

# Domestic Trade and Inequality

Farid Farrokhi<sup>1</sup> and David Jenkins<sup>\*2</sup>

<sup>1</sup>Penn State

<sup>2</sup>Copenhagen Business School

June 2, 2015

Inequality has long fascinated economists, and growing income inequality has been heatedly discussed in public forums recently.<sup>1</sup> In this paper, we add to this discussion by documenting some new facts about the interaction of geography with inequality, and developing a model which can rationalize what we find. After estimating the model on American data, we will discuss the effect of changes in geography on inequality. These changes might be a new highway, or population movement due to climate change.

Our first empirical contribution is to document facts about inequality and geography. We assign to each American city a measure of remoteness meant to capture its distance from all other cities. We then show that this measure correlates negatively with the skill premium, the ratio of the mean wage of the highly educated to the mean wage of the less educated. That is, highly educated workers make less relative to less educated workers in remote cities. On the other hand, conditional on population and other controls, the more remote a city is, the higher its share of highly educated workers, which we call the college share. As far as we know, our paper is the first to document these facts.

At least since the development of new economic geography, economists have examined the role of trade costs in determining the location of manufacturing centers **CITE KRUGMAN, allen arkolakis**. On the other hand, recent models of spatial inequality have treated cities as isolated. Workers in a city can interact with each other, but there is no interaction between cities. This literature, then, has no way to discuss the interaction of geography with inequality, since the geographic location of a city relative to other cities is irrelevant. By including intra-city trade in an model of inequality, we bridge this gap in the literature.

In our model, we have a continuum of locations, workers, and firms. Workers come in two types, skilled and unskilled, and each worker has an ideosyncratic

---

<sup>\*</sup>We thank Sina Smid for research assistance, and participants in seminars at Copenhagen Business School and Cardiff University for helpful suggestions.

<sup>1</sup>It is not every year that a 700-page collection of charts and academic theory on inequality is a New York Times bestseller.

utility from living in each location. A worker decides where to live taking wages as given. A firm also takes local wages as given, and produces a tradable good using skilled and unskilled labor as inputs. The sole difference between the two types of labor is that skill labor benefits from agglomeration, while unskilled labor does not.<sup>2</sup>

The key in our model is to generate skill premia which differ by location. In particular, we want less remote cities to have higher skill premia. Consider a city near other cities. The access to cheap tradable goods makes this location attractive to live in. This leads the city to have a relatively high population. Through agglomeration forces, skilled workers become relatively productive. Thus, a firm in this city chooses a relatively high share of skilled workers as inputs. The more skilled workers are attracted to the city, the more difficult it is to attract the marginal skilled worker, since she has a strong preference to live elsewhere. A high skill premium is necessary in this city to increase the relative supply and decrease the relative demand for skilled workers.

To clarify this mechanism, consider software engineers and gas station attendants. There are more software engineers per gas station attendant in Silicon Valley than in Minneapolis. Agglomeration effects make software engineers much more productive in Silicon Valley, but gas station attendants are almost the same productivity everywhere. Thus, Silicon Valley demands relatively more software engineers. In order to both increase Silicon Valley relative supply of software engineers and reduce relative demand, we need to make the software engineer wages in Silicon Valley relatively high compared with gas station attendants.

The literature on the skill premium has found a number of patterns. To the extent that we are able to measure the relevant quantities, all of these facts are consistent with our data. The literature has shown convincingly that the skill premium is higher in cities with larger populations (Davis and Dingel, 2012). The relationship between the skill premium and city size has become stronger over time (Baum-Snow and Pavan, 2013; Lindley and Machin, 2014). In addition, larger cities have been shown to have a higher share of college-educated worker, another pattern which has strengthened over time (Moretti, 2008; Lindley and Machin, 2014). Areas of denser economic activity have higher skill premiums (Combes et al., 2012).

