Bridges*

Roc Armenter, Miklós Koren and Dávid Krisztián Nagy†

May 6, 2013

Abstract

We build a continuous-space theory of trade in which people in a region agglomerate to exploit trading opportunities with another region. The regions are separated by a river, which can be crossed anywhere, but more cheaply at bridges. In the model, most trade takes place via bridges, leading to a key prediction that population density declines with distance to the bridge. We derive additional predictions about the spatial distribution of population and test them on high-resolution population density data around six major American rivers. The data are mostly consistent with our model. More generally, our results suggest that economies of density arising from transport infrastructure can help explain why and where people agglomerate.

*We thank Treb Allen, Costas Arkolakis, Jon Vogel and Esteban Rossi-Hansberg for useful comments and Bálint Menyhért for excellent research assistance. Koren acknowledges funding from European Research Council Starting Grant No. 313164. The paper does not represent the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

†Armenter: Federal Reserve Bank of Philadelphia. Koren: Central European University, IECERS and CEPR. Nagy: Princeton University.

1 Introduction

People agglomerate in cities to exploit spatial externalities and other economies of density. What are the source of economies of density? Where do cities emerge in space? Some locations can be clearly linked to natural advantages, whereas others seem to be the outcome of historical accidents.

We build a continuous-space theory of trade to explain why and where people agglomerate. There are two forces of agglomeration in our model. First, people move close to trading opportunities to minimize transportation cost. Second, they strive to exploit economies of density in transportation technology. To understand the second motive, consider choosing a location next to a river dividing two productive regions. If the river is easily navigable, boats may provide a suitable means of transport between the regions. As economies of scale in boating are small, traders have no incentive to agglomerate and can trade from small villages along the river. By contrast, if the river is less navigable, one has to build a bridge to cross it. Bridges bring about clear economies of density as locations close to the bridge will have lower trade costs with the other side. Traders agglomerate near the bridge, and a trading city emerges.

In our model, people choose their location on a homogeneous plane divided by a single linear feature (a "river"). The two sides of the river differ in comparative advantage, providing an incentive to trade across the river. Trading, however, is costly. The cost increases in the distance travelled, and crossing the river entails additional costs. The river can be crossed in two ways: by boat at any point, and on existing bridges for a lower cost. For a given set of bridges, we study the patterns of specialization and the distribution of population (and economic activity) in space.

Our model leads to a number of equilibrium predictions about rivers, bridges, and population density. First, population density declines with distance to the river. Second, along the river, population density declines with distance to closest bridge. Third, population is more clustered along the river than far from the river. Fourth, along the river, the population densities on the right and the left bank are positively correlated. Fifth, this correlation is smaller within the neighborhood of bridges.

Our work is motivated by the historical relevance of crossing points, often referenced in city names, such as Ox*ford* and Cam*bridge*. Although the emergence of such crossing points is not exogenous, they may lead to further and faster economic development and agglomeration. Writing about the Upper Black Eddy–Milford bridge on the Delaware, built in 1842, Dale (2003, p. 43) concludes that "[t]his new crossing brought additional business to this part of the river valley. It gave farmers and small industrualists in the area quick access to the Delaware Canal in Pennsylvania. And this increased use brought additional funds in the form of dividends to the stockholders of the Upper Black Eddy-Milford Bridge. By 1844, business in the now growing town of Milford included three stores, three taverns, twelve to fiteen mechanics' shops, a flour mill, and two new sawmills that made lumber trade, here, an especially important business. The town also had many non-commercial structures, including forty-five homes, two churches, and a fine school. Upper Black Eddy on the Pennsylvania side of the river directly opposite Milford was a favorite stop for timber raftsmen in the early days. By the mid-nineteenth century the bridge brought even more business. Upper Black Eddy was booming, too. It had forty houses, three hotels, and several stores and shops."

To evaluate the model more systematically, we test its predictions on six major North American rivers: the Hudson, the Delaware, the Mississippi, the Missouri, the Ohio, and the Tennessee. In doing so, we rely on high-resolution population density data, the precise path of rivers and the locations of bridges.

First, we estimate how population density varies with distance to the river and distance to the nearest bridge. It declines with distance to the Hudson, the Mississippi, the Missouri, the Ohio, and the Tennessee, but not with distance to the Delaware. Except for the Hudson, population density declines with distance to the nearest bridge on the other five rivers.

Second, we check whether population is more clustered at the river. We calculate the coefficient of variation of population density within 10 miles of the river, and find that it is higher than between 20 and 30 miles from the river. That is, there is more variation in population along the river than inland, as predicted by the model.

Third, we measure the correlation of population densities between the left and the right bank with and without taking bridges into account. For all six rivers, the correlation between the two banks is strongly positive (with an average of 0.48), suggesting that people agglomerate near the same points on either side of the river. When looking at bridges only, however, we find that correlations are substantially lower (average 0.37) between the two sides of the bridge. This is consistent with the model, where population density is a decreasing function of trade costs. Moving away from a bridge along the river, trade costs increase both on the left and on the right bank of the river, leading to a comovement in population density across the two banks. Therefore, starting from a bridge, the longer the interval over which we calculate the correlation coefficient, the larger value we find.

Our paper is related to two strands of the literature. First, it is related to an increasing number of papers which model space as ordered and continuous – a much more realistic assumption than the ones used in classical economic geography models. Rossi-Hansberg (2005), Desmet and Rossi-Hansberg (2012), and Coşar and Fajgelbaum (2012) characterize the spatial distribution of economic activity over a line segment in Ricardian models with agglomeration externalities. Fabinger (2011) and Allen and Arkolakis (2013), on the other hand, examine the implications of neoclassical models with CES preferences on the geographic distribution of economic activity. All of these papers come to the conclusion that *lower-dimensional trade barriers and trade infrastructure* – ports in Coşar and Fajgelbaum (2012), borders in

Rossi-Hansberg (2005) and Fabinger (2011), and highways and waterways in Allen and Arkolakis (2013) – might have a significant impact on how population, income, and other relevant economic variables are distributed in space. In most of the cases, trade infrastructure has an agglomeration-creating effect since people want to exploit spatial proximity to trading opportunities, while trade barriers repel agglomeration in these models.[1] Our main contribution to this literature is that, to the best of our knowledge, we are the first to study the role of bridges, or other point-like transport infrastructure, in creating agglomeration.

The second literature related to this paper studies the role of transport infrastructure in development in more empirical settings. Donaldson (2012) and Donaldson and Hornbeck (2012) come to the conclusion that the expansion of railroads in the 19th century was a crucial determinant of local development in India and the U.S., respectively. Baum-Snow et al. (2012) and Duranton et al. (2012), on the other hand, find that highways have been playing an important role in city development in China and the U.S. The fact that these effects are likely to be long-lasting is pointed out by Bleakley and Lin (2012), who find that pre-19th-century portage sites remain population centers, despite the fact that their advantage in transportation have been obsolete for long. Relative to this second strand of the literature, we identify a new mechanism for agglomeration. In these papers, transport infrastructure is assumed to reduce trade costs, but is not a source of agglomeration itself. In our model, bridges not only make trade between two regions cheaper, but also serve as focal points of agglomeration.

The structure of the paper is as follows. Section 2 describes the model together with the set of predictions that the model provides, while Section 3 presents the data, the empirical strategy, and the results. Section 4 concludes.

[1] Rossi-Hansberg (2005), however, points to a case in which more trade restrictions on the border are responsible for creating agglomeration.

