The Global Spatial Distribution of Economic Activity:
Nature, History, and the Role of Trade

August 27, 2015

Vernon Henderson, London School of Economics
Tim Squires, Amazon.com
Adam Storeygard, Tufts University
David Weil, Brown University and NBER[1]

Abstract

We study the distribution of economic activity, as proxied by lights at night, across 250,000 grid cells of average area 560 square kilometers. We first document that nearly half of the variation in lights across these cells can be explained by a parsimonious set of physical geography attributes, which we divide into two categories: those primarily important for trade with the rest of the world, and those primarily important for agriculture. A full set of country indicators only explains a further 10%. We then show that the agriculture variables are relatively more powerful in explaining the location of economic activity within developed countries that slowed their population growth (and urbanized) earlier, while the trade variables have relatively more explanatory power in developing countries whose population distribution was in flux later. We interpret this in the context of a two-region model in which two technological shocks occur: one increasing agricultural productivity, and the other decreasing transportation costs. Our results are consistent with a world in which the biggest agricultural productivity increases occurred after the onset of mechanized transport in the developing world, but before them (or concurrent with them) in the developed world.

Key words: Agriculture, physical geography, development
JEL Codes: O13, O18, R12

[1] We thank Alex Drechsler, Joshua Herman, Young Min Kim, Kevin Proulx, Sameer Sarkar, Yang Shen, and Sanjay Singh for excellent research assistance, and seminar participants at Berkeley, LSE, and Williams for helpful comments and suggestions. Storeygard thanks Deborah Balk, Marc Levy, Glenn Deane and colleagues at CIESIN for conversations on related work in 2004-2006. Financial support from the World Bank's Knowledge for Change Program and from a World Bank-funded project at LSE and Oxford (1134034) on Policy Research on Urban Development in Africa is gratefully acknowledged.

1. Introduction

Three broad forces shape the spatial distribution of population as well as economic activity. The first is the physical setting, often called "first nature."[2] Some places are simply more amenable to human habitation and economic production than others. Overlaying these natural characteristics, a second broad force results from the interplay of agglomeration and congestion. For many economic activities, productivity is increased by being close to other people, through a combination of reduced transaction costs, gains from trade, specialization, and spillovers. At the same time, increased physical concentration of economic activity introduces congestion costs as well as raising transport costs for commodities that are produced using land (e.g. food -- see Gollin and Rogerson, 2014). The resulting system of cities represents a balance of these two forces, in which regions of agglomeration are dispersed over space. The third important force is historical persistence. Once established, agglomerations tend to persist, even if the forces that created them are no longer in operation.

In this paper, we begin by examining the predictive power of first-nature characteristics for the distribution of economic activity. Our primary dependent variable is information from satellite observations of lights at night at the quarter degree (longitude/latitude) scale, for roughly 250,000 grid squares per year. Our measures of first nature include data on climate and characteristics of the land surface, natural water bodies, and plant life (temperature, precipitation, elevation and ruggedness, coasts, navigable rivers, natural ports, and biomes). We are particularly interested in the relative importance of characteristics related to the possibility of trade (such as being located near a natural harbor) versus those associated with agricultural productivity.

We then turn our focus to understanding how the currently observed spatial distribution of population and economic activity *within countries* reflects the historical arc of economic

[2] As far as we are aware, this terminology was introduced to the economics literature by Krugman (1991), based on Cronon (1991).

development. We present a simple theoretical model that incorporates the effects of the two most important changes in the determinants of this spatial distribution: rising agricultural productivity and declining transport costs. Rising agricultural productivity lowers the fraction of the population that works producing food and so raises urbanization. Falling trade costs weaken the link between city locations and regions of high agricultural productivity. The time pattern in which these changes arrive will affect how first-nature characteristics are mapped into the eventual distribution of economic activity. To test the model, we consider different ways of splitting the countries in the world in samples of "early agglomeration" and "late agglomeration" sub-samples. Specifically, we let the data choose the cutoff values of different variables (education and urbanization in 1950) so that countries are sorted into one bin or another. We show that in "early agglomeration" countries, first-nature characteristics related to food production are more important in determining current settlement than are those associated with trade, relative to the case of "late agglomeration" countries.

The effects of first-nature geography on economic activity and population density are typically studied at the level of countries. Such an approach misses the obvious point that the average value of a geographic characteristic in a country is not equal to the average value of that characteristic in the places where people actually live and work. Furthermore, most existing research takes a partial view of the processes described above. For example Davis and Weinstein (2002), Bleakley and Lin (2012), and Michaels and Rauch (2013) all establish the importance of persistence of city locations in specific historical settings. But these papers do not establish, for example, how much of the overall variation in location of population observed today is explained by history. Similarly, Mellinger, Gallup, and Sachs (2000) and Rappaport and Sachs (2003) investigate the role of coasts, for both productive and amenity reasons; Nunn and Puga (2012) look at the effect of terrain ruggedness; and Nordhaus (2006) and Nordhaus and Chen (2009) look at a whole suite of geographic factors using coarse subnational data. But these latter papers are generally silent on the role of history in determining the spatial patterns observed.[3]

[3] See also Gennaioli et al. (2013, 2014), who regress subnational income and growth on geographic factors along with institutions, population and human capital measures, for a sample that covers much of the world but largely

Our work also has several similarities to Mesbah *et al.* (2014). They estimate the year in which a given half-degree grid cell passed various urbanization rate thresholds. Their urban and rural population data are gridded estimates for the past 2,000 years from Klein Goldenwijk *et al*. (2011). Mesbah *et al*. regress the date of urbanization on a cultivation suitability index, distance to coast, a river navigability proxy, frost, and elevation, finding significant predictive power for all of these variables except elevation. We view their work as complementary to ours, in that they examine the determinants of early urbanization and we examine the effect of early urbanization, along with other factors, on outcomes today.

2. Data

In order to carry out this exercise, we need measures of economic activity and several components of physical geography, all available on a global scale.

Our proxy for economic activity is night lights (Figure 1). Unlike Henderson, Storeygard and Weil (2012) and nearly all quantitative work on lights, we use the radiance-calibrated version of the data (Elvidge, et al. 1999; Ziskin et al. 2010). In normal operations, the light detection sensor is very good at detecting low levels of light in small cities. However, the strong amplification that enables this detection also saturates the sensor in the most brightly lit places, including the centers of most of the largest 100 cities in the United States, so that their values are top coded. The 2010 Global Radiance Calibrated Nighttime Lights dataset we use combines the high magnification regime for low light places with a lower magnification regime for more brightly lit places. Thus nearly all topcoding is removed, with minimal loss of information about low light places. The lights data are distributed as a grid of pixels of dimension 0.5 arc-minute resolution (1/120 of a degree of longitude/latitude, or approximately 1 square kilometer at the equator).[4]

excludes Africa. Related work in the trade literature (e.g. Allen and Arkolakis 2013) have used a more structural approach and focused on the United States, where data on subnational trade flows are available.

[4] Available at http://ngdc.noaa.gov/eog/dmsp.html.

Our other variables of interest are all reported at several different geographic scales, ranging from 1/120 of a degree to 1/2 degree. For analysis, we convert them all to a grid of 0.25-degree squares, with each square covering approximately 770 square kilometers at the equator, decreasing with the cosine of latitude. This scale is a compromise between the fine detail observed at the native resolution of several datasets and the computational practicality of coarser cells. It also allows us to be less concerned about spatial autocorrelation than we would be at finer scales, and to reduce true spillovers as well. At this resolution we already have well over 200,000 grid squares.

Variables originally reported at scales smaller than 0.25 degrees are aggregated with an appropriate function. In the case of continuous variables, values for our grid squares represent the mean or sum of all input cells falling within them, as appropriate. So for example, the night lights measure for each quarter-degree grid square is the sum of the 900 component raw lights pixels. In the case of categorical variables, we assign the modal value. For variables originally reported at scales larger than 0.25 degrees, each 0.25-degree grid square receives the value of the larger input cell into which it falls (i.e. the original value is autoreplicated).

We use lights as the measure of economic activity because it is measured consistently worldwide at the same spatial scale. Alternatively, we could have considered population. There are two main sources of global population data. Landscan[5] uses other geographic data to interpolate population within census geographic units, which has the potential to bias our estimates. The Gridded Population of the World (GPW)[6] uses population data exclusively, assuming uniform population density within enumeration units larger than its native (2.5 arc-minute) grid resolution. On average, this means that population estimates are more heavily smoothed in poorer countries with lower statistical capacity, as well as in more sparsely populated regions. This could also bias our results.

[5] http://web.ornl.gov/sci/landscan/
[6] http://sedac.ciesin.columbia.edu/

Apart from measurement issues, the lights data are conceptually different from population, because they reflect intensity of economic activity, which is a combination of population and income per capita. Assuming a reasonable degree of population mobility within countries in the long run, in the presence of country fixed effects in our analysis below, lights variation will mostly reflect variation in the spatial distribution of population.

To analyze the variation in economic activity across locations, we then define three sets of variables, while acknowledging that the boundaries between the sets are somewhat permeable. The general intent is to distinguish variables relating to agricultural productivity from those relating to costs of trade. With that in mind, we first define a set of two variables, a malaria index and ruggedness, which we didn't put in either camp because they affect both. Malaria affects human ability to live in an area regardless of the economic activities they perform, and ruggedness increases the cost of both trade and agriculture. The index of the stability of malaria transmission, based on species-specific measures of human biting rates and climate predictors of mosquito survival, is from Kiszewski et al. (2004). Ruggedness is based on Nunn and Puga (2012). We correct the Nunn and Puga measure to account for the fact that two east-west neighboring cells at high latitudes are closer than two east-west neighboring cells at low latitudes, biasing their measure downward at high latitudes.[7]

We consider six continuous agriculture variables, temperature, precipitation, length of growing period, land suitability for agriculture, elevation, latitude, and a set of 14 biome indicators. The temperature variable is a long run (1960-1990) average of UEA CRU et al. (2013) based on Mitchell and Jones (2005) and precipitation is the Wilmott and Matsuura (2012) measure averaged over the same period. Length of growing period, in days, is from FAO/IIASA (2011).

[7] Applying this corrected measure to the main regression in Nunn and Puga (2012) leads to virtually no change in the point estimate of the variable of interest and an approximately 15% increase in its standard error. We also area-weight the average to follow Nunn and Puga. In practice, area weighting has minimal impact within our small units.

Land suitability is the predicted value of the propensity of a given parcel of land to be under cultivation based on four measures of climate and soil, from Ramankutty et al. (2002).[8]

Elevation, in meters, from Isciences (2008). While high elevation locations often have poor transport, we believe that once distance to various types of water transport (see below) and ruggedness are controlled for, it is best interpreted as an agricultural variable. Furthermore, while ruggedness and malaria had similar effects on economic activity across different sample splits discussed below, elevation, like many trade and agricultural variables, did not. In practice, the choice of whether to place elevation in the agriculture category or the "both" category is inconsequential to the main results. Finally, we control for latitude which has affected agriculture even net of climate, because the diffusion of domesticatable plant and animal species happens more easily within narrow latitude bands.

Biomes are mutually exclusive regions encoding the dominant natural vegetation is expected in an area, based on research by biologists. The distribution of 14 biomes is from Olson et al (2001). We combine "tropical and subtropical dry broadleaf forests" with "tropical and subtropical coniferous forests", and "tropical and subtropical grasslands and savannas and shrublands" with "flooded grasslands and savannas" because each pair is broadly similar, and because the second member of each pair contains less than 1% of cells globally. We exclude areas covered by permanent ice.

Our five trade variables focus on access to water transport. We calculate Euclidean distances in kilometers from cell centroids to the nearest coast, navigable river, and major lake using the

[8] Because several variables are only defined or reported for grid squares containing land, and different datasets have different effective definitions of the land surface, as noted below, values for some variables are imputed (or "grown") as the mean (continuous) or mode (categorical) of their eight 0.25-degree grid square neighbors. This process is repeated up to two times until nearly all cells containing land based on our coastline dataset have values for all variables. Between the two iterations, interpolated values assigned to cells containing no land are dropped, so that imputation cannot occur across large water bodies. The only land cells without data following this spatial interpolation process are small islands. Land suitability, biomes, temperature and precipitation are grown twice, and length of growing season is grown once.

Fuller isohedral map projection, and great circle distances to the nearest natural harbor.[9] Our specifications include indicators for the presence of each of these 4 within 25 km of a cell centroid, as well as a continuous measure of distance to the coast. In each case, we take a more systematic approach to characterizing the universe of waterbodies than previous work. Vector coastline data are from NOAA (2011; "low" resolution), based on Wessell and Smith (1996). The same data are also gridded at 0.5 arc minutes in order to determine the fraction of these 0.5 minute cells in a quarter-degree grid square that fall on land. Our universe of rivers is those in size categories 1-5 (on a scale of 1-7) of the river and lake centerline dataset from Natural Earth (2012). We restrict to river segments that are navigable, having determined the navigability of each river using a variety of text sources.[10] Lakes data are from the Global Lakes and Wetlands Database produced by the World Wildlife Fund and the Center for Environmental Systems Research, University of Kassel (Lehner and Döll 2004).[11] We define 29 lakes as major based on their their surface area is greater than 5000 square kilometers, having excluded four that were wholly created by dams. Natural harbors data are port locations digitized from US Navy (1953), restricting to ports defined there as natural harbors.