These stylized facts have generated a number of theories to help understand them (Davis and Dingel, 2012, 2014; Baum-Snow and Pavan, 2012; Combes et al., 2012). These theories abstract from costly trade, treating trade between cities as either non-existent or frictionless.<sup>3</sup> The style of our modeling exercise below has more in common with recent forays of trade economists into economic geography (Allen and Arkolakis, 2014; Desmet et al., 2014). These theories model costly trade, and focus on the spatial location of economic activity. They have, however, only one type of labor, so they cannot analyze the interaction between geography and inequality. In an international trade context, Fujita and

<sup>2</sup>We allow there to be a Hicks-neutral agglomeration force as well.

<sup>3</sup>Davis and Dingel (2012) does contain an extension where trade costs are treated in the limit as they go to zero.

Thisse (2006) study inequality with costly trade, but their model has immobile unskilled workers.

**ADD THE PhD guy, and maybe a moretti paper or two more**

## 1 Documenting inequality and geography

In this section, we describe our data sources, give our definitions of measures of geography and inequality, and present the empirical findings which motivate our modeling exercise.

### 1.1 Data sources

Our empirical section is based on data from the 2000 American census public use micro-data (IPUMS) 5% sample.<sup>4</sup> In this cut of the data, we use individuals older than 24 and younger than 65 with reported income, giving us observations on over four million workers distributed across the United States.

We want to compare inequality in different locations. For our purposes, a location will be a US county. There are more than 3000 counties in the United States. In order to think about the geography of counties, we need to know the precise location of each county in space. We download a population weighted centroid for each county from the Missouri census data center (<http://mcdc2.missouri.edu/websas/geocorr2k.html>).

### 1.2 Important concepts

Our main measure of geography is remoteness, a concept we borrow from the international trade literature.<sup>5</sup> The particular functional form we use is based on a suggestion in Head (2003). We label each county with a number  $i = 1 \dots N$ . The population of county  $i$  is  $P_i$ , and the geodesic distance, i.e. the distance as the crow flies, between county  $i$  and county  $j$  is  $d_{i,j}$ . The remoteness of county  $i$   $R_i$  is given by:

$$R_i = \frac{1}{\sum_{j \neq i} \frac{P_j}{d_{i,j}}}$$

In words, an county which is close to other counties with high populations will have low remoteness. We want this measure to reflect the ease with which a

---

<sup>4</sup>We clean the data using modified replication code from Baum-Snow and Pavan (2013), which is available on Nathaniel Baum-Snow's website.

<sup>5</sup>It is a famous empirical regularity that the trade of two countries is proportional to the product of their national products divided by the physical distance between them. This relationship is known as the naive gravity equation. The adjective naive makes an appearance because such a gravity equation implies that trade between two countries is unaffected by what takes place in a third country. For example, the trade between the United States and Mexico is unaffected by the rapid increase in trade between China and the United States. Remoteness in an international trade context is the national-product-weighted sum of the distance between a country and all other countries. Multiplying the gravity equation by a remoteness term allows third countries to influence bilateral trade relationships.

Statistic	Mean	Standard Dev
Remoteness	$2.4 \times 10^{-6}$	$1.1 \times 10^{-6}$
Population	$2.3 \times 10^5$	$5.3 \times 10^5$
Skill premium	1.58	0.15
College share	0.37	0.26
Census observations	$4.3 \times 10^6$	
County observations	815	

Table 1: Data summary statistics

location can trade with other locations. This measure does not, of course, take into account geographic features such as rivers, mountains, roads, and airports, but physical distance is still an important part of transportation cost.

We have two measures of inequality, the college wage premium and the college population share. Both are measured in a way standard in the labor literature, and are calculated in each county. A worker is highly educated if he has a four-year college degree. The college premium is the mean wage of highly-educated workers in an county divided by the mean wage of other workers. The college population share is the population of highly-educated workers in an county divided by the population of less-educated workers. When calculating all means, we use census population weights.