2 A model of trading across a river

We model production and trade in continuous space. There are a continuum of workers freely choosing location on a plane that is separated by a river. They produce two goods, using land and their labor. The two sides of the river differ in relative productivities, leading to Ricardian gains from trade across the river. There are no gains from within-region trade. (Most of our results survive if all agents specialize fully, such as in Allen and Arkolakis, 2013.) Transportation is costly, giving incentives to move close to trade opportunities.

We study how the relative price of the two goods varies in space, and the patterns of specialization and agglomeration. For a fixed set of bridges, we derive a handful of predictions on the equilibrium distribution of population around the river and bridges.

2.1 Geography

Space is continuous. We concentrate on a compact and connected subset S of the sphere that represents the globe. A circle segment called the *river* divides S into two parts: the Home country (H) and the Foreign country (F) – see Figure 1.[2] Locations (i.e., points in S) are indexed by the triplet (C, ℓ, h), where $C \in \{H, F\}$ is the country which the location belongs to, ℓ is the distance of the location from the river, and h is the distance of the location from an arbitrarily chosen circle $h = 0$ that is perpendicular to the river. (In other words, h represents the river mile.) For simplicity, we refer to ℓ as "longitude," and to h as "latitude." Finally, there is a finite set of latitudes $h_1, \ldots, h_B$ at which *bridges* span the river.

[2] A circle on the sphere is equivalent to a straight line on the plane: it is the shortest path between any two points that lie on it. Also notice that the river is assumed to have zero width. However, this assumption is without loss of generality because the only relevant geographical feature of the river in our model is the cost of crossing it, which is given exogenously.

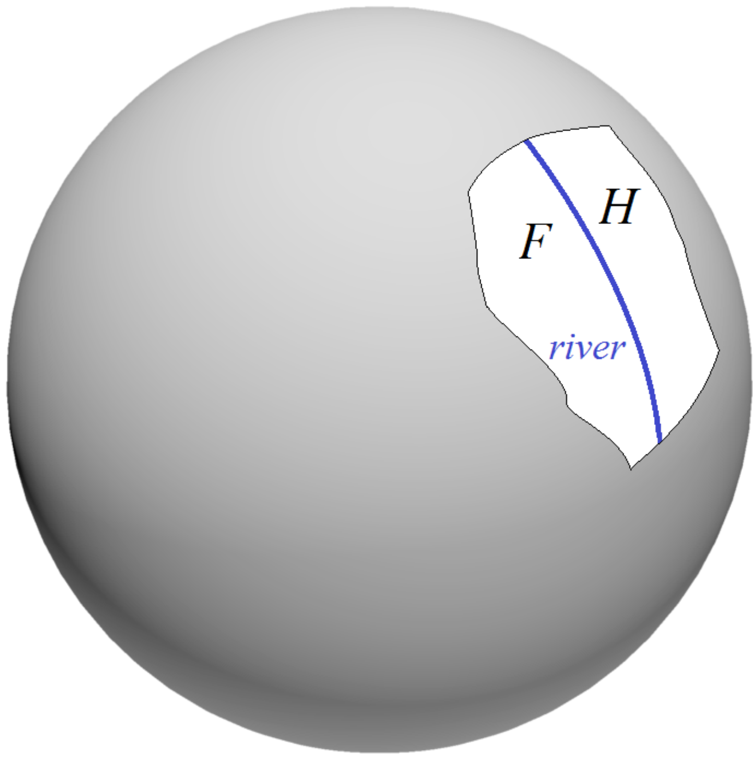

Figure 1: Geography of the river

There are two goods, denoted by X and M. Shipment of goods is costly. Land shipping of good i involves an iceberg cost of $e^{t_i d}$, where t_i is a positive constant, and d is total distance traveled. Crossing the river entails additional costs. The river can be crossed in two ways: (1) by boat at any point, at an iceberg cost of $e^{\tau_i^0}$, or (2) through bridge $b \in \{1, \ldots, B\}$, at an iceberg cost of $e^{\tau_i^b}$, where $\tau_i^0, \tau_i^b > 0$, and the value of τ_i^b can potentially vary with b.

Productivity is distributed evenly within countries, but can differ across countries. Let a_i^C be the unit cost of production in sector $i \in \{X, M\}$ of country $C \in \{H, F\}$. Then the autarky price of the X-good relative to the M-good is given by $p_A^C = \frac{a_X^C}{a_M^C}$ in country C.

Finally, the spatial distribution of factor endowments is as follows. There is a mass of N^C workers in country C, each of them supplying one unit of labor inelastically. Workers are freely mobile within countries, but cannot migrate from one country to the other. Also, each location (C, ℓ, h) is endowed with a strictly positive amount of land $\lambda(C, \ell, h)$. Land is owned by local landlords.

2.2 Consumption

Workers have Cobb–Douglas preferences over goods X and M, spending half of their income on each good. Therefore, the representative worker at location (C, ℓ, h) has indirect utility

$$u(C, \ell, h) = \frac{w(C, \ell, h)}{P_X(C, \ell, h)^{\frac{1}{2}} P_M(C, \ell, h)^{\frac{1}{2}}}, \tag{1}$$

where $w(C, \ell, h)$ is the wage at (C, ℓ, h), and $P_X(C, \ell, h)$ and $P_M(C, \ell, h)$ are the local prices of the X- and M-goods, respectively. Within each country, workers move to the location at which their indirect utility is largest.

Landlords have the same preferences as workers. Landlords are immobile, and do not work. We assume that the number of landlords is small enough that we can approximate total population by the number of workers at each location.

2.3 Production

Both goods are produced under constant returns to scale, using labor and land. The production function is Cobb–Douglas with an α share of labor in both sectors. Both sectors are characterized by perfect competition at each location. Therefore, a firm that operates in the i-sector at (C, ℓ, h) solves the problem

$$\max_{n_i(C,\ell,h)} P_i(C, \ell, h) \frac{n_i(C, \ell, h)^{\alpha}}{a_i^C} - w(C, \ell, h) n_i(C, \ell, h) - r(C, \ell, h),$$

where $n_i(C,\ell,h)$ is labor usage *per unit of land*, and $r(C,\ell,h)$ is the rent per unit unit of land.

The first-order condition to the firm's maximization problem implies

$$n_i(C,\ell,h) = \alpha^{\frac{1}{1-\alpha}} \left(a_i^C\right)^{-\frac{1}{1-\alpha}} \left[\frac{P_i(C,\ell,h)}{w(C,\ell,h)}\right]^{\frac{1}{1-\alpha}}. \tag{2}$$

Hence, if a good is produced at location (C,ℓ,h), then the mass of workers in the good's production is positively linked to the good's local price relative to the nominal wage.

2.4 Equilibrium

Now we define a competitive equilibrium in this economy.

