameters.

3.1 Model Economy

We consider a model economy on a latitude and longitude grid, as shown in Figure 1, where each dot corresponds to a location. We assume a transport cost for each location and compute the least cost route of traveling between each pair of locations (d_{ni}).[9] Before the transport infrastructure improvement, we assume the same transport cost for traveling across each point on the grid (equal to the estimated cost of land travel of 7.9 in Donaldson 2014).

Figure 1: Model Economy



Note: Grid of locations in latitude and longitude space and the route of the transport infrastructure improvement.

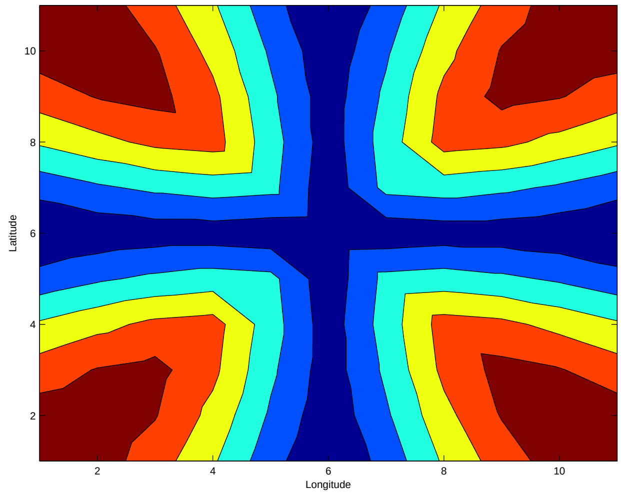
We suppose that the transport infrastructure improvement involves the construction of a road/railroad that directly reduces the cost of traveling across each point on its route to 1 (equal to the estimated cost of rail travel in Donaldson 2014). This transport cost improvement also indirectly reduces the transport cost of traveling between other bilateral pairs of locations to the extent that the least cost route between these locations involves traveling along the road/railroad. In Figure 1, we show the assumed route of the road/railroad by the horizontal and vertical lines. In Figure 2, we show the resulting overall reduction in average transport costs for each location as a contour plot. Blue (cold) colors correspond to lower values (larger reductions) and red (hot) colors correspond to higher values (smaller reductions). As apparent from the figure, those locations directly along the route of the road/railroad experience the largest reductions in average transport costs. But neighboring locations close to the road/railroad also experience larger reductions in average transport costs than those further away from the road/railroad.

We assume that productivity, amenities and land supply are unchanged before and after the transport improvement. Each location is assumed to have a land area (H_n) of 100 meters squared.[10] We allow both productivity and amenities to differ randomly across locations. For each location, we draw a realization

[9] We compute a measure of the lowest cost route effective distance following Donaldson (2014). Denoting the transport costs for a pair of neighboring locations by c_1 and c_2, the accumulative cost for orthogonal links is $a = (c_1 + c_2)/2$, while the accumulative cost for diagonal links is $a = \left(\left(2(c_1 + c_2)^2\right)^{0.5}\right)/2$. The transport costs for a pair of non-neighboring locations are the sum of these transport costs between neighboring locations along the least cost route between that pair of locations.

[10] While it is straightforward to allow land area (H_n) to vary across locations $n \in N$, such differences in land area enter the model is the same way as differences in amenities (B_n).

Figure 2: Relative Reduction in Transport Costs



Note: Contours for average reductions in transport costs to other locations from the transport infrastructure improvement.

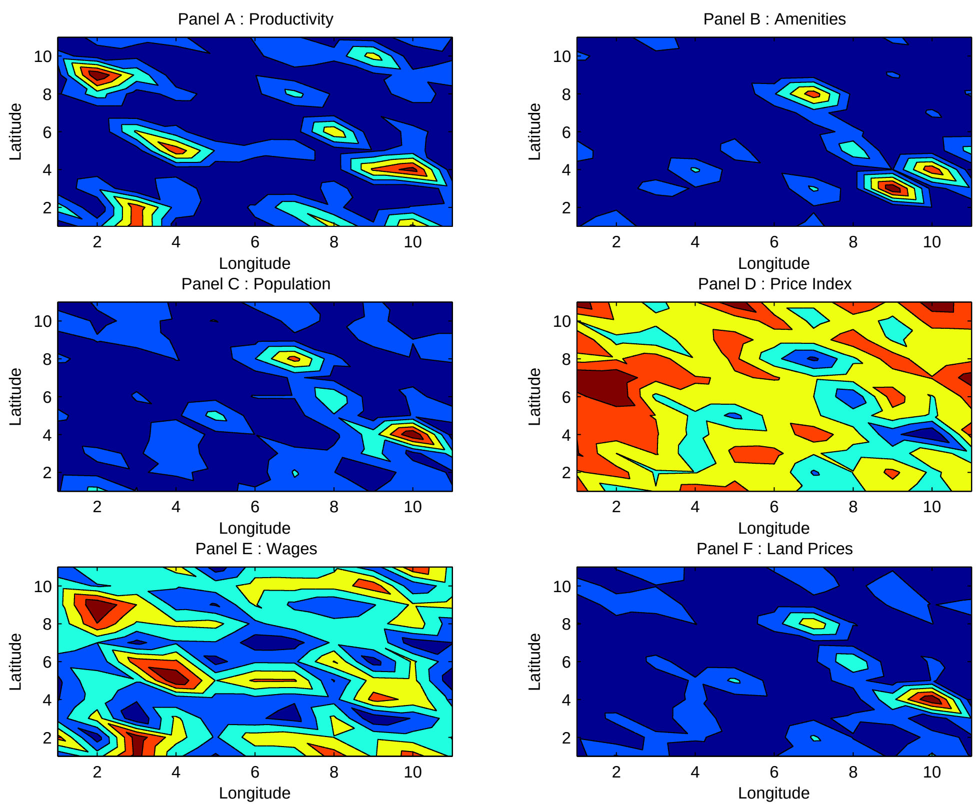
for the Fréchet scale parameter for productivity (A_n) and a realization for the Fréchet scale parameter for amenities (B_n) from independent standard log normal distributions.

We choose central values for the model's parameters based on the existing empirical literature. First, we set the share of land in residential consumption expenditure ($1-\alpha$) to 25 percent, which is in line with the housing expenditure share in Davis and Ortalo-Magné (2011). Second, we set the elasticity of substitution (σ) equal to four, which is consistent with the estimates using plant-level U.S. manufacturing data in Bernard, Eaton, Jensen, and Kortum (2003). Third, the Fréchet shape parameter for productivity (θ) corresponds to the elasticity of trade flows with respect to trade costs. We assume a value of four for θ as a central value for the trade elasticity in the empirical trade literature (e.g. Simonovska and Waugh 2014), which ensures that the condition for the integral in the price index to converge ($\theta > \sigma - 1$) is satisfied. Fourth, we assume a constant elasticity relationship between trade costs and distance ($d_{ni} = \text{dist}_{ni}^{\phi}$), and suppose that the elasticity of trade costs with respect to distance (ϕ) is one third, which lies within the range of existing empirical estimates (see for example Hummels 2007, Limao and Venables 2001). These values for θ and ϕ imply an elasticity of trade with respect to distance ($\theta \times \phi$) of around one, which is consistent with the gravity equation estimates reviewed in Head and Mayer (2014). Fifth, the Fréchet shape parameter for migration decisions (ϵ) corresponds to the elasticity of population with respect to real income. We assume a benchmark value of four for ϵ and explore the robustness of the model's quantitative properties to alternative values for this parameter.

3.2 Initial Equilibrium

Using the assumed parameters $\{\alpha, \sigma, \phi, \theta, \epsilon\}$ and location characteristics $\{H_n, A_n, B_n\}$, we solve for the general equilibrium of the model using the system of equations (15)-(17). In Figure 3, we show the distribution of economic activity across locations in the initial equilibrium before the transport improvement. In each panel, we display contours of economic activity across the latitude and longitude grid. Again blue (cold) colors denote lower values and red (hot) colors denote higher values.[11] Panels A and B show the realizations of productivity (A_n) and amenities (B_n) that are randomly distributed across locations. Panels C and D show the equilibrium distribution of population (L_n) and consumption price indices (P_n) across locations. From Panels A-C, population is concentrated close to locations with high productivities and high amenities. From Panels A-D, price indices are low in locations close to concentrations of economic activity (population), because of transport costs. Panels E and F show the equilibrium distribution of wages (w_n) and land prices (r_n). From Panels A, C and E, wages are high in locations where productivity (A_n) is high relative to the endogenous supply of labor (L_n). From Panels A, B, C and F, land prices are high in locations with large populations (L_n) relative to the common supply of land ($H_n = H$).

Figure 3: Initial Equilibrium



Note: Contours for levels of economic activity across the latitude and longitude grid.

3.3 Reduction in Transport Costs

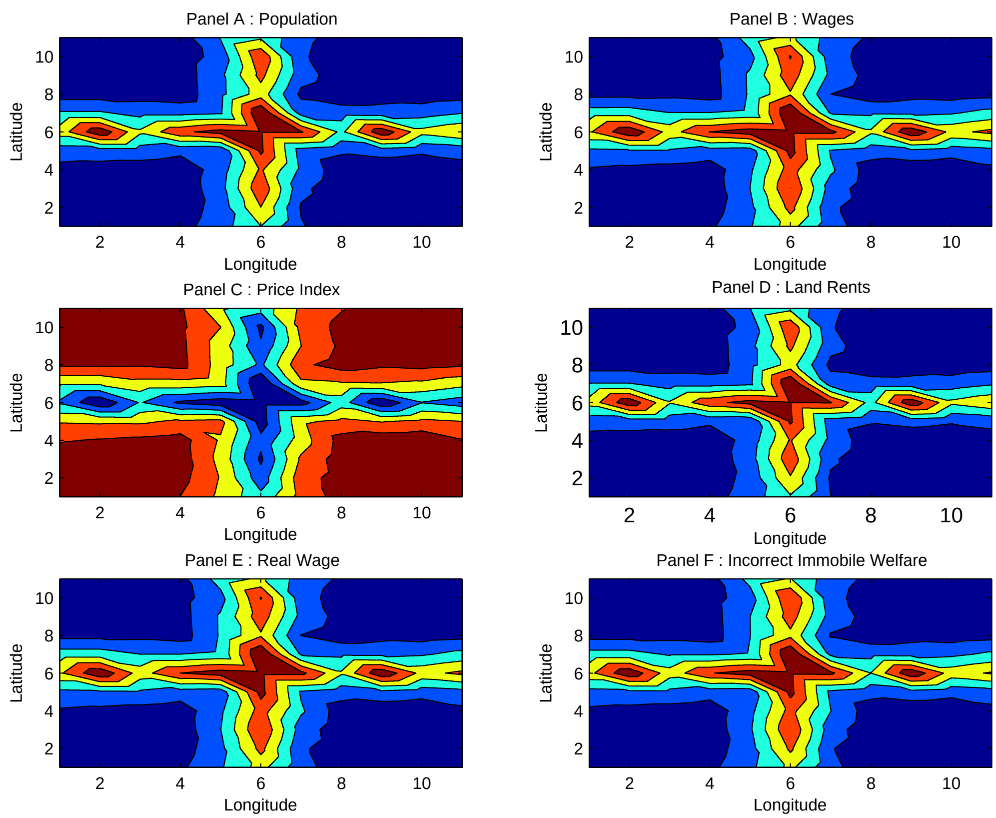
In Figure 4, we show the impact of the reduction in transport costs on the spatial distribution of economic activity, by displaying the relative change in each measure of economic activity ($\hat{x} = x'/x$) between the new equilibrium (denoted by a prime) and the initial equilibrium. Again blue (cold) colors denote lower values and red (hot) colors denote higher values. From Figure 2, locations close to the route of the

[11] Specifically, we construct a three-dimensional surface through the values for economic activity at each point on the latitude and longitude grid using linear (triangular) interpolation. The figures show the contours for this three-dimensional surface.

road/railroad experience the largest reductions in transport costs. Therefore, in Figure 4, we find that these locations experience the largest increases in population (Panel A) and wages (Panel B). Nevertheless, the contours for these relative increases in population and wages do not perfectly coincide with the contours for the mean reduction in transport costs in Figure 2. The reason is that the economic impact of a given transport cost reduction between a pair of locations depends on the economic characteristics of those locations (e.g. productivity and amenities). These economic characteristics differ across locations depending on the stochastic variation in productivity and amenities, as shown in Figure 3.

The direct effect of the larger reduction in transport costs for locations close to the route of the road/railroad is a larger reduction in the consumption price index for these locations. But there is also an indirect effect of the larger increase in wages for these locations that raises the consumption price index. Nevertheless, the direct effect dominates, so that locations close to the route of the road/railroad experience larger reductions in consumption price indices (Panel C). The increase in both population and wages in these locations also leads to an increase in the price of land (Panel D).