Table 1 contains some descriptive statistics. Figure 1 shows how our measures vary across the United States. In each of these figures, the blue background is remoteness, with deeper blues signifying higher remoteness. Remoteness has an obvious geographic pattern, lowest on the Northern East and Southern West Coasts. Dark dots in the other figures signify counties with the highest levels of the variable listed under the figure. The college wage premium is higher on the East Coast and in the South than it is in the Midwest and the Pacific Northwest. It is more difficult to see an obvious geographic pattern in college share and population, except that both are higher in New England, the Pacific Northwest, and the coast of California.

### 1.3 Empirical results

In this section we document the covariance of our measure of geography, remoteness, with our measures of inequality, the college wage premium and college share. Regressions of the skill premium on remoteness and other variables are found in Table 2. Remoteness covaries negatively with the skill premium, with an elasticity of 5%. We find that the population of a county varies positively with the skill premium, a result that has been emphasized in other studies. That the coefficient on population loses significance with state fixed effects is a provocative result. The covariance of remoteness and the skill premium is the same order of magnitude as the covariance of remoteness with population found in the literature, but in the opposite direction.

In Table 3 we report coefficient estimates from regressions of college popu-

	(1)	(2)	(3)
VARIABLES	l_wage_prem	l_wage_prem	l_wage_prem
l_remote	-0.0562*** (0.00558)	-0.0367*** (0.0113)	-0.0310** (0.0125)
l_pop		0.00809** (0.00391)	0.00432 (0.00431)
Constant	-0.275*** (0.0728)	-0.113 (0.112)	0.0495 (0.123)
State FE	no	no	yes
Observations	815	815	815
$R^2$	0.119	0.124	0.307

Robust standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 2: Regressions of log skill premium on log geography variables

	(1)	(2)	(3)
VARIABLES	l_pop_share	l_pop_share	l_pop_share
l_remote	-0.348*** (0.0288)	-0.182*** (0.0536)	-0.287*** (0.0682)
l_pop		0.0692*** (0.0190)	0.0485** (0.0236)
Constant	-5.701*** (0.375)	-4.312*** (0.529)	-5.489*** (0.675)
State FE	no	no	yes
Observations	815	815	815
$R^2$	0.151	0.163	0.314

Robust standard errors in parentheses  
\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table 3: Regressions of log college share on log geography variables

lation ratio on the same regressors. We find that remoteness covaries positively with the college population ratio. Nearly everything said about the was before, the magnitude of the population effect and the remoteness effect are similar. The one difference is the population remains significant after including state fixed effects.

The takeaway from the empirical section is that we have identified a new relationship in the data. Relatively remote counties in the United States have relatively low college wage premia, and relatively low college population shares. As has been found in the previous literature we find that relatively high population counties also have relatively high college wage premia and relatively high college population shares. These two facts motivate the theory developed in the next section.

## 2 Theory

In the last section, we presented evidence that geography plays a role in shaping the college wage premium and college population shares. In this section we develop a theoretical model to help us understand the patterns we found: that the college wage premium and college population share both decrease with the remoteness of a location while the college share increases.

### 2.1 Environment and Technology

The model is static, with a continuum of locations  $J$ , a continuum of skilled workers  $L_H$ , and a continuum of unskilled workers  $L_L$ . Workers can choose to reside and work in any (single) location. Firms in each location can produce a location-specific variety of a tradable final good with a constant elasticity production function using the two types of labor as inputs. There is free entry of firms in each location. Both workers and firms are price takers.

### 2.2 Worker's problem

The utility of a worker in location  $i$  is:

$$Q(i)u(i)\epsilon(i) \tag{1}$$

where  $Q(i)$  is utility from tradeables,  $u(i)$  is the utility from exogenous amenities, and  $\epsilon^h(i)$  is worker  $h$ 's exogenous preference for location  $i$ . Tradeable goods are differentiated by the origin of production, and  $q(j, i)$  is consumer's consumption in  $i$  from goods originated in location  $j$ . The aggregator  $Q(i)$  in the utility function with elasticity of substitution  $\sigma > 0$  is:

$$Q(i) = \left[ \int_J q(j, i)^{\frac{\sigma-1}{\sigma}} dj \right]^{\frac{\sigma}{\sigma-1}},$$