Definition 1. *An equilibrium is a set of functions P_X, P_M, n_X, n_M, n, λ_X, λ_M, w and r, as well as real wages u^H and u^F such that*

(1) utility of workers is maximized and equalized across locations:

$$\frac{w(C,\ell,h)}{P_X(C,\ell,h)^{\frac{1}{2}} P_M(C,\ell,h)^{\frac{1}{2}}} = u^C$$

for all $C \in \{H,F\}$, ℓ, and h,

(2) profits are maximized and driven down to zero:

$$P_i(C,\ell,h)\frac{n_i(C,\ell,h)^{\alpha}}{a_i^C} - w(C,\ell,h)\,n_i(C,\ell,h) - r(C,\ell,h) = 0$$

for all $C \in \{H,F\}$, ℓ, and h,

(3) local land markets clear:

$$\lambda_X(C,\ell,h) + \lambda_M(C,\ell,h) = \lambda(C,\ell,h)$$

for all $C \in \{H,F\}$, ℓ, and h, where $\lambda_i(C,\ell,h)$ denotes local land usage by *sector i,*

(4) local and global labor markets clear:

$$\begin{aligned} \frac{\lambda_X(C,\ell,h)\,n_X(C,\ell,h)+\lambda_M(C,\ell,h)\,n_M(C,\ell,h)}{\lambda(C,\ell,h)} &= n(C,\ell,h) \\ \int_C \lambda(C,\ell,h)\,n(C,\ell,h)\,ds &= N^C \end{aligned}$$

for all $C \in \{H, F\}$, ℓ, and h,

(5) there is no arbitrage possibility within countries:

$$P_i(C,\ell,h) \le e^{t_i d[(C,\ell,h),(C,\ell',h')]} P_i(C,\ell',h')$$

for all (C,ℓ,h) and (C,ℓ',h'), where $d[(C,\ell,h),(C,\ell',h')]$ denotes the distance between the two locations, and we have equality if (C,ℓ',h') ships good i through (C,ℓ,h),

(6) there is no arbitrage possibility over the river:

$$P_i(C,0,h) \le e^{\tau_i^0} P_i(C',0,h)$$

for all C, C' and h, and we have equality if country C' exports good i at h by boat,

(7) there is no arbitrage possibility over bridges:

$$P_i(C,0,h_b) \le e^{\tau_i^b} P_i(C',0,h_b)$$

for all C, C' and $b \in \{1,\ldots,B\}$, and we have equality if country C' exports good i through bridge b,

(8) trade is balanced between each pair of locations.

Let us introduce the notation $p(C,\ell,h) = \frac{P_X(C,\ell,h)}{P_M(C,\ell,h)}$, that is, $p(C,\ell,h)$ is the *relative price* of the X-good at location (C,ℓ,h). What is the pattern of specialization in equilibrium? By constant returns to scale, this only depends on the relationship between the equilibrium relative price and the autarky relative price. In particular,

- (C, ℓ, h) is fully specialized in good X if $p(C, \ell, h) > p_A^C$, implying $n(C, \ell, h) = n_X(C, \ell, h)$,
- (C, ℓ, h) is fully specialized in good M if $p(C, \ell, h) < p_A^C$, implying $n(C, \ell, h) = n_M(C, \ell, h)$,
- if (C, ℓ, h) is incompletely specialized, then $p(C, \ell, h) = p_A^C$, and $n(C, \ell, h) = n_X(C, \ell, h) = n_M(C, \ell, h)$ by (2).

Also notice that, due to trade costs, any location that is incompletely specialized is necessarily in autarky: a consumer at such a place would never find it optimal to buy any of the two goods from another location.

We assume that Home has a comparative advantage in X. In other words, no Home location specializes in good M, and no Foreign location specializes in good X. Given the trade costs, a sufficient condition for this is

$$p_A^H < p_A^F e^{-\max\left\{\max_b\left(\tau_X^b + \tau_M^b\right), \tau_X^0 + \tau_M^0\right\}}.$$

Also notice that there can be no within-country trade in equilibrium: locations that are in autarky do not trade at all, whereas locations specialized in the country's export good only trade with the other country.

We then have the following proposition that is a generalization of Proposition 2 in Coşar and Fajgelbaum (2012).

Proposition 1. *In any equilibrium, each country C is a union of two disjoint sets ("regions") T^C and A^C such that*

(i) all locations in region T^C trade with the other country,

(ii) all locations in region A^C that are not on the boundary of T^C are in autarky, and

(iii) locations in A^C *that are on the boundary of* T^C *are indifferent between trade and autarky.*

Moreover, for each country C *and latitude* h*, there exists a longitude* $\widehat{\ell}(C,h)$ *such that* $(C,\ell,h) \in T^C$ *for all* $\ell < \widehat{\ell}(C,h)$*, and* $(C,\ell,h) \in A^C$ *for all* $\ell > \widehat{\ell}(C,h)$*.*

Proof. See Appendix. □

Figure 2 is a graphical illustration of Proposition 1. As one can see, locations in the trading region T^C are closer to the river than locations in the autarky region A^C for each latitude.

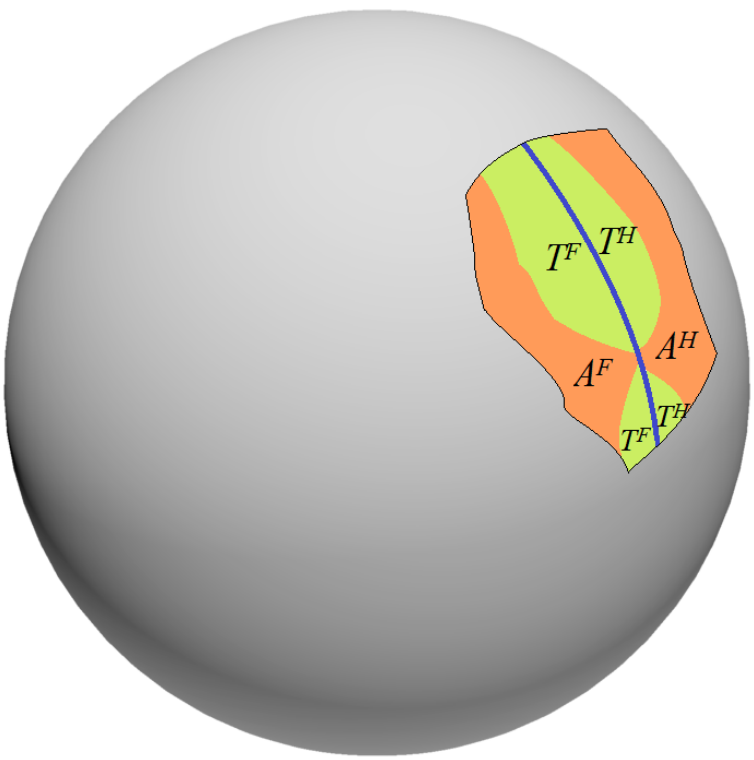

Figure 2: Spatial specialization

Combining equations (1) and (2), and using the equalization of utility in equilibrium, we can relate the equilibrium spatial distribution of population to the equilibrium spatial distribution of relative prices:

$$n\left(H,\ell,h\right)=n_X\left(H,\ell,h\right)=\alpha^{\frac{1}{1-\alpha}}\left(u^H a_X^H\right)^{-\frac{1}{1-\alpha}}p\left(H,\ell,h\right)^{\frac{1}{2(1-\alpha)}} \tag{3}$$

and

$$n\left(F,\ell,h\right)=n_M\left(F,\ell,h\right)=\alpha^{\frac{1}{1-\alpha}}\left(u^F a_M^F\right)^{-\frac{1}{1-\alpha}}p\left(F,\ell,h\right)^{-\frac{1}{2(1-\alpha)}}. \tag{4}$$

That is, within-country differences in population density are solely driven by differences in the relative price. At Home, locations that have a high p offer a high price of the export good and a low price of the import good. Hence, many people decide to move to these locations. On the contrary, a location with a high p is not attractive in the Foreign country; thus, such locations are characterized by low population density in equilibrium.

