Web Appendix (not for publication)

Online Appendix B. The Model

Setup

We consider a country consisting of two regions, which we call coast (c) and interior (i) for exposition. Workers in each region potentially produce food (f) with decreasing returns and a manufactured good (m) with external economies of scale subject to congestion. In both sectors, workers are paid their average product. We assume that (in the "long run") workers are free to move between regions and among sectors such that utility is equalized. The economy is closed to the outside world.

For either region $r \in \{c, i\}$, in the food sector average product is $A_f L_{fr}^{-\beta}$ and total production is $A_f L_{fr}^{1-\beta}$, where $A_f > 0$ reflects productivity and L_{fr} is the amount of labor in the food sector in the region. Food sector productivity is the same across regions. Decreasing marginal productivity of labor in agriculture, due to a fixed supply of land, is reflected in the parameter $1 > \beta > 0$. The urban sector produces the manufactured good. Average product per unit of labor in the urban sector is $A_{mr}(v + L_{mr}^{\epsilon})$, where the v allows nonzero output by the marginal worker as $L_{mr} \to 0$, and $\epsilon > 0$ represents agglomeration economies increasing productivity in the presence of more workers. Each worker is endowed with one unit of time, to be used for labor and commuting in the city as in standard urban models (see Duranton and Puga, 2004, for a review), so labor supplied per worker is $1 - tL_{mr}$, where $0 < t << 1$ represents unit-distance commuting costs in the city.[1] Average product per worker in the city is thus

[1] Following Duranton and Puga (2004), each worker is endowed with 1 unit of time, and lives on a lot of fixed size 1, with zero opportunity cost, in a two sided linear city. Working time is $1\text{-}4tu$ where u is distance from the city center and $4t$ unit commuting costs. Given a wage rate w, income after commuting is $w(1\text{-}4tu)$. Residents pay rents that differ by distance from the city center, and rental income is redistributed as an (equal) dividend to all city residents. Since the lots have zero opportunity costs, rent at the city edge ($u_{max}=L/2$) is zero. Net income of the person at the city edge before rent income transfers is $w(1\text{-}4tu_{max}) = w(1\text{-}2tL)$ and that equals net income of any other person $w(1\text{-}4tu)\text{-}R(u)$, where $R(u)$ is rent at location u. Thus $R(u) = 4tw(L/2\text{-}u)$, so integrating across the whole city, total rents are twL^2, and rents remitted per worker are twL. Thus, total (labor plus rental dividend) income per person in the city net of rent and commuting time is $W=w(1\text{-}tL)$ and total income is $w(1\text{-}tL)L$. This corresponds to effective labor supply in the city being $L(1\text{-}tL)$.

$A_{mr}(v + L_{mr}^{\epsilon})(1 - tL_{mr})$. The size of the manufacturing labor force that maximizes this is a solution to $\varepsilon L_{mr}^{\varepsilon-1} - (1+\varepsilon)tL_{mr}^{\varepsilon} - vt = 0$. As long as $\epsilon < 1$, any L_{mr} fulfilling this expression will be a unique maximum, but further restrictions on the parameter space are required to guarantee the existence of an interior maximum.

Food, which is traded costlessly between regions as in standard new economic geography models, is the numeraire good. Preferences are such that each worker consumes a fixed amount of food λ, and spends the remainder of her value of net average product on the manufactured good. Welfare for any person in region r is then equivalent to consumption of the manufactured good, $(W_r - \lambda)/p_{mr}$, where W_r is net wage income and p_{mr} is the price of the manufactured good in region r .

A fixed national population of workers L is free to move between sectors and regions, so that

$$L_r = L_{fr} + L_{mr}, \qquad r = c,\ i \tag{B1a}$$

$$L = L_c + L_i. \tag{B1b}$$

Real income equalization across sectors within each region (assuming both sectors exist in the region) implies:

$$A_f L_{fr}^{-\beta} = p_{mr} A_{mr}(v + L_{mr}^{\epsilon})(1 - tL_{mr})\ ,\ r = c,\ i\,. \tag{B2}$$

Free migration equalizes per person welfare (i.e. manufacturing consumption), across regions so that:

$$\frac{A_f L_{fc}^{-\beta} - \lambda}{p_{mc}} = \frac{A_f L_{fi}^{-\beta} - \lambda}{p_{mi}}\,. \tag{B3}$$

The model is closed by imposing equilibrium in goods markets. How that is done depends on whether there is inter-regional trade or not and whether regions absolutely specialize or not. There are three different types of closure relating to three types of equilibria.

Autarkic equilibrium

If there is no trade between regions, clearing of the manufacturing good market in each region requires total regional demand equals regional supply, or:

$$L_r(A_f L_{fr}^{-\beta} - \lambda) = p_{mr} A_{mr}(v + L_{mr}^{\epsilon})(1 - tL_{mr})\, L_{mr}\ , \qquad \text{(B4a)}$$

or alternatively, using the agricultural market

$$\lambda L_r = A_f L_{fr}^{1-\beta}\ . \qquad \text{(B4b)}$$

Given $L, A_f, A_{mi}, A_{mc}, \beta, \epsilon, \lambda, t, v$, the eight equations implied in (B1)-(B3) and (B4b) specify equilibrium in the distribution of labor and the price of the manufactured good wherever it is produced ($L_{mc}, L_{mi}, L_{fc}, L_{fi}, L_c, L_i,\ p_{mc}, p_{mi}$).[2]