We assume that worker location preferences  $\epsilon^h(i)$  are independent across workers and locations, and have a Type II Extreme Value distribution,

$$\Pr(\epsilon(i) \leq x) = \exp(-x^{-\theta})$$

Workers are endowed with exogenous skill  $s \in \{H, L\}$ . A worker's wage in location  $i$  is denoted by  $w_s(i)$ , and her budget constraint is

$$w_s(i) = \int_J p(j, i) q(j, i) dj \quad (2)$$

where  $p(j, i)$  is price of good  $j$  in destination  $i$ , and  $J$  is the given set of locations. The *only* place skill type appears in the worker's problem, is through the wage's effect on the budget constraint. Otherwise workers of both types have the same preferences.

An worker has two types of decision to make. She decides where to live, and how much of each good to consume. Given a choice of location, the second problem is standard. A worker of type  $s$  in location  $i$  spends  $x_s(j, i)$  on goods produced in  $j$ ,

$$x_s(j, i) \equiv p(j, i) q_s(j, i) = \left[ \frac{p(j, i)}{P(i)} \right]^{1-\sigma} w_s \quad (3)$$

where  $P(i)$  is the CES price index,

$$P(i) = \left[ \int_J p(j, i)^{1-\sigma} dj \right]^{\frac{1}{1-\sigma}} \quad (4)$$

The second decision a worker makes is where to live. In order to characterize this decision, we introduce congestion forces as in Allen and Arkolakis (2014):

$$u(i) = \bar{u}(i) n(i)^\gamma \quad (5)$$

Here,  $n(i)$  is total population in location  $i$  and  $\gamma < 0$  is the degree of congestion effect. This specification is isomorphic to a setting where preferences are Cobb-Douglas between tradeables and housing in which housing supply is inelastic. A worker with skill level  $s$  faces the following discrete choice problem over locations:

$$\max_{i \in J} \frac{w_s(i)}{P(i)} u(i) \epsilon(i)$$

Using the properties of the Type-II extreme value distribution, we can characterize the share of type- $s$  labor in location  $i$  as:<sup>6</sup>

---

<sup>6</sup>Incidentally this equation gives the model its first testable prediction:

$$\frac{\frac{w_H(i)}{w_L(i)}}{\frac{n_H(i)}{n_L(i)}^{\frac{1}{\theta}}} = \frac{\frac{w_H(j)}{w_L(j)}}{\frac{n_H(j)}{n_L(j)}^{\frac{1}{\theta}}}$$

A higher ratio of skilled workers to unskilled workers implies a higher skill premium in wages.

$$\frac{n_s(i)}{n_s(j)} = \left( \frac{w_s(i)u(i)/P(i)}{w_s(j)u(j)/P(j)} \right)^\theta \quad (6)$$

where  $n_s(i)$  is population of workers with type  $s$  in location  $i$ . The elasticity of relative labor supply to relative wages is:

$$\frac{\partial(n_s(i)/n_s(j)) / (n_s(i)/n_s(j))}{\partial(w_s(i)/w_s(j)) / (w_s(i)/w_s(j))} = \theta$$

The variance of  $\epsilon(i)$  across both workers and locations is decreasing in  $\theta$ . A large  $\theta$  implies that individual unobserved preferences for location are similar across locations. Small changes in wages, prices, or amenities induce large movements of workers. Another way of putting it is that the supply curve of workers to a location is flat. When  $\theta$  is small, workers have widely varying preferences over locations, so that large changes in wages, prices, or amenities are necessary to induce movement. That is, the supply curve of workers to a location is steep.