Figure 4: Relative Changes ($\hat{x} = x'/x$) Following the Transport Improvement



Note: Contours for relative changes in economic activity following the transport infrastructure improvement.

While the increase in wages and the reduction in consumer price indices raise real wages, the increase in land prices has the opposite effect of reducing real wages. On net, we find that locations close to the route of the road/railroad experience larger increases in real wages (Panel E) as higher real wages have to be paid to attract additional workers with lower realizations of idiosyncratic tastes for these locations.

In contrast, average utility conditional on residing in each location (not shown in the figure) takes into account both real wages and average idiosyncratic tastes. As discussed above, average utility is equalized across all locations in the spatial equilibrium and is equal to expected utility for the economy as a whole. We find that the transport improvement raises this common level of expected utility by 9.5 percent. This common change in expected utility takes into account both the change in the domestic trade share and the change in population (28). To provide a point of comparison, Panel F displays the increase in welfare that would be computed by a policy analyst, who falsely assumed that labor is immobile across locations and calculated the welfare gains from the transport improvement based solely on the change in the domestic trade share (31). Whereas the true change in expected utility is the same for all locations, those locations close to the route of the road/railroad experience the largest increases in this incorrect measure of welfare based on the (false) assumption of labor immobility.

3.4 Treatment Effects of the Transport Improvement

A growing empirical literature uses quasi-experimental variation in transport infrastructure investments to estimate the reduced-form impact of these investments on the spatial distribution of economic activity (see for example Duranton and Turner 2012). In our model economy in Figure 1, the route for the transport infrastructure improvement was exogenously assigned. Therefore we use this quasi-experimental variation to estimate the impact of this transport infrastructure improvement on the spatial distribution of economy activity within the model. Under exogenous assignment, the causal impact of the transport infrastructure improvement can be estimated using the following "differences-in-differences" specification:

$$\ln Y_{nt} = \vartheta_n + \beta \mathbb{I}_{nt} + d_t + u_{nt}, \tag{33}$$

where n indexes locations and t indexes periods (before and after the transport improvement); Y_{nt} is an economic outcome of interest (e.g. population); $\mathbb{I}_{nt}$ is an indicator variable that is one if a location is treated with transport infrastructure and zero otherwise; treatment is defined in terms of a location being directly affected by the transport infrastructure improvement; ϑ_d are location fixed effects; d_t are period fixed effects; and u_{nt} is a stochastic error. The inclusion of both sets of fixed effects ensures a "differences-in-differences" interpretation, where the first difference is between treated and untreated locations and the second difference is before and after the transport improvement.

Taking differences in (33) before and after the transport infrastructure improvement, we obtain the following "long differences" specification:

$$\triangle \ln Y_{nt} = \nu + \beta \mathbb{I}_{nt} + e_{nt}, \tag{34}$$

where the location fixed effects have now differenced out and with only two periods the change in the period fixed effects is captured in the regression constant ν.

In Table 1, we report the results of estimating the long differences specification (34) for the transport infrastructure improvement shown in Figure 1. Consistent with the reorientation of the spatial distribution of economic activity shown in Figure 4, we find positive average treatment effects for population and wages, a negative average treatment effect for the price index, and positive average treatment effects for land rents, real wages and the incorrect measure of welfare based on the (false) assumption of labor immobility. However, as is also apparent from Figure 4, these estimated average treatment effects mask considerable heterogeneity in the impact of the transport improvement. Among the treated locations, those locations closest to the intersection of the horizontal and vertical lines, experience the largest reductions in transport costs and hence the largest increases in the concentration of economic activity. Among the untreated locations that are not directly affected by the transport infrastructure, many are indirectly affected by it because it reduces transport costs along the least cost route to other locations.

Table 1: Treatment Effects of the Transport Improvement

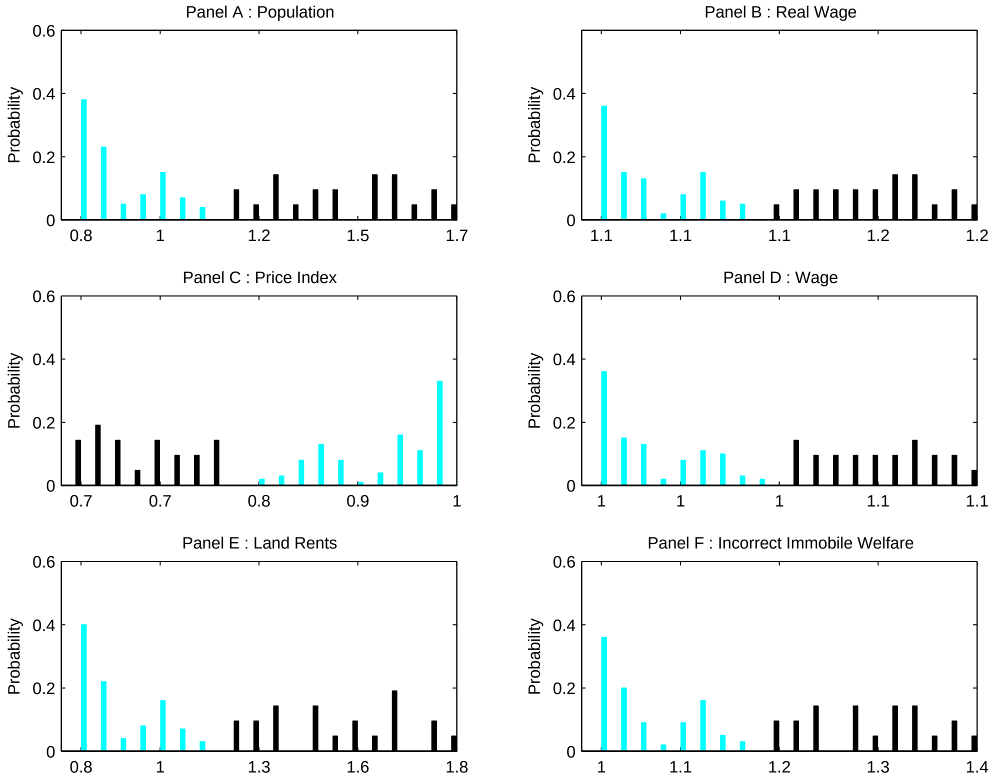
Economic Outcome	Treatment Effect
Population	0.4282
	(0.0224)
Wage	0.0793
	(0.0042)
Price Index	-0.2062
	(0.0108)
Land Rents	0.5075
	(0.0266)
Real Wages	0.1071
	(0.0056)
Immobile Welfare	0.2141
	(0.0112)

Note: Table reports the results of the estimating the long differences specification (34) for the impact of the transport infrastructure improvement on each of the economic outcomes for treated relative to untreated locations. A separate regression is estimated for each economic outcome. Standard errors in parentheses.

In Figure 5, we provide further evidence on these heterogeneous treatment effects by displaying the distributions of the relative changes ($\hat{x} = x'/x$) in the economic outcomes shown in Figure 4 as histograms across twenty equally-spaced bins. We show the distributions for treated locations (in black) and untreated locations (in light blue) separately. As apparent from the figure, we find considerable heterogeneity among both groups of locations. Heterogeneity in access to transport infrastructure among the treated locations and in the indirect effects of the transport infrastructure among the untreated locations imply that the smallest positive changes for the treated locations are close to the largest positive changes for the untreated locations. For example, for population, the relative change among treated locations varies from 1.7 to less than 1.2, while the relative change among untreated locations varies from 0.8 to more than 1.1. Taken together, these results highlight that a road/railroad connection between any two locations affects all other locations, to a degree that varies depending on the change in the transport network and the general equilibrium reallocation of economic activity that it induces.

As discussed above, the true relative change in welfare as a result of the transport infrastructure (equation (28)) is the same across all locations. Therefore, the treatment effect for true welfare (not reported in

Figure 5: Distribution of Relative Changes ($\hat{x} = x'/x$) Following the Transport Improvement



Note: Histogram of relative changes in economic activity following the transport infrastructure improvement.

Table 1) is zero, because there is no differential change between treated and untreated locations. Furthermore, the distribution of relative changes in true welfare in Figure 5 is degenerate at 1.095. In contrast, a policy analyst who computed the relative change in welfare under the false assumption of labor immobility (equation (31)) would estimate a substantial positive treatment effect of 0.2141 for this incorrect measure of welfare in Table 1. Additionally, this policy analyst would find substantial heterogeneity across locations in the welfare effects of the transport improvement, ranging from around 1 to 1.4 in Panel F of Figure 5. These results suggests that not controlling for factor mobility across locations (and the resulting changes in the price of the immobile factor of production land) can lead to quantitatively substantial discrepancies between the true and measured welfare gains from transport infrastructure improvements.

Finally, another challenge faced by the reduced-form regression specification is that the relative comparison between treated and untreated locations does not distinguish reallocation effects from the creation of new economic activity. Comparing the distributions of treatment effects for population in Figure 5 (which range up to 70 percent) to the true common change in welfare across all locations (of 9 percent), it is apparent that the reallocation effects are large relative to the welfare effect. As argued by Fogel (1964), large-scale reallocations of economic activity as a result of a new transport technology need not necessarily imply welfare gains of the same magnitude. Although labor mobility is imperfect, the equalization of expected utility across locations implies that the welfare effects of the transport improvement on treated locations are shared with the economy as a whole, dampening the magnitude of these welfare effects.

3.5 Size of Goods and Labor Market Frictions

We now explore how the size of goods and labor market frictions influences the impact of the transport improvement. We undertake a grid search over values for the Fréchet shape parameter for productivity θ (that determines the substitutability of locations in goods production) and the Fréchet shape parameter for idiosyncratic tastes ϵ (that determines the substitutability of locations in worker utility). We consider values for $\{\theta,\, \epsilon\}$ ranging from 3.1 to 5.1 (which satisfy $\theta > \sigma - 1$) and hold the other parameters $\{\alpha,\, \sigma,\, \phi\}$ and productivity, amenities and land supply $\{A_n,\, B_n,\, H_n\}$ constant at their calibrated values. For each parameter combination, we solve for the initial and final equilibria before and after the transport improvement, and estimate the reduced-form regression (34) for the average impact on treated relative to untreated locations (β).

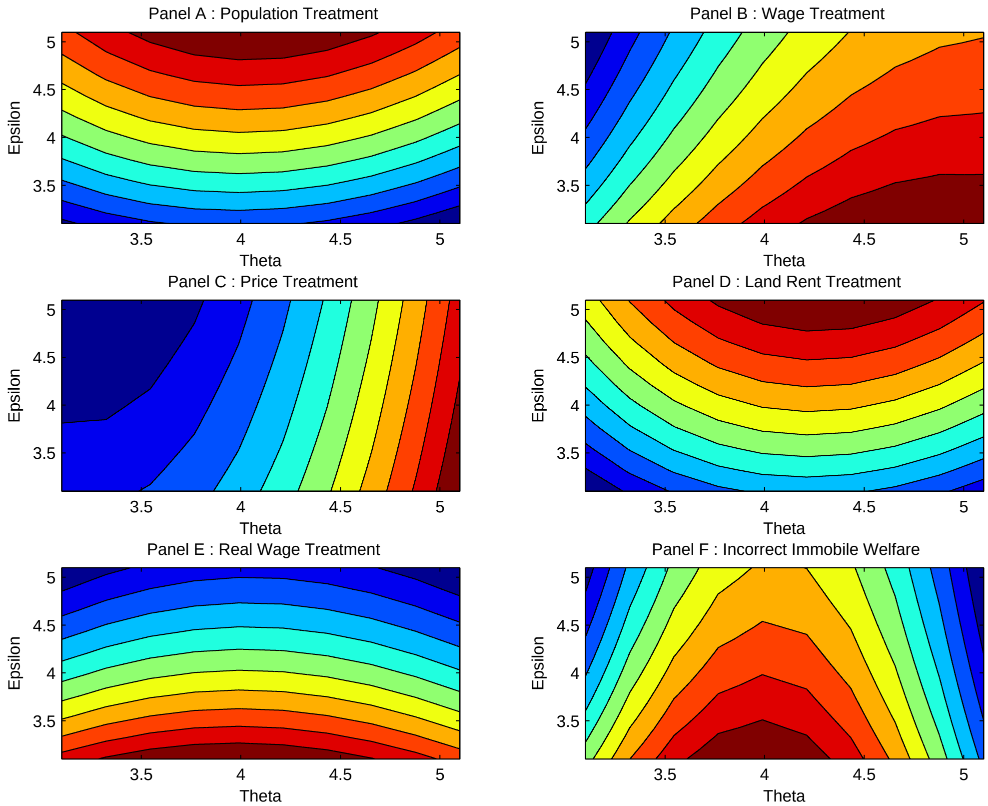
In Figure 6, we display contour plots for the average treatment effects for each economic outcome in the $\{\theta,\, \epsilon\}$ parameter space. Again red (hot) corresponds to higher values and blue (cold) corresponds to lower values. As shown in Panel A, the average treatment effect for population is larger for higher values of ϵ and intermediate values of θ. The reason is as follows. For higher values of ϵ, there is less dispersion in idiosyncratic tastes and more labor mobility, so that the transport improvement leads to a greater reallocation of population. For intermediate values of θ, the change in transport costs has the largest effects on the relative attractiveness of locations, so that the transport improvement again leads to a greater reallocation of population. In contrast, for low values of θ all locations have similar levels of access to goods from other locations, while for high values of θ all locations are largely closed to goods trade.