Trade equilibrium with both regions producing manufactures

If transport costs are sufficiently low, both regions can trade *and* produce manufactures if they have differential comparative advantage. If they are identical and have sufficient manufacturing scale beyond the point where average product is maximized, then there will be no trade. We generally designate one region to be slightly better at manufacturing, in order to allow trade equilibria when trade costs are sufficiently low. We maintain the assumption that food can move costlessly between regions, and further assume that there is an iceberg cost τ that applies to movement of the manufactured good between regions. Trade will occur when the autarky price ratio of manufactured goods is outside the range $(1 - \tau,\ \frac{1}{1-\tau})$. When there is trade, and no absolute specialization in either region, the within-region goods market clearing conditions (B4a) and (B4b) are replaced by an inter-regional goods market clearing condition and an arbitrage condition. We specify food market equilibrium and leave the manufactured good as a residual:

$$\lambda L = A_f L_{fc}^{1-\beta} + A_f L_{fi}^{1-\beta}\ . \qquad \text{(B5)}$$

Assigning manufacturing comparative advantage to the coastal region, in practice it will always be the manufacturing exporter in this class of equilibria, although we check for equilibria where the interior is exporting manufactures as well. The prices of the manufactured good in the two regions are related by an arbitrage condition:

[2] To see that these represent eight equations, note that (B1a), (B2) and (B4b) each must be fulfilled for each region.

$$p_{mi} = p_{mc}(1 - \tau)^{-1} \,, \tag{B6}$$

for τ the trade cost. Given $L, A_f, A_{mi}, A_{mc}, \beta, \epsilon, \lambda, t, \tau, v$, the eight equations embedded in (B1)-(B3), (B5) and (B6) specify an equilibrium in the distribution of labor and the price of the manufactured good in the two regions.

Specialization equilibrium

Finally, there are equilibria where all manufactured goods are produced in one region, Since that one region can be either the coast or the interior, we consider the two corresponding types of specialized equilibria in the solution mechanism below. It can be defined by slightly adjusting the trade equilibria without specialization above, setting manufacturing employment in one region to zero and removing equation (B2) for that region.

Solving the model

For any given set of parameters, we solve the model as follows. We have 3 types of possible equilibria: autarkic, trade without absolute specialization, and trade with absolute specialization, with each of the last two available in two variants, one for each region exporting manufactures. We pick an allocation of population to the interior region (with the coastal population being the remainder of national population) and suspend equation (B3) (equalizing welfare across regions). We then use the remaining equations in each type of equilibria to solve for all remaining variables. From these we calculate the consumption per worker in each region (the LHS and RHS to (B3)). Then, for each equilibrium type, we plot these two regional consumptions as a function of (say) interior population. Their intersections are equilibria.

We limit attention to stable equilibria, subject to two stability conditions. Type 1 stability is with respect to small changes in the population allocation across regions, assuming within-region labor markets and all goods markets always clear ("instantly"). Equilibria are stable as long as per-person manufacturing consumption in the interior (coast) is a declining (increasing) function of L_i (i.e., there are overall diseconomies to regional size). Type 2 stability is with respect to perturbations within regions, focused in particular on adding a small number of workers to a non-existent or small manufacturing sector within a region. For example, we perturb a small number of workers out of food production in, say, the interior region and move them into manufacturing in the interior. We keep regional populations fixed, but allow intra- and inter-regional goods markets and coastal labor markets to clear ("instantly"). Equilibria are unstable if interior manufacturing workers then have higher consumption than interior food workers. They are stable if the reverse is the case. This condition implicitly assumes slower than instant adjustment in inter-regional labor markets. We note however that in practice in all examples we solved, in this type of experiment, under stability, interior food workers have the highest welfare (manufacturing consumption) of workers anywhere and interior manufacturing workers the lowest; and vice versa under instability.

Details of the solution method as applied to the examples below are given below. In general, for any τ there will either be an autarky or non-specialization equilibrium but not both, with higher τ having autarky. There may or may not be specialization equilibrium in one or both regions, with the likelihood of stable specialization equilibria enhanced as τ falls.

Analysis of possible equilibria

As with many similar models, there is no closed form solution. We illustrate the relevant properties with several examples. Our baseline parameter set is

$\{L = 10,000,000;\ v = 0.5;\ \epsilon = 0.08; t = 7\times10^{-8}; \beta = 0.25; \lambda = 0.018; A_f = 1; A_{mi} = 1;\ A_{mc} = 1.01\}$.

With these parameters, average manufacturing product peaks at a city population of about 969,100. Note the regions are not precisely symmetrical, so that if trade is feasible, it will occur because the coast has a slight comparative advantage in manufacturing production (with a higher A_m). We consider all the specialization and non-specialization equilibria that exist and are stable for values of transport costs, τ, from 0.99 to 0.01. There are two types of specialization equilibria: the coast producing only food and the interior producing only food.

Non-specialization implies autarky at high τ; while at lower τ, when stable non-specialized equilibria exist, they are trade equilibria.

Our focus is on how these patterns change in the transition from low ($A_f = 1$) to high ($A_f = 1.5$) agricultural productivity. When A_f is low, at least with non-specialized manufacturing, there is insufficient manufacturing employment to support a city populous enough to exploit scale economies in any one region. When A_f is high, much less labor is needed to produce the required food, so there is a lot more manufacturing employment to allocate between the two regions.