Using equations (3) and (4), the model generates two predictions on the distribution of population, summarized in Propositions 2 and 3.[3]

Proposition 2 (Concentration at the river)**.** *Take a country C, and restrict attention to a "rectangular" subset of locations $\left\{(C,\ell,h):\underline{\ell}\leq\ell\leq\overline{\ell},\underline{h}\leq h\leq\overline{h}\right\}\subset C$. Then average population density of locations at distance ℓ from the river is at least as high as average population density of locations at distance $\ell'>\ell$ from the river.*

Proposition 3 (Concentration at bridges)**.** *In any country C, take two locations (C,ℓ,h) and (C,ℓ',h') which trade over the same bridge. Then $n\left(C,\ell,h\right)>n\left(C,\ell',h'\right)$ if and only if (C,ℓ,h) is closer to the bridge than (C,ℓ',h'). As a consequence, locations at bridges over which trade takes place are the only local maxima of $n\left(C,\ell,h\right)$ if boat trade is prohibitively costly.*

[3]The Appendix contains the proofs of these propositions.

2.5 Random variations in productivity

This section presents a generalization of the model in which productivity is not necessarily evenly distributed within countries. We do this to account for idiosyncratic variation in population density in equilibrium.

Let $a_i(C, \ell, h)$ be the unit cost of production in sector $i \in \{X, M\}$ at location (C, ℓ, h). We assume that $a_i(C, \ell, h)$ are random variables, each with marginal cdf $G_i^C(\cdot)$, but not necessarily independent. That is, our specification allows for both spatial and cross-industry correlation of productivity draws. Finally, we assume that $G_M^H(\cdot)$ and $G_X^F(\cdot)$ are such that Home locations specialize in the X-good, and Foreign locations specialize in the M-good with probability one.[4]

Under these assumptions, equations (3) and (4) can be written as

$$n(H, \ell, h) = n_X(H, \ell, h) = \alpha^{\frac{1}{1-\alpha}} \left(u^H\right)^{-\frac{1}{1-\alpha}} a_X(H, \ell, h)^{-\frac{1}{1-\alpha}} p(H, \ell, h)^{\frac{1}{2(1-\alpha)}} \tag{3'}$$

and

$$n(F, \ell, h) = n_M(F, \ell, h) = \alpha^{\frac{1}{1-\alpha}} \left(u^F\right)^{-\frac{1}{1-\alpha}} a_M(F, \ell, h)^{-\frac{1}{1-\alpha}} p(F, \ell, h)^{-\frac{1}{2(1-\alpha)}} . \tag{4'}$$

Equations (3') and (4') imply that Propositions 2 and 3 still hold in expectation, i.e., if one replaces "population density" at a given location, $n(C, \ell, h)$, by "expected population density" at the location, $\mathbf{E}n(C, \ell, h)$.

Generalizing the distribution of productivitycomes at the expense of more restrictions on geographical structure. First, we assume that the two countries are mirror images of each other, that is, (1) $(H, \ell, h) \in H$ if and only if $(F, \ell, h) \in F$ for all ℓ and h, (2) Home and Foreign are endowed with the same number of workers ($N^H = N^F$), (3) the distribution of land is such that $\lambda(H, \ell, h) = \lambda(F, \ell, h)$ for all ℓ and h. Second, we assume that the

[4]The easiest way to satisfy this restriction is to assume that $a_M(H, \ell, h)$, or $a_X(F, \ell, h)$, or both, are "very large" with probability one.

productivity of the good the country specializes in (good X at Home, and good M in Foreign) is drawn from the same distribution in the two countries, that is, $G_X^H(\cdot) = G_M^F(\cdot)$. Third, we assume that trade over every bridge has the same iceberg cost: $\tau_i^b = \tau_i$, and boat trade is prohibitively costly.

Under these assumptions, one can show that the distribution of relative prices along the river takes the form

$$\begin{aligned} p(H,0,h) &= e^{2p-(t_X+t_M)\delta(h)} \\ p(F,0,h) &= e^{2p+\tau_X+\tau_M+(t_X+t_M)\delta(h)}, \end{aligned}$$

where p is a constant, and $\delta(h)$ denotes the distance of location $(0,h)$ from the closest bridge. Denote $\tau = \frac{\tau_X+\tau_M}{2}$ and $t = \frac{t_X+t_M}{2}$. Then (3') and (4') yield, in logs,

$$\begin{aligned} \log n(H,0,h) &= \frac{1}{1-\alpha}\left[\log\alpha - \log u^H + p - \log a_X(H,0,h) - t\delta(h)\right] \\ \log n(F,0,h) &= \frac{1}{1-\alpha}\left[\log\alpha - \log u^F - p - \tau - \log a_M(F,0,h) - t\delta(h)\right]. \end{aligned}$$

Therefore,

$$\mathbf{Cov}\left[\log n(H,0,h), \log n(F,0,h)\right] = \frac{1}{(1-\alpha)^2}\left[C_{LR} + t^2\mathbf{Var}\left[\delta(h)\right]\right]$$

where $C_{LR} = \mathbf{Cov}\left[\log a_X(H,0,h), \log a_M(F,0,h)\right]$ is the covariance between productivity realizations of the two banks of the river. Now since

$$\mathbf{Var}\left[\log n(H,0,h)\right] = \frac{1}{(1-\alpha)^2}\left[\sigma^2 + t^2\mathbf{Var}\left(\delta(h)\right)\right] = \mathbf{Var}\left[\log n(F,0,h)\right]$$

where σ^2 is the common variance of $\log a_X(H,0,h)$ and $\log a_M(F,0,h)$, we obtain that the correlation between left- and right-bank log population density is

$$\rho = \frac{C_{LR} + t^2\mathbf{Var}\left[\delta(h)\right]}{\sigma^2 + t^2\mathbf{Var}\left[\delta(h)\right]} = 1 - \frac{\sigma^2 - C_{LR}}{\sigma^2 + t^2\mathbf{Var}\left[\delta(h)\right]}. \tag{5}$$

This equation allows us to provide the following two predictions on the distribution of population along the river.

Proposition 4 (Left- and right-bank density positively correlated)**.** *If left- and right-bank log productivities are positively correlated or uncorrelated, then the correlation between left- and right-bank log population density is positive.*

Proof. If log productivities are positively correlated or uncorrelated, then $C_{LR} \geq 0$. Then equation (5) immediately implies

$$\rho \geq 1 - \frac{\sigma^2}{\sigma^2 + t^2 \mathbf{Var}\left[\delta\left(h\right)\right]} > 0.$$

□

Proposition 5 (Lower correlation at bridges)**.** *The correlation between left- and right-bank population density is lower at (trading) bridges than in general.*

Proof. Calculating the correlation coefficient at trading bridges only, we find

$$\rho_{bridges} = 1 - \frac{\sigma^2 - C_{LR}}{\sigma^2} = \frac{C_{LR}}{\sigma^2}$$

because $\delta\left(h\right) \equiv 0$, hence $\mathbf{Var}\left[\delta\left(h\right)\right] = 0$ in this case. $\rho_{bridges} < \rho$ follows from comparing this to equation (5). □

The intuition for Proposition 5 is as follows. As we move away from a bridge along the river, trade costs increase by as much on the left bank as on the right bank of the river. This leads to a comovement in the terms of trade (p in Home, and p^{-1} in Foreign) on the two banks, and hence to a comovement in Home and Foreign population density (which are increasing power functions of the terms of trade). This comovement in densities acts against the variation caused by fluctuations in productivity. Thus, starting from a bridge, the longer the interval over which we calculate the correlation coefficient, the larger value we find for ρ.

Notice that the existence of bridges over which trade takes place is crucial for this result: the above mentioned comovement in trade costs is absent whenever people trade by boat, or are in autarky. Hence, the fact that this prediction is verified in the data can be taken as a clear indication that bridges matter for the spatial distribution of economic activity.

3 Rivers and population density

We test the model's predictions on six major North American rivers, the Delaware, the Hudson, the Mississippi, the Missouri, the Ohio and the Tennessee.