Table 1, columns 1 and 2 report summary statistics for all of these variables.

3. The measure of economic activity and initial results

3.1 The measure of economic activity and specification

[9] All available GIS software of which we are aware performs this calculation in the plane, and thus requires choosing a projection (see Tobler (2002) for a critique). No projection preserves distance in general, and many, including the Plate Carrée implicitly used in most economics research, can induce substantial error. Spherical point-to-point distances, in contrast, can be calculated easily in many software packages. We use Fuller's icosahedral projection, which we believe is relatively well-suited for the task, and has not previously been used for such quantitative purposes in any literature of which we are aware.

[10] Full list available upon request

[11] http://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-large-lake-polygons-level-1, accessed 2014/8/14

As seen in Figure 1, the lights data convey a great deal of information about the location of economic activity. At the country level, the cross-sectional correlation between emitted light and GDP is 0.90. More importantly for our purposes, the lights map out the location of economic activity within countries. Note that lights represent total economic activity, which is a combination of the number of people and the activity level person. Lights are comparably bright in northern India and the eastern United States, because while economic activity per person is lower in India, population density is higher in many places.

The land area falling within each grid square varies, both because some are partially covered by water or permanent ice, and because the surface distance between lines of longitude varies with the cosine of latitude. To deal with the first problem, we normalize our lights measure by land area, to yield the intensity of light emitted from each grid square, so we are looking ultimately at the allocation of the intensity of economic activity. In particular we sum the lights in each grid square. For grid squares that are only land, we are summing over 900 pixels. For cells with not entirely land, we inflate the sum by the inverse of the fraction of the 900 pixel falling on land. This is equivalent to averaging over land cells only.[12]

One notable limitation of the lights data is that they are censored. 59% of our grid squares emit too little light for the satellite to detect. Since nearly all grid squares contain population and thus presumably emit some level of light, we consider this a censoring problem. Values as high as 1 or 2 are almost exclusively interpreted as noise and recoded to zero, in initial processing by NOAA. The lowest non-zero value of the sum of lights adjusted by fraction of pixels over land is 3.05. We assign this value to all , to avoid any jumps in the data when we move from so called zero to minimal readings.[13]

[12] While cell area varies with latitude, the light readings are densities.
[13] Alternatively, we could estimate a Tobit model, which is the traditional way to capture censoring. OLS avoids the Tobit error structure and provides a more intuitive measure of goodness of fit, which is our focus. Estimated coefficients for the analogous tobit models (with and without country fixed effects) are almost exclusively of the same sign and are mostly larger in magnitude.

The base formulation for grid square i in country c in time t is thus

$$ln(light_{ict}) = X_{ict}\beta + \varepsilon_{ict} \qquad \text{if } light_{ict} \geq 3.05 \tag{1}$$
$$= 1.115 \qquad \text{otherwise}$$

where $light_{ict} = \dfrac{900 \sum_{j \in i} light_{jct}}{\sum_{j \in i} land\ pixels}$

We also consider the intensive and extensive margins separately. Figure 2 plots the distribution of the dependent variable excluding the bottom code.

We emphasize two further points about equation 1. First, it is a very simple functional form. With such a large number of covariates, a 2nd order Taylor series has hundreds of terms, which improves the fit but limits interpretation substantially. Second, we consider multiple error structures. We show all results with and without country fixed effects. This distinction is critical: the fixed effects regressions rely on within-country variation and account for the allocation of activity within a country, in a context where we expect a higher variance across countries than across grid cells within a country. Errors are clustered by 3-by-3 squares of cells to account for spatial autocorrelation. Conley standard errors, used in alternative specifications below, tend to be smaller.

3.2 Basic results

Columns 3 and 4 of Table 1 report the results of a regression of our lights variable on the full suite of physical geography characteristics without and with country fixed effects. The coefficients with and without fixed effects are generally of similar magnitudes and are of the same sign for all covariates except the mangroves indicator, which applies to 0.4% of the sample. Because of the high potential for co-linearity among the right hand side variables, it may be that looking at the change in a particular coefficient in comparing the specification with country fixed effects to the specification without them is not very informative. As an alternative, we created fitted values from the specifications in columns (3) and (4), in the latter case suppressing the country fixed effects (i.e. setting all the country dummies to zero when forming fitted values). The correlation of the fitted values is 0.860, suggesting that the two specifications provide very similar predictions of which regions have high light density. In other words, the geographic forces that drive the allocation of economic activity within and across countries are similar. In Figures 3a and 3b we plot the fitted values of lights under the two specifications. The absolute scales differ because when we omit the FEs themselves, the FE predictions are all relative to the base country. Nonetheless, the relative variation in lights within continents and countries is similar in the two figures.

In columns 3 and 4, coefficients on covariates are largely in the expected direction. Most biomes have significantly more lights than deserts (the reference biome); only boreal forests, tundra, and perhaps surprisingly, tropical grasslands, have significantly less. Being near the coast, lakes, navigable rivers and natural harbors is associated with more lights, as is a longer growing season and higher agricultural suitability. Net of growing season and land suitability and the biomes, higher temperatures and lower precipitation are associated with more lights, perhaps because of their residential consumer amenity value. In an alternative specifications excluding growing season, land suitability, and the biomes (not shown), precipitation has a positive effects overall as might be expected based on agricultural productivity. When entered in quadratic form (not shown), both temperature and precipitation increase lights at a decreasing rate. In the base formulation, net of ruggedness, higher elevation is associated with more lights.

As reported in Table A1, using column 3 as an example, Conley (2008) standard errors using a kernel of radius 40 km (similar to clustering for immediate queen neighbors) are slightly larger are smaller than the ones in column 3 by 10-15%. At longer cutoff distances such as 100 km, Conley standard errors increase further, but most t-statistics are still very large. In any case, our ultimate goal is not to establish precise marginal effects, and we believe that our cluster design, which is substantially easier to compute for our many variants, yields reasonable estimates of standard errors.

The most important numbers for our exposition in Table 1, columns 3 and 4, are the R-squared values. These 23 variables account for 46 percent of the variation in lights globally in column 3. We consider it remarkable that such a parsimonious specification can account for so much of the variation in global economic activity, without explicit regard to agglomeration or history. Of course to the extent that grid cells with better characteristics have neighbors with better characteristics (which we will see later is highly likely) and there are agglomeration forces, some of the light intensity associated with more lights in better places represents agglomeration. And of course there are within-grid square agglomeration forces. In short, these coefficients are reduced form estimates, which in part capture forces of agglomeration (see Section 7). In column 4 we add country fixed effects. While on their own these account for 34 percent of light variation, they only increase the column 4 R-squared by 11 percentage points relative to column 3. Country-level variation adds relatively little once physical geography factors are accounted for. Conversely, the geographic factors add 23 percentage points in explaining variation, on top of the fixed effects.

Table A2 reports OLS estimates of the effects of the same variables on the extensive and intensive margin of lights. Sign patterns for covariates across margins and with and without FEs are mostly the same but there are some distinct differences both across margins and FEs, especially for the 3 tropical biomes measures and elevation. Higher elevation increases the probability of being lit but is associated with lower light intensity, conditional on being lit.

Table 2 reports R-sq for a variety of specifications, exploring the role of different variable sets in explaining variation in lights in more detail. Column 1 excludes country fixed effects and column 2 includes them. The first row repeats the Rsq's from Table 1. Rows 2 and 3, report the extensive and intensive margin Rsq's, respectively. Although the R-squared values are not strictly comparable across margins, it is nonetheless striking that the extensive margin, a linear probability model, has a relatively large R-squared value of 0.39 without fixed effects and 0.48 with them. As shown in rows 4 and 5, country fixed effects alone capture differing levels of economic development (and underlying cultural and institutional differences), explaining more than 20 percent of both extensive and intensive margin variation.

In rows 6-8, we start to explore the relative role of trade and agricultural variables. Row 6 shows that the two base controls on their own explain little and that FE's explain a lot. In row 7, agricultural variables on their own have high explanatory power, 0.44 without fixed effects, and 0.56 with them. Row 8 suggests that the short list of trade variables on their own explain much less of lights variation and add little to the explanatory power of country fixed effects. However, as we will show, these relative contributions vary importantly between early- and late-developing countries.

4. Model

As suggested in the introduction, the effect of physical geography is modulated by history. Changes in productive technologies certainly have changed the importance (or relative prices) of different first nature factors. We focus on two critical changes. Over the last two centuries, the cost of transporting goods has fallen dramatically worldwide and even within countries. Within low productivity agricultural developing countries, by say 1950 and often much earlier, transport costs had fallen dramatically with the building of colonial rails and roads and the use of trucks. Donaldson's (2015) paper on rails in India under the Raj is instructive as well as work on rails in Africa by Jedwab and Moradi (2014 and 2015). Second there was the agricultural revolution in today's developed countries inducing structural transformation with many models reviewed in Desmet and Henderson (2015). That transformation released labor from agriculture to agglomerate in cities. That transformation in today's developing countries has been slower, occurring in Latin America after 1950 and still on-going in most of Asia and Sub-Saharan Africa today.

We develop a model in which the order in which these two changes occurred may influence the possible spatial distributions of economic activity. Consider small- medium size countries which have say a coastal and a hinterland region.. As we will argue below, those countries which experienced the agricultural revolution before much of the dramatic drop in transport costs had local agglomerations in all regions develop early. Trade was still difficult allowing local manufacturing to thrive in larger agricultural regions with population freed up to move into the local city. We will argue that these agglomerations in large size agricultural regions regions then persisted after the period of most dramatic drop in transport costs. In contrast, today's developing countries started with semi-autarkic regions, each with large agricultural populations and small urban ones. When these countries experienced substantial drops in transport costs, before they underwent structural transformation, that allowed these small urban populations to attain more efficient scale by agglomerating into one major, say coastal city within the country. Lowered transport costs allowed trade of manufactures across the regions (as opposed to just local production). Then once structural transformation starts to occur, these large agglomerations persist and grow, with little big city development in the hinterlands..

How do model this and how do we implement the ideas empirically? We start with the model.

4.1 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.

For either region $r = c, i$, 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 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 $\beta > 0$. Cities produce manufacturing. Average product per unit of labor in the urban sector producing manufactures is $A_{mr}(v + L_{mr}^{\epsilon})$, where the v allows labor output by the marginal labor entrant as $L_{mr} \to 0$. We will allow manufacturing productivity to potentially differ across regions, so there is comparative advantage. Each worker is endowed with one unit of time, to be used for labor and commuting in the city as in classic urban models (see Duranton and Puga, 2004, for a review), so labor supplied is $(1 - tL_{mr})$, where $0 < t << 1$ represents unit-distance commuting costs in the city. People live on fixed-size lots (with zero opportunity cost) in cities paying differential rents according to distance and commuting from the city center. Rental income is redistributed as an (equal) dividend to all city residents. Average product in the city is thus $A_{mr}(v + L_{mr}^{\epsilon})(1 - tL_{mr})^{1+\epsilon}$, and total production of the manufactured good is $A_{mr}(v + L_{mr}^{\epsilon})(1 - tL_{mr})^{1+\epsilon}L_{mr}$.[14] The size of the manufacturing labor force that maximizes the average product of labor in manufacturing is the solution to $\varepsilon = tL_{mr}[1 + v(1 + \varepsilon)L_{mr}^{-\varepsilon}]$. Since the right hand side of this equation can be shown to be increasing in L_{mr} , it has a unique solution, which can be shown to be a maximum.

Food, which is traded costlessly between regions as in the standard NEG model, is the numeraire good. Preferences are such that each worker consumes a fixed amount of food γ , and spends the remainder of her value of net average product on the manufactured good. Welfare for any person in region r is then is equivalent to consumption of the manufactured good, $(w_r - \gamma)/p_{mr}$, where w_r is the real 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

$$L_r = L_{fr} + L_{mr}, \quad r = c, i \tag{2a}$$

[14] Following Duranton and Puga (2004), in a linear city, where each worker is endowed with 1 unit of time and working time is *1-4tu* where *u* is distance from the city center and *4t* unit commuting costs, it is easy to derive expressions for city labor force (by integrating over the two halves of the city), for the city rent gradient (equating rent plus commuting costs for a person at *u* with that of a person at the city edge where rents are 0, so they are equally well off in equilibrium), and for total rents, each as a function of population. Any resident's income based on wages times working time after subtracting rents paid and giving their share in total urban rent income is given by the expression in the text.

$$L = L_c + L_i. \tag{2b}$$

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

$$A_f L_{fr}^{-\beta} = p_{mr} A_{mr}(c + L_{mr}^{\epsilon})(1 - tL_{mr})^{1+\epsilon}\,,\ r = c,\ i\,. \tag{3}$$

Free migration equalizes per person welfare, or manufacturing consumption across regions (assuming both regions are populated), so that:

$$\frac{A_f L_{fc}^{-\beta} - \gamma}{p_{mc}} = \frac{A_f L_{fi}^{-\beta} - \gamma}{p_{mi}}\,. \tag{4}$$

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

4.2 Autarkic equilibrium

If there is no trade between regions, clearing of the manufacturing good market in each regions requires:

$$L_r(A_f^{-\beta} - \gamma) = p_{mr} A_{mr}(c + L_{mr}^{\epsilon})(1 - tL_{mr})^{1+\epsilon} L_{mr}\,, \tag{5a}$$

or alternatively, using the agricultural market

$$\gamma L_r = A_f L_{fr}^{1-\beta}\,. \tag{5b}$$

Given $L, A_f, A_{mr}, \beta, \epsilon, \gamma, t$, the eight equations implied in (2)-(4) and (5b) specify equilibrium in the distribution of labor and the price of the manufactured good wherever it is produced (

$L_{mc}, L_{mi}, L_{fc}, L_{fi}, L_c, L_i,\ p_{mc}, p_{mi}$).[15]

4.3 Trade equilibrium with both regions producing manufactures

If transport costs are sufficiently low, we can have trade with manufactures in both regions if the regions have differential comparative advantage. If they are identical and have sufficient manufacturing scale (say, beyond the point where average product is maximized) then there would be no trade. We will generally allow 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 by any region, the two within-region goods market clearing conditions 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:

$$\gamma L = A_f L_{fc}^{1-\beta} + A_f L_{fi}^{1-\beta}. \tag{6}$$

Given comparative advantage in the coastal region, our non-specialization equilibria have the manufacturing export region being the coast, although we look for equilibria where the interior is the exporting region. The price of the manufactured good on the coast is p_{mc}, and it is related to the interior price by an arbitrage condition:

$$p_{mi} = p_{mc}(1-\tau)^{-1}, \tag{8}$$

where τ is an iceberg trade cost between the two regions. Given $L, A_f, A_m, \beta, \epsilon, \gamma, t, \tau$, the eight equations embedded in (2)-(4), (7) and (8) specify an equilibrium in the distribution of labor and the price of the manufactured good.