Figure B1 shows the stable specialization and non-specialization equilibria when $A_f = 1$ for different costs of trade, τ, as graphed against the population of the interior region. The two outer prongs correspond to the two sets of specialization equilibria: one where the interior produces only food and one where the coast does that. When do these specialization equilibria exist and when are they stable? They are (type 2) unstable when trade costs are high ($\tau > 0.4$). In that case, workers who begin manufacturing in the region with no existing manufacturing will be better off, because high trade costs make them profitable in their home market despite the limited scale. When τ is lower, the scale effect advantage of the existing manufacturing sector in the foreign region dominates any trade cost advantage in starting manufacturing in the home region to sell in the home market. Starting a small scale manufacturing operation is not profitable for those workers. In our example with $A_f = 1$, the allocation of workers to manufacturing in the specialized region is less than the city size that maximizes average product. Thus stable specialization equilibria persist as τ falls to 0.

The middle prong represents non-specialized equilibria. At high τ, they are autarkic and stable. While manufacturing scale is low in both regions, trade is too costly for workers to profitably move to take advantage of scale economies in one region. As trade costs fall, it becomes potentially profitable to trade. However, once it is profitable to trade it is also profitable to enhance manufacturing scale in one region relative to the other. Thus at these parameter values, the only stable equilibria when τ is low have manufacturing located in a single region, with the other region producing only food.

In Figure B2 we turn to our second case, where agricultural productivity A_f has risen to 1.5, allowing more workers to enter manufacturing. As in the low agricultural productivity case, specialization equilibria are not stable at the highest τ . As in the previous case, when trade costs are high, the only non-specialized equilibrium is autarky. However, what is new in this case is that, as trade costs fall and autarky becomes unstable, a trade equilibrium without specialization now becomes stable. In this equilibrium, both regions have enough manufacturing scale that shifting a small number of manufacturing workers one way or another is not profitable for workers. Given our assumption of a slight coastal comparative advantage in manufacturing, the coast exports manufactured goods in this equilibrium, but such goods are produced in both regions.

Figure B3 shows the existence of the different classes of equilibria in (A_f, τ) parameter space for values of A_f from 1 to 2.3 and τ from 0.01 to 0.6. In the upper left of the figure, where agricultural productivity is high and trade costs are low, the only equilibrium is one in which manufacturing takes place in both regions (area G). At this equilibrium, the coast has a comparative advantage in manufacturing, so it exports this good. Maintaining low trade costs but lowering agricultural productivity, there are a series of different equilibrium configurations: in area B, there are two equilibria: one with trade and manufacturing taking place in both regions, and one with trade and manufactures produced only on the coast. In area C, there are three equilibria: the two just listed as well as one in which manufactures are produced only in the interior. For even lower agricultural productivity, maintaining low trade costs, are only two equilibria, those where manufactures are produced in a single region (area D). Raising trade costs (i.e. moving from left to right in the figure), autarky appears as a possible equilibrium, although there are also possible equilibria where manufactures are produced in one or both regions (areas E and F). Finally, with high enough trade costs, the only possible equilibrium is autarky (area A).

Figure B3 demonstrates the path dependence we explore empirically in the paper. Circa 1800, all of the world was in the high trade cost and low agricultural productivity autarkic equilibrium at the lower right in area A. In the countries that developed early, the historical paths of rising agricultural productivity and transport costs that did not fall too quickly maintained the economy in areas A, E, or F, in all of which there was a stable autarkic equilibrium. Although we do not model persistence explicitly, our assumption is that in such areas, an economy that was formerly at the autarkic equilibrium will remain there. Finally, with further declines in transport costs, these early developing economies moved into areas B, C, or G, all of which lack an autarky equilibrium, but do feature an equilibrium with non-specialized trade. Again, in the absence of an explicit model of persistence, we assume that when the autarkic equilibrium disappeared, these economies naturally moved to the non-specialized trade equilibrium, which featured very similar distributions of population and economic activity between regions.

Today's developing world followed a different path, with the fall in transport costs relative to the rise in agricultural productivity occurring earlier. In this case, countries entered area D in the figure, where neither autarky nor non-specialized trade is a stable equilibrium. Manufactures will be produced in only one region, and it is natural to expect that this would be the coast, where productivity is higher. Once this specialization equilibrium has been entered, persistence would mean that further increases in agricultural productivity (moving into areas C and B) would not move the economy away from the equilibrium in which manufactures are produced in only one region. Thus, by the time of the mid- to late-20th century agricultural revolution in the developing world, urbanization (and manufacturing activity) is more concentrated in a smaller set of regions with better access to world markets.

Algorithm to solve the model

The algorithm begins by creating a vector of all possible interior populations and a corresponding vector of all possible coastal populations based on L and L_i such that:

$$L_c = L - L_i$$

To find equilibria, we cycle through these vectors in a loop. As a result, the following steps are carried out for a fixed population allocation between the interior and coast.