3.1 Mapping model to data

In the model, each location is on either side of the river is characterized by two coordinates: its distance from the river (longitude) and its distance along the river from a chosen rivermile (latitude). In reality, rivers are not straight lines. To calculate these two relevant coordinates, we proceed as follows.

Let river $R : \mathbb{R} \to \mathbb{R}^2$ be a parametric curve mapping rivermiles into points ons the plane. $R(0)$ is the vector of geocoordinates of the river's mouth, $R(1)$ is the geocoordinate of the first rivermile, etc. For any point (x, y), we can determine the river-coordinates as follows

$$\ell(x, y, R) \equiv \min_m d[(x, y), R(m)],$$
$$h(x, y, R) \equiv \arg\min_m d[(x, y), R(m)],$$

where $d : \mathbb{R}^4 \to \mathbb{R}^+$ measures the distance between a pair of points.

That is, distance is measured as distance to the nearest point of the river, and h is measured as the rivermile of this neaerst point. For straight rivers, these measures exactly correspond to the ones used in the model.

Note that the $(x, y) \to (\ell, h)$ mapping is not a bijection. While there is only one nearest point with probability one, there may be multiple (x, y) points on the plane for which m is the closest rivermile.

The use of this mapping is illustrated in Figure 3, which plots population density on the left and right bank of the Delaware as a function of rivermiles. The high-density areas of Philadelphia (mostly right bank) and Trenton (mostly left bank) are clearly visible.

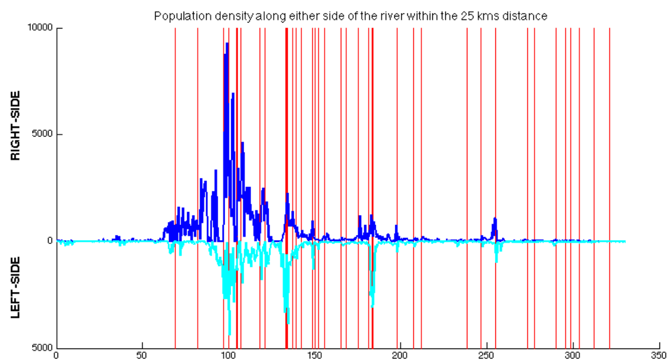

Figure 3: Bridges and population density on the two banks of the Delaware

3.2 Data

To measure population density, we use Version 1 of the Global Rural-Urban Mapping Project population density grid. This dataset provides population count (and density) estimates for each 30 arc-second by 30 arc-second gridcell of the U.S. (The are of these gridcells is around 0.25 km^2.) We use the values from year 2000.

We take the geocoordinates of rivers from the ESRI Map of U.S. Major Waters, containing polygons of 29,167 water surfaces, including rivers and lakes. We selected the Delaware, the Hudson, the Mississippi, the Missouri, the Ohio and the Tennessee. After making the necessary topological corrections (connecting segments of the river and intermittent lakes), we determine the left and right bank of each river. For the Delaware, we exclude Philadelphia and for the Hudson, we exclude New York City from the analysis. Our results are stronger with these cities included.

The location of major bridges comes from Wikipedia. For the Delaware, we have more detailed data on bridges, including year of construction. Figure 4 shows the location of the rivers, the bridges, and the population density map of the United States.

Figure 4: Map of the six rivers, their bridges, and population density

3.3 Testing the five predictions

Table 1: Population density and distance to the river

	Population density		
River	0-10mi	10-20mi	20-30mi
Delaware*	467	592	566
Hudson**	585	236	245
Mississippi	137	81	35
Missouri	136	84	37
Ohio	452	196	173
Tennessee	247	155	97

* excluding Philadelphia
** excluding NYC

Table 1 shows how population density varies with distance to the six major rivers. We measure population density within 10-miles bands along the river. On five out of the six rivers, population density between 10 and 20 miles from the river is strictly lower than within 0 and 10 miles, and on four rivers, it further reduces as we move farther away to 20 to 30 miles. This is consistent with the model, where trading opportunities on the other side of the river lead to a density gradient.

The exceptions are the Delaware and the Hudson, where population density does not show a clear declining pattern. The two rivers are very close to each other, and their 30-mile neighborhood may be affected by the metropolitan areas of New York City and Philadelphia.

Table 2: Population density and distance to the nearest bridge

River	Correlation of log population density with distance to nearest bridge
Delaware*	-0.353
Hudson**	0.177
Mississippi	-0.429
Missouri	-0.501
Ohio	-0.504
Tennessee	-0.404

* excluding Philadelphia
** excluding NYC

Table 2 shows the correlation coefficient of log population density along the river with distance to the nearest bridge. With the exception of the Hudson, the other five rivers display very strong negative correlation. In the model, as bridges are the focal points of agglomeration, population density falls with distance, consistently with the facts.

Figures B1 through B6 (in the Appendix) plot the distribution of population density near and away from bridges. For each river, the red line plots the kernel density of log population densities at gridcells that have a bridge within 3 miles. (We calculated average population density between 0 and 10 miles from the river.) The blue line plots the kernel estimate of log population densities for gridcells more than 3 miles from a bridge.

For all rivers except the Hudson, the distribution of population densities near bridges is shifted to the right. That is, average population density is higher within 3 miles of the bridge than outside this distance. (This is consistent with the correlations reported in Table 2.)

We also see that there is a large variation in population densities both near and away from bridges. In particular, some locations without a bridge are as densely populated as some of those with one. This suggests that building bridges involves nontrivial costs, and not every community can overcome these costs, severly limiting their access to the other side of the river.

Table 3: Clustering close and far from the river

	Coefficient of variation of population density (left bank)		
River	0-10mi	10-20mi	20-30mi
Delaware*	1.509	1.745	0.847
Hudson**	1.586	2.265	1.461
Mississippi	3.256	3.161	1.564
Missouri	4.011	3.176	2.008
Ohio	2.509	2.259	1.419
Tennessee	1.886	1.854	0.990

* excluding Philadelphia
** excluding NYC

Table 3 displays a measure of clustering at various distances to the river. We calculate the coefficient of variation of population density. This measure is high when population density varies a lot, between, say, a large a city and sparse surroundings. It is low when many small cities or towns are roughly evenly distributed in space. Note that the coefficient of variation is unaffected by the overall mean population density, which we have reported in Table 1.

As we move farther away from the river, the coefficient of variation tends to fall. This is in line with our theory, where the agglomerating force of bridges can only be felt close to the river, and not farther from where multiple bridges are equally easily accessible.

Table 4: Correlation between the two banks of the river

River	Correlation of population density between left and right bank
Delaware*	0.567
Hudson**	0.416
Mississippi	0.558
Missouri	0.317
Ohio	0.483
Tennessee	0.510

* excluding Philadelphia
** excluding NYC

Table 4 reports the correlation of population density between the left and the right bank of the river. The model predicts that population is going to cluster on both sides of the bridge, leading to positive correlation across the two banks. On all six rivers, the correlation is highly positive, with an average of 0.48. Part of this correlation is driven by the mere presence of bridges. Bridges are surrounded by people on either side of the river, whereas areas far from bridges tend to be more sparse. Table 5 measures this correlation at the bridges. More specifically, we ask how population on the two sides of the bridge is correlated. As predicted by the model, these correlations are positive, but smaller than the unconditional correlations.[5]

4 Conclusion

We built a continuous-space theory of trade to explain why and where people agglomerate. We tested the equilibrium predictions of our model on data

[5]The results are very similar if we use log population density when calculating correlations.