---

[15] To see that these represent eight equations, note that 2a, 3 and 5b each must be fulfilled for each region.

4.4 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 (3) for that region.

4.5 Solving the model

For any set of parameters and value of transport costs, we solve the model as follows. We have 4 types of possible equilibria: autarkic, trade without absolute specialization, and absolute specialization by one region or by the other. We pick an allocation of population to the interior region (with the coastal population being the reminder of national population and suspend equation (4) (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 (4)). Then, for each type of equilibria, for each allocation of population between regions we plot these two regional consumptions. If they intersect that is a potential equilibrium. Details of the actual algorithm used are in Appendix xx.

Actual equilibria are subject to two stability conditions. Type 1 stability is with respect to small changes in regional population allocation, assuming within region labor markets and 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 designed to deal with whether if we add a small number of workers to a non-existent or small manufacturing sector in a region that induces further agglomeration. For example, we take 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 have goods markets clear and have coastal labor markets clear. Equilibria are unstable if interior manufacturing workers then have the highest consumption of any workers in the country. They are stable if interior food workers then have the highest welfare (manufacturing consumption) of workers anywhere and interior manufacturing workers the lowest. Type 2 stability is crucial in determining if a region with no manufacturing would remain that way if by chance a small number of workers started manufacturing there. [will need footnote on whether the cases described for type 2 stability cover the outcomes we encounter]

Details of the solution method as applied in solving examples are given in the Appendix, including stability implementation. 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 existence of stable specialization equilibria enhances as τ falls.

4.6 Analysis of possible equilibria

As in many of these models, closed form solutions and their properties are elusive. We illustrate the basic issues with some examples. For these we start with

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

For these parameters average product in cities 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 then look for all the specialization and non-specialization equilibria that exist and are stable for different values of transport costs, τ, starting at 0.9 and going to almost 0. There are two sets of specialization equilibria: the coast only producing food and the interior only producing food. For non-specialization, at high τ these are regions are in autarky while at lower τ, if equilibria exist and are stable, they are trade equilibria.

The key analysis looks at how these patterns vary between a situation of low agricultural productivity which in this case has $A_f = 1$ and high agricultural productivity where we set $A_f = 1.5,$ for example. In the former case, at least with non-specialized manufacturing, there is relatively little manufacturing employment to support a city of sufficient size to exploit scale economies in any one region. In the second case, with much less labor needed to produce the required food, there is a lot more manufacturing employment to spread around.

We start with $A_f = 1.$ we show the solution method for the different types of equilibrium in Figure 4. There, for one value of $\tau = zz$ we plot the consumption per worker in each region (not imposing equation (4)) as a function of interior population. Where the curves cross is a potential equilibrium.

Then Figure 5 shows the specialization and non-specialization stable solutions for different costs of trade, τ, as graphed against the population of the interior region for each solution. These are so called pitchfork figures. As noted above there are two sets of specialization equilibria for the outer prongs: one where the interior produces only food and one where it is the coast doing that. For specialization equilibria there are none with this figure with low $A_f = 1$ at high costs of trade. Say the interior produces no manufacturing. Then because trade costs are so high if we move a small number of workers to start manufacturing in the interior that will be profitable despite the limited scale. Such an equilibrium is type 2 unstable. As τ falls, given limited scale effects in any one city, the advantage for scale effects of having all manufacturing on the coast dominates, so starting a small scale operation in the interior is not profitable for those workers. Note the allocation of workers to manufacturing in the coast is less that the city size which maximizes average product. Thus specialization equilibria start at lower values of τ, and then their existence persists persists as τ falls to 0. For the middle prong, at high τ, the equilibria are autarkic ones. We have no trade and limited manufacturing scale in each region., but trade is too costly for it to be profitable to workers to enhance scale in one the two regions, so these equilibria meet type 2 stability. As trade costs fall, it becomes potentially profitable to trade noting that the coast has a comparative advantage in manufacturing. The difficulty is that once it is profitable to trade it is also profitable to enhance manufacturing scale in one region relative to the other. Thus it turns out in this figure there are no stable trade equilibria at all and no autarkic equilibrium at lower values of τ .

In Figure 6 we turn to our second case, post an "agricultural revolution" where rises to $A_f = 1.5$. Now there are lots of manufacturing workers to go around. We use the same 3 prong diagram. As before specialization equilibria do not exist at the highest τ . However in this case at intermediate and then lower values of τ the large specialized city in one region or the other that produces manufacturing very competitively relative to starting up at a small scale in the other region. We note in these models urban scale economies vs commuting diseconomies play out so that output per worker rises very quickly from low scale, the rise slows, then output per worker peaks and then the decline is very slow past the peak (and there is always the option of having a second city in one region or the other not explicitly introduced here).

What about the middle prong? Again at high costs of trade there is autarky. Then at some point as trade costs fall, with comparative advantage slightly on the coast, it pays to have the coast relatively specialized and exporting some of the interior's manufacturing consumption to it, with the interior exports food. These trade equilibria persist as τ falls. Both region have a reasonable manufacturing scale so that shifting a small number of manufacturing workers one way or another is not profitable for workers, and stability holds.

Given these two figures, our story which we explore empirically in the next section, is as follows. Our world starts in the 18th century with high trade costs and low agricultural productivity. Regions are in autarky within countries, since as both figures suggest, specialization equilibria cannot persist. Then in today's developed world there is an agricultural revolution. In these countries history starts with non-specialization in Figure 5 and with the revolution both regions expand urban production to achieve reasonable urban scale as in Figure

6. After this agricultural revolution then transport costs fall in the late 19th and early 20th century everywhere. In today's developed world, this simply moves along the country along the middle prong in Figure 6, into a situation of inter-regional trade and relative specialization, which allows the regions to exploit their comparative advantage. Note with scale effects, the country benefits from both trade and factor movements; they are not perfect substitutes as in the old 2x2x2 Heckscher-Ohlin trade models. But the interior with its agricultural comparative advantage has an urban agglomeration as well as the coast.

For the developing world, it is different. There with persistent low agricultural productivity, in Figure 5, transport costs fall. As they fall, at some point, non-specialization equilibria disappear, because with low costs of trade and low urban scale, having industry in both regions is not an equilibrium. One region or the other specializes in manufacturing and the other does only food production. Since the coast has a modest comparative advantage, type 2 instability plays out for it first, in which case one could argue it will be the coast region which contains manufacturing. For these countries the agricultural revolution comes later (say post World War II). At this point they start increased urbanization and the development of manufacturing as in Latin America and parts of East Asia and then later the rest of Asia. Relatively small coastal cities become huge urban agglomerations, while hinterland regions do not have large agglomerations.

In our data we will see this as inducing a stronger correlation between agricultural fundamentals and lights concentration in the developed world than the less developed world, while the opposite will be the case for transport costs.

5. Empirical specification

Operationalizing our model requires defining factors that affect trade and those that affect agriculture, as well as defining the set of early-agglomerating countries and the set of late-agglomerating countries. Above in the data section we discussed the variables that relate primarily to trade and to agriculture.

To categorize countries that agglomerated late and those that agglomerated early, we rely primarily on human capital, which allowed farmers to take advantage of higher-yield technologies. The basic issue is illustrated in Figure 7, which plots the UK (early advanced technology) adult literacy rate against time compared to that for India. The UK had over 50% literacy by the 17th century, and following a rapid rise after 1820, over 75% by 1870. Thus the UK achieved a massive increase in human capital before the precipitous decline in the global freight cost index of the late 19th and early 20th century. In 1951 India's literacy rate was still under 20% and only then started to rise quickly, achieving 50% in the 1990s. This is our theme. There are a set of higher technology countries that urbanize and agglomerate earlier before the radical decline in freight costs. And then there are a set where higher effective technology, agglomeration, and urbanization are delayed until after 1950, well after the largest declines in transport costs.

We operationalize our human capital measure using average national average years of schooling in the adult population in 1950, the earliest year with comprehensive data, from Barro and Lee (2010). Alternatively, we use an agglomeration outcome, the level of urbanization in 1950, again the earliest year with comprehensive data, from United Nations (2014). Figure 8 shows the cumulative distribution of these two measures in 1950 weighted by national populations. They closely track each other and indeed the partitions we make and the regression results are similar across the two measures.

Our theory provides no guidance on the precise distinction between early and late spatially transforming countries. We thus follow Durlauf and Johnson (1995), letting the data tell us the cutoff at which the overall unexplained variance, summed across the "early and "late" samples, is minimized. We carry this out for four separate exercises, one with and one without country fixed effects for each of our two proxies, education and urbanization.

Figure 9a provides an illustration of the approach for the education proxy without fixed effects. The vertical axis represents the sum of squared residuals (SSR) summed across two regressions carried out with the same specification on two separate samples. The horizontal axis specifies the cutoff level of education defining the early and late samples. SSR is minimized (and therefore explained variance is maximized, at a cutoff level of 3.6 years of education in 1950. In the fixed effects regressions in Figure 9b, explained variance is maximized at 3.0 years of education. Figures 9c and 9d show the analogous information for the urbanization proxy. An urbanization level of 44 percent is our cutoff without fixed effects; adding fixed effects reduces it slightly to 38 percent.

Regardless of the proxy we use, we end up with a similar split of the sample. Table 3 cross-tabulates the full sample and the lit sample using the two types of cutoffs, separately with and without country fixed effects. The off-diagonal cells, those that are characterized as being in early agglomeration countries using the urbanization criterion and late agglomeration countries using the education criterion, or vice versa, represent less than 5 percent of the full sample and less than 10% of the lit sample. It is also the case that the chosen cutoffs split the samples quite evenly. The countries and their categorization in each of the four versions are listed in Table A3. [16]

A related theory of differential spatial development might instead focus on the new world versus the old world, because the new world's economic geography was altered so radically by the Columbian Exchange. Table 4 explores the various ways one might divide the sample. It reports R-sq from regressions of the form

$$ln(light_{ict}) = X_{ict}\beta + D\_split_c + D\_split_c X_{ict}\beta_d + \varepsilon_{ict} \qquad (11)$$

[16] Because some countries lack a measure of urbanization in 1950, and several more lack an education measure, the sample sizes in these two differential exercises are smaller by 0.13% and 6.27%, respectively, than those in Tables 1 and 2. However, the overall R-sq change by less than 1% with fixed effects, and 0.1% without fixed effects.

where D_split_c is a dummy variable (or a set of dummy variables) indicating, for example, whether a country is in the high education category. From Table 2, the overall Rsq's without splits are 0.459 and 0.569 without and with FE's, respectively.

The first three rows show splits by education, urbanization, and hemisphere, respectively Each necessarily increases explanatory power, but the magnitudes of the gains are modest. In rows 4 and 5, combining the hemisphere split with education and urbanization splits adds a little more. However, as we note below, ultimately we focus on the education and urbanization splits rather than hemispheres, because we find that the overall pattern across these splits hold within both the old and new worlds. This suggests that the story in our model is not purely a distinction between the old and new worlds.

6. Differential results

6.1 Explanatory power

Table 5 reports our main results, the contribution to lights variation within the early versus late agglomeration samples of different blocks of variables, with and without country fixed effects. We discuss results for the education split shown in the left half of the table. Results for the urbanization split, on the right side, are very similar. We start the discussion by looking at three sets of results at the top of the table: the top panel labelled "Both margins" and the next two panels, labelled respectively "Differentials for both margins" and "Summary….", all without fixed effects. The top panel shows the contribution to Rsq's for low and high education countries for each block of variables on their own. To help sort the numbers, we net out the base variables and then calculate a double difference. First in the panel, "Differentials for both margins", we show the explanatory power of the agricultural and trade variables net of the base variables. In the high education countries, the additional explanatory power of the agricultural variables is more than that of the trade variables. In the low education countries, it is the trade variables that offer relatively more explanatory power. So, for example, in the education split without fixed effects, agriculture adds 0.54 to explanatory power relative to the base for high education countries but only 0.27 for low education countries. In contrast, trade relative to the base adds 0.062 to explanatory power for high education countries but a higher 0.17 for low education countries. This pattern holds for all such pairwise comparisons in that panel: for the education split with FE's and for urbanization split with and without FE's. Finally in the "Summary" first row, we double difference, to show the relative advantage of agriculture over trade variables in explaining lights variation for high versus low education countries. The patterns are clear across both education and urbanization splits. Agriculture is relatively more important for early relative to later developing countries. Without fixed effects for both education and urbanization the double differential is over 0.31 and for with fixed effects for both it is over 0.17. This pattern is our basic result upon which we will build.