First, we create another vector of all possible L_{fi} values, ranging from 0 (no agriculture in the interior) to the entire interior population (everyone is employed in agriculture in the interior). Then, a corresponding vector is created of L_{fc} values. This vector is calculated based on food needs of the entire population, solving the following equation based on the text:

$$L_{fc} = \left(\frac{L\lambda - A_f L_{fi}^{1-\beta}}{A_{fc}} \right)^{\frac{1}{1-\beta}}$$

We subsequently cycle through these vectors in another loop, nested within the previous one. Consequently, the following steps are carried out for fixed regional agricultural labor forces *and* regional population allocations.

Within these two loops, we begin to find equilibria. If $L_{fc} < L_c$ and $L_{fi} < L_i$ (both regions have some manufacturing labor force), then we calculate L_{mc} and L_{mi} using the following equation:

$$L_{mr} = L_r - L_{fr}$$

Now that we have L_i, L_c, L_{fi}, L_{fc}, L_{mi}, and L_{mc}, we calculate prices in each region based on the average product of agriculture and manufacturing in each region (so that wages are equalized across sectors within each region):

$$p_{mr} = \frac{A_f L_{fr}^{-\beta}}{A_{mr} L_{mr}^{\epsilon} (1 - t L_{mr})^{1+\epsilon}}$$

Next, we determine which of the two regions is exporting manufactured goods. This can be determined by checking which of the two regions produces less food than its population requires. Then, we check if the inter-regional goods market clears by checking if prices in the exporting region are equal to prices in the importing region, adjusted for the iceberg trade cost τ. For most allocations of L_{fc} and L_{fi} this condition is not met, and the algorithm simply ends at this point and starts at the next allocation of L_{fc} and L_{fi}.

However, if this condition is met, manufacturing consumption per capita is calculated for each region. In the exporting region, manufacturing consumption is calculated by subtracting the quantity of manufactured goods that are exported from the total quantity of manufactured goods produced in the region, divided by the region's population. The quantity of exported manufactured goods is determined utilizing the fact that the inter-regional goods market clears. As a result, exported manufactured goods necessarily equals the quantity of imported food divided by the price of manufactured goods in the region. The quantity of imported food is determined by the gap in the region's food needs and food production in the region.

In the importing region, manufacturing consumption per capita is equal to total manufactured goods produced plus the quantity of imported manufactured good, divided by regional population. Analogously to the previous case, the value of imported manufactured goods is determined by the quantity of exported food divided by price of manufactured goods in that region.

Manufacturing consumption in each region is not necessarily equal at this point. As such, this data point is recorded as a "possible equilibrium," where every equilibrium condition is met *except* that manufacturing consumption is equal across regions. If manufacturing consumption is also equal across regions, then this data point is recorded as an "equilibrium."

If $L_{fc} = L_c$ or $L_{fi} = L_i$, then we have a corner solution where one region has no manufacturing labor force. In this case, L_{mc} and L_{mi} are calculated just like before. Prices in the region that has a manufacturing labor force are calculated using the average products of agriculture and manufacturing just like above. However, prices in the region that has no manufacturing labor force are now determined solely by adjusting the other region's prices by the iceberg trade cost. Next, manufacturing consumption per capita is calculated for each region. The region that has a manufacturing labor force obviously exports manufactured goods in this case. Manufacturing consumption per capita in this region is equal to total manufactured goods minus exported manufactured goods (determined just as before) divided by regional population. In the region with no manufacturing labor force, manufacturing consumption is just equal to imported manufactured goods divided by regional population.

This data point is recorded as a possible equilibrium. If manufacturing consumption is also equal across regions, then this data point is recorded as an equilibrium. This ends the loop through possible values of L_{fc} and L_{fi}.

Next we address the endogenous no-trade equilibria where, as the name might imply, there is no trade between the two regions. First we check if the current (fixed) population allocation is feasible in that each region can feed itself without any trade. Then, we calculate L_{fc} and L_{fi} based on each region's individual food needs, remembering that there is no trade between regions. As a result:

$$L_{fi} = \left(\frac{L_i \lambda}{A_f}\right)^{\frac{1}{1-\beta}}$$

L_{mc} and L_{mi} are then calculated using regional population and agricultural labor force. Prices for each region are also calculated based on the average product of agriculture and manufacturing so that wages are equalized across sectors in a region. Next, we check if prices are in the "no-trade band" where no amount of trade is profitable, or:

$$\frac{p_{mi}}{1-\tau} \geq p_{mc} \text{ and } \frac{p_{mc}}{1-\tau} \geq p_{mi}$$

If this condition is met, then there is no incentive for trade between regions. We then calculate manufacturing consumption per capita in each region as total manufactured goods divided by regional population. Since the regional populations are still fixed (i.e. no mobility between regions), manufacturing consumption is not necessarily equal across regions. This data point is recorded as a possible equilibrium; every equilibrium condition is met *except* that manufacturing consumption is equal across regions. If manufacturing consumption is also equal across regions, then this data point is recorded as an equilibrium. This ends the loop through possible values of L_i.

This ends the procedure for calculating equilibria for a set of parameters. Next, we check the stability of all "full equilibria," where all markets clear and manufacturing consumption is equal across regions. We define two types of stability. "Type 1 stability" occurs when there is no incentive to move between regions. "Type 2 stability" occurs when there is no incentive to move industries *within* regions (i.e. move from agriculture to manufacturing).

To check type 1 stability, we take 100 people from the coast and move them to the interior. We allow for all other markets to clear, but manufacturing consumption is not equal between regions (i.e. calculate the resulting “partial equilibrium”). If the people who moved have lower consumption than before, then the equilibrium passes the stability check.