As shown in Panel B, the average treatment effect for wages is larger for smaller values of ϵ and higher values of θ. For smaller values of ϵ, there is more dispersion in idiosyncratic tastes and less labor mobility, so that larger wage changes are required to induce workers to reallocate between locations. For higher values of θ, the goods market effects of the transport improvement are more localized, which in turn leads to larger wage changes. As shown in Panel C, the average treatment effect for the price index displays a similar pattern as for wages. The reduction in the price index in treated relative to untreated locations is larger for high values of ϵ and low values of θ. On the other hand, as shown in Panel D, the average treatment effect for land prices shows a similar pattern as for population. The increase in land prices in treated relative to untreated locations is larger for higher values of ϵ and intermediate values for θ, which reflects the larger population reallocations for these parameter values shown in Panel A.

As shown in Panel E, the average treatment effect for real wages is larger for lower values of ϵ and intermediate values for θ. For lower values of ϵ, there is more dispersion in idiosyncratic tastes, and hence less labor mobility to arbitrage away real wage differences. For intermediate values of θ, the change in transport costs has the largest effects on the relative attractiveness of locations, as discussed above.

Across all of the parameter combinations shown in Figure 6, the average treatment effect for the true measure of welfare (equation (28)) is zero, because expected utility conditional on living in a location is the

Figure 6: Comparative Statics for Average Treatment Effects



Note: Contours for average treatment effect from the transport improvement for each measure of economic activity across the parameter grid $\epsilon \in [3.1, 5.1]$ and $\theta \in [3.1, 5.1]$.

same for treated and untreated locations. In contrast, as shown in Panel F, a policy analyst who computed the average treatment effect for welfare under the false assumption of labor immobility (equation (31)) would estimate a larger average treatment effect for smaller values of ϵ (more dispersion in idiosyncratic tastes and less labor mobility) and intermediate values of θ (for which the transport improvement has the largest effects on the relative attractiveness of locations). Therefore the bias in computing the welfare gains from the transport improvement under the (false) assumption of labor immobility varies systematically with the degree of goods and labor market frictions.

In general, lower levels of labor mobility imply larger treatment effects for wages, but smaller treatment effects for population and land prices. For example, for our baseline value of productivity dispersion ($\theta = 4$) and $\epsilon = 3.1$ (lower labor mobility), we find average treatment effects for population (0.3798), wages (0.0867) and land rents (0.4664). In contrast, for $\theta = 4$ and $\epsilon = 5.1$ (higher labor mobility), we find average treatment effects for population (0.4733), wages (0.0725) and land rents (0.5458).

3.6 Recovering the Parameters and Unobserved Location Characteristics

We now examine whether data on population and wages before and after the transport improvement can be used to estimate the model's parameters $\{\theta, \epsilon\}$. First, we use our calibrated parameters $\{\alpha, \sigma, \phi, \theta, \epsilon\}$

to solve for the equilibrium distribution of economic activity before and after the transport improvement. Second, we suppose that a policy analyst does not know the true values of the parameters $\{\theta, \epsilon\}$, but can correctly calibrate the remaining parameters $\{\alpha, \sigma, \phi\}$, knows trade costs and land area $\{d_{ni}, H_n\}$, and observes population and wages $\{L_n, w_n\}$ before and after the transport improvement.

To estimate $\{\theta, \epsilon\}$, we use the identifying assumption that the unknown values of productivity and amenities $\{A_n, B_n\}$ are unchanged before and after the transport improvement. We combine this identifying assumption with the result in Proposition 3 that there is a one-to-one mapping from the parameters $\{\alpha, \sigma, \phi, \theta, \epsilon\}$ and the data $\{L_n, w_n, d_{ni}, H_n\}$ to unobserved productivity and amenities $\{A_n, B_n\}$.

We estimate $\{\theta, \epsilon\}$ using the generalized method of moments (GMM). For each candidate parameter value, we solve for the implied change in productivity and amenities $\{A_i, B_i\}$ for each location before and after the transport improvement, and compute the sum of squared values of these changes. Under our identifying assumption, the GMM objective function defined by the sum of these squared changes should be equal to zero for the true parameter values. We therefore undertake a grid search over the parameter space $\{\theta, \epsilon\}$ to minimize the GMM objective function:

$$\Lambda = \begin{pmatrix} \triangle \ln a_i & \triangle \ln b_i \end{pmatrix} \times \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} \triangle \ln a_i \\ \triangle \ln b_i \end{pmatrix}, \tag{35}$$

where we weight the moment conditions equally using the identity matrix.

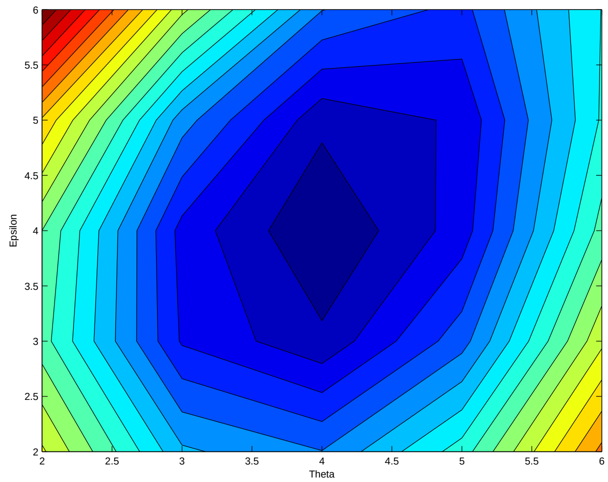
In Figure 7, we display the GMM objective (35) as a contour plot in the $\{\theta, \epsilon\}$ parameter space. Again red (hot) corresponds to higher values and blue (cold) corresponds to lower values. As apparent from the figure, the GMM objective function is well behaved in the parameter space with a unique global minimum at the true value of the model's parameters $\{4, 4\}$. Using these estimated parameter values, we recover the correct values of unobserved productivity and amenities $\{A_n, B_n\}$ for each location.

Intuitively, as shown in the previous subsection, the parameters $\{\theta, \epsilon\}$ have different implications for the impact of the transport improvement on population and wages $\{L_n, w_n\}$. Therefore the changes in population and wages following the transport improvement, together with our identifying assumption that productivity and amenities $\{A_n, B_n\}$ are unchanged, can be used to estimate the parameters $\{\theta, \epsilon\}$.

4 Agglomeration Forces

In this section, we examine the implications of introducing agglomeration forces in our setting with both goods and labor market frictions. These agglomeration forces take the form of pecuniary externalities as a result of transport costs, increasing returns to scale and love of variety, as in the new economic geography literature following Krugman (1991a,b), Krugman and Venables (1995) and Helpman (1998), and synthesized in Fujita, Krugman, and Venables (1999). This literature typically restricts attention to stylized settings with a small number of symmetric locations and assumes either perfect labor mobility, perfect labor immobility or a mechanical relationship between migration flows and relative wages. In contrast, we consider a rich geography with a large number of asymmetric locations, and incorporate both goods and labor market frictions.

Figure 7: Monte Carlo Results



Note: Contours for generalized method of moments (GMM) objective function of sum of squared deviations of changes in productivity and amenities across the parameter grid $\epsilon \in [3.1, 5.1]$ and $\theta \in [3.1, 5.1]$.

We show that the general equilibrium of the model can be represented in a similar way to the constant returns model in the previous section, but with the key difference that the measure of goods produced by a location depends on population. In an international trade context, population is exogenous and this difference between the two models is inconsequential for their counterfactual predictions. However, in our setting with imperfect labor mobility, changes in trade costs lead to reallocations of population. In the increasing returns model, these population reallocations lead to changes in the measure of goods produced by each location, which affect bilateral trade shares. In contrast, in the constant returns model, the measure of goods produced by each location is an exogenous primitive of the model. As a result, even if the two models are calibrated to same initial equilibrium with the same trade and population supply elasticities, they have different counterfactual predictions for the effects of reductions in trade costs from this common initial equilibrium. Therefore they have different counterfactual predictions for the welfare gains from these reductions in trade costs.

4.1 Consumer Preferences

Preferences are again defined over goods consumption (C_n) and residential land use (H_{Un}) and take the same form as in (1). The goods consumption index (C_n), however, is now defined over the endogenous measures of horizontally differentiated varieties supplied by each location (M_i):

$$C_n = \left[\sum_{i \in N} \int_0^{M_i} c_{ni}(j)^{\rho} \, dj \right]^{\frac{1}{\rho}}, \tag{36}$$

where trade between locations i and n is again subject to iceberg variable trade costs of $d_{ni} \geq 1$.

4.2 Production

Varieties are produced under conditions of monopolistic competition and increasing returns to scale. To produce a variety, a firm must incur a fixed cost of F units of labor and a constant variable cost in terms of labor that depends on a location's productivity A_i. Therefore the total amount of labor ($l_i(j)$) required to produce $x_i(j)$ units of a variety j in location i is:

$$l_i(j) = F + \frac{x_i(j)}{A_i}. \tag{37}$$

Profit maximization and zero profits imply that equilibrium prices are a constant mark-up over marginal cost:

$$p_{ni}(j) = \left(\frac{\sigma}{\sigma-1}\right)\frac{d_{ni}w_i}{A_i}, \tag{38}$$

and equilibrium employment for each variety is equal to a constant:

$$l_i(j) = \bar{l} = \sigma F. \tag{39}$$

Given this constant equilibrium employment for each variety, labor market clearing implies that the total measure of varieties supplied by each location is proportional to the endogenous supply of workers choosing to locate there:

$$M_i = \frac{L_i}{\sigma F}. \tag{40}$$

4.3 Expenditure Shares and Price Indices

Using the CES expenditure function, equilibrium prices (38) and labor market clearing (40), the share of location n's expenditure on goods produced in location i is:

$$\pi_{ni} = \frac{L_i\left(\frac{d_{ni}w_i}{A_i}\right)^{1-\sigma}}{\sum_{k\in N} L_k\left(\frac{d_{nk}w_k}{A_k}\right)^{1-\sigma}}, \tag{41}$$

where the elasticity of trade with respect to trade costs is now determined by the elasticity of substitution ($\sigma - 1$). Furthermore, trade shares now depend directly on population (L_i) because this determines the endogenous measure of varieties produced by a location through the labor market clearing condition (40).

Using equilibrium prices (38), labor market clearing (40), the trade share (41) and $d_{nn} = 1$, the consumption goods price index can be written as:

$$P_n^{1-\sigma} = \frac{L_n}{\sigma F \pi_{nn}}\left(\frac{\sigma}{\sigma-1}\frac{w_n}{A_n}\right)^{1-\sigma}, \tag{42}$$

which again depends directly on population (L_n) through the endogenous measure of varieties.

4.4 Residential Choices and Income

Residential choices take a similar form as in section 2. Using the Fréchet distribution of idiosyncratic shocks to amenities, the probability that a worker chooses to live in location $n \in N$ is:

$$\frac{L_n}{\bar{L}} = \frac{B_n \left(v_n / P_n^{\alpha} r_n^{1-\alpha}\right)^{\epsilon}}{\sum_{k \in N} B_k \left(v_k / P_k^{\alpha} r_k^{1-\alpha}\right)^{\epsilon}}, \tag{43}$$

where the elasticity of population with respect to real income is again determined by the Fréchet shape parameter for consumer tastes ϵ. Expected worker utility is:

$$\bar{U} = \delta \left[\sum_{k \in N} B_k \left(v_k / P_k^{\alpha} r_k^{1-\alpha}\right)^{\epsilon}\right]^{\frac{1}{\epsilon}}, \tag{44}$$

where $\delta = \Gamma\left((\epsilon - 1)/\epsilon\right)$; $\Gamma(\cdot)$ is the Gamma function; and $\epsilon > 1$. The Fréchet distribution of utility again implies that expected utility conditional on residing in location n is the same across all locations n and equal to expected utility for the economy as a whole.

Expenditure on land in each location is redistributed lump sum to the workers residing in that location, which implies that total income (v_n) equals labor income plus expenditure on residential land (as in (12)). Land market clearing implies that the equilibrium land rent again can be determined from the equality of land income and expenditure (as in (14)).