The bottom of Table 5 shows two further sets of results. First, we split overall effects into the intensive and extensive margins. The distribution of effects between country sets and variable sets are similar across the two margins. Finally in the last panel, we show the relative advantage of agriculture over trade variables in explaining lights variation for high versus low education countries, within new and old world countries. Patterns are similar to those for both worlds combined.

6.2 Marginal effects

Table 5 emphasized the overall explanatory power of groups of trade and agricultural variables in the two samples. We now consider their relative marginal effects. If marginal effects of trade variables, relative to marginal effects of agricultural variables, are stronger in late agglomerator countries than in early agglomerator countries, this is consistent with the explanatory power results. We first estimate equation (11) in general form with a full set of interactions. Results, shown in Table A4 and A5 for education and urbanization, respectively, generally show a pattern analogous to the explanatory power results. The (uninteracted) marginal effects of the agriculture variables are usually of the same sign as their interactions with the early agglomerator indicator, implying that they have a stronger effect in the early agglomerator countries. Conversely, the (uninteracted) marginal effects of the trade variables are usually of the opposite sign as their interactions with the early agglomerator indicator, implying that they have a weaker effect in the early agglomerator countries.

To test this idea more formally, we consider the following variant of equation (11):

$$ln(light_{ict}) = X^{B}_{ict}\beta_B + X^{A}_{ict}\beta_A + X^{T}_{ict}\beta_T + D\_split_c(\alpha X^{A}_{ict}\beta_A + \gamma X^{T}_{ict}\beta_T) + \varepsilon_{ict} \tag{12}$$

In (12), "B" refers to the 2 base covariates, "A" to agriculture, and "T" to trade. The common (constrained) deviation of effects for early agglomerators are α and γ for the sets of agricultural and trade variables, respectively. Table 6 reports non-linear least squares estimates of equation (12). Across the board (education, urbanization, FE's and no FE's) patterns are similar. The deviation recorded is for high education or high urbanization countries. α is positive and significant, meaning that marginal effects of agricultural variables are larger in absolute value for high education (urbanization) countries compared to low education (urbanization) countries. For trade the opposite is the case. γ is negative and significant, meaning that marginal effects of trade variables are smaller in absolute value for high education (urbanization) countries compared to low education (urbanization) countries. Thus not only is agriculture relatively more important than trade in explaining lights variation for high compared to low education or urbanization countries, but marginal effects of agriculture compared to trade variables are relatively stronger for high versus low education or urbanization countries. .

In Figure 10, we perform an interesting counterfactual. We ask how much different areas of high education countries gain or lose because they *have* high not low education coefficients. Correspondingly we ask how areas of low education countries would change if they *were given* high education coefficients rather than low. The analysis is based on the estimations of equation (11) reported in Table A4, without and with FE's. For both sets of countries, we do the same calculation. We first predict for each country's grid squares what its outcome would be if it had high education coefficients and then what it would be for low education coefficients. We then calculate the difference between these two predictions, $X_{ict}\widehat{\beta}_d$, take the mean worldwide difference and plot each grid square's deviation from this mean. Figure 10 allows us visually to compare low education sub-Saharan Africa (excluding South Africa) with high education Europe. In Europe (agriculturally rich) interior areas are more populated than they would be with low education coefficients, while coastal areas are less populated with high education coefficients than they would be with low education ones. For Africa the quantitative results are similar as they should be, but the interpretation is for the context. Africa interiors would be more populated if they were given high education coefficients rather than low, while coastal areas would be less populated with high education coefficients than under low education coefficients. So the color patterns are the same for high and low education places, consistent with the model and the experiment just carried out.

7. Spatial spillover and correlation issues

Both the lights and the physical geography characteristics predicting them are highly spatially correlated. To the extent that this is manifested in spatially correlated errors, we have accounted for this by clustering errors and, alternatively, calculating Conley standard errors. However, direct agglomeration of lights and spillover effects of the X's, such as a port on its hinterland, are also possible. We have tried to minimize the extent to which these affect our results by aggregating individual light pixels to much larger grid squares. While minimizing spillovers across cells, this essentially internalizes the agglomeration externalities. The estimated coefficients are thus reduced form, reflecting not just raw agricultural and trade effects but also endogenous agglomeration. For example, a natural harbor may represent a natural trade advantage, which induces clustering around it, but because of scale economies, the intensity of lights in a cell containing a harbor reflects both the basic harbor advantage and the induced agglomeration from scale economies focused on a harbor.

Separating these three phenomena (correlated errors, spillovers, and agglomeration) is notoriously difficult (e.g. Gibbons, Overman, and Patacchini 2015).

One solution is focus on the reduced form, where we add as covariates the trade and agriculture determinants of neighbors' lights which also accounts for spillovers from the covariates themselves. One can also add in neighbors of neighbors in a spatially lagged construct. But one may want to try to uncover the spillover effects from just lights themselves (pure agglomeration), although it is hard to do counterfactuals with the results (exogenously vary an endogenous variable).

One way to estimate these effects, common in the literature, is to add neighbors' lights as a covariate and instrument for it, using second order neighbors' Xs. This method relies on spillovers attenuating fully beyond immediate neighbors.

However, the physical geography characteristics we use are highly spatially correlated. The simple correlation coefficient between cells and their neighbors for each variable are given in Table 7, with neighbors defined on a rook basis (the N, S, E and W neighbors of a grid square sharing a finite border) and on a queen basis, which adds the NE, NW, SE and SW neighbors sharing a corner. In either case, fifteen of these have a spatial coefficient in excess of 0.958 and several ones are over 0.99). Given these correlations, we cannot credibly separate own and neighbor effects let alone instrument for them.

As an alternative, we impose the structure of the spatial autoregressive model

$$ln\ L_i = \rho W_i ln\ L_i + X_i \beta + \varepsilon_i \tag{13}$$

where the weights matrix W_i is 1 for queen neighbors and 0 for all other squares. Note in this we assume neighbors' X and ε have no effect on outcomes. We then estimate this model in the traditional fashion where

$$ln\ L_i = (I - \rho W_i)^{-1} X_i \beta + \varepsilon_i \qquad (14)$$

Results are in Appendix Table A6. The coefficient of interest is ρ, the effect of neighbor's lights on own lights. The estimates for the full sample in columns 1 and 2, 0.99 with rook contiguity and 1.001 with queen, imply extreme spillovers. This is part a symptom of the huge patches of contiguous unlit areas. However, when we run the same specification on lit areas, the coefficient is still above 0.6, whereas the agglomeration literature suggests a value of 0.10 at most. Our attempts at instrumenting yielded similar implausible results. We did not further pursue this line of investigation.

7. Conclusion

References

Allen, Treb and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics* 129(3): 1085-1140.

Barro, Robert and Jong-Wha Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics*, 104: 184-198, April.

Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127: 587-644.

Broadberry, Stephen and O'Rourke, Kevin H. (2010) The Cambridge Economic History of Modern Europe: Volume 1, 1700-1870. Cambridge: Cambridge UP

Buringh, Eltjo and Jan Luiten Van Zanden (2009). Charting the "Rise of the West": Manuscripts and Printed Books in Europe, A Long-Term Perspective from the Sixth through Eighteenth Centuries. *The Journal of Economic History*, 69, pp 409-445.

Central Intelligence Agency. 2013. The World Factbook 14. Washington: Central Intelligence Agency. https://www.cia.gov/library/publications/the-world-factbook/ Accessed 9 August 2015

Chen, Xi and William Nordhaus. 2011. “Using luminosity data as a proxy for economic statistics.” *Proceedings of the National Academy of Sciences* 108(21): 8589-8594.

Conley, Timothy G. 2008. Spatial Econometrics. *The New Palgrave Dictionary of Economics*, Second Edition, Eds .Steven N. Durlauf and Lawrence E. Blume.

Costinot, Arnaud, Dave Donaldson and Cory Smith. (forthcoming). “Evolving Comparative Advantage and the Impact of Climate Change on Agricultural Markets: Evidence from 1.7 million Fields Around the World.” *Journal of Political Economy*.

Cronon W. 1991. *Nature’s Metropolis: Chicago and the Great West*. New York: W.W. Norton.

Davis, Donald R. and David E. Weinstein. 2002. “Bones, Bombs, and Break Points: The Geography of Economic Activity.” *American Economic Review* 92(5): 1269-1289, December.

Desmet K and J.V. Henderson (2015). The geography of development within countries. In G. Duranton, J. V. Henderson and W. Strange (eds.), *Handbook of Regional and Urban Economics*, 5.

Donaldson, Dave (forthcoming) “Railroads of the Raj: Estimating the Impact of Transportation Infrastructure,” *American Economic Review*.

Durlauf, Steven N & Johnson, Paul A, 1995. "Multiple Regimes and Cross-Country Growth Behaviour," *Journal of Applied Econometrics*, 10(4): 365-384, Oct.-Dec.

Elvidge, C.D., Baugh, K.E., Kihn, E.A., Kroehl, H.W, Davis, E.R. 1997. “Mapping of city lights using DMSP Operational Linescan System data.” *Photogrammetric Engineering and Remote Sensing* 63: 727-734.

Elvidge, C.D., Baugh, K.E., Dietz, J., Bland, T., Sutton, P.C., Kroehl, H.W. 1999. “Radiance Calibration of DMSP-OLS Low-Light Imaging Data of Human Settlements.” *Remote Sensing of Environment* 68(1): 77-88.

FAO/IIASA, 2011. Global Agro-ecological Zones (GAEZ v3.0). FAO Rome, Italy and IIASA, Laxenburg, Austria. http://gaez.fao.org

Gollin, Douglas, and Richard Rogerson. 2014. “Productivity, transport costs and subsistence agriculture.” *Journal of Development Economics* 107: 38-48, March.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2013. “Human Capital and Regional Development.” *Quarterly Journal of Economics* 128(1): 105-164.

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez De Silanes, and Andrei Shleifer. 2014. “Growth in Regions.” *Journal of Economic Growth* 19: 259-309

Gibbons, S., H.G.Overman and E. Patacchini. 2015. Spatial methods. In G. Duranton, J. V. Henderson and W. Strange (eds.) *Handbook of Urban and Regional Economics* Volume 5.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. “Measuring Economic Growth from Outer Space.” *American Economic Review* 102(2): 994-1028.

ISciences LLC (2008) Elevation and Depth v2 2000 (dataset) http://geoserver.isciences.com:8080/geonetwork/srv/en/resources.get?id=208&fname=ElevationGridv2.zip&access=private

Jedwab, Remi and Alexander Moradi (forthcoming) “The Permanent Economic Effects of Transportation Revolutions in Poor Countries: Evidence from Africa,” Review of Economics and Statistics.

Klein Goldewijk, Kees, Arthur Beusen, Gerard Van Drecht, Martine De Vos. 2011. “The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years” Global Ecology & Biogeography 20(1):73–86, January.

Krugman, Paul. 1991. “First Nature, Second Nature, and Metropolitan Location” NBER Working Paper 3740.

Lehner, B. and Döll, P. 2004. “Development and validation of a global database of lakes, reservoirs and wetlands.” *Journal of Hydrology* 296(1-4): 1-22.

Mellinger, Andrew, Jeffrey D. Sachs, and John Luke Gallup. 2000. “Climate, Coastal Proximity, and Development,” in Clark, Gordon L., Maryann P. Feldman, and Meric S. Gertler, eds., *The Oxford Handbook of Economic Geography*, 169-194.

Mesbah J. Motamed, Raymond J.G.M. Florax and William A. Masters. 2014. “Agriculture, Transportation and the Timing of Urbanization: Global Analysis at the Grid Cell Level.” *Journal of Economic Growth* 19(3): 339-368, September.

Michaels, Guy and Ferdinand Rauch. 2013. “Resetting the Urban Network: 117-2012.” BREAD Working Paper 405, November.

Ministry of Human Resource Development. 1987. A Hand Book of Educational and Allied Statistics. New Delhi.

Mitchell, Timothy D. and Philip D. Jones. 2005. “An improved method of constructing a database of monthly climate observations and associated high-resolution grids.” *International Journal of Climatology* 25(6): 693–712.

Mohammed, S.I.S. and J.G. Williamson 2004. Freight rates and productivity gains in British tramp shipping 1869–1950. *Explorations in Economic History* 41: 172–203

Natural Earth. 2012. Natural Earth version 2.0.0. http://www.naturalearthdata.com/, Accessed 2013.

NOAA National Geophysical Data Center (2011). Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) Version 2.2.0. http://www.ngdc.noaa.gov/mgg/shorelines/shorelines.html.

Nordhaus, William. 2006. “Geography and Macroeconomics: New Data and New Findings,” *Proceedings of the National Academy of Sciences* 103(10): 3510-3517, March.

Nordhaus, William and Xi Chen. 2009. “Geography: Graphics and Economics.” *The B.E. Journal of Economic Analysis and Policy* 9(2): 1-12.