To check type 2 stability, we take 100 people from agriculture and move them into manufacturing *in a single region,* which puts the labor market in that particular region in disequilibrium. We then calculate the agricultural labor force in the other region based on the food needs of the entire population. We hold regional populations constant, so manufacturing labor forces are determined by regional population minus agricultural labor force. Next, since we still allow the labor market in the other region (where people did not initially change sectors) to be in equilibrium, we calculate prices in that region using the average product of manufacturing and agriculture like in earlier steps. Prices in the region where the labor market is in disequilibrium are then determined by prices in the other region adjusted for the iceberg trade cost. Wages in agriculture and manufacturing are finally calculated for that region. If the manufacturing consumption of manufacturing workers rises above that of their neighbors in agriculture within the region, then the equilibrium fails the stability check. The stability test is passed if manufacturing consumption of manufacturing workers in, for example, the interior is below food workers in the interior. We note that under stability in this experiment in all examples, food workers in the interior have the highest consumption nationally and manufacturing interior workers the lowest, and vice versa for instability.

Reference

Duranton, Gilles, and Diego Puga,“Micro-foundations of urban agglomeration economies,” in *Handbook of Regional and Urban Economics*, Volume 5, Chapter 48, 2063-2117, Gilles Duranton, J. Vernon Henderson and William C. Strange eds. (Elsevier, 2004).

Figure B1. Equilibria with low agricultural productivity ($A_f = 1$)

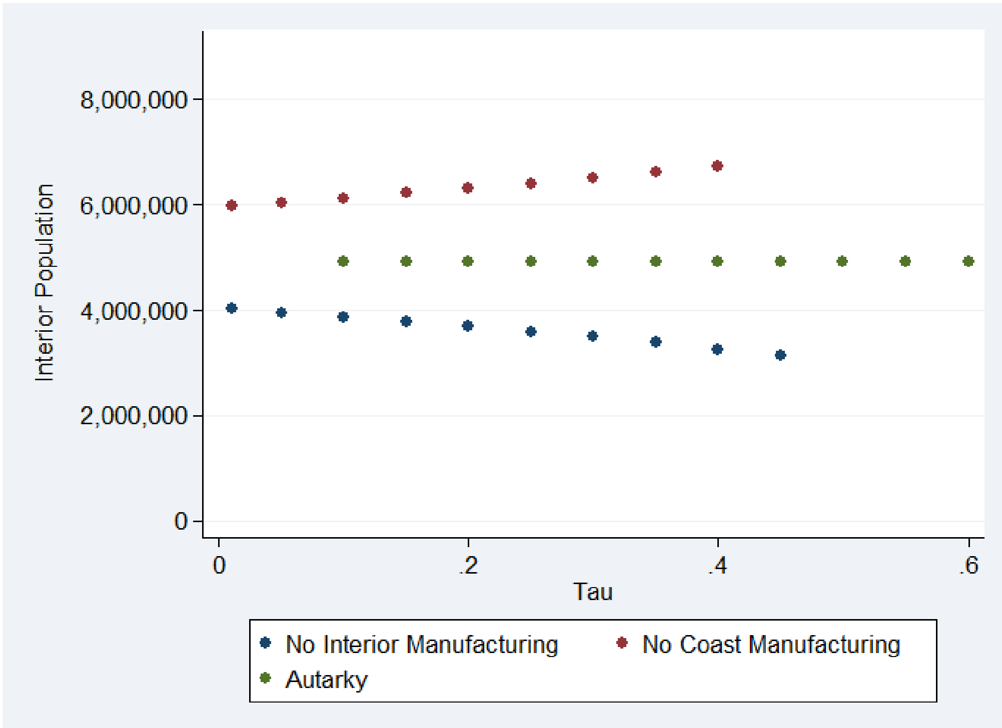


Figure B2. Equilibria with high agricultural productivity ($A_f = 1.5$)