4.5 General Equilibrium

The general equilibrium of the model can be represented by the measure of workers (L_n) in each location $n \in N$, the share of each location's expenditure on goods produced by other locations (π_{ni}) and the wage in each location (w_n). Using labor income (13), the trade share (41), residential choice probabilities (43) and land market clearing (14), the equilibrium triple $\{L_n, \pi_{ni}, w_n\}$ solves the following system of equations for all $i, n \in N$. First, each location's income must equal expenditure on the goods produced in that location:

$$w_i L_i = \sum_{n \in N} \pi_{ni} w_n L_n. \tag{45}$$

Second, location expenditure shares are:

$$\pi_{ni} = \frac{L_i \left(\frac{d_{ni} w_i}{A_i}\right)^{1-\sigma}}{\sum_{k \in N} L_k \left(\frac{d_{nk} w_k}{A_k}\right)^{1-\sigma}}. \tag{46}$$

Third, residential choice probabilities imply:

$$\frac{L_n}{\bar{L}} = \frac{B_n A_n^{\alpha\epsilon} H_n^{\epsilon(1-\alpha)} \pi_{nn}^{-\frac{\alpha\epsilon}{\sigma-1}} L_n^{-\left(\epsilon(1-\alpha) - \frac{\alpha\epsilon}{\sigma-1}\right)}}{\sum_{k \in N} B_k A_k^{\alpha\epsilon} H_k^{\epsilon(1-\alpha)} \pi_{kk}^{-\frac{\alpha\epsilon}{\sigma-1}} L_k^{-\left(\epsilon(1-\alpha) - \frac{\alpha\epsilon}{\sigma-1}\right)}}. \tag{47}$$

4.6 Existence and Uniqueness

We now show that there exists a unique general equilibrium that solves the system of equations (45)-(47). Using the requirement that labor income for each location equals expenditure on goods produced in that location (45), we obtain one system of equations for the wages and populations of locations $n \in N$ as a function of parameters and a transformation of expected utility ($\bar{W}$):

$$\bar{W}^{1-\sigma} \frac{1}{\sigma F} \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} = \frac{w_n^{\sigma} A_n^{1-\sigma}}{\sum_{k \in N} d_{kn}^{1-\sigma} B_k^{\frac{\sigma-1}{\alpha\epsilon}} H_k^{\frac{(\sigma-1)(1-\alpha)}{\alpha}} w_k^{\sigma} L_k^{1-(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}. \tag{48}$$

where $\bar{W} = \left[\alpha^{\epsilon} \left(\frac{1-\alpha}{\alpha}\right)^{\epsilon(1-\alpha)} \left(\bar{U}/\delta\right)^{\epsilon} \left(\bar{L}\right)^{-1} \right]^{1/\alpha\epsilon}$, as shown in the web appendix.

Using the price index (42), we obtain a second system of equations for each location linking wages and populations of locations $n \in N$ as a function of parameters and the transformation of expected utility ($\bar{W}$):

$$\bar{W}^{1-\sigma} \frac{1}{\sigma F} \left(\frac{\sigma}{\sigma - 1} \right)^{1-\sigma} = \frac{w_n^{1-\sigma} B_n^{\frac{1-\sigma}{\alpha\epsilon}} H_n^{\frac{(1-\sigma)(1-\alpha)}{\alpha}} L_n^{-(1-\sigma)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)}}{\left[\sum_{k \in N} L_k \left(\frac{d_{nk} w_k}{A_k} \right)^{1-\sigma} \right]}. \tag{49}$$

Under the assumption that trade costs are symmetric ($d_{nk} = d_{kn}$), these two wage systems imply the following closed-form solution linking the endogenous variables for each location $n \in N$:

$$w_n^{1-2\sigma} A_n^{\sigma-1} B_n^{-\frac{\sigma-1}{\alpha\epsilon}} H_n^{-\frac{(\sigma-1)(1-\alpha)}{\alpha}} L_n^{(\sigma-1)\left(\frac{1}{\alpha\epsilon}+\frac{1-\alpha}{\alpha}\right)} = \kappa, \tag{50}$$

where κ is a scalar. If equation (50) holds, then any functions w_n and L_n that satisfy the system of equations (48) will also satisfy the system of equations (49) (and vice versa). In the proposition below, we prove that equation (50) is the unique relationship between w_n and L_n that satisfies both systems of equations. Substituting this relationship (50) into (49), we obtain the following system of equations that uniquely determines the equilibrium population of each location as a function of the parameters of the model:

$$L_n^{\tilde{\sigma}\gamma_1} A_n^{-\tilde{\sigma}(\sigma-1)} B_n^{-\frac{\tilde{\sigma}\sigma}{\alpha\epsilon}} H_n^{-\frac{\tilde{\sigma}\sigma(1-\alpha)}{\alpha}} = \bar{W}^{1-\sigma} \left[\sum_{k \in N} \frac{1}{\sigma F} \left(\frac{\sigma d_{nk}}{\sigma - 1} \right)^{1-\sigma} A_k^{\tilde{\sigma}\sigma} B_k^{\frac{\tilde{\sigma}(\sigma-1)}{\alpha\epsilon}} H_k^{\frac{\tilde{\sigma}(\sigma-1)(1-\alpha)}{\alpha}} \left(L_k^{\tilde{\sigma}\gamma_1} \right)^{\frac{\gamma_2}{\gamma_1}} \right], \tag{51}$$

$$\tilde{\sigma} = \frac{\sigma - 1}{2\sigma - 1},$$

$$\frac{1-\tilde{\alpha}}{\tilde{\alpha}} = \left(\frac{1}{\alpha\epsilon} + \frac{1-\alpha}{\alpha} \right), \qquad \tilde{\alpha} = \frac{\alpha}{1 + \frac{1}{\epsilon}},$$

$$\gamma_1 = \sigma \left(\frac{1-\tilde{\alpha}}{\tilde{\alpha}} \right),$$

$$\gamma_2 = 1 + \frac{\sigma}{\sigma - 1} - (\sigma - 1) \left(\frac{1-\tilde{\alpha}}{\tilde{\alpha}} \right),$$

where expected utility ($\bar{W}$) is implicitly determined by the requirement that the labor market clears across all locations: $\sum_{n \in N} L_n = \bar{L}$. The condition for there to exist a unique stable equilibrium is:

$$\sigma \left(1 - \tilde{\alpha}\right) > 1, \qquad \Leftrightarrow \qquad \frac{\gamma_2}{\gamma_1} < 1. \tag{52}$$

In the special case of the model in which there is no dispersion in idiosyncratic shocks to amenities ($\epsilon \to \infty$), this condition for a unique equilibrium reduces to the condition in the new economic geography model of Helpman (1998) for the case of two regions and perfect labor mobility of $\sigma(1-\alpha) > 1$.

Proposition 4 *Assume $\sigma(1-\tilde{\alpha}) > 1$. Given the land area, productivity and amenity parameters {H_n, A_n, B_n} and bilateral trade frictions {d_{ni}} for all locations $n, i \in N$, there exist unique equilibrium populations (L_n^*), trade shares (π_{ni}^*) and wages (w_n^*).*

Proof. See the web appendix. ■

Having determined unique equilibrium populations (L_n) from the system of equations (51), we can solve for equilibrium wages (w_n) as a function of populations from the closed-form solutions (50), and we can solve for equilibrium trade shares (π_{ni}) as a function of wages from the expenditure shares (41).

Intuitively, as population concentrates in a location, this expands the measure of varieties produced by that location, which in the presence of trade costs makes that location a more attractive residence (an agglomeration force). However, as population concentrates in a location, this also bids up land prices (a dispersion force). As long as the parameter inequality (52) is satisfied, the dispersion force dominates the agglomeration force, which ensures the existence of a unique equilibrium distribution of economic activity.

4.7 Comparative Statics

Despite the introduction of agglomeration forces in a setting with a large number of asymmetric locations, the model continues to admit closed-form expressions for the comparative statics of the endogenous variables with respect to the relative value of location characteristics. To characterize these comparative statics, we re-write the system of equations for equilibrium populations (51) as the following implicit function:

$$\begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_N \end{pmatrix} = \begin{pmatrix} \Omega_1^I \\ \vdots \\ \Omega_N^I \end{pmatrix} - \begin{pmatrix} \Omega_1^{II} \\ \vdots \\ \Omega_N^{II} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{53}$$

$$\begin{aligned}
\Omega_n^I &= L_n^{\tilde{\sigma}\gamma_1} A_n^{-\tilde{\sigma}(\sigma-1)} B_n^{-\frac{\tilde{\sigma}\sigma}{\alpha\epsilon}} H_n^{-\frac{\tilde{\sigma}\sigma(1-\alpha)}{\alpha}}, \\
\Omega_n^{II} &= \sum\nolimits_{k\in N} \Omega_{nk}^{II}, \\
\Omega_{nk}^{II} &= \bar{W}^{1-\sigma} \frac{1}{\sigma F} \left(\frac{\sigma d_{nk}}{\sigma-1}\right)^{1-\sigma} A_k^{\tilde{\sigma}\sigma} B_k^{\frac{\tilde{\sigma}(\sigma-1)}{\alpha\epsilon}} H_k^{\frac{\tilde{\sigma}(\sigma-1)(1-\alpha)}{\alpha}} \left(L_k^{\tilde{\sigma}\gamma_1}\right)^{\frac{\gamma_2}{\gamma_1}},
\end{aligned}$$

where Ω_n^{II} has an interpretation as a *market access* term that captures the goods market access of each location (depending on trade costs d_{nk}) to the characteristics of other locations.

The implicit function (Ω_n) is monotonically decreasing in the productivities, amenities and land supplies of all locations and monotonically increasing in the trade costs between all pairs of locations, as shown in the web appendix, where we report the closed-form solutions for each of these derivatives. An implication is that the equilibrium population of each location (L_n) depends solely on the *relative* rather than the absolute levels of these characteristics. The reason is that proportionate changes in the absolute level of these characteristics for all locations simply lead to a change in the common level of expected utility ($\bar{W}$), so as to ensure that the labor market clears, while leaving the relative levels of population across the locations unchanged. The implicit function (Ω_n) is also monotonically increasing in own population (L_n) and monotonically decreasing in the population of other locations (L_k for $k \neq n$). Therefore the system of equations for equilibrium populations (53) satisfies gross substitution and yields the following unambiguous comparative static predictions.

Proposition 5 *Assume $\sigma\,(1-\tilde{\alpha}) > 1$. A location n's equilibrium population (L_n) is increasing in its productivity (A_n), amenities (B_n) and land supply (H_n) relative to the values of these characteristics for all other locations $k \neq n$ and decreasing in its trade costs (d_{nk}) relative to the trade costs for all other locations $k \neq n$.*

Proof. See the web appendix. ■

In these comparative statics, the effects of changes in location characteristics on equilibrium populations are shaped by both goods and labor market frictions $\{\sigma - 1, \epsilon\}$. These parameters enter γ_1, γ_2 and Ω_{2nk} in the implicit function (53) and hence shape the sensitivity of equilibrium populations to changes in productivity (A_n), amenities (B_n), land supplies (H_n) and trade costs (d_{ni}). Intuitively, locations with higher productivity, more attractive amenities, larger land supplies and lower trade costs attract larger populations, but the trade elasticity $\sigma - 1$ and the population supply elasticity ϵ determine the sensitivity of equilibrium populations to differences in these characteristics.

4.8 Recovering Location Fundamentals

Given values for the model's parameters $\{\alpha,\,\theta,\,\epsilon\}$, a parameterization of bilateral trade costs $\{d_{ni}\}$ and data on populations, wages and land supplies $\{L_n,\,w_n,\,H_n\}$, we now show that the solution to the general equilibrium of the model again can be used to recover the unobserved location characteristics of amenities (B_n) and productivities (A_n).

Proposition 6 *Given the model parameters $\{\alpha,\,\sigma,\,\epsilon\}$, a parameterization of bilateral trade costs $\{d_{ni}\}$ and data on populations, wages and land supplies $\{L_n,\,w_n,\,H_n\}$, there exist unique values of amenities (B_n) and productivities (A_n) that are consistent with the data up to a normalization that corresponds to a choice of units in which to measure amenities and productivities.*

Proof. See the web appendix. ■

From Propositions 3 and 6, the constant and increasing returns models can be both calibrated to replicate the same data on populations, wages and land supplies $\{L_n,\,w_n,\,H_n\}$. In the constant returns model, the elasticity of trade with respect to variable trade costs is determined by the shape parameter of the productivity distribution ($\theta^N > \sigma^N - 1$), where the subscript N (neoclassical) indicates the constant returns model. In contrast, in the increasing returns model, the trade elasticity is dictated by the elasticity of substitution between variables ($\sigma^G - 1$), where the superscript G indicates the increasing returns to scale (new economic geography) model. Therefore calibrating both models to the same initial equilibrium and trade elasticities requires different structural parameters for the elasticity of substitution ($\sigma^G - 1 = \theta^N > \sigma^N - 1$). Furthermore, population directly affects the trade shares in the increasing returns model (46), but does not directly affect the trade shares in the constant returns model (16). Therefore calibrating both models to the same initial equilibrium also requires assuming different unobserved productivities in the two models, as summarized in the following proposition.