Table 5: Correlation between the two sides of bridges

River	Correlation of population density on either side of bridge
Delaware*	0.399
Hudson**	0.251
Mississippi	0.551
Missouri	0.251
Ohio	0.465
Tennessee	0.281

* excluding Philadelphia
** excluding NYC

from six major American rivers, finding that spatial patterns of population density are consistent with our model.

Taken together with agglomeration externalities (not currently modeled), our theory can relate to the question of whether and how trade causes development. There are two puzzling facts about trade and development. First, the macro correlations between trade and development (even those using plausibly exogenous variation in trade, as Feyrer, 2009a and b) are much larger than model-based meausures of gains from trade (Alvarez and Lucas, 2007, Arkolakis, Costinot and Rodríguez-Clare, 2012). Second, land-locked countries are much less developed than coastal countries, even though transportation costs make up only a small fraction of broader trade costs (Anderson and van Wincoop, 2004).

Our theory has the potential to explain these facts because trade increases the incentives to agglomerate, which leads to external effects. These external effects represent a multiplier of trade on development (consistent with Fact 1). They are also stronger in coastal countries, where ports provide a natural focal point of agglomeration (consistent with Fact 2).

In future work, we intend to estimate the agglomeration effect of bridges. The crucial identification concern is that both the location of bridges and population density are correlated with unobserved local amenities. We plan to use variation in building costs (geographical and hydrological measures) and transit traffic demand to instrument bridge location.

References

[1] Allen, Treb and Arkolakis, Costas, 2013. "Trade and Topography of the Spatial Economy." Working paper.

[2] Alvarez, Fernando and Lucas, Robert E. Jr., 2007. "General Equilibrium Analysis of the Eaton–Kortum Model of International Trade." Journal of Monetary Economics, 54(6): 1726–1768.

[3] Anderson, James E. and van Wincoop, Eric, 2004. "Trade Costs." Journal of Economic Literature, 42(3): 691–751.

[4] Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, 2012. "New Trade Models, Same Old Gains?" American Economic Review, 102(1): 94–130.

[5] Baum-Snow, Nathaniel, Brandt, Loren, Henderson, J. Vernon, Turner, Matthew A., and Zhang, Qinghua, 2012. "Roads, Railroads and Decentralization of Chinese Cities." Working paper.

[6] Bleakley, Hoyt and Jeffrey Lin, 2012. "Portage and Path Dependence." Quarterly Journal of Economics, 1–58.

[7] Coşar, A. Kerem and Pablo D. Fajgelbaum, 2012. "Internal Geography, International Trade, and Regional Outcomes." Working paper.

[8] Dale, Frank T., 2003. "Bridges over the Delaware River: A History of Crossings." Rutgers University Press.

[9] Desmet, Klaus and Rossi-Hansberg, Esteban, 2012. "Spatial Development." Working paper.

[10] Donaldson, Dave, 2012. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." American Economic Review, forthcoming.

[11] Donaldson, Dave and Hornbeck, Richard, 2012. "Railroads and American Economic Growth: A 'Market Access' Approach." Working paper.

[12] Duranton, Gilles, Morrow, Peter M., and Turner, Matthew A., 2012. "Urban Growth and Transportation." Review of Economic Studies, 79(4): 1407–1440.

[13] Fabinger, Michal, 2011. "Trade and Interdependence in a Spatially Complex World." Working paper.

[14] Feyrer, James, 2009a. "Distance, Trade, and Income – The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment." NBER Working Paper 15557.

[15] Feyrer, James, 2009b. "Trade and Income – Exploiting Time Series in Geography." NBER Working Paper 14910.

[16] GRUMP, 2000. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Population Density Grid. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University.

[17] Rossi-Hansberg, Esteban, 2005. "A Spatial Theory of Trade." American Economic Review, 95(5): 1464–1491.

Appendix

A Proofs

We first state the following lemma that we use in the proofs of Propositions 1 and 2.

Lemma 1. *Take two locations (C,ℓ,h) and (C,ℓ',h) such that $\ell'>\ell$. Then $p(C,\ell',h)\leq p(C,\ell,h)$ if $C=H$, and $p(C,\ell',h)\geq p(C,\ell,h)$ if $C=F$.*

Lemma 1. We prove the lemma for $C=H$; the proof for $C=F$ involves the exact same steps. Notice first that $p(H,\ell,h)\geq p_A^H$ and $p(H,\ell',h)\geq p_A^H$ by the assumption that no Home location specializes in good M. If $p(H,\ell',h)=p_A^H$, then the result is immediate. So suppose $p(H,\ell',h)>p_A^H$. Then (H,ℓ',h) is fully specialized in X, and trades with the Foreign country. As a consequence, there must exist a location at the river $\left(H,0,\widehat{h}\right)$ such that (H,ℓ',h) trades through it.

Equilibrium condition (5) then implies

$$P_X\left(H,0,\widehat{h}\right)=e^{t_X d\left[\left(H,0,\widehat{h}\right),(H,\ell',h)\right]}P_X(H,\ell',h)$$

and

$$P_M\left(H,0,\widehat{h}\right)=e^{-t_M d\left[\left(H,0,\widehat{h}\right),(H,\ell',h)\right]}P_M(H,\ell',h).$$

Dividing these two equations yields

$$p\left(H,0,\widehat{h}\right)=e^{(t_X+t_M)d\left[\left(H,0,\widehat{h}\right),(H,\ell',h)\right]}p(H,\ell',h). \tag{6}$$

Similarly, by equilibrium condition (5),

$$P_X\left(H,0,\widehat{h}\right)\leq e^{t_X d\left[\left(H,0,\widehat{h}\right),(H,\ell,h)\right]}P_X(H,\ell,h)$$

and

$$P_M\left(H,0,\widehat{h}\right)\leq e^{-t_M d\left[\left(H,0,\widehat{h}\right),(H,\ell,h)\right]}P_M(H,\ell,h),$$

irrespectively of whether $\left(H,0,\widehat{h}\right)$ and (H,ℓ,h) trade or not. Dividing the last two inequalities, we get

$$\begin{aligned} p\left(H,0,\widehat{h}\right) &\leq e^{(t_X+t_M)d\left[\left(H,0,\widehat{h}\right),(H,\ell,h)\right]}p\left(H,\ell,h\right) \\ &\leq e^{(t_X+t_M)d\left[\left(H,0,\widehat{h}\right),(H,\ell',h)\right]}p\left(H,\ell,h\right), \end{aligned}$$

where the second inequality follows from $\ell' > \ell$. Combining this with equation (6) and cancelling $e^{(t_X+t_M)d\left[\left(H,0,\widehat{h}\right),(H,\ell',h)\right]}$ on both sides yields the result. $\square$

Now we are ready to prove Propositions 1 to 3.

Proposition 1. Define

$$T^C = \left\{(C,\ell,h) : p\left(C,\ell,h\right) \neq p_A^C\right\},$$

and

$$A^C = \left\{(C,\ell,h) : p\left(C,\ell,h\right) = p_A^C\right\}.$$

$C = T^C \cup A^C$ and $T^C \cap A^C$ follow directly from the definitions. To see (i), notice that $p\left(C,\ell,h\right) \neq p_A^C$ necessarily implies that location (C,ℓ,h) is fully specialized, hence it trades with the other country.

For (ii), suppose that a location (H,ℓ,h) from the interior of A^H is not in autarky, thus it trades with the Foreign country. Then there must exist another location $(H,\ell',h') \in A^H$ such that location (H,ℓ,h) trades through it. By equilibrium condition (5), this implies

$$p\left(H,\ell',h'\right) = p\left(H,\ell,h\right)e^{(t_X+t_M)d[(C,\ell,h),(C,\ell',h')]} > p\left(H,\ell,h\right),$$

which contradicts the fact that $p\left(H,\ell',h'\right) = p\left(H,\ell,h\right) = p_A^H$. The argument is similar for $C = F$.