Nunn, Nathan and Diego Puga. 2012. “Ruggedness: The Blessing of Bad Geography in Africa.” *The Review of Economics and Statistics* 94(1): 20-36.

Olson, D.M., E. Dinerstein, E.D. Wikramanayake, N.D. Burgess, G.V.N. Powell, E.C. Underwood, J.A. D'Amico, I. Itoua, H.E. Strand, J.C. Morrison, C.J. Loucks, T.F. Allnutt, T.H. Ricketts, Y. Kura, J.F. Lamoreux, W.W. Wettengel, P. Hedao, and K.R. Kassem. 2001. “Terrestrial ecoregions of the world: A new map of life on earth.” *Bioscience* 51(11): 933-938.

Ramankutty, Navin, Jonathan A. Foley, John Norman and Kevin McSweeney. 2002. “The global distribution of cultivable lands: current patterns and sensitivity to possible climate change.” *Global Ecology and Biogeography* 11(5): 377-392.

Rappaport, Jordan and Jeffrey Sachs. 2003. "The United States as a Coastal Nation." *Journal of Economic Growth* 8: 5-46.

Rauch J. 1993. "Does History Matter Only When it Matters Little?" *Quarterly Journal Of Economics* 108(3): 843-67

UNESCO. 1957. *World Illiteracy at Mid-Century: A Statistical Study*. Paris, UNESCO

United States Navy (1953) *World Port Index 1953*. Hydrographic Office, US Navy

University of East Anglia Climatic Research Unit; Jones, P.D.; Harris, I. (2013): CRU TS3.20: Climatic Research Unit (CRU) Time-Series (TS) Version 3.20 of High Resolution Gridded Data of Month-by-month Variation in Climate (Jan. 1901 - Dec. 2011). NCAS British Atmospheric Data Centre. http://catalogue.ceda.ac.uk/uuid/2949a8a25b375c9e323c53f6b6cb2a3a

United Nations. 2014. *World Urbanization Prospects, the 2014 revision*. New York: United Nations.

Wessel, P., and W. H. F. Smith. 1996. "A Global Self-consistent, Hierarchical, High-resolution Shoreline Database." *Journal of Geophysical Research* 101(B4): 8741-8743.

Willmott, C. J. and K. Matsuura (2012) Terrestrial Precipitation: 1900-2010 Gridded Monthly Time Series (V 3.02), http://climate.geog.udel.edu/~climate/html_pages/Global2011/Precip_revised_3.02/README.GlobalTsP2011.html. Accessed 25 January 2013

World Bank. 2015. World Development Indicators, http://data.worldbank.org/data-catalog/world-development-indicators Accessed 9 August 2015.

Ziskin, Daniel, Kimberly Baugh, Feng Chi Hsu, Tilottama Ghosh, Chris Elvidge. 2010. "Methods Used For the 2006 Radiance Lights." Proceedings of the 30th Asia-Pacific *Advanced Network Meeting* 131-142.

Appendix. 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:

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

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):

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

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:

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

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 manufacturing consumption for manufacturing workers rises above that is agriculture and in the other region, then the equilibrium fails the stability check. The stability test is passed if manufacturing consumption of manufacturing wages in the interior is below all other workers and that of food workers in the interior is above coastal workers. There are no type 2 stability situations which do not meet one of these two criteria.

Table 1: Summary Statistics and Baseline Regression Results

	Summary Statistics		Baseline Regression Results	
	mean, (sd)	min, max	w/o country FEs	w/ country FEs
	(1)	(2)	(3)	(4)
ln(light/ land ratio)	3.501	1.115		
	(3.176)	15.783		
Base Covariates				
ruggedness	2766.746	0	-8.70e-06***	-1.82e-05***
	(4838.898)	95814.383	(1.99e-06)	(1.68e-06)
malaria index	1.912	0	-0.0348***	-0.0486***
	(5.279)	38.081	(0.00260)	(0.00243)
Agricultural Covariates				
tropical moist forest	0.116	0	-0.0406	-0.203***
	(0.321)	1	(0.0756)	(0.0661)
tropical dry forest	0.022	0	0.936***	0.228***
	(0.147)	1	(0.0943)	(0.0804)
temperate broadleaf	0.104	0	1.771***	1.294***
	(0.305)	1	(0.0707)	(0.0648)
temperate conifer	0.033	0	0.769***	0.159**
	(0.178)	1	(0.0822)	(0.0776)
boreal forest	0.167	0	-0.442***	-1.246***
	(0.373)	1	(0.0769)	(0.0812)
tropical grassland	0.12	0	-0.862***	-0.0548
	(0.325)	1	(0.0561)	(0.0487)
temperate grassland	0.077	0	0.710***	0.922***
	(0.267)	1	(0.0648)	(0.0568)
montane grassland	0.033	0	0.631***	0.762***
	(0.179)	1	(0.0806)	(0.0721)
tundra	0.124	0	-0.874***	-1.438***
	(0.329)	1	(0.0863)	(0.0895)
Mediterranean forest	0.024	0	0.837***	1.373***
	(0.153)	1	(0.0926)	(0.0877)
mangroves	0.004	0	0.418**	-0.0312
	(0.064)	1	(0.180)	(0.154)
desert	0.176	0	(reference)	(reference)
	(0.381)	1		
temperature	9.970	-22.286	0.174***	0.124***
	(13.788)	30.366	(0.00337)	(0.00387)
precipitation	60.678	0.387	-0.00896***	-0.0112***
	(59.29)	921.909	(0.000446)	(0.000468)
growing days	139.306	0	0.00985***	0.00856***
	(99.067)	366	(0.000284)	(0.000283)
land suitability	0.273	0	2.690***	2.184***
	(0.319)	1	(0.0552)	(0.0528)

Continued on next page

Table 1 – *Continued from previous page*

	Summary Statistics		Baseline Regression Results	
	mean, (sd)	min, max	w/o country FEs	w/ country FEs
	(1)	(2)	(3)	(4)
abs(latitude)	38.34	0	0.114***	0.0371***
	(20.923)	75	(0.00251)	(0.00336)
elevation	601.505	-187.341	0.000501***	8.02e-05***
	(788.097)	6169.01	(2.41e-05)	(2.58e-05)
Trade Covariates				
$\mathbb{1}$ (coast)	0.098	0	0.421***	0.414***
	(0.297)	1	(0.0396)	(0.0323)
distance to coast	485.226	0	-0.000665***	-0.000677***
	(480.554)	2273.801	(2.81e-05)	(3.24e-05)
$\mathbb{1}$ (harbor<25km)	0.027	0	1.625***	1.414***
	(0.163)	1	(0.0677)	(0.0575)
$\mathbb{1}$ (river<25km)	0.027	0	0.764***	0.690***
	(0.163)	1	(0.0662)	(0.0601)
$\mathbb{1}$ (big lake<25km)	0.127	0	0.317***	0.179***
	(0.333)	1	(0.0272)	(0.0244)
Observations	243,985	243,985	243,985	243,985
Non-zero observations	98,941	98,941	98,941	98,941
R-squared			0.459	0.568

Notes: Clustered standard errors in parentheses in columns (3) and (4). *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Summary of overall Rsq's

	(1) No country FEs	(2) With country FEs
(1) All variables, both margins (N = 243,985)	0.459	0.568
(2) All variables, extensive margin (LPM)	0.385	0.475
(3) All variables, intensive margin (OLS)	0.262	0.359
(4) Country fixed effects, extensive margin		0.269
(5) Country fixed effects, intensive margin		0.219
(6) Base variables (malaria, ruggedness)	0.020	0.350
(7) Agriculture variables, both margins (plus base)	0.437	0.554
(8) Trade variables, both margins (plus base)	0.072	0.367

Table 3: Cell counts in 4 way splits: Education vs. Urbanization

No country FEs (ed cutoff = 3.6; urban cutoff = 0.44)			With country FEs (ed cutoff = 3; urban cutoff = 0.38)		
Whole Sample	high urban	low urban	Whole Sample	high urban	low urban
high educ	119,824	7,399	high educ	123,908	3,879
low educ	1,681	99,713	low educ	6,214	94,616
Lit Sample	high urban	low urban	Lit Sample	high urban	low urban
high educ	43,444	5,840	high educ	46,614	3,196
low educ	961	44,360	low educ	3,790	41,005

Table 4: Various splits; overall explanatory power

Rsq's. Regression on all covariates, split indicator(s), split indicator(s)*all covariates

Split indicator(s)	(1) No country FEs	(2) With country FEs
Low-high education	0.492	0.581
Low-high urbanization	0.483	0.579
New-old world	0.488	0.584
Low-high education, new-old world	0.516	0.596
Low-high urbanization, new-old world	0.502	0.593

Table 5: Explanatory differentials of trade and agriculture
for high/low education & urbanization countries

	Rsq's, Education				Rsq's, Urbanization			
	No FEs		With FEs		No FEs		With FEs	
	High	Low	High	Low	High	Low	High	Low
Both margins								
Base	0.008	0.052	0.379	0.291	0.008	0.058	0.352	0.346
Agriculture plus base	0.546	0.326	0.638	0.445	0.529	0.356	0.611	0.489
Trade plus base	0.070	0.218	0.418	0.399	0.073	0.212	0.389	0.438
Differentials, for both margins								
Agriculture minus base	0.538	0.274	0.259	0.154	0.521	0.298	0.259	0.143
Trade minus base	0.062	0.166	0.039	0.108	0.065	0.154	0.037	0.092
Summary, high[(agric-base)-(trade-base)] - low[(agric-base)-(trade-base)]								
	No FEs		With FEs		No FEs		With FEs	
	0.368		0.174		0.312		0.171	
Extensive margin LPM								
Base	0.006	0.046	0.281	0.248	0.006	0.047	0.264	0.285
Agriculture plus base	0.489	0.268	0.555	0.371	0.475	0.280	0.533	0.401
Trade plus base	0.058	0.171	0.328	0.324	0.065	0.163	0.306	0.350
Intensive margin OLS								
Base	0.011	0.050	0.253	0.171	0.008	0.063	0.240	0.214
Agriculture plus base	0.251	0.184	0.366	0.255	0.239	0.213	0.344	0.298
Trade plus base	0.082	0.163	0.281	0.262	0.078	0.166	0.271	0.297
New World - Old World High vs. Low Splits								
Summary, high[(agric-base)-(trade-base)] - low[(agric-base)-(trade-base)]								
	New World		Old World		New World		Old World	
	No FEs	With FEs	No FEs	With FEs	No FEs	With FEs	No FEs	With FEs
	0.329	0.279	0.377	0.097	0.342	0.237	0.285	0.113

Notes: Education cutoffs are 3.6 (no FEs) and 3 (with FEs), and urbanization cutoffs are 0.44 (no FEs) and 0.38 (with FEs).

Table 6: Differential group marginal effects

	Dependent variable: lrad2010land_ csd			
	Education		Urbanization	
	No FEs	With FEs	No FEs	With FEs
	(1)	(2)	(3)	(4)
ruggedness	-9.44e-06***	-1.46e-05***	-9.53e-06***	-1.70e-05***
	(2.09e-06)	(1.78e-06)	(1.97e-06)	(1.69e-06)
malaria index	-0.0429***	-0.0443***	-0.0403***	-0.0436***
	(0.00261)	(0.00262)	(0.00252)	(0.00235)
tropical moist forest	-0.406***	-0.440***	-0.346***	-0.326***
	(0.0703)	(0.0660)	(0.0676)	(0.0629)
tropical dry forest	0.757***	0.140*	0.717***	0.182**
	(0.0890)	(0.0815)	(0.0882)	(0.0777)
temperate broadleaf	1.253***	1.044***	1.359***	1.184***
	(0.0640)	(0.0633)	(0.0618)	(0.0633)
temperate conifer	0.610***	0.118	0.719***	0.158**
	(0.0703)	(0.0722)	(0.0697)	(0.0731)
boreal forest	-0.354***	-1.049***	-0.274***	-1.061***
	(0.0684)	(0.0826)	(0.0664)	(0.0846)
tropical grassland	-0.760***	-0.103**	-0.727***	-0.0583
	(0.0500)	(0.0483)	(0.0479)	(0.0457)
temperate grassland	0.464***	0.671***	0.538***	0.776***
	(0.0543)	(0.0558)	(0.0540)	(0.0562)
montane grassland	0.562***	0.737***	0.502***	0.714***
	(0.0756)	(0.0757)	(0.0730)	(0.0721)
tundra	-0.548***	-1.086***	-0.462***	-1.139***
	(0.0745)	(0.0891)	(0.0723)	(0.0922)
Mediterranean forest	0.548***	1.086***	0.655***	1.195***
	(0.0746)	(0.0807)	(0.0758)	(0.0823)
mangroves	-0.930***	-1.042***	-0.327**	-0.532***
	(0.180)	(0.154)	(0.166)	(0.149)
temperature	0.139***	0.113***	0.135***	0.117***
	(0.00338)	(0.00411)	(0.00306)	(0.00411)
precipitation	-0.00709***	-0.00996***	-0.00704***	-0.0101***
	(0.000370)	(0.000437)	(0.000356)	(0.000436)