There is a lot of intuition for the way the model works here already. A location with a higher relative wage for skilled workers will attract more workers, but it won't attract all the skilled workers. Small changes in relative wages will induce small changes in relative populations.<sup>7</sup>

Define the real-wage index  $W_s$ :

$$W_s \equiv \frac{1}{N_s} \int_{j=0}^{\|J\|} (w_s(j)u(j)/P(j))^\theta dj$$

Here  $N_s$  is the exogenously given total population of type- $s$  workers. This index is a per-capita measure of welfare for workers of type  $s$ . Again using properties of the Type-II extreme value distribution, we derive the relationship for all  $i \in J$ :

$$W_s = \left( \frac{w_s(i)}{P(i)} u(i) \right)^\theta \frac{1}{n_s(i)} \quad (7)$$

That is, this relationship between real wages and population is constant across locations.

## 2.3 Labor Demand

Each location has a representative firm with a CES production function using skilled and unskilled workers,

$$A(i) \left[ \beta_H(i) n_H(i)^{\frac{\epsilon-1}{\epsilon}} + \beta_L(i) n_L(i)^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}}$$

---

<sup>7</sup>In a model without preference heterogeneity, in equilibrium all populated locations must offer the same utility within skill type. This means that a location which deviates and offers a slightly higher wage will attract *all* workers.



where  $A(i)$  is total factor productivity in location  $i$ .  $\varepsilon > 0$  is the elasticity of substitution between high and low skill workers.  $\beta_H(i) > 0$  and  $\beta_L(i) > 0$  are factor intensities. In the model, differences in factor productivities (and agglomeration) are the *only* things which separate skilled and unskilled labor. Other models in the inequality literature, such as **PHD guy, who else?**, build in more fundamental differences in skilled and unskilled preferences or taste for amenities. By cost minimization, the unit cost of production equals

$$\frac{c(i)}{A(i)}, \text{ where } c(i) = \left[ \beta_H(i)^\varepsilon w_H(i)^{1-\varepsilon} + \beta_L(i)^\varepsilon w_L(i)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \quad (8)$$

Share of spending of producers on high skill workers, denoted by  $b(i)$ , is given by

$$b(i) = \frac{\beta_H(i)^\varepsilon w_H(i)^{1-\varepsilon}}{\beta_H(i)^\varepsilon w_H(i)^{1-\varepsilon} + \beta_L(i)^\varepsilon w_L(i)^{1-\varepsilon}} \quad (9)$$

Following the literature, we distinguish two sorts of agglomeration forces. First

$$A(i) = \bar{A}(i)n(i)^\alpha \quad (10)$$

with  $\alpha > 0$ . This agglomeration force changes productivity of both low and high skill workers. A Krugman type monopolistic competition generates similar force through endogenous measure of firms. In addition, there is overwhelming evidence that agglomeration forces are stronger for high skill workers **CITATION**, suggesting that high skill individuals are more productive through exchanging ideas with each other. Motivated by these findings, we let a high skill individual be more productive through concentration of high skill workers,

$$\begin{aligned} \beta_H(i) &= \bar{\beta}_H(i)n_H(i)^\varphi \\ \beta_L(i) &= \bar{\beta}_L(i) \end{aligned} \quad (11)$$

As was suggested above, the intuition for the mechanisms in the model is now available. Larger cities will have more productive skilled workers, but unskilled workers will be exactly as productive everywhere. Firms in relatively large cities will therefore demand a relatively high share of skilled workers. In order to attract these workers, the location will need a relatively high skill premium.

Finally, markets are perfectly competitive. Let  $d(j, i)$  be the trade costs of shipping a good from  $j$  to  $i$ . Price of a good produced in location  $j$  and consumed in location  $i$ ,  $p(j, i)$ , equals  $c(j)d(j, i)/A(j)$ .