For (iii), we first prove that $p\left(C,\cdot,\cdot\right)$ is a continuous function. By equilibrium condition (5), we have

$$p(C,\ell,h) \leq p(C,\ell',h')\, e^{(t_X+t_M)d}$$

and

$$p(C,\ell',h') \leq p(C,\ell,h)\, e^{(t_X+t_M)d}$$

for any (C,ℓ,h) and (C,ℓ',h'), where $d := d\left[(C,\ell,h),(C,\ell',h')\right]$. Combining these two inequalities yields

$$e^{-(t_X+t_M)d} \leq \frac{p(C,\ell,h)}{p(C,\ell',h')} \leq e^{(t_X+t_M)d}.$$

Hence, in the limit as $(C,\ell',h') \to (C,\ell,h)$ (and thus $d \to 0$), we obtain $\frac{p(C,\ell,h)}{p(C,\ell',h')} \to 1$, that is, $p(C,\ell',h') \to p(C,\ell,h)$. This proves that $p(C,\cdot,\cdot)$ is continuous.

Now pick a location $(H,\ell,h) \in A^H$ that is on the boundary of T^H; the proof is similar for $C = F$. Clearly, location (H,ℓ,h) weakly prefers autarky over trade as $p(H,\ell,h) = p_A^H$. Assume that (H,ℓ,h) strictly prefers autarky over trade; this means $p(H,\ell,h) > p(H,\ell',h')\, e^{-(t_X+t_M)d[(C,\ell,h),(C,\ell',h')]}$ for all trading locations $(H,\ell',h') \in T^H$. However, by the compactness and connectedness of C, there exists a sequence of locations $\left\{(H,\ell_m,h_m) \in T^H\right\}$ converging to (H,ℓ,h). By continuity of $p(C,\cdot,\cdot)$, there must exist a large enough m such that $p(H,\ell_m,h_m) > p(H,\ell',h')\, e^{-(t_X+t_M)d[(C,\ell,h),(C,\ell',h')]}$, implying that (H,ℓ_m,h_m) prefers autarky over trade. But this contradicts the fact that $(H,\ell_m,h_m) \in T^H$.

For the final part, let $\widehat{\ell}(C,h) = \sup_\ell \left\{(C,\ell,h) \in C : p(C,\ell,h) \neq p_A^C\right\}$ if there exists an ℓ such that $p(C,\ell,h) \neq p_A^H$, and $\widehat{\ell}(C,h) = 0$ otherwise. Then Lemma 1 implies that $p(H,\ell,h) > p_A^H$ if $\ell < \widehat{\ell}(H,h)$, hence $(H,\ell,h) \in T^H$; and $p(H,\ell,h) = p_A^H$ if $\ell > \widehat{\ell}(H,h)$, hence $(H,\ell,h) \in A^H$. For Foreign, $\ell < \widehat{\ell}(F,h)$ implies $p(F,\ell,h) < p_A^F$, so $(F,\ell,h) \in T^F$; and $\ell > \widehat{\ell}(F,h)$ implies $p(F,\ell,h) = p_A^F$, so $(F,\ell,h) \in A^F$. This concludes the proof. □

Proposition 2. Average population density at distance ℓ from the river is

$$\int_{\underline{h}}^{\overline{h}} n\left(C,\ell,h\right) dS\left(C,\ell,h\right),$$

and average population density at distance ℓ' is

$$\int_{\underline{h}}^{\overline{h}} n\left(C,\ell',h\right) dS\left(C,\ell',h\right).$$

Suppose $C = H$. Then, by Lemma 1, $p\left(C,\ell,h\right) \geq p\left(C,\ell',h\right)$, which, together with equation (3), implies $n\left(C,\ell,h\right) \geq n\left(C,\ell',h\right)$ for all $h \in \left[\underline{h},\overline{h}\right]$. As a consequence,

$$\int_{\underline{h}}^{\overline{h}} n\left(C,\ell,h\right) dS\left(C,\ell,h\right) \geq \int_{\underline{h}}^{\overline{h}} n\left(C,\ell',h\right) dS\left(C,\ell',h\right).$$

Now suppose $C = F$. Then Lemma 1 implies $p\left(C,\ell,h\right) \leq p\left(C,\ell',h\right)$, so by equation (4), $n\left(C,\ell,h\right) \geq n\left(C,\ell',h\right)$ for all $h \in \left[\underline{h},\overline{h}\right]$. Hence,

$$\int_{\underline{h}}^{\overline{h}} n\left(C,\ell,h\right) dS\left(C,\ell,h\right) \geq \int_{\underline{h}}^{\overline{h}} n\left(C,\ell',h\right) dS\left(C,\ell',h\right).$$

□

Proposition 3. If $C = H$, and (C,ℓ,h) and (C,ℓ',h') both trade over bridge b, then we have

$$p\left(C,0,h_b\right) = p\left(C,\ell,h\right) e^{(t_X+t_M)d[(C,\ell,h),(C,0,h_b)]}$$

and

$$p\left(C,0,h_b\right) = p\left(C,\ell',h'\right) e^{(t_X+t_M)d[(C,\ell',h'),(C,0,h_b)]}$$

by equilibrium condition (5).

Then the fact that (C,ℓ,h) is closer to the bridge than (C,ℓ',h') is equivalent to $p(C,\ell,h) > p(C,\ell',h')$, which, by equation (3), is equivalent to $n(C,\ell,h) > n(C,\ell',h')$.

If $C = F$, equilibrium condition (5) yields $p(C,\ell,h) < p(C,\ell',h')$ if and only if (C,ℓ,h) is closer to the bridge than (C,ℓ',h'), which is equivalent to $n(C,\ell,h) > n(C,\ell',h')$ by equation (4). □

B Additional Figures

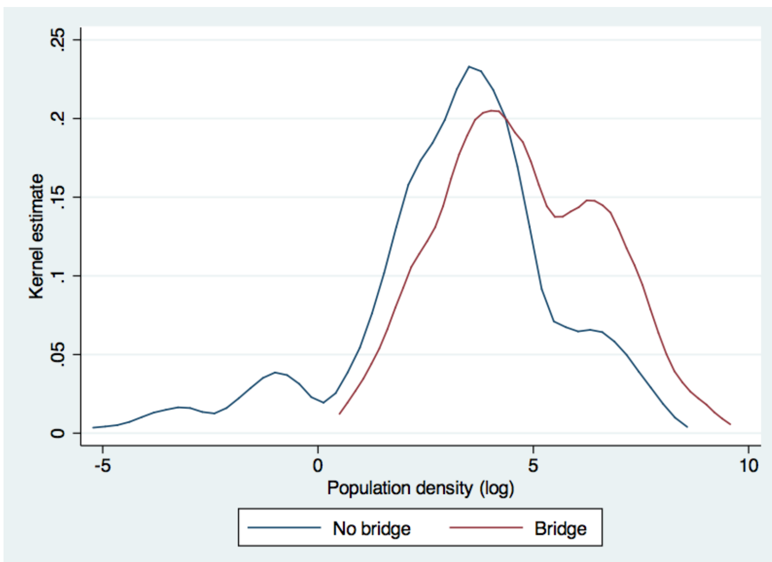

Figure B1: The distribution of population densities near and far of bridges: Delaware