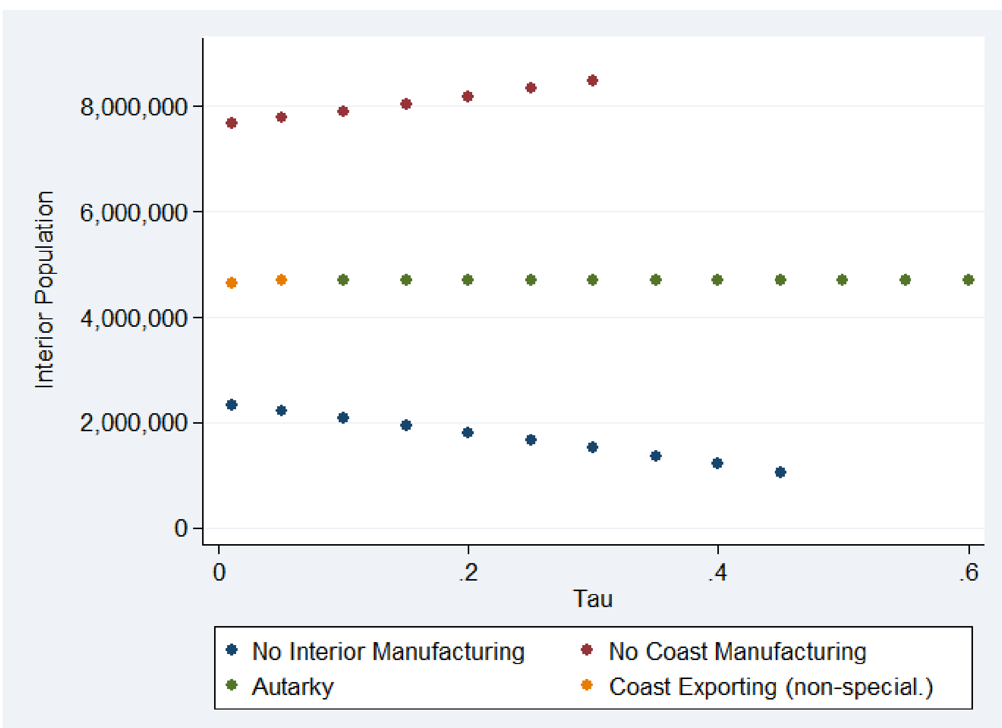
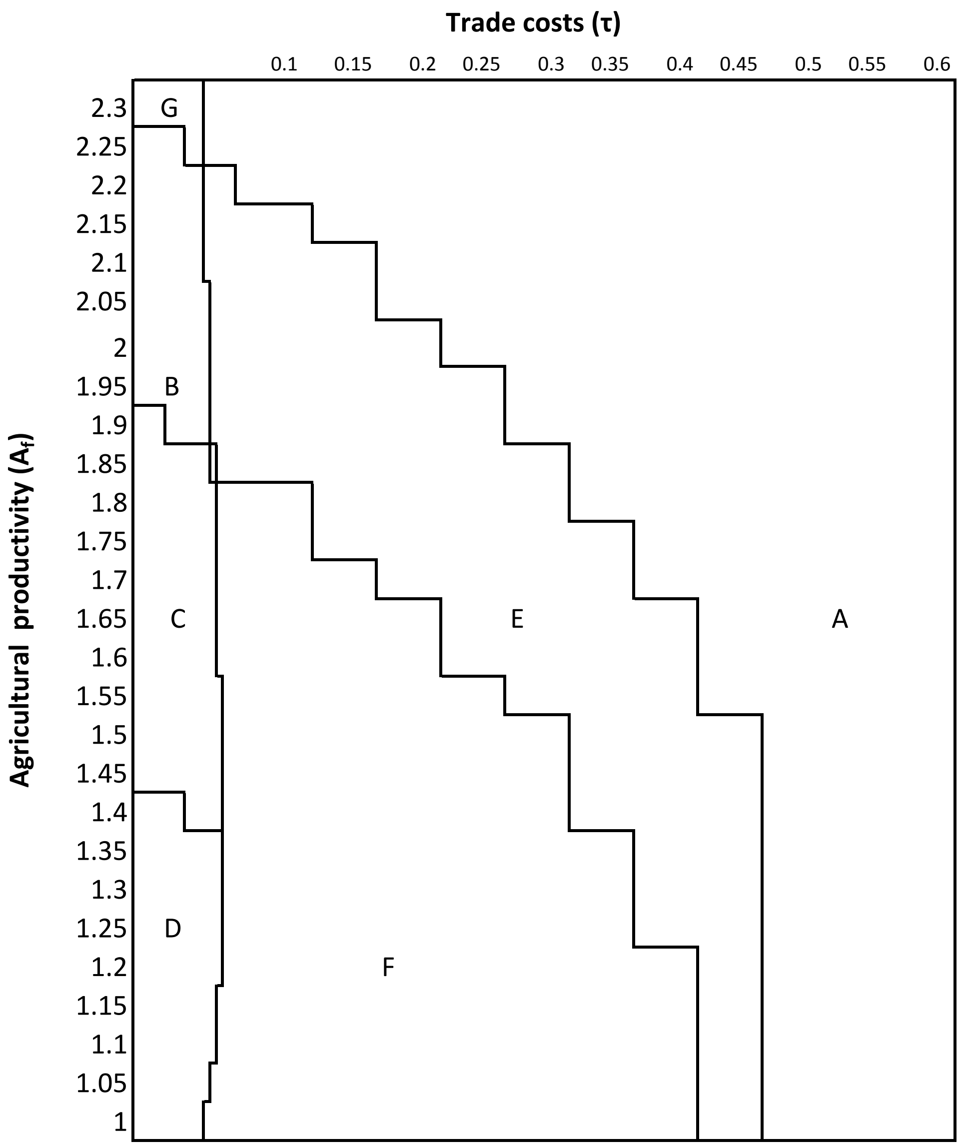


Figure B3. Equilibria in agricultural productivity-transport cost ($A_f - \tau$) phase space



Key (number of equilibria in bold)

A **1:** autarky

B **2:** trade with manufacturing in both regions or solely in coast

C **3:** trade with manufacturing in both regions, solely in coast, or solely in interior

D **2:** trade with manufacturing solely in coast or solely in interior

E **2:** autarky or manufacturing solely on coast

F **3:** autarky or trade with manufacturing solely in coast or solely in interior