Proposition 7 *Given the parameters {α, ϵ} and $\theta^N = \sigma^G - 1$, and given a parameterization of bilateral trade costs {d_{ni}}, the constant returns model (superscript N) and increasing returns model (superscript G) both can be calibrated to the same data on populations, wages and land supplies {L_n, w_n, H_n} in an initial equilibrium. This calibration involves different structural parameters ($\sigma^N \neq \sigma^G$) and productivities ($A_n^N \neq A_n^G$) but the same amenities ($B_n^N = B_n^G$) in the two models.*

Proof. See the web appendix. ■

Given the different structural parameters and productivities, the constant and increasing returns models both rationalize the same data on the endogenous variables of the model as an equilibrium.

4.9 Counterfactuals

The system of equations for general equilibrium (45)-(47) again provides an approach for undertaking model-based counterfactuals that uses only parameters and the values of endogenous variables in an initial equilibrium. Denoting the relative value of variables in the counterfactual and initial equilibria by a hat ($\hat{x} = x'/x$), we can solve for the counterfactual effects of a change in trade costs, productivity or amenities using:

$$\hat{w}_i \hat{L}_i Y_i = \sum_{n \in N} \pi'_{ni} \hat{w}_n \hat{L}_n Y_n, \tag{54}$$

$$\pi'_{ni} = \frac{\pi_{ni} \left(\hat{d}_{ni} \hat{w}_i / \hat{A}_i \right)^{1-\sigma} \hat{L}_i}{\sum_{k \in N} \pi_{nk} \left(\hat{d}_{nk} \hat{w}_k / \hat{A}_k \right)^{1-\sigma} \hat{L}_k}, \tag{55}$$

$$\lambda'_n = \frac{\hat{B}_n \hat{A}_n^{\alpha\epsilon} \hat{\pi}_{nn}^{-\frac{\alpha\epsilon}{\sigma-1}} \hat{\lambda}_n^{-\left(\epsilon(1-\alpha) - \frac{\alpha\epsilon}{\sigma-1}\right)} \lambda_n}{\sum_{k \in N} \hat{B}_k \hat{A}_k^{\alpha\epsilon} \hat{\pi}_{kk}^{-\frac{\alpha\epsilon}{\sigma-1}} \hat{\lambda}_k^{-\left(\epsilon(1-\alpha) - \frac{\alpha\epsilon}{\sigma-1}\right)} \lambda_k}. \tag{56}$$

where $Y_i = w_i L_i$ again denotes labor income and $\lambda_n = L_n/\bar{L}$ again denotes the population share of each location in the initial equilibrium.

Comparing the counterfactual systems in the constant returns model ((23)-(25)) and the increasing returns model ((54)-(56)), the dependence of the measures of varieties on populations in the increasing returns model is reflected in both the trade shares (in the terms in $\hat{L}_i$ in (55)) and the residential choice probabilities (in the different exponents on $\hat{L}_i$ in (56) compared to (25)). This dependence of the measure of varieties on the endogenous populations of locations in the increasing returns model implies different counterfactual predictions for the impact of changes in trade costs from the constant returns model. These differences exist even if the two models are calibrated to the same initial equilibrium $\{w_n, L_n, \pi_{ni}\}$, the same trade elasticity ($\theta^N = \sigma^G - 1$), and the same values of the other model parameters.

Proposition 8 *Suppose that the constant and increasing returns to scale models are calibrated to the same data on populations, wages and land supplies $\{L_n, w_n, H_n\}$ in an initial equilibrium with the same trade elasticity $\theta^N = \sigma^G - 1$ and the same values of the other model parameters. Even when calibrated in this way, the two models imply different counterfactual predictions for the effects of a reduction in trade costs on population, wages, trade shares and welfare $\{L_n, w_n, \pi_{ni}, \bar{U}\}$.*

Proof. See the web appendix. ■

In an international trade context, in which population is immobile between locations, these two models imply the same counterfactual predictions for the effects of a reduction in trade costs on wages, trade shares and welfare (see Arkolakis, Costinot, and Rodriguez-Clare 2012).[12] In contrast, in a setting in which labor is imperfectly mobile across locations, the reallocation of population across locations in response to the reduction in trade costs leads to different counterfactual predictions in the two models.

4.10 Welfare Gains from Trade

We now examine the implications of the introduction of agglomeration forces for the welfare gains from trade. Using the residential choice probabilities (43), expected utility (44), income equals expenditure (12), land market clearing (14) and the goods price index (42), expected utility for each location can be re-written solely in terms of its domestic trade share and population and model parameters:

$$\bar{U} = \frac{\delta B_n^{\frac{1}{\epsilon}} A_n^{\alpha} \left(\frac{1}{\pi_{nn}}\right)^{\frac{\alpha}{\sigma-1}} H_n^{1-\alpha} L_n^{-\left(\frac{1}{\epsilon}+(1-\alpha)-\frac{\alpha}{\sigma-1}\right)}}{\alpha \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \left(\frac{\sigma}{\sigma-1}\right)^{\alpha} (\sigma F)^{\frac{\alpha}{\sigma-1}} \left(\bar{L}\right)^{-\frac{1}{\epsilon}}}, \tag{57}$$

where the condition for the existence of a unique equilibrium $\sigma\left(1-\tilde{\alpha}\right) > 1$ implies that the expected utility for each location is decreasing in its population ($\frac{1}{\epsilon} + (1-\alpha) > \frac{\alpha}{\sigma-1}$). The domestic trade share (π_{nn}), population (L_n), the trade elasticity ($\sigma - 1$), the population supply elasticity (ϵ) and the share of

[12] When labor is immobile between locations, both the constant and increasing returns models satisfy the macro restrictions in Arkolakis, Costinot, and Rodriguez-Clare (2012): balanced trade (R1); aggregate profits are a constant share of aggregate revenues (R2); and a CES import demand system with a constant elasticity of trade with respect to variable trade costs (R3). In contrast, when labor is mobile between locations, the import demand system no longer has a constant elasticity.

tradables in expenditure (α) are again sufficient statistics for the welfare gains from trade:

$$\frac{\bar{U}^T}{\bar{U}^A} = \left(\frac{1}{\pi_{nn}^T}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{L_n^A}{L_n^T}\right)^{\frac{1}{\epsilon}+(1-\alpha)-\frac{\alpha}{\sigma-1}}. \tag{58}$$

In this expression for the welfare gains from trade in the increasing returns model (58), the exponent on relative populations now has an additional term ($-\alpha/(\sigma-1)$) that captures the impact of population on the endogenous measure of varieties (absent in the constant returns model in the previous section). Furthermore, from Proposition 8, the two models have different counterfactual predictions for the effects of reductions in trade costs on domestic trade shares and populations, even when calibrated to the same initial equilibrium. Therefore the two models have different implications for the welfare gains from reductions in trade costs, as explored further below.

In the special case in which labor is perfectly *mobile* across locations (no dispersion or idiosyncratic utility, which corresponds to $\epsilon \to \infty$), the expression for the common level of welfare across locations (27) simplifies to become:

$$\bar{U} = \frac{A_n^\alpha \left(\frac{1}{\pi_{nn}}\right)^{\frac{\alpha}{\sigma-1}} H_n^{1-\alpha} L_n^{-\left((1-\alpha)-\frac{\alpha}{\sigma-1}\right)}}{\alpha \left(\frac{1-\alpha}{\alpha}\right)^{1-\alpha} \left(\frac{\sigma}{\sigma-1}\right)^{\alpha} (\sigma F)^{\frac{\alpha}{\sigma-1}}}. \tag{59}$$

Therefore the domestic trade share (π_{nn}), population (L_n), the trade elasticity ($\sigma - 1$), and the share of tradables in expenditure (α) are again sufficient statistics for the common welfare gains from trade:

$$\frac{\bar{U}^T}{\bar{U}^A} = \left(\frac{1}{\pi_{nn}^T}\right)^{\frac{\alpha}{\sigma-1}} \left(\frac{L_n^A}{L_n^T}\right)^{\left((1-\alpha)-\frac{\alpha}{\sigma-1}\right)}, \tag{60}$$

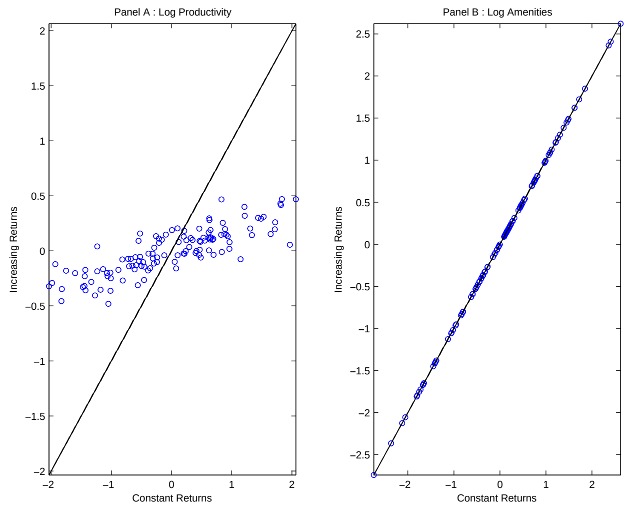
In contrast, in the opposite polar extreme of perfect labor *immobility*, the level of welfare and the welfare gains from trade are both in general different across locations. In this case, a location's domestic trade share (π_{nn}), the trade elasticity ($\sigma - 1$), and the share of tradables in expenditure (α) are sufficient statistics for its welfare gains from trade:

$$\frac{\bar{U}_n^T}{\bar{U}_n^A} = \left(\frac{1}{\pi_{nn}^T}\right)^{\frac{\alpha}{\sigma-1}} \neq \frac{\bar{U}_k^T}{\bar{U}_k^A}, \qquad n \neq k. \tag{61}$$

4.11 Quantitative Analysis

To examine the quantitative implications of introducing agglomeration forces, we return to the model economy in section 3, and calibrate the increasing returns model to the same initial equilibrium as the constant returns model. We assume the same values for the share of land in consumption expenditure ($1 - \alpha = 0.25$), the elasticity of population supply with respect to real income ($\epsilon = 4$), the elasticity of trade costs with respect to distance ($\phi = 1/3$), and the elasticity of trade with respect to trade costs in the two models ($\theta^N = \sigma^G - 1$). We begin with the values of population, wages and land supply $\{L_n, w_n, H_n\}$ from the initial equilibrium of the constant returns model. We first calibrate the model with increasing returns to exactly replicate these endogenous variables as an equilibrium (as in Proposition 7). Starting from this common initial equilibrium, we next examine the impact of the transport improvement shown in Figure 1 on the spatial distribution of economic activity (as in Proposition 8).[13]

Figure 8: Productivity and Amenties (Calibrated to Same Initial (w_n, L_n, H_n))



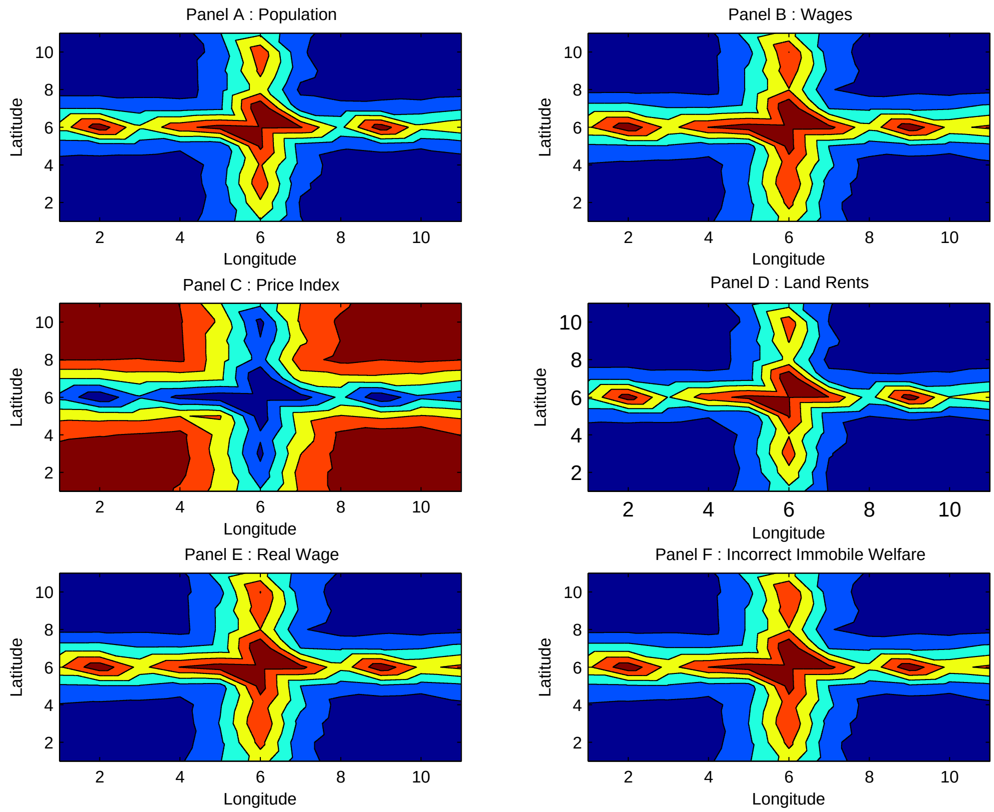
Note: Increasing returns model calibrated to the same initial equilibrium (wages, population and land area) as the constant returns model, with the same trade elasticity ($\theta = \sigma - 1$), migration elasticity (ϵ), land share ($1 - \alpha$) and elasticity of trade costs with respect to distance (ϕ), which implies different elasticities of substitution and initial productivities in the two models.