Continued on next page

Table 6 – *Continued from previous page*

	Education		Urbanization	
	No FEs	With FEs	No FEs	With FEs
	(1)	(2)	(3)	(4)
growing days	0.00901***	0.00821***	0.00853***	0.00810***
	(0.000263)	(0.000289)	(0.000251)	(0.000284)
land suitability	2.348***	1.955***	2.313***	1.971***
	(0.0501)	(0.0559)	(0.0485)	(0.0532)
abs(latitude)	0.107***	0.0380***	0.101***	0.0382***
	(0.00265)	(0.00310)	(0.00235)	(0.00313)
elevation	0.000429***	5.03e-05*	0.000380***	8.67e-05***
	(2.34e-05)	(2.57e-05)	(2.14e-05)	(2.53e-05)
$\mathbb{1}$ (coast)	1.127***	1.109***	0.861***	0.946***
	(0.0737)	(0.0667)	(0.0670)	(0.0604)
distance to coast	-0.00149***	-0.00126***	-0.00133***	-0.00119***
	(3.65e-05)	(4.72e-05)	(3.37e-05)	(4.49e-05)
$\mathbb{1}$ (harbor<25km)	2.379***	2.292***	2.160***	2.088***
	(0.115)	(0.102)	(0.104)	(0.0901)
$\mathbb{1}$ (river<25km)	1.244***	1.172***	1.218***	1.051***
	(0.109)	(0.101)	(0.103)	(0.0968)
$\mathbb{1}$ (big lake<25km)	0.454***	0.377***	0.501***	0.314***
	(0.0495)	(0.0441)	(0.0459)	(0.0405)
above cut	-3.954***	-74.26***	-3.780***	24.30***
	(0.142)	(0.158)	(0.134)	(2.468)
α	0.368***	0.167***	0.375***	0.119***
	(0.0209)	(0.0258)	(0.0191)	(0.0237)
γ	-0.761***	-0.712***	-0.740***	-0.644***
	(0.0164)	(0.0177)	(0.0188)	(0.0201)
Observations	228,690	228,690	243,661	243,661
R-squared	0.486	0.574	0.474	0.571

Notes: Above cut is an indicator, which equals 1 if average years of education is greater than or equal to 3.6 (column 1) or 3 (column 2), or the fraction of urban population is greater than or equal to 0.44 (column 3) or 0.38 (column 4). Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Correlations with the average value of neighbors
as defined by rook or queen contiguity

	(1) Rook	(2) Queen
mangroves	0.6336	0.6486
ruggedness	0.7550	0.7611
$\mathbb{1}$ (harbor<25km)	0.8222	0.7843
$\mathbb{1}$ (big lake<25km)	0.8815	0.8585
$\mathbb{1}$ (river<25km)	0.8914	0.8735
$\mathbb{1}$ (coast)	0.9050	0.8772
tropical dry forest	0.9057	0.9001
temperate conifer	0.9201	0.9120
montane grassland	0.9291	0.9235
tropical moist forest	0.9595	0.9586
Mediterranean forest	0.9674	0.9639
temperate grassland	0.9683	0.9635
tropical grassland	0.9721	0.9699
temperate broadleaf	0.9737	0.9699
tundra	0.9753	0.9701
boreal forest	0.9761	0.9709
malaria index	0.9876	0.9799
land suitability	0.9892	0.9818
elevation	0.9923	0.9895
precipitation	0.9943	0.9903
growing days	0.9989	0.9985
temperature	0.9994	0.9988
abs(latitude)	1	1
distance to coast	1	1

Figure 1. Global nighttime lights in 2010

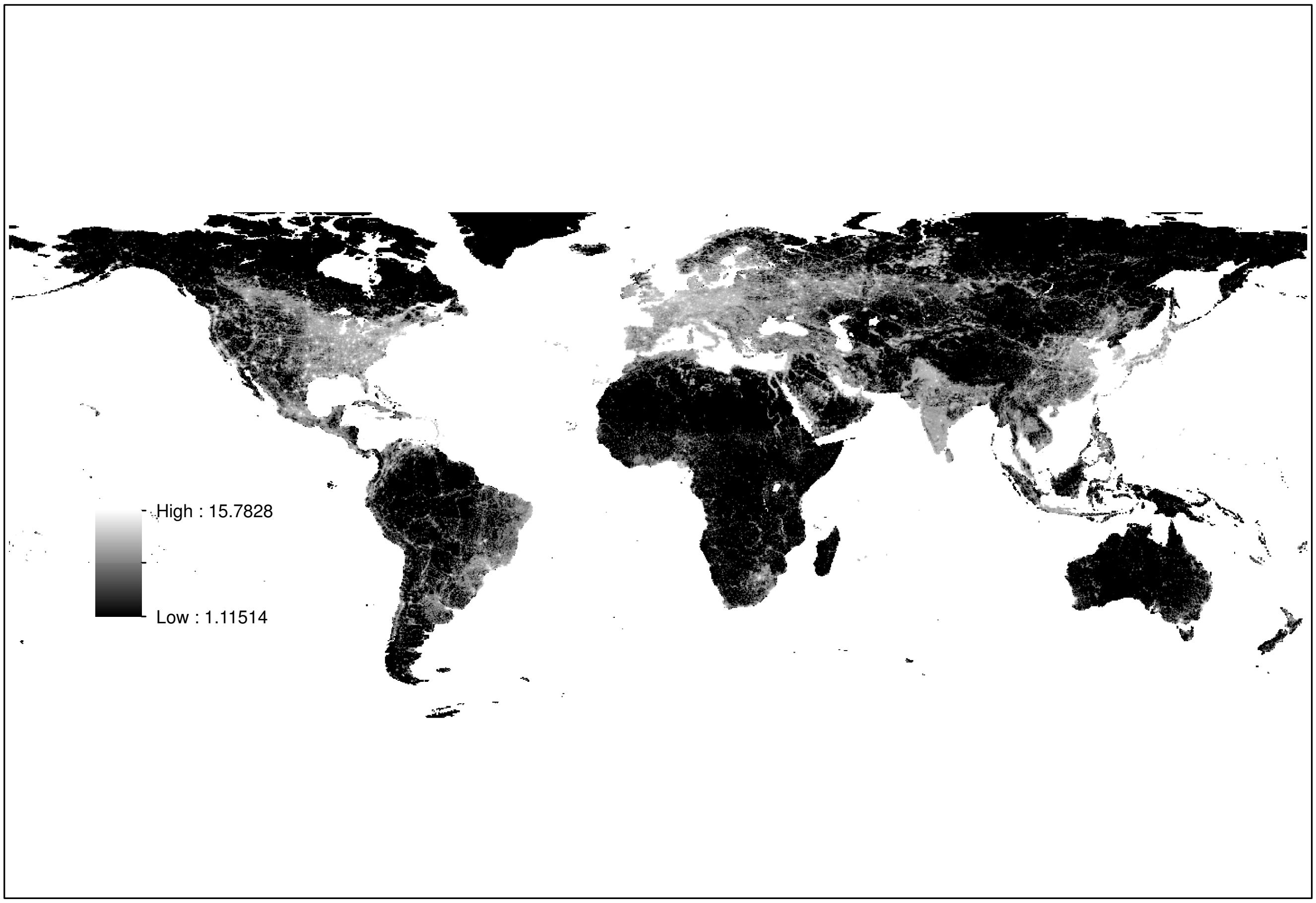


Figure 2. Distribution of lights in lit areas

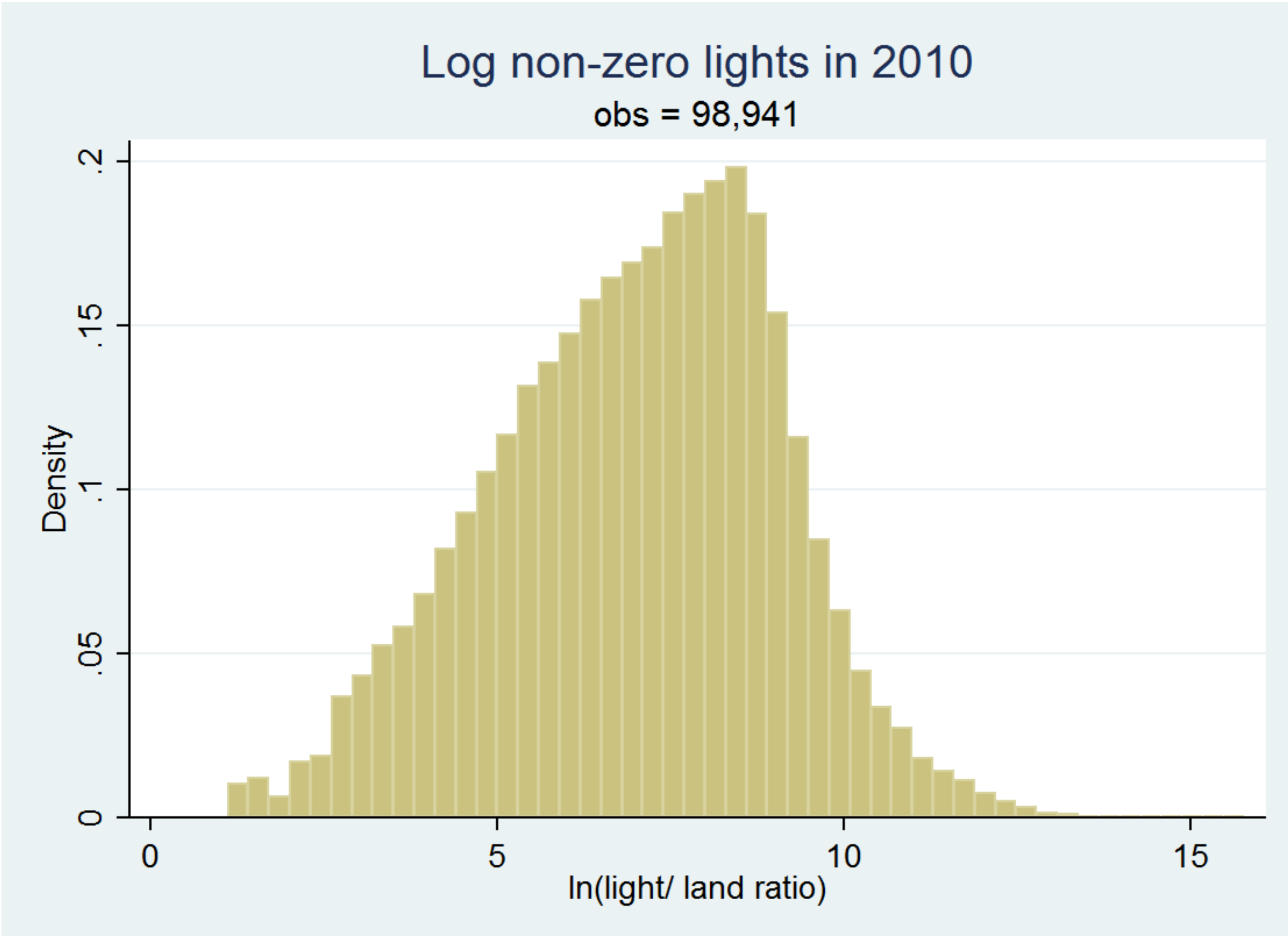
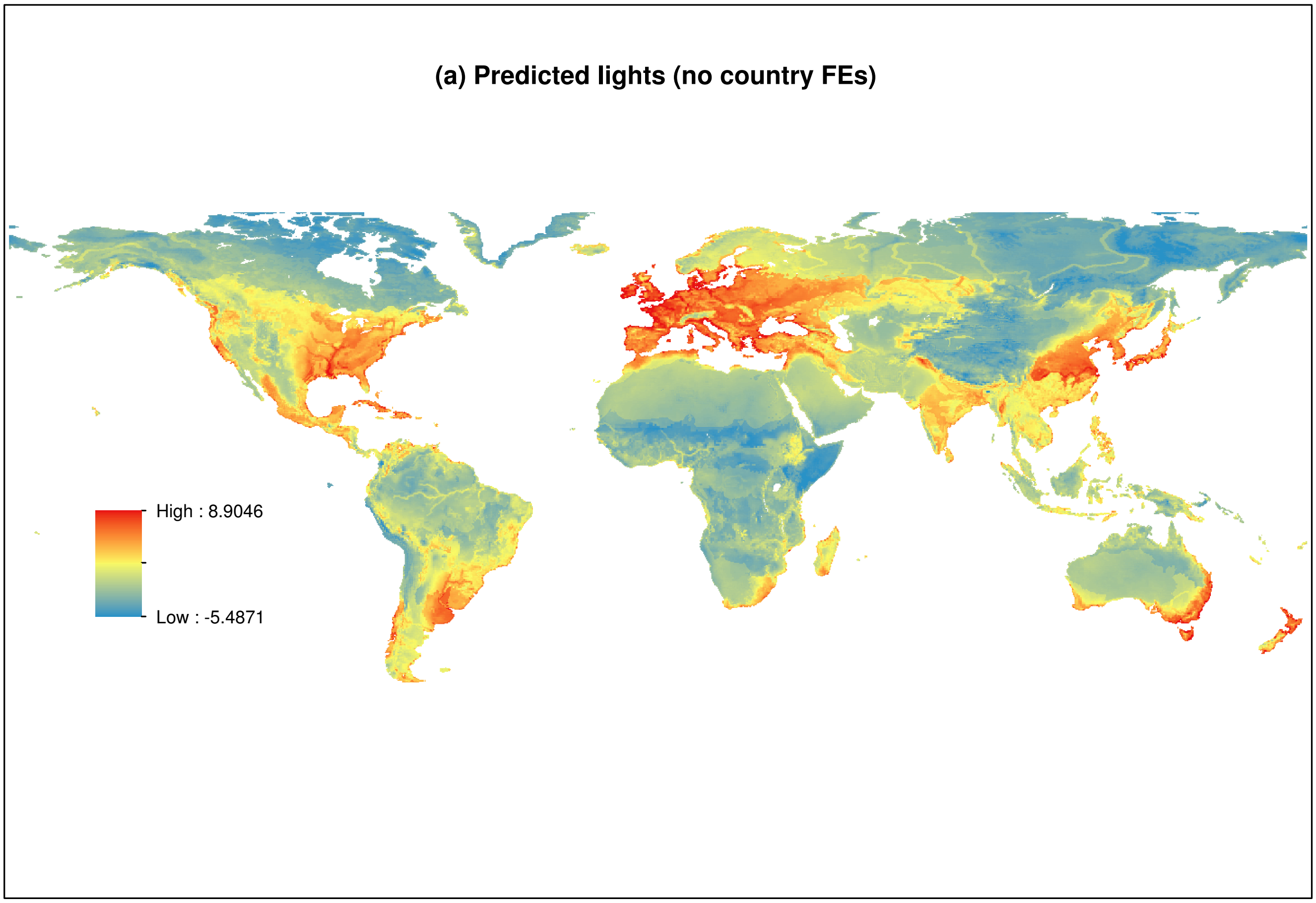