G **1**: trade with manufacturing in both regions

Figure C1. Log Nonzero Lights in 2010

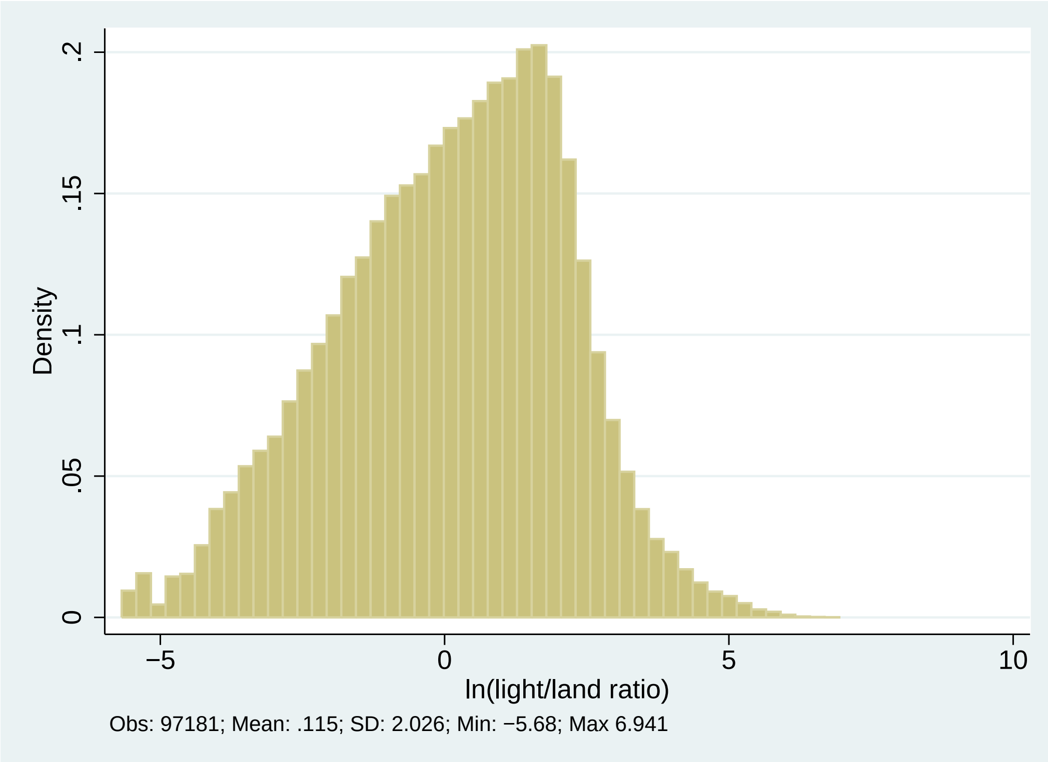


Table C1. Intensive and Extensive Margin R-squared

	No FE	FE
Both margins	0.467	0.577
Extensive margin, LPM	0.390	0.480
Intensive margin, OLS	0.258	0.360
Country FE, extensive margin		0.272
Country FE, intensive margin		0.229
Base, both margins	0.020	0.355
Agriculture, both margins	0.450	0.566
Trade, both margins	0.066	0.370

Notes: Each number represents an R^2 value from a separate regression on all geographic variables or the subset shown. In extensive margin rows, the dependent variable is a dummy for lights being positive, and the sample is the full global sample. In the intensive margin rows, the dependent variable is ln(lights), and the sample is grid squares with positive light values. FE stands for country fixed effects.

Table C2. Intensive and Extensive Margin Coefficients

	no country FEs		w/country FEs	
	Extensive	Intensive	Extensive	Intensive
ruggedness (000s)	-0.000441	-0.0282***	0.000933***	-0.0220***
	(0.000269)	(0.00189)	(0.000303)	(0.00211)
malaria index	-0.00644***	-0.0453***	-0.00497***	-0.0368***
	(0.000437)	(0.00428)	(0.000443)	(0.00288)
tropical moist forest	-0.00415	-0.190***	0.0389***	-0.404***
	(0.00980)	(0.0545)	(0.0109)	(0.0570)
tropical dry forest	0.0606***	-0.120*	0.179***	-0.0685
	(0.0113)	(0.0615)	(0.0127)	(0.0611)
temperate broadleaf	0.157***	0.603***	0.178***	0.753***
	(0.00926)	(0.0515)	(0.00988)	(0.0542)
temperate conifer	0.0466***	0.135**	0.100***	0.386***
	(0.0118)	(0.0604)	(0.0120)	(0.0627)
boreal forest	-0.180***	-0.423***	-0.0813***	-0.313***
	(0.0122)	(0.0676)	(0.0118)	(0.0675)
tropical grassland	0.0184**	-0.220***	-0.0878***	-0.870***
	(0.00755)	(0.0568)	(0.00877)	(0.0568)
temperate grassland	0.184***	0.147***	0.152***	0.0127
	(0.00892)	(0.0500)	(0.00991)	(0.0525)
montane grassland	0.0806***	0.391***	0.0776***	0.0231
	(0.0122)	(0.0645)	(0.0133)	(0.0671)
tundra	-0.230***	-0.594***	-0.176***	-0.334***
	(0.0134)	(0.103)	(0.0131)	(0.0960)
Mediterranean forest	0.218***	0.443***	0.0760***	0.256***
	(0.0130)	(0.0697)	(0.0123)	(0.0643)
mangroves	-0.0495**	-0.310***	0.0321	-0.365***
	(0.0198)	(0.102)	(0.0224)	(0.111)
temperature (deg. C)	0.0226***	0.0196***	0.0262***	0.0882***
	(0.000602)	(0.00445)	(0.000515)	(0.00345)
precipitation (mm/month)	-0.00160***	-0.00568***	-0.00118***	-0.00343***
	(0.0000647)	(0.000378)	(0.0000604)	(0.000349)
growing days	0.00111***	0.00520***	0.00134***	0.00445***
	(0.0000398)	(0.000227)	(0.0000385)	(0.000215)
land suitability	0.353***	0.802***	0.418***	0.740***
	(0.00776)	(0.0411)	(0.00784)	(0.0396)
abs(latitude)	0.00880***	0.00343	0.0176***	0.0520***
	(0.000498)	(0.00324)	(0.000375)	(0.00211)
elevation (km)	0.0455***	-0.270***	0.0907***	0.0391
	(0.00410)	(0.0270)	(0.00385)	(0.0250)
coast	-0.0000640	0.545***	-0.00200	0.561***
	(0.00440)	(0.0308)	(0.00517)	(0.0337)
distance to coast (000s km)	-0.0969***	-0.288***	-0.102***	-0.408***
	(0.00516)	(0.0328)	(0.00450)	(0.0289)
harbor < 25km	0.137***	0.524***	0.152***	0.452***
	(0.00740)	(0.0394)	(0.00824)	(0.0441)
river < 25km	0.113***	0.320***	0.129***	0.329***
	(0.00840)	(0.0467)	(0.00884)	(0.0528)
lake < 25km	0.0626***	0.644***	0.0756***	0.682***
	(0.0121)	(0.0776)	(0.0125)	(0.0814)
N	242184	97181	242184	97181