In Figure 8, we compare the calibrated values of productivity and amenities in the increasing returns and constant returns models required to rationalize the initial distribution of economic activity shown in Figure 3. We display the values in the increasing returns model on the vertical axis and the values in the constant returns model on the horizontal axis. As apparent from the figure, less dispersion in productivity is required in the increasing returns model to explain the same dispersion in economic activity across locations. This property reflects two features of the increasing returns model. First, productivity enters the trade shares (46) in the increasing returns to scale model with the exponent $\sigma - 1$ (compared to an exponent of one for the trade shares (16) in the constant returns model). Second, some of the concentration of economic activity in the increasing returns model is explained by agglomeration forces from an endogenous measure of varieties (leaving less to be explained by exogenous differences in productivity than in the constant returns model). As also apparent from the figure, the two models rationalize the common initial equilibrium with the same calibrated amenities, as shown formally in Proposition 7 above.

In Figure 9, we show the impact of the reduction in transport costs by displaying the relative change in each measure of economic activity ($\hat{x} = x'/x$) between the new equilibrium in the increasing returns model (denoted by a prime) and the common initial equilibrium. Again blue (cold) colors denote lower values and red (hot) colors denote higher values. We find a similar qualitative pattern for the effects of the transport improvement as in the constant returns model (comparing with Figure 9 in the constant

[13] An alternative approach is to calibrate the two models to have the same common structural parameters $\{\alpha, \sigma, \phi, \epsilon\}$ and location characteristics $\{A_n, B_n, H_n\}$, which implies different trade elasticities and different initial spatial distributions of economic activity in the two models. Also in this case, we find quantitatively relevant differences between the two models, with larger predicted impacts of the transport improvement in the increasing returns model than in the constant returns model.

Figure 9: Impact of Transport Improvement (Calibrated to Same Initial (w_n, L_n, H_n))



Note: Increasing returns model calibrated to the same initial equilibrium (wages, population and land area) as the constant returns model, with the same trade elasticity ($\theta^N = \sigma^G - 1 > \sigma^N - 1$) and the same values of other model parameters.

returns model). Locations close to the route of the road/railroad experience the largest increases in population (Panel A), wages (Panel B), land rents (Panel D), real wages (Panel E) and welfare under the (false) assumption of labor immobility (Panel F). These locations also experience the largest reductions in price indices (Panel C).

Although the qualitative pattern is similar, we find that the quantitative magnitudes are substantially larger in the increasing returns model than in the constant returns model. In Table 2, we report the results of estimating the long differences specification (34) for the transport infrastructure improvement in the increasing returns model. Consistent with the reorientation of the spatial distribution of economic activity shown in Figure 9, we find positive average treatment effects for population and wages, a negative average treatment effect for the price index, and positive average treatment effects for land rents, real wages and the incorrect measure of welfare based on the (false) assumption of labor immobility. However, the estimated magnitude of these average treatment effects is substantially and statistically significantly larger in the increasing returns model. For example, we find average treatment effects of around 60 percent for population (compared to around 40 percent in the constant returns model) and around 18 percent for wages (compared to around 8 percent in the constant returns model).

Again we find that lower levels of factor mobility imply larger treatment effects for wages, but smaller

Table 2: Treatment Effects of the Transport Improvement (Calibrated to Same Initial (w_n, L_n, H_n))

Economic Outcome	Treatment
Population	0.6177
	(0.0334)
Wage	0.1830
	(0.0099)
Price Index	-0.2288
	(0.0124)
Land Rents	0.8008
	(0.0432)
Real Wages	0.1544
	(0.0083)
Immobile Welfare	0.1931
	(0.0104)

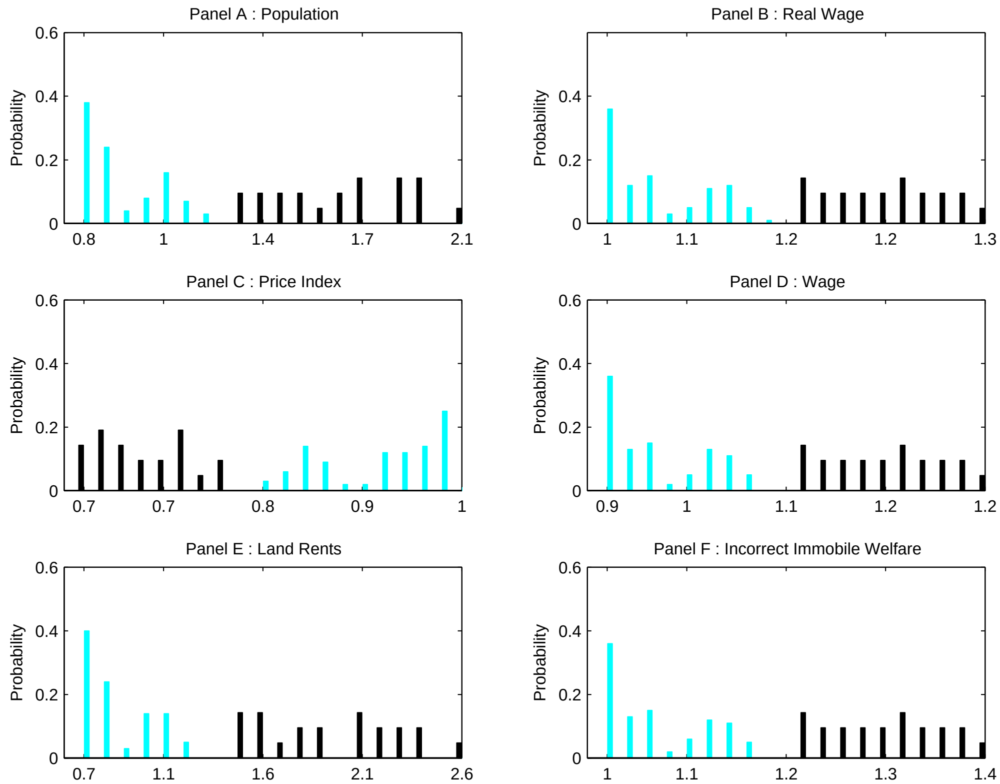
Note: Increasing returns model calibrated to the same initial equilibrium (wages, population and land area) as the constant returns to scale model, with the same trade elasticity ($\theta^N = \sigma^G - 1 > \sigma^N - 1$) and same values of other model parameters. Table reports the results of the estimating the long differences specification (34) for each economic outcome. A separate regression is estimated for each economic outcome. Standard errors in parentheses.

treatment effects for population and land prices. For example, for our baseline value of productivity dispersion ($\theta = 4$) and $\epsilon = 3.1$ (lower labor mobility), we find average treatment effects for population (0.5205), wages (0.1766) and land rents (0.6971). In contrast, for $\theta = 4$ and $\epsilon = 5.1$ (higher labor mobility), we find average treatment effects for population (0.7183), wages (0.1899) and land rents (0.9082).

As in our earlier analysis for the constant returns model, we again find that these average treatment effects mask considerable heterogeneity in the impact of the transport improvement among treated and untreated locations. In Figure 10, we illustrate this heterogeneity by displaying the distributions of the relative changes ($\hat{x} = x'/x$) in the economic outcomes as histograms across twenty equally-spaced bins. We again show the distributions for treated locations (in black) and untreated locations (in light blue). The largest increases in population (Panel A), wages (Panel D) and land rents (Panel E) are more than 100 percent, around 20 percent and more than 150 percent respectively in the increasing returns model, compared to less than 70 percent, around 10 percent and around 80 percent respectively in the constant returns model. We also find larger increases in real wages and in the measure of welfare based on the (false) assumption of labor immobility in the increasing returns model. The magnitudes of the differences for real wages and immobile welfare are smaller, which is consistent with our earlier findings of larger reallocation effects than welfare effects.

These findings of larger average treatment effects than in the constant returns model reflect the endogenous changes in the measures of varieties produced by each location in the increasing returns model. As population increases in treated locations, this expands the measure of varieties produced, which further increases the attractiveness of these treated locations. In contrast, as population declines in untreated locations, this reduces the measure of varieties produced, which further reduces the desirability of these

Figure 10: Impact of Transport Improvement (Calibrated to Same Initial Equilibrium (w_n, L_n, H_n))



Note: Increasing returns to scale model calibrated to the same initial equilibrium (wages, population and land area) as the constant returns to scale model, with the same trade elasticity ($\theta^N = \sigma^G - 1 > \sigma^N - 1$) and values of other model parameters.

untreated locations. As long as the stability condition $\sigma(1 - \tilde{\alpha}) > 1$ is satisfied, there is a unique equilibrium distribution of economic activity across locations, but these agglomeration forces magnify the impact of the transport improvement on the relative attractiveness of locations.

Although the reallocation effects of the transport improvement are larger than the welfare effects, we find that the differences in counterfactual predictions between the two models are also relevant for the evaluation of these welfare effects. We find that the transport improvement leads to a relative change in welfare of 1.1004 in the agglomeration model compared to 1.0925 in the constant returns model (a difference of around 8 percent). Again the true welfare effect of the transport improvement under imperfect factor mobility (1.1004 for all locations) differs substantially from the measured welfare effect under the (false) assumption of labor immobility, which ranges from 1-1.4 in Panel F of Figure 10.

5 Regions and Countries

In this section, we generalize the analysis further to introduce an additional form of imperfect labor mobility. We distinguish between regions and countries, where labor is *imperfectly mobile* across regions within countries, but is *completely immobile* between countries. We show that the general equilibrium of the model can be characterized and counterfactuals can be undertaken using a directly analogous approach to before. We examine the implications of imperfect factor mobility within countries for the measurement of countries' welfare gains from trade.

5.1 Preferences, Endowments and Technology

We consider a world economy consisting of many (potentially asymmetric) countries indexed by $j \in J$. We allow each country to consist of many (potentially asymmetric) regions indexed by $i, n \in N^j$, such that the world economy comprises $N = \{N^1, \ldots, N^J\}$ regions. Between countries, labor is completely immobile. Within countries, workers have heterogeneous tastes for regions, as modeled in the previous two sections. We first develop this extension of our baseline constant returns model from Section 2, but later report results for the same extension of our increasing returns model from Section 4.

5.2 General Equilibrium

The general equilibrium of the model again can be represented by the measure of workers (L_n), the trade share (π_{ni}) and the wage (w_n) for each location $n, i \in N^j$ and each country $j \in J$. Using labor income (13), the trade share (7), the price index (8), residential choice probabilities (10) and land market clearing (14), this equilibrium triple $\{L_n, \pi_{ni}, w_n\}$ solves the following system of equations for all $i, n \in N$. First, each location's income must equal expenditure on the goods produced in that location:

$$w_i L_i = \sum_{j \in J} \sum_{n \in N^j} \pi_{ni} w_n L_n. \tag{62}$$

Second, regional expenditure shares are:

$$\pi_{ni} = \frac{A_i \left(d_{ni} w_i\right)^{-\theta}}{\sum_{j \in J} \sum_{k \in N^j} A_k \left(d_{nk} w_k\right)^{-\theta}}. \tag{63}$$

Third, residential choice probabilities imply:

$$\frac{L_n}{\bar{L}^j} = \frac{B_n \left(\frac{A_n}{\pi_{nn}}\right)^{\frac{\alpha\epsilon}{\theta}} \left(\frac{L_n}{H_n}\right)^{-\epsilon(1-\alpha)}}{\sum_{k \in N^j} B_k \left(\frac{A_k}{\pi_{kk}}\right)^{\frac{\alpha\epsilon}{\theta}} \left(\frac{L_k}{H_k}\right)^{-\epsilon(1-\alpha)}}, \tag{64}$$

where the only difference is that the residential choice probabilities ($L_n/\bar{L}^j$) now apply country by country, so that the summation in the denominator of (64) is across regions k within a given country j.