## 2.4 Equilibrium Definition

A *spatial equilibrium* is a set of  $w_H(i)$ ,  $w_L(i)$ ,  $n_H(i)$ , and  $n_L(i)$  such that:<sup>8</sup>

---

<sup>8</sup>There is a bit hidden in this definition, as we have already used perfect competition to give us goods prices  $p(j, i)$  given wages, the consumer's optimization problem to give us demand  $x(j, i)$  given prices and wages, and the firm's optimization problem to give us relative labor demand (9)

1. Labor shares in locations are according to (7).
2. Demand and supply of skilled and unskilled labor equal in each location.
3. Goods market clear

$$w_H(i)n_H(i) = b(i) \int_J \sum_{s \in \{H,L\}} n_s(i)x_s(j,i) dj \quad (12)$$

where  $x_s(j,i)$  is given by (3), and  $b(i)$  is given by (9).

4. Feasible allocation of labor,  $\int_J n_H(j) dj = N_H$  and  $\int_J n_L(j) dj = N_L$ .

## 2.5 Equilibrium Analysis

Before we derive a small number of equations characterizing equilibrium, we first discuss features of the model equilibrium. In particular, we show that without both heterogeneous location preferences as well as agglomeration forces, our model would not generate skill premia location specific skill premia in wages. However, we need no other type of skill specific preferences or productivity to generate this result.

Perfect competition gives us that the income share of skilled labor equals firm's optimal input share on skilled labor:

$$\frac{w_H(i)n_H(i)}{w_H(i)n_H(i) + w_L(i)n_L(i)} = \frac{\beta_H(i)^\varepsilon w_H(i)^{1-\varepsilon}}{\beta_H(i)^\varepsilon w_H(i)^{1-\varepsilon} + \beta_L(i)^\varepsilon w_L(i)^{1-\varepsilon}}$$

With some algebra, and substituting in agglomeration, we get:

$$\omega(i) \equiv \frac{w_H(i)}{w_L(i)} = \frac{\bar{\beta}_H(i)}{\bar{\beta}_L(i)} \left[ \frac{n_H(i)}{n_L(i)} \right]^{-1/\varepsilon} n_H(i)^{\varphi/\varepsilon} \quad (13)$$

where  $\omega(i)$  is the skill premium in location  $i$ .

Using equation (7), the supply side of labor markets imply:<sup>9</sup>

$$\omega(i) = \left[ \frac{W_H}{W_L} \frac{n_H(i)}{n_L(i)} \right]^{\frac{1}{\theta}} \quad (14)$$

Labor markets clear when skill premia,  $\{\omega(i)\}_{i \in J}$ , satisfy the pairs of demand by (13) and supply by (14):

$$n_L(i) = \left( \left( \frac{W_H}{W_L} \right)^{\frac{1}{\theta}} \frac{\bar{\beta}_L(i)}{\bar{\beta}_H(i)} \right)^{\frac{\theta\varepsilon}{\theta+\varepsilon}} n_H(i)^{1-\frac{\theta\varepsilon}{\theta+\varepsilon}} \quad (15)$$

---

<sup>9</sup>As far as we are aware, other models of inequality in the literature assume fundamental differences in preferences between skilled and unskilled workers. To derive (14) we are using the fact that skilled and unskilled workers face the same amenities, as well as the same prices.

The share of skilled population  $n_H/n_L$  is constant across locations if skilled agglomeration is zero,  $\varphi = 0$ . From (14) we see that if the share of skilled population is constant everywhere, so is the skill premium. Also from (14), we see that if  $\theta \rightarrow \infty$  the skill premium is constant across (populated) locations.<sup>10</sup> In order to have equilibria with varying skill premia, we need *both*  $\theta$  to be finite and  $\varphi > 0$ . In words, we need both variance in unobserved location preferences and the agglomeration force for skilled workers. Large cities demand more skilled workers due to agglomeration, but when unobserved preferences matter, skilled workers do not fully adjust to a wage increase as some of them are very attached to a particular city. In this respect, our model departs from a standard Rosen-Roback (or Allen-Arkolakis) model by highlighting the interaction of agglomeration and the labor supply elasticity.

## 2.6 Equilibrium equations

In this section we characterize our equilibrium with only two integral equations. These integral equations will be the basis of our empirical exercise.