Figure 3. Worldwide predicted lights



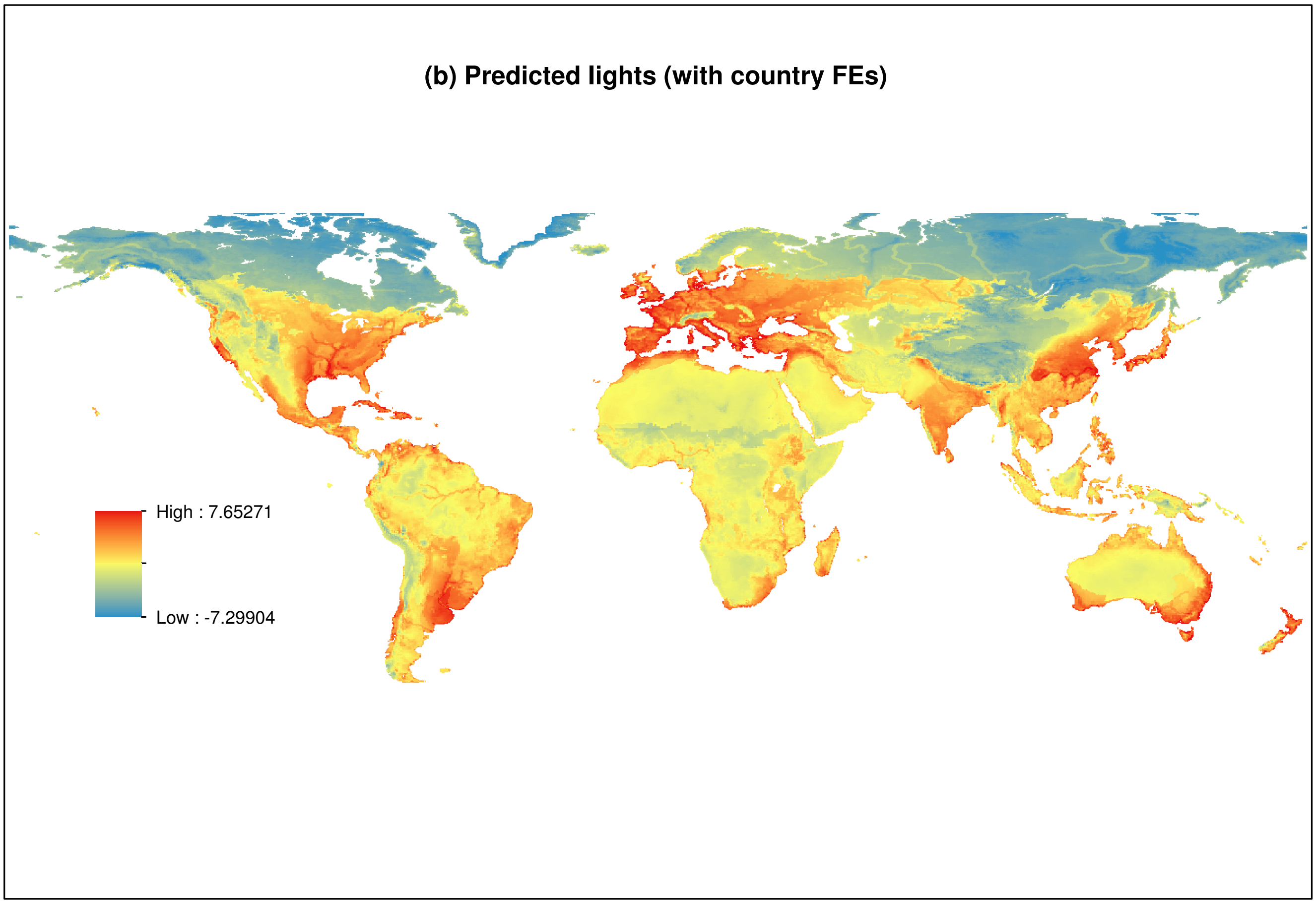

(b) Predicted lights (with country FEs)
High : 7.65271
Low : -7.29904

Figure 4. Model solution method

Figure 5. Equilibria with low agricultural productivity

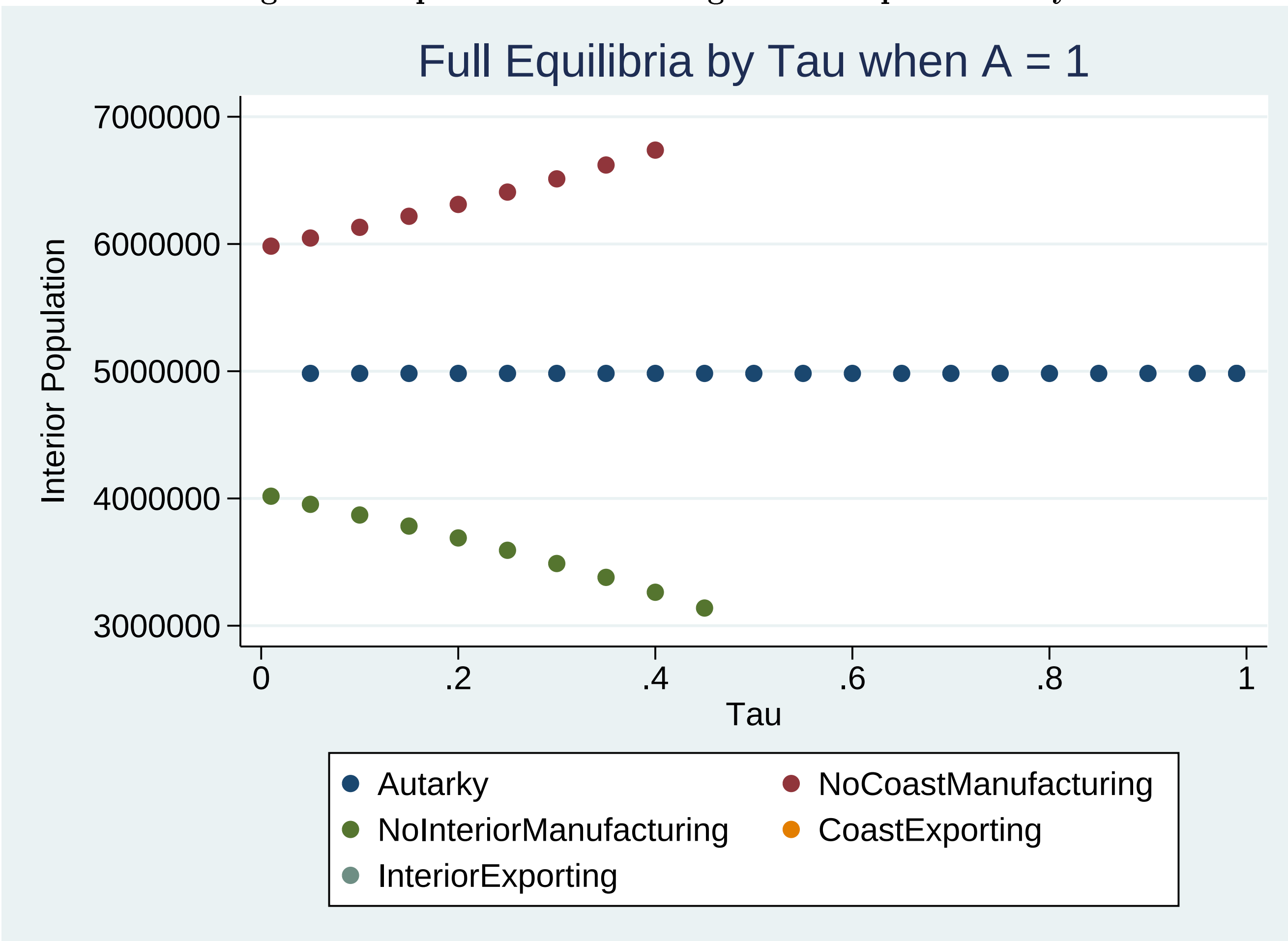


Figure 6. Equilibria with high agricultural productivity

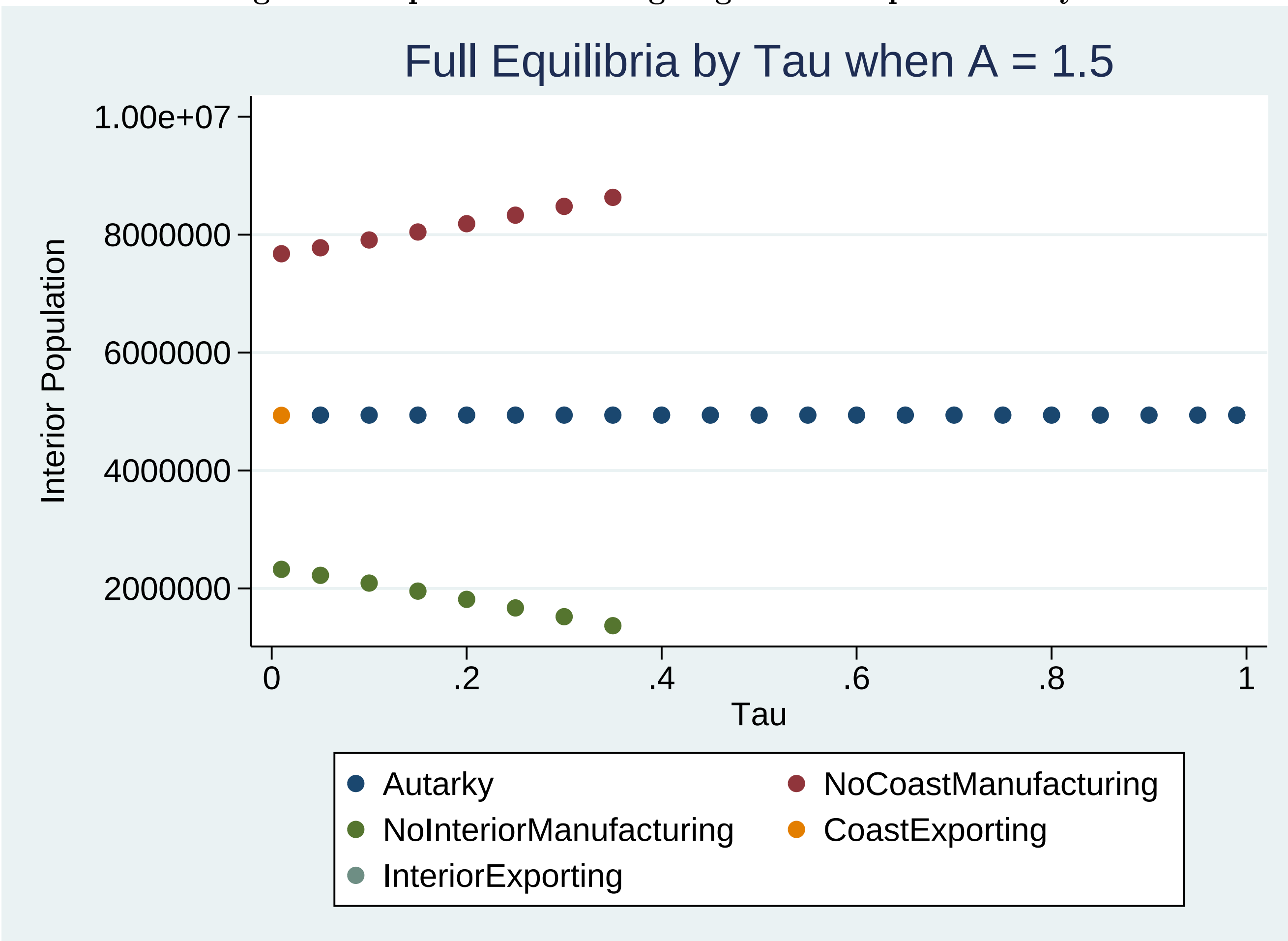
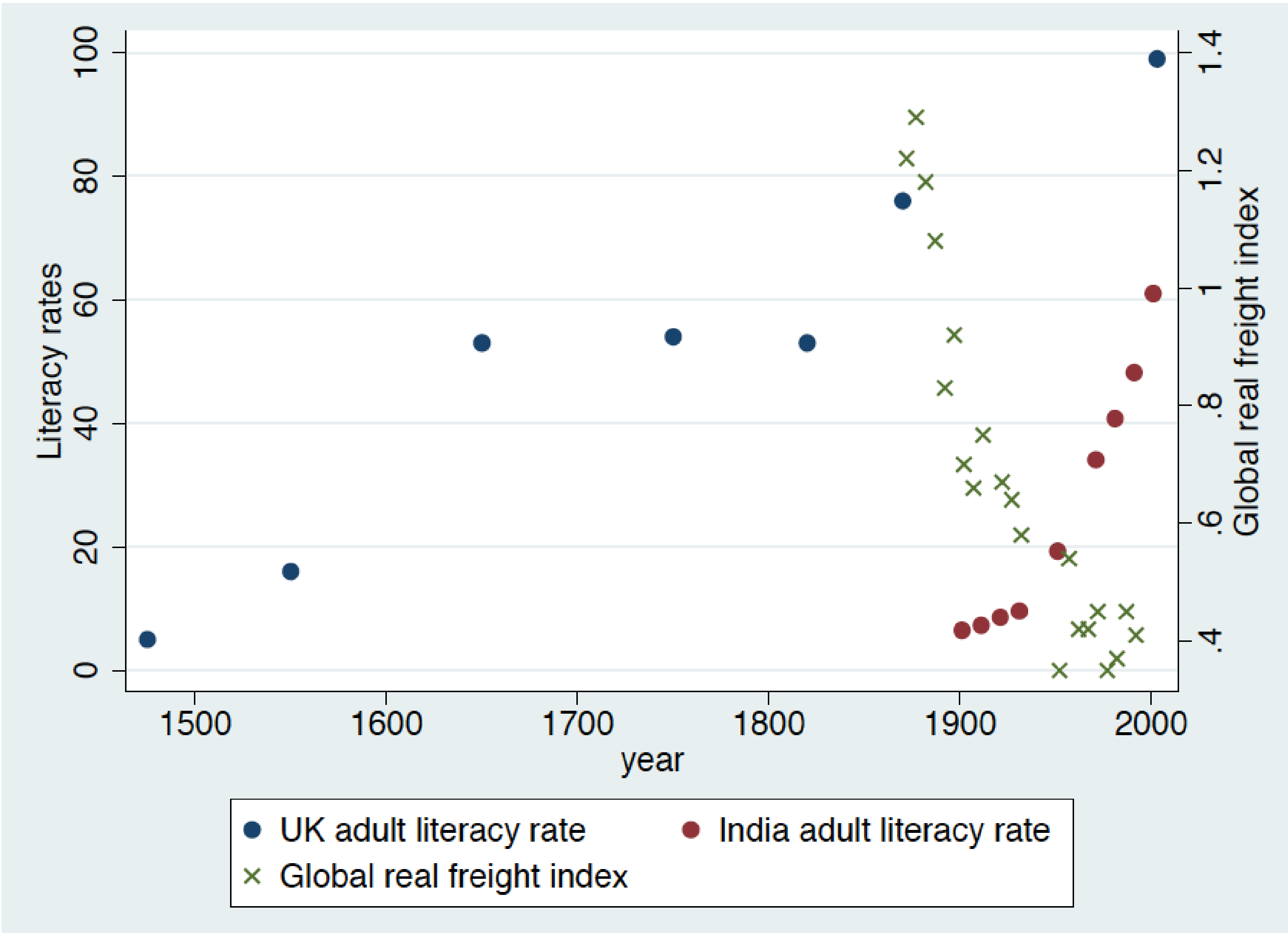