5.3 Counterfactuals

The system of equations for general equilibrium (62)-(64) again provides an approach for undertaking model-based counterfactuals that uses only parameters and the values of endogenous variables in an initial equilibrium. Denoting the relative value of variables in the counterfactual and initial equilibria by a hat ($\hat{x} = x'/x$), we can solve for the counterfactual effects of a change in trade frictions, productivity or amenities using:

$$\hat{w}_i \hat{L}_i Y_i = \sum_{j \in J} \sum_{n \in N^j} \pi'_{ni} \hat{w}_n \hat{L}_n Y_n, \tag{65}$$

$$\pi'_{ni} = \frac{\pi_{ni} \hat{A}_i \left(\hat{d}_{ni} \hat{w}_i\right)^{-\theta}}{\sum_{j \in J} \sum_{k \in N^j} \pi_{nk} \hat{A}_k \left(\hat{d}_{nk} \hat{w}_k\right)^{-\theta}}, \tag{66}$$

$$\lambda_n^{j\prime} = \frac{\hat{B}_n \hat{A}_n^{\frac{\alpha\epsilon}{\theta}} \hat{\pi}_{nn}^{-\frac{\alpha\epsilon}{\theta}} \left(\hat{\lambda}_n^j\right)^{-\epsilon(1-\alpha)} \lambda_n^j}{\sum_{k \in N^j} \hat{B}_k \hat{A}_k^{\frac{\alpha\epsilon}{\theta}} \hat{\pi}_{kk}^{-\frac{\alpha\epsilon}{\theta}} \left(\hat{\lambda}_k^j\right)^{-\epsilon(1-\alpha)} \lambda_k^j}, \tag{67}$$

where $Y_i = w_i L_i$ denotes labor income and $\lambda_n^j = L_n / \bar{L}^j$ denotes the population share of each location $n \in N^j$ in the initial equilibrium in country j. Again the only difference from before is that the residential choice probabilities ($\lambda_n^j = L_n / \bar{L}^j$) apply country by country, so that the summation in the denominator of (67) is across regions k within a given country j.

5.4 Welfare Gains from Trade

As in the specification of the model without the distinction between regions and countries, the change in the domestic trade share (π_{nn}), change in population (L_n), trade elasticity (θ), population supply elasticity (ϵ) and consumption goods share (α) are sufficient statistics for the welfare gains from trade. Since the welfare gains from trade are a log linear function of the changes in domestic trade shares and populations, this result not only holds for each location n within a country j, but also holds for the geometric mean of locations within that country:

$$\frac{\bar{U}_n^{jT}}{\bar{U}_n^{jA}} = \frac{\bar{U}^{jT}}{\bar{U}^{jA}} = \left(\frac{1}{\tilde{\pi}_{nn}^T}\right)^{\frac{\alpha}{\theta}} \left(\frac{\tilde{L}_n^A}{\tilde{L}_n^T}\right)^{\frac{1}{\epsilon}+(1-\alpha)}, \qquad \forall\, n \in N^j, \tag{68}$$

where the tilde denotes a geometric mean, such that $\tilde{L}_n^A = \left[\prod_{n \in N^j} L_n^A\right]^{\frac{1}{|N|}}$.

We compare this true measure of the welfare gains from trade with two alternative measures. First, we consider a policy analyst who computes welfare for each region under the (false) assumption that labor is immobile across regions (equation (32) for each region). Second, we consider a policy analyst who aggregates regions within countries and computes welfare treating each country as a single location (equation (32) for each country). The country-level measure of the welfare gains from trade is thus:

$$\frac{\bar{U}^{jT}}{\bar{U}^{jA}} = \left(\frac{1}{\dot{\pi}^{jT}}\right)^{\frac{\alpha}{\theta}}, \tag{69}$$

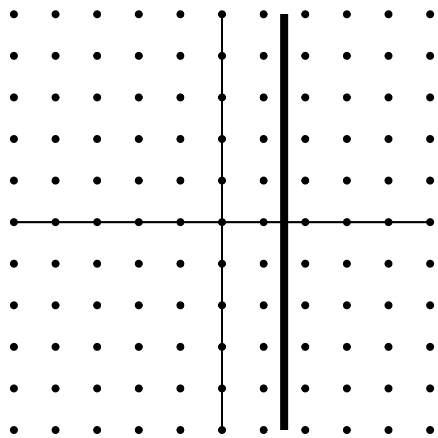
where the dot denotes a country-level variable, such that $\dot{\pi}^{jT}$ is country j's domestic trade share.

Comparing these two expressions, the aggregate domestic trade share for a country ($\dot{\pi}^{jT}$) in (69) involves a different aggregation of regional domestic trade shares from the geometric mean ($\tilde{\pi}_{nn}^T$) in (68). Furthermore, the ratio of the geometric means of regional populations in the closed and open economies ($\tilde{L}_n^A / \tilde{L}_n^T$) in (68) need not necessarily equal one. Therefore, the extent to which the country-level measure of the welfare gains from trade (69) approximates the true regional measure of the welfare gains from trade (68) will depend on whether the aggregate domestic trade share ($\dot{\pi}^{jT}$) is close to the geometric mean of regional domestic trade shares ($\tilde{\pi}_{nn}^T$) and the extent to which population reallocations across regions following the opening of trade change the geometric mean of regional populations ($\tilde{L}_n^A \neq \tilde{L}_n^T$).

5.5 Quantitative Analysis

To examine the quantitative implications of the distinction between regions and countries, we return to the model economy in section 3. We assume that the model economy consists of two countries (East and West), with the border between them shown by the thick vertical line in Figure 11. We suppose that population is distributed between the two countries in proportion to their share of the economy's total land area. We assume the same values for the trade elasticity ($\theta^N = \sigma^G - 1 > \sigma^N - 1$) and other model parameters in the constant and increasing returns models.

Figure 11: Model Economy with East and West

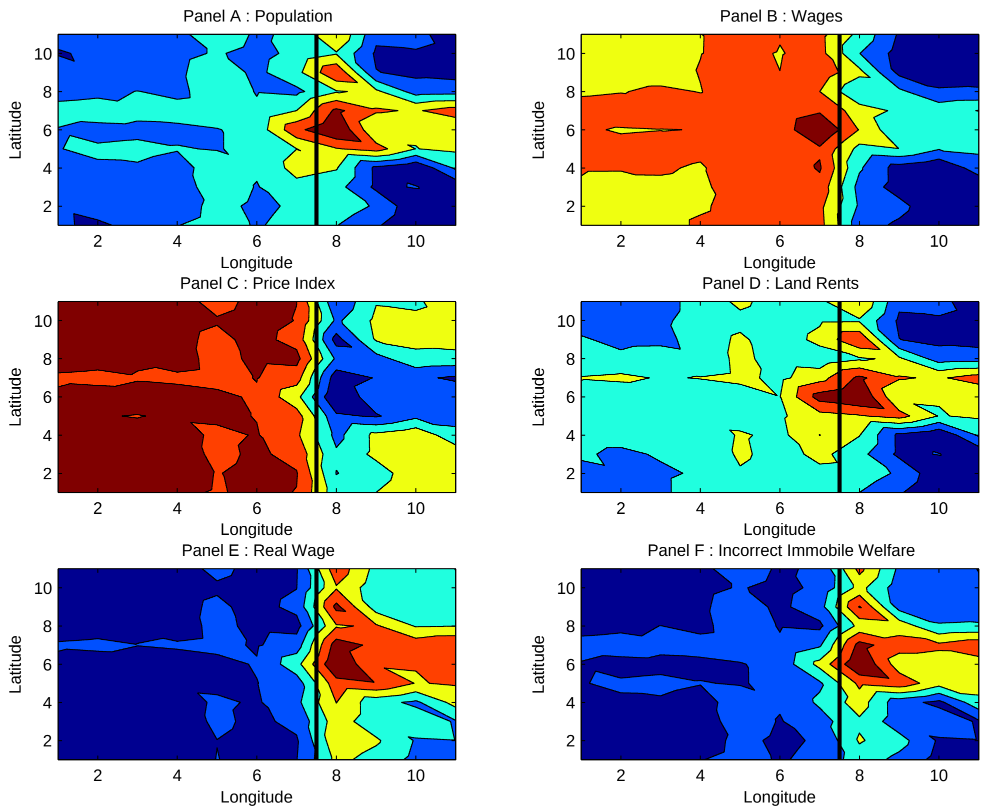


Note: Grid of locations in latitude and longitude space, the route of the transport infrastructure improvement (thin lines), and the border between East and West (thick lines).

We suppose that the two countries are initially closed to trade and solve for the equilibrium spatial distribution of economic activity in the constant returns model. We next calibrate the increasing returns model to exactly replicate this spatial distribution of economic activity as an equilibrium in the closed economy. We then open the closed economy to trade in both models and examine the implications for the spatial distribution of economic activity within countries and for expected utility in each country. We undertake two separate counterfactuals in both models. In a first exercise, we assume that the transport infrastructure shown by the thin lines in Figure 11 exists in the closed economy before the opening of trade (but is closed off at the border). In a second exercise, we assume that the transport infrastructure does not exist in the closed economy but is constructed in response to the opening of trade (e.g. the construction of road and rail networks to connect with ports and border crossings).

We begin with our first counterfactual, in which the transport infrastructure exists in the closed economy. In Figure 12, we show the impact of the opening of trade on the spatial distribution of economic activity in East and West in the constant returns model. We display the relative change in each measure of economic activity ($\hat{x} = x'/x$) between the open economy (denoted by a prime) and the closed economy. Again blue (cold) colors denote lower values and red (hot) colors denote higher values. The thick black vertical line denotes the border between East and West. We find a similar qualitative pattern for the impact of the opening of trade in the increasing returns model (but the quantitative magnitudes are again larger). Locations close to the opened border typically experience the largest increases in population (Panel A), wages (Panel B), land rents (Panel D), real wages (Panel E) and welfare under the (false) assumption of labor immobility (Panel F). These locations also typically experience the largest reductions in price indices (Panel C). Of all locations along the opened border, those closest to the route of the transport infrastructure experience the largest changes in levels of economic activity.

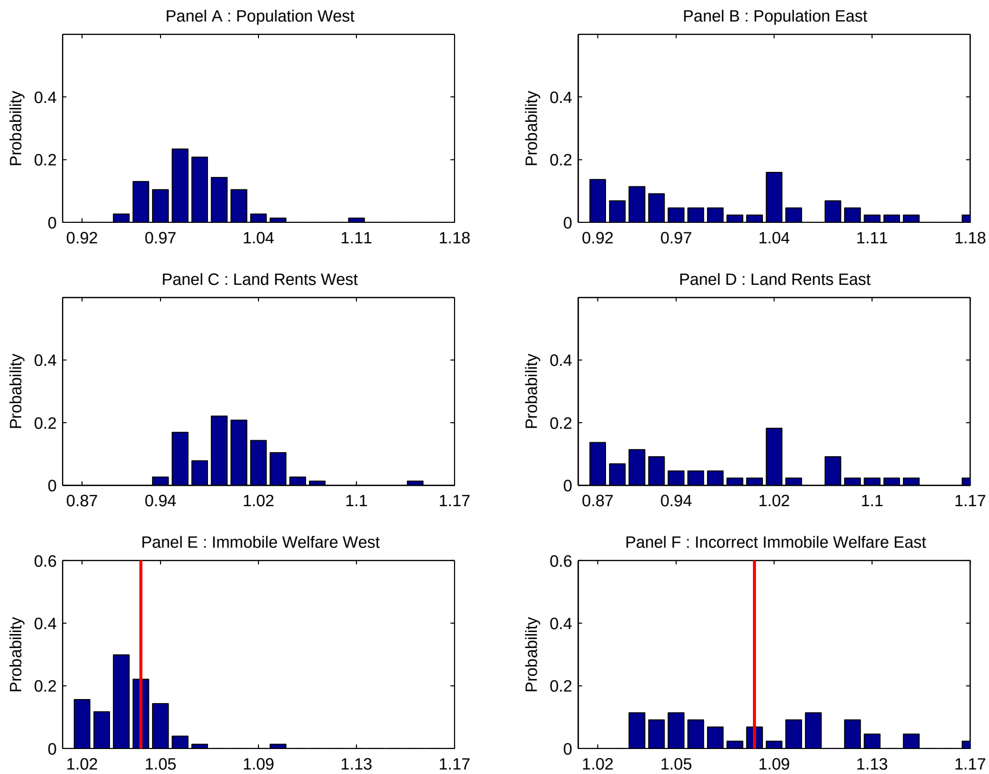
Figure 12: Impact of Opening of Trade Between East and West



Note: Results from the constant returns model under the assumption that the transport infrastructure exists in the closed economy and is unaffected by the opening of trade between East and West. Thick vertical line denotes the border.