By simply inverting the definition (9) of input share  $b(i)$ , we get total income in  $i$  equal to  $\frac{1}{b(i)} w_H(i) n_H(i)$ . Using the goods market clearing condition, equation (12), we derive:

$$w_H(i) n_H(i) b(i)^{-1} = \int_J \left[ \frac{c(i) d(i, j)}{A(i) P(j)} \right]^{1-\sigma} w_H(j) n_H(j) b(j)^{-1} dj$$

Equations (8) and (9) further imply that

$$\begin{aligned} c(i) &= \tilde{c}(i) w_H(i) \\ \text{where } \tilde{c}(i) &= \left[ \bar{\beta}_H(i) n_H(i)^\varphi \right]^{\frac{\epsilon}{1-\epsilon}} \left[ b(i) \right]^{\frac{-1}{1-\epsilon}} \end{aligned} \quad (16)$$

Replacing  $c(i)$  with  $\tilde{c}(i)$  in the above integral equation and also replacing the price index using (7), after some algebra we get:

$$\begin{aligned} A(i)^{1-\sigma} \tilde{c}(i)^{\sigma-1} n_H(i) w_H(i)^\sigma b(i)^{-1} \\ = W_H^{\frac{1-\sigma}{\theta}} \int_J d(i, j)^{1-\sigma} u(j)^{\sigma-1} n_H(j)^{(\theta+1-\sigma)/\theta} w_H(j)^\sigma b(j)^{-1} dj \end{aligned} \quad (17)$$

Again substituting the price index from (7) and  $\tilde{c}(i)$  into (4), we get another integral equation:

$$\begin{aligned} u(i)^{1-\sigma} n_H(i)^{(\sigma-1)/\theta} w_H(i)^{1-\sigma} = \\ W_H^{\frac{1-\sigma}{\theta}} \int_J d(j, i)^{1-\sigma} A(j)^{\sigma-1} \tilde{c}(j)^{1-\sigma} w_H(j)^{1-\sigma} dj \end{aligned} \quad (18)$$

---

<sup>10</sup>More intuition can be gleaned from (15). If  $\theta \rightarrow 0$ , we see that the share of skilled population will also be constant, as no one cares about wages, only their location preference shocks which are symmetric.

Equations 17–18 give us two systems of integral equations. If  $\varphi = 0$  and  $\theta = \infty$ , the integral equations collapse to those in Allen and Arkolakis (2014).

We can reduce our system further using a trick from Allen and Arkolakis (2014). If either of the systems of integral equation hold allong with the following relation, both systems of integral equations must hold:

$$\frac{w_H(i)^\sigma n_H(i) A(i)^{1-\sigma} \tilde{c}(i)^{\sigma-1}}{b(i)} = \lambda w_H(i)^{1-\sigma} u(i)^{1-\sigma} n_H(i)^{(\sigma-1)/\theta} \quad (19)$$

where  $\lambda > 0$  is some constant.

## 2.7 Solution Algorithm

1. Guess  $n_H(i)$  for all  $i$ .
2. Compute  $W_H/W_L$  according to (??). Then plug it in (15) to find  $n_L(i)$ .
3. Calculate skill premia,  $\omega(i)$ , according to (14).
4. Compute  $b(i) = 1/(1 + \frac{n_L(i)}{n_H(i)} \frac{1}{\omega(i)})$
5. Find  $\tilde{c}(i)$  according to (16).
6. Calculate  $w_H(i)$  according to (19) (up to a scale imposed by  $\lambda$ )
7. Let  $f(i) = w_H(i)^{1-\sigma}$ ,  $\kappa = W_H^{1-\sigma}$ , and

$$K(j, i) = n_H(i)^{(1-\sigma)/\theta} u(i)^{\sigma-1} d(j, i)^{1-\sigma} A(j)^{\sigma-1} \tilde{c}(j)^{1-\sigma}$$