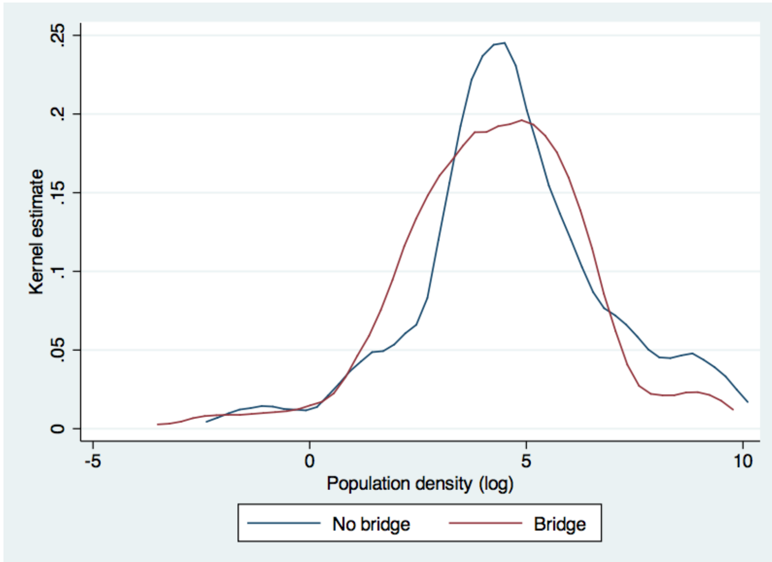

Figure B2: The distribution of population densities near and far of bridges: Hudson

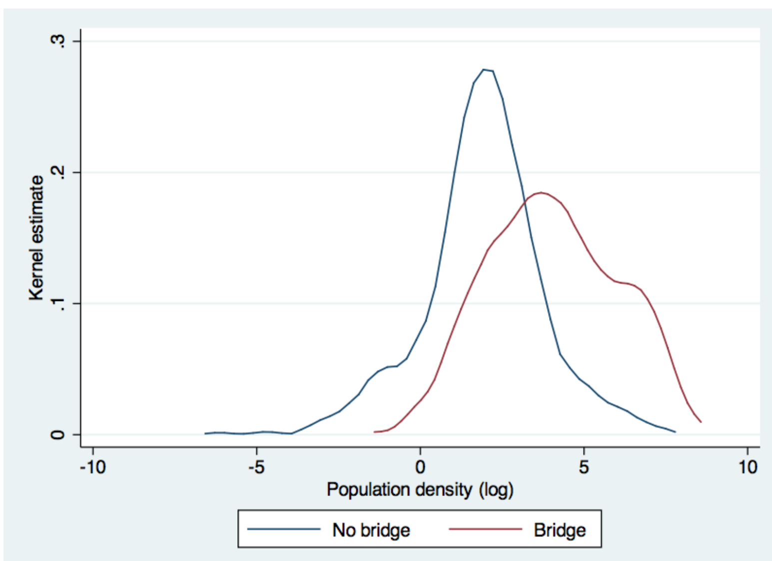

Figure B3: The distribution of population densities near and far of bridges: Mississippi

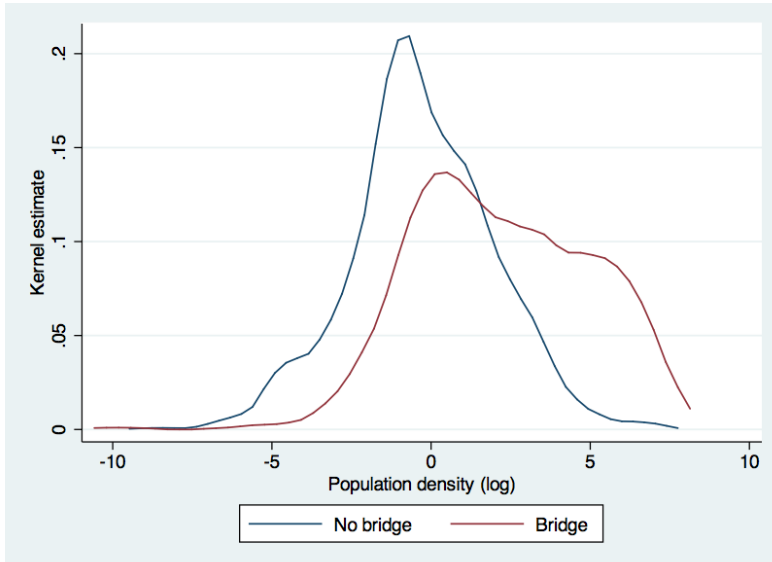

Figure B4: The distribution of population densities near and far of bridges: Missouri

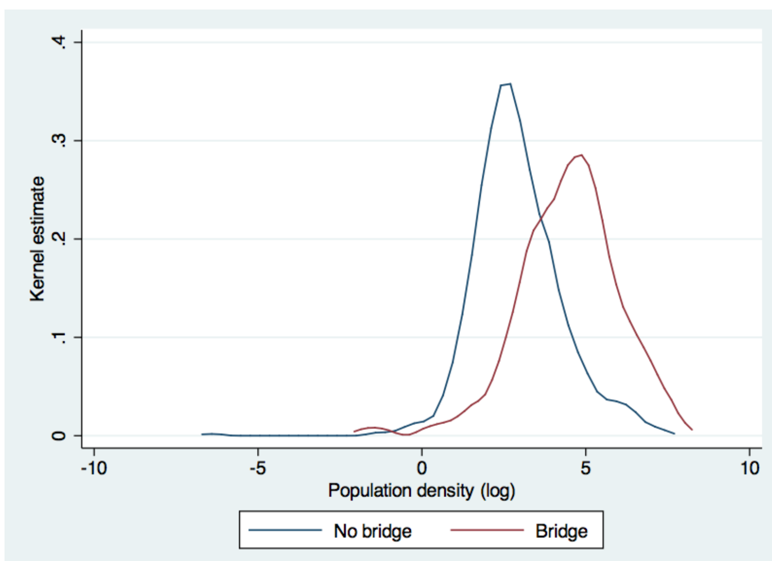

Figure B5: The distribution of population densities near and far of bridges: Ohio

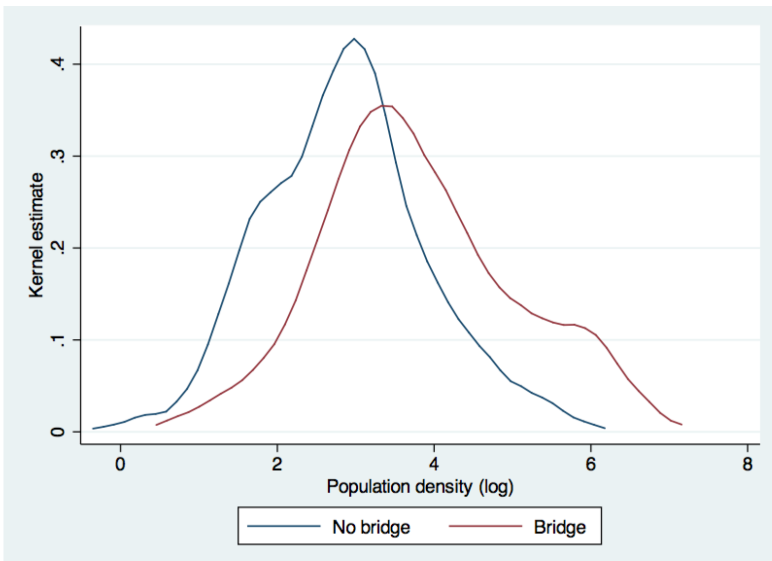

Figure B6: The distribution of population densities near and far of bridges: Tennessee