In Figure 13, we show the distribution of relative changes in population (Panels A and B), land rents (Panels C and D), and welfare under the (false) assumption of labor immobility (Panels E and F). The left panels (A, C and E) show outcomes for West, while the right panels (B, D and F) show outcomes for East. We find that the opening of trade leads to substantial reallocations of economic activity within countries. Population changes range from reductions of 8 percent to rises of 18 percent, while land rent changes vary from reductions of 13 percent to rises of 17 percent. In general, the dispersion of population and land rent changes is larger in East than in West, which in part reflects its smaller size and relative scarcity of transport infrastructure (only one East-West line). Again we find that the true changes in welfare are smaller than the reallocation effects (1.0771 for East and 1.0384 for West). Although these true changes in welfare are the same across regions within each country (as shown by the vertical red lines in Panels E and F), a policy analyst who measured the welfare gains from trade at the regional level under the (false) assumption that labor is immobile across regions would incorrectly find substantial variation in welfare effects (ranging from 1.02 to 1.17 in Panels E and F).

Figure 13: Impact of Opening of Trade Between East and West



Note: Results from the constant returns model under the assumption that the transport infrastructure exists in the closed economy and is unaffected by the opening of trade between East and West.

In Table 3, we compare the true welfare gains from trade (68) to those measured by a policy analyst who aggregates regions to the country-level (69). The first two columns report results for the constant returns to scale model (for East and West). The second two columns report results for the increasing returns model (for East and West). In each column, Panel A reports results for our counterfactual in which the transport infrastructure exists in the closed economy (as shown in Figures 11 and 12). In contrast, Panel B reports results for our counterfactual in which the transport infrastructure does not exist in the closed economy and is constructed in response to the opening of trade.

Comparing the first two columns to the second two columns, we again find that the differences in counterfactual welfare effects between the constant and increasing returns to scale models are smaller than the differences in counterfactual reallocation effects. From Panel A, we find that the country-level measure of the welfare gains from trade provides a good approximation to the true welfare gains in our first counterfactual with constant transport infrastructure (and a much better approximation than the regional measure assuming factor immobility in Panels E and F of Figure 13). From Panel B, we find that the country-level measure of the welfare gains from trade provides a much less good approximation to

Table 3: True and Measured Welfare Effects of Trade Liberalization

	Constant Returns	Constant Returns	Increasing Returns	Increasing Returns
	West	East	West	East
Panel A: Constant Transport Infrastructure				
True Expected Utility	1.0384	1.0771	1.0387	1.0776
Measured Country Welfare	1.0376	1.0746	1.0382	1.0767
Panel B: Changing Transport Infrastructure				
True Expected Utility	1.1436	1.1076	1.1510	1.1150
Measured Country Welfare	1.0376	1.0746	1.0391	1.0792

Note: Results from the constant returns model with both constant and changing transport infrastructure.

the true welfare gains in our second counterfactual in which the opening of trade leads to a change in transport infrastructure. The reason is that the transport infrastructure that connects the two countries also facilities internal trade between regions within each country. These changes in internal trading opportunities between regions generate welfare gains that are not fully captured in each country's domestic trade share with itself.

6 Conclusions

We develop a quantitative spatial model that incorporates a rich geography of trade costs and imperfect labor mobility across locations. We allow locations to differ from one another in terms of their productivity, amenities and transport infrastructure. We provide micro foundations for imperfect labor mobility based on heterogeneity in worker tastes for each location. Despite the large number of asymmetric locations and the presence of goods and labor market frictions, the model remains highly tractable and amenable to both analytical and quantitative analysis. We provide general conditions for the existence and uniqueness of the spatial distribution of economic activity. We provide unambiguous comparative statics for the effect of location characteristics on population in that location and all other locations. We show that there is a one-to-one mapping from the model's parameters and data on wages, population, land area and trade costs to the unobserved characteristics of locations (productivity and amenities). Therefore the model can be inverted to recover these unobserved characteristics from the endogenous variables.

Our quantitative spatial model provides a useful complement to reduced-form regressions in analyzing the impact of transport infrastructure improvements. These reduced-form regressions abstract from general equilibrium effects and face the challenge of distinguishing between the creation and reallocation of economic activity. They also mask considerable heterogeneity in the impact of transport improvements within the groups of treated and untreated locations. Even if a location is not directly treated with transport infrastructure, it can indirectly benefit because of the resulting reduction in transport costs along the least cost route to other locations. We show that the average treatment effect of the transport improvement depends in a quantitatively relevant way on the degree of both goods and labor market frictions.

Given data before and after a transport improvement, we show how these different predictions for the impact on endogenous variables can be used to estimate the degree of goods and labor market frictions under the assumption of constant unobserved location characteristics.

In an international trade context, where labor is immobile across countries, the constant and increasing returns models considered in this paper have the same counterfactual predictions for the effects of reductions in trade costs. In contrast, when labor is imperfectly mobile across locations, we show that the two models have different counterfactual predictions for the impact of these reductions in trade costs, even when calibrated to the same initial equilibrium and trade elasticity. As trade costs fall, population reallocates across locations. In the increasing returns model, these population reallocations directly affect trade shares (whereas they do not in the constant returns to scale model), which leads to different counterfactual predictions in the two models for wages, trade shares, populations and welfare. We show that these differences in counterfactual predictions can be quantitatively relevant for empirically plausible reductions in trade costs.

We show that reallocations of economic activity across locations play an important role in understanding the welfare gains from trade. To the extent that some locations experience larger reductions in trade costs than others, populations reallocates to these locations, until the price of land adjusts such that all locations experience the same welfare gains from trade. We find that these reallocations of economic activity are large relative to the overall welfare gains from trade. Failing to take them into account when measuring welfare at the regional level can lead to large discrepancies from the true welfare gains from trade. In contrast, measuring the welfare gains from trade at the country-level provides a much closer approximation to the common welfare gains from trade across regions within each country, as long as the opening of trade does not lead to endogenous changes in transport infrastructure (e.g. the construction of ports, roads and railroads) that alter internal trading opportunities.

References

Allen, T., and C. Arkolakis (2014): "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, forthcoming.

Alvarez, F., and R. E. Lucas (2007): "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *Journal of Monetary Economics*, 54, 1726–1768.

Arkolakis, C., A. Costinot, and A. Rodriguez-Clare (2012): "New Trade Models, Same Old Gains," *American Economic Review*, 102(1), 94–130.

Artuc, E., S. Chaudhuri, and J. McLaren (2010): "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 100(3), 1008–1045.

Baum-Snow, N. (2007): "Did Highways Cause Suburbanization?," *Quarterly Journal of Economics*, 122(2), 775–780.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–1290.

Blanchard, O., and L. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 23(1), 1–76.

Bound, J., and H. J. Holzer (2000): "Demand Shifts, Population Adjustments and Labor Market Outcomes During the 1980s," *Journal of Labor Economics*, 18(1), 20–54.

Busso, M., J. Gregory, and P. Kline (2013): "Assessing the Incidence and Efficiency of a Prominent Place Based Policy," *American Economic Review*, 103(2), 897–947.

Caliendo, L., and F. Parro (2012): "Estimates of the Trade and Welfare Effects of NAFTA," *NBER Working Paper*, 18508.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2014): "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," NBER Working Paper.

Coşar, K., and P. Fajgelbaum (2013): "Internal Geography, International Trade, and Regional Specialization," *NBER Working Paper*, 19697.

Costinot, A., D. Donaldson, and I. Komunjer (2012): "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas," *Review of Economic Studies*, 79(2), 581–608.

Courant, P. N., and A. V. Deardorff (1992): "International Trade with Lumpy Countries," *Journal of Political Economy*, 100(1), 198–210.

——— (1993): "Amenities, Nontraded Goods, and the Trade of Lumpy Countries," *Journal of Urban Economics*, 34(2), 299–317.

Davis, D. R., and D. E. Weinstein (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, 92(5), 1269–1289.

Davis, M. A., and F. Ortalo-Magné (2011): "Housing Expenditures, Wages, Rents," *Review of Economic Dynamics*, 14(2), 248–261.

Dekle, R., J. Eaton, and S. Kortum (2007): "Unbalanced Trade," *American Economic Review*, 97(2), 351–355.

Desmet, K., D. K. Nagy, and E. Rossi-Hansberg (2014): "The Geography of Development: Evaluating Migration Restrictions and Coastal Flooding," Princeton University, mimeograph.

Desmet, K., and E. Rossi-Hansberg (2014): "Spatial Development," *American Economic Review*, forthcoming.

Donaldson, D. (2014): "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, forthcoming.

Duranton, G., and M. A. Turner (2012): "Urban Growth and Transportation," *Review of Economic Studies*, 79(4), 1407–1440.

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

Eaton, J., S. Kortum, B. Neiman, and J. Romalis (2011): "Trade and the Global Recession," *NBER Working Paper*, 16666.

Faber, B. (2014): "Integration and the Periphery: The Unintended Effects of New Highways in a Developing Country," *Review of Economic Studies*, forthcoming.

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

Fogel, R. (1964): *Railroads and American economic growth: Essays in econometric history.* Johns Hopkins University Press, Baltimore (MD).

Fujita, M., P. Krugman, and A. J. Venables (1999): *The Spatial Economy: Cities, Regions and International Trade.* MIT Press, Cambridge.

Grogger, J., and G. Hanson (2011): "Income Maximization and the Selection and Sorting of International Migrants," *Journal of Development Economics*, 95(1), 42–57.

Hanson, G. H. (1996): "Localization Economies, Vertical Organization, and Trade," *American Economic Review*, 86(5), 1266–1278.

——— (1997): "Increasing Returns, Trade, and the Regional Structure of Wages," *Economic Journal*, 107(1), 113–133.

——— (2005): "Market Potential, Increasing Returns, and Geographic Concentration," *Journal of International Economics*, 67(1), 1–24.

Head, K., and T. Mayer (2014): "Gravity Equations: Workhorse, Toolkit, and Cookbook," in *Handbook of International Economics*, ed. by E. Helpman, G. Gopinath, and K. Rogoff, vol. 4, pp. 131–196. Elsevier, North Holland, Amsterdam.

Head, K., and J. Ries (2001): "Increasing Returns versus National Product Differentiation as an Explanation for the Pattern of U.S.-Canada Trade," *American Economic Review*, 91, 858–876.

Helpman, E. (1998): "The Size of Regions," in *Topics in Public Economics: Theoretical and Applied Analysis*, ed. by D. Pines, E. Sadka, and I. Zilcha. Cambridge University Press, Cambridge.

Hsieh, C.-T., and R. Ossa (2011): "A Global View of Productivity Growth in China," *NBER Working Paper*, 16778.

Hummels, D. (2007): "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, 21(3), 131–154.

Jones, R. (1967): "International Capital Movements and the Theory of Tariffs and Trade," *Quarterly Journal of Economics*, 81, 1–38.

Kennan, J., and J. R. Walker (2011): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, 79(1), 211–251.

Krugman, P. (1991a): *Geography and Trade.* MIT Press, Cambridge.

——— (1991b): "Increasing Returns and Economic Geography," *Journal of Political Economy*, 99(3), 483–499.

Krugman, P., and A. J. Venables (1995): "Globalisation and Inequality of Nations," *Quarterly Journal of Economics*, 60, 857–80.

Limao, N., and A. J. Venables (2001): "Infrastructure, Geographical Disadvantage, Transport Costs and Trade," *World Bank Economic Review*, 15(3), 451–479.

Markusen, J. R. (1983): "Factor Movements and Commodity Trade as Complements," *Journal of International Economics*, 14(34), 341–356.

McFadden, D. (1974): "The Measurement of Urban Travel Demand," *Journal of Public Economics*, 3(4), 303–328.

Michaels, G. (2008): "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics*, 90(4), 683–701.

Moretti, E. (2011): "Local Labor Markets," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter, vol. 4b, pp. 1238–1303. Elsevier North Holland, Amsterdam.

Mundell, R. A. (1957): "International Trade and Factor Mobility," *American Economic Review*, 47, 321–337.

Ossa, R. (2011): "Trade Wars and Trade Talks with Data," *NBER Working Paper*, 17347.

Puga, D. (1999): "The Rise and Fall of Regional Inequalities," *European Economic Review*, 43, 303–334.

Ramondo, N., A. Rodríguez-Clare, and M. Saborio (2012): "Scale Effects and Productivity Across Countries: Does Country Size Matter?," *NBER Working Paper*, 18532.

Redding, S. J., and D. M. Sturm (2008): "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review*, 98(5), 1766–1797.

Redding, S. J., and A. J. Venables (2004): "Economic Geography and International Inequality," *Journal of International Economics*, 62(1), 53–82.

Rossi-Hansberg, E. (2005): "A Spatial Theory of Trade," *American Economic Review*, 95(5), 1464–1491.

Simonovska, I., and M. E. Waugh (2014): "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 92(1), 34–50.