Then, rewrite (18) as follows:

$$f(i) = \kappa \int_J K(j, i) f(j) \, dj$$

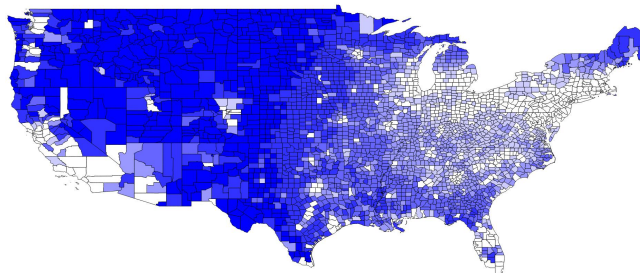
In iteration  $t$ , update  $f^{(t)}(i)$  according to

$$f^{(t+1)}(i) = \frac{\int_J K(j, i) f^{(t)}(j) \, dj}{\int_J \int_J K(j, i) f^{(t)}(j) \, dj \, di} \quad (20)$$

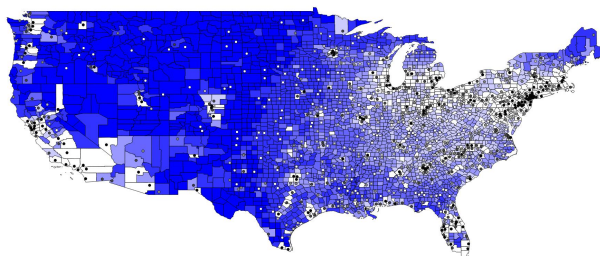
8. Find  $\lambda$  such that  $N_H = \int_J n_H(i) \, di$ .

## References

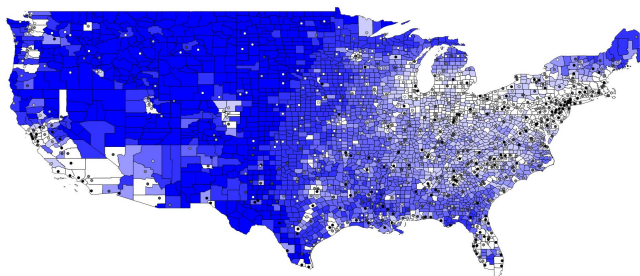
- Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.
- Baum-Snow, N. and Pavan, R. (2012). Understanding the city size wage gap. *The Review of economic studies*, 79(1):88–127.
- Baum-Snow, N. and Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, 95(5):1535–1548.
- Combes, P.-P., Duranton, G., Gobillon, L., and Roux, S. (2012). Sorting and local wage and skill distributions in france. *Regional Science and Urban Economics*, 42(6):913–930.
- Davis, D. R. and Dingel, J. I. (2012). A spatial knowledge economy. Technical report, National Bureau of Economic Research.
- Davis, D. R. and Dingel, J. I. (2014). The comparative advantage of cities. Technical report, National Bureau of Economic Research.
- Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2014). The geography of development: Evaluating migration restrictions and coastal flooding.
- Fujita, M. and Thisse, J.-F. (2006). Globalization and the evolution of the supply chain: Who gains and who loses? *International Economic Review*, 47(3):811–836.
- Head, K. (2003). Gravity for beginners. *University of British Columbia*, 2053.
- Lindley, J. and Machin, S. (2014). Spatial changes in labour market inequality. *Journal of Urban Economics*, 79:121–138.
- Moretti, E. (2008). Real wage inequality. Technical report, National Bureau of Economic Research.



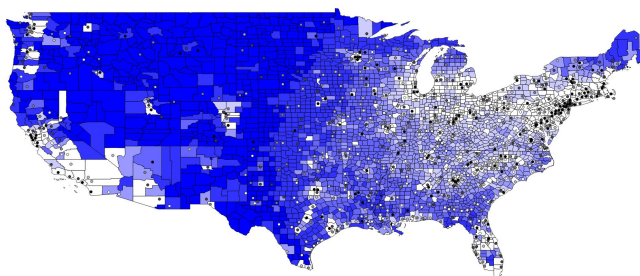
(a) Remoteness



(b) Population



(c) Skill Premium



(d) College Share

Figure 1: Counties by data feature