Notes: Each column reports OLS coefficient estimates from a separate regression. In extensive margin columns, the dependent variable is a dummy for lights being positive, and the sample is the full global sample. In the intensive margin columns, the dependent variable is ln(lights), and the sample is grid squares with positive light values. FEs stands for country fixed effects. Standard errors, clustered by 3x3 sets of grid squares, are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table C3. Full Nonlinear Differential Coefficient Results

	Education	Urbanization	GDP per capita
agriculture differential (α)	0.332***	0.194***	0.254***
	(0.0238)	(0.0209)	(0.0239)
trade differential (γ)	-0.650***	-0.393***	-0.526***
	(0.0178)	(0.0218)	(0.0321)
ruggedness (000s)	0.00507**	0.000734	0.0112***
	(0.00217)	(0.00209)	(0.00253)
malaria index	-0.0809***	-0.0749***	-0.0553***
	(0.00296)	(0.00277)	(0.00276)
tropical moist forest	-1.697***	-1.453***	-1.110***
	(0.0762)	(0.0743)	(0.0761)
tropical dry forest	0.0264	0.260***	0.385***
	(0.0893)	(0.0887)	(0.0910)
temperate broadleaf	1.080***	1.400***	0.747***
	(0.0668)	(0.0679)	(0.0765)
temperature conifer	0.626***	0.804***	0.278***
	(0.0745)	(0.0772)	(0.0858)
boreal forest	-1.203***	-1.175***	-2.233***
	(0.0794)	(0.0788)	(0.0983)
tropical grassland	-1.566***	-1.569***	-1.394***
	(0.0537)	(0.0526)	(0.0540)
temperate grassland	0.146**	0.321***	-0.330***
	(0.0582)	(0.0603)	(0.0693)
montane grassland	0.570***	0.355***	0.183**
	(0.0768)	(0.0811)	(0.0861)
tundra	-1.775***	-1.877***	-3.193***
	(0.0900)	(0.0919)	(0.115)
Mediterranean forest	0.261***	0.522***	-0.0133
	(0.0813)	(0.0878)	(0.0879)
mangroves	-1.936***	-1.438***	-1.289***
	(0.156)	(0.143)	(0.153)
temperature (deg. C)	-0.0547***	-0.0594***	-0.0807***
	(0.00153)	(0.00155)	(0.00175)
precipitation (mm/month)	-0.00856***	-0.00928***	-0.00913***
	(0.000392)	(0.000400)	(0.000426)
growing days	0.00658***	0.00605***	0.00604***
	(0.000260)	(0.000265)	(0.000285)
land suitability	2.550***	2.594***	2.629***
	(0.0551)	(0.0570)	(0.0621)
abs(latitude)	-0.0310***	-0.0360***	-0.0167***
	(0.00121)	(0.00121)	(0.00145)
elevation (km)	-0.599***	-0.605***	-0.684***
	(0.0195)	(0.0201)	(0.0228)
coastal	-0.0753	-0.143***	-0.00128
	(0.0700)	(0.0543)	(0.0664)
distance to coast (000s km)	-2.003***	-1.611***	-1.801***
	(0.0373)	(0.0383)	(0.0396)
harbor < 25km	3.011***	2.665***	2.435***
	(0.121)	(0.103)	(0.105)
river < 25km	1.251***	1.102***	1.036***
	(0.110)	(0.0940)	(0.112)
lake < 25km	0.170	0.523***	0.132
	(0.148)	(0.139)	(0.163)
N	227032	241995	180912

Notes: Each column reports non-linear least squares estimates of equation (3), for education, urbanization and GDPpc split variables. Standard errors, clustered by 3x3 sets of grid squares, are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.