Figure 7. Timing of transport cost changes versus changes in effective technology



Note: Freight rates exclude periods including world war years. Sources: Freight rates: Mohammed and Williamson (2004)' UK Literacy 1475-1750 (midpoints of reported periods 1451-1500, 1501-1600, 1601-1700, 1701-1800): Broadberry and O'Rourke (2010), 1820-1870: Buringh and Van Zanden (2009), 2003: CIA (2015), India: 1901-1951: UNESCO (1957), 1971: Ministry of Human Resource Development (1987), 1981-2001: World Bank (2015).

Figure 8. Cumulative distribution of education and urbanization in 1950

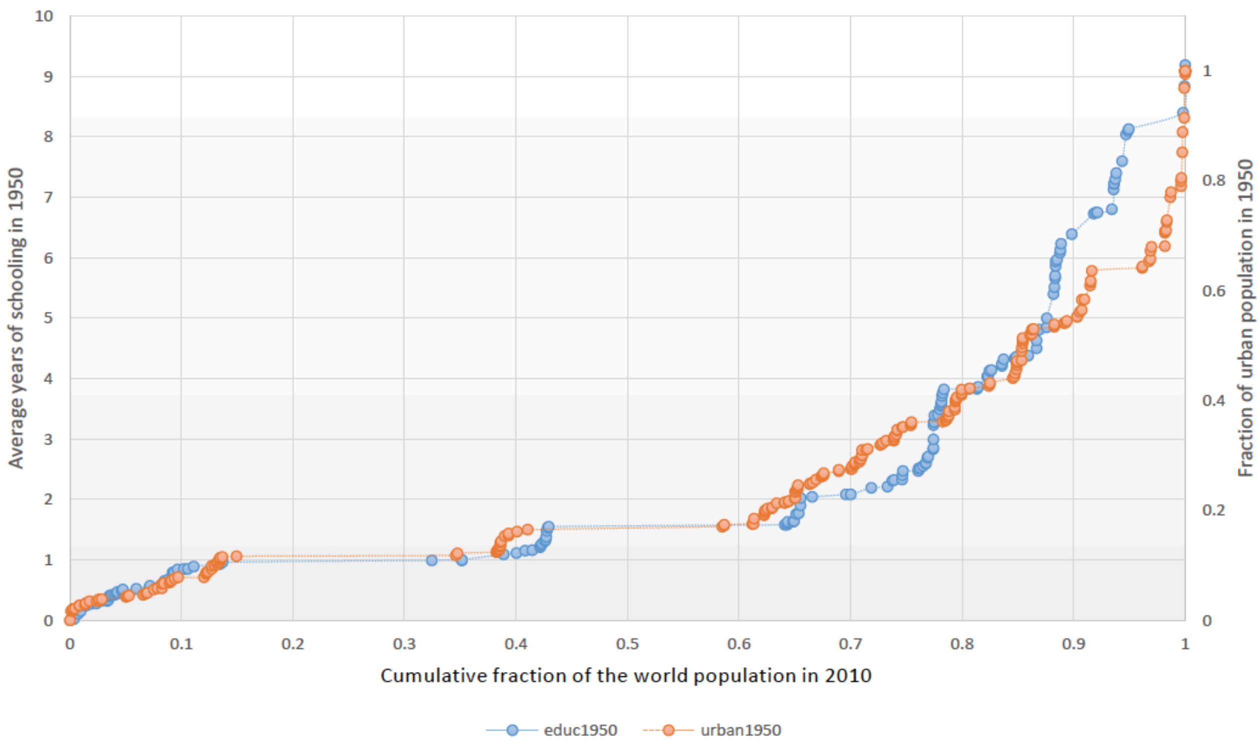
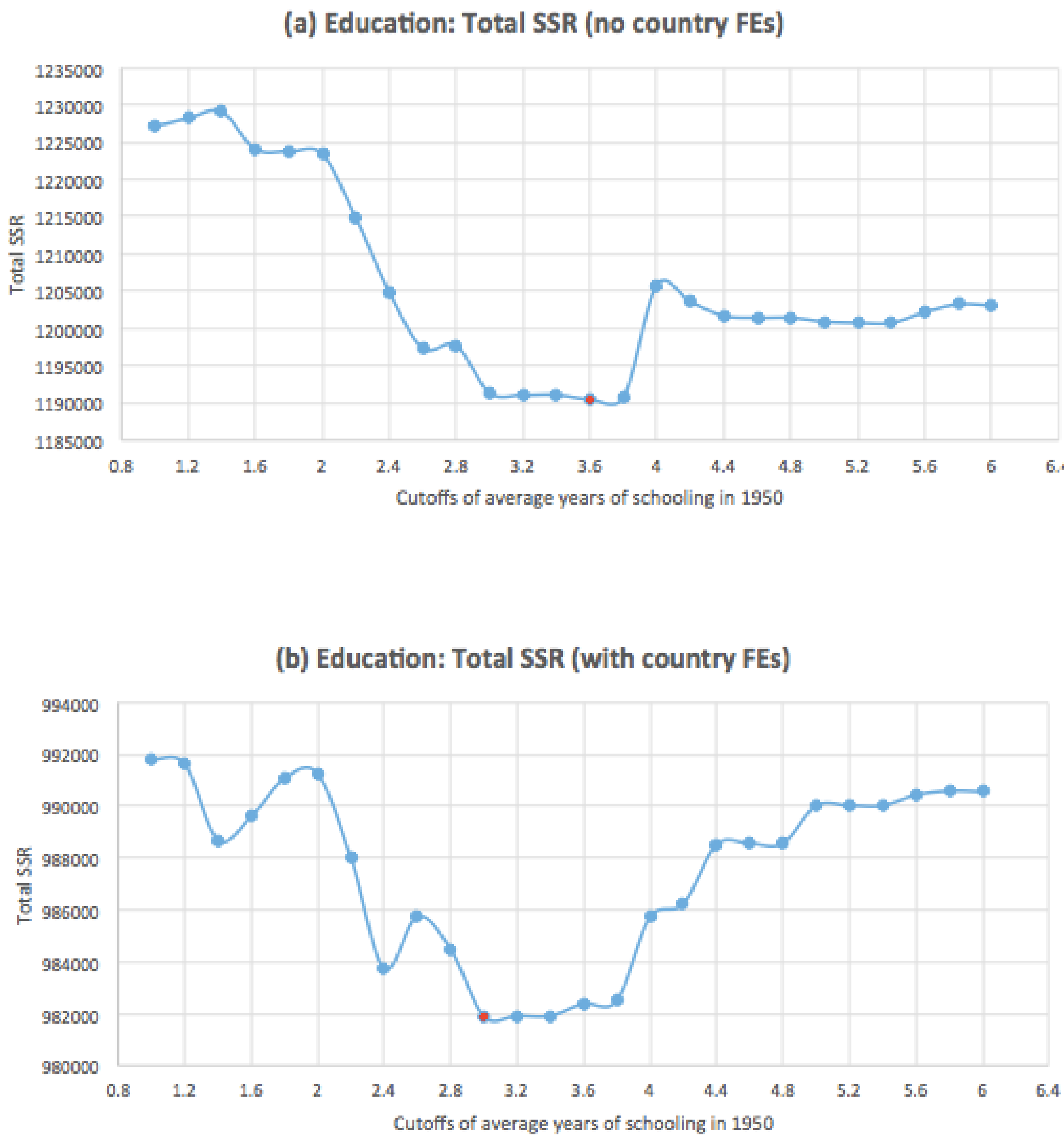
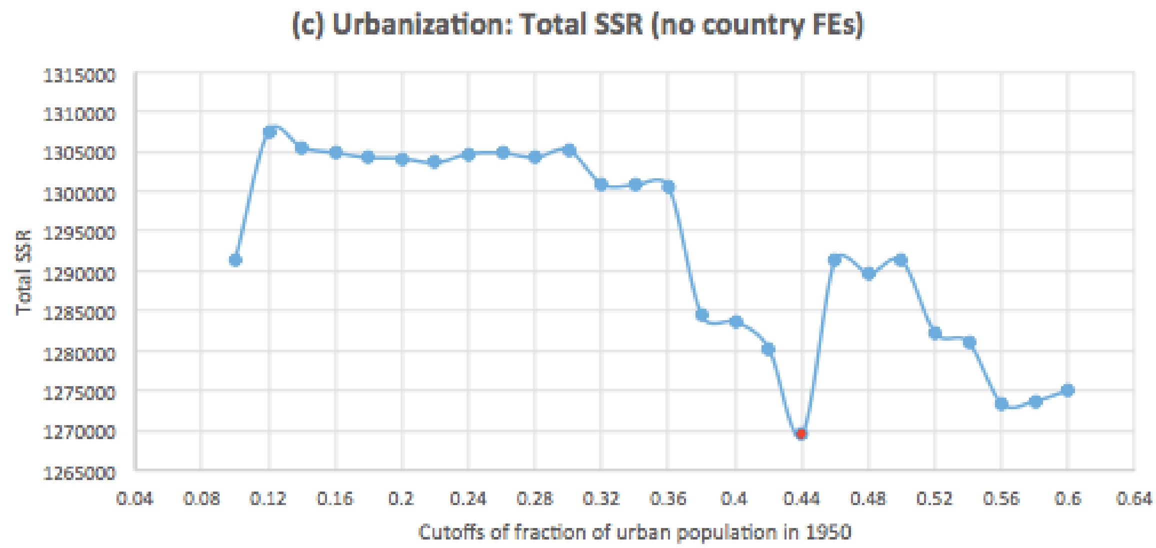


Figure 9. Cutoffs to minimize sum of squared residusla




(c) Urbanization: Total SSR (no country FEs)
Total SSR
1315000
1310000
1305000
1300000
1295000
1290000
1285000
1280000
1275000
1270000
1265000
0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4 0.44 0.48 0.52 0.56 0.6 0.64
Cutoffs of fraction of urban population in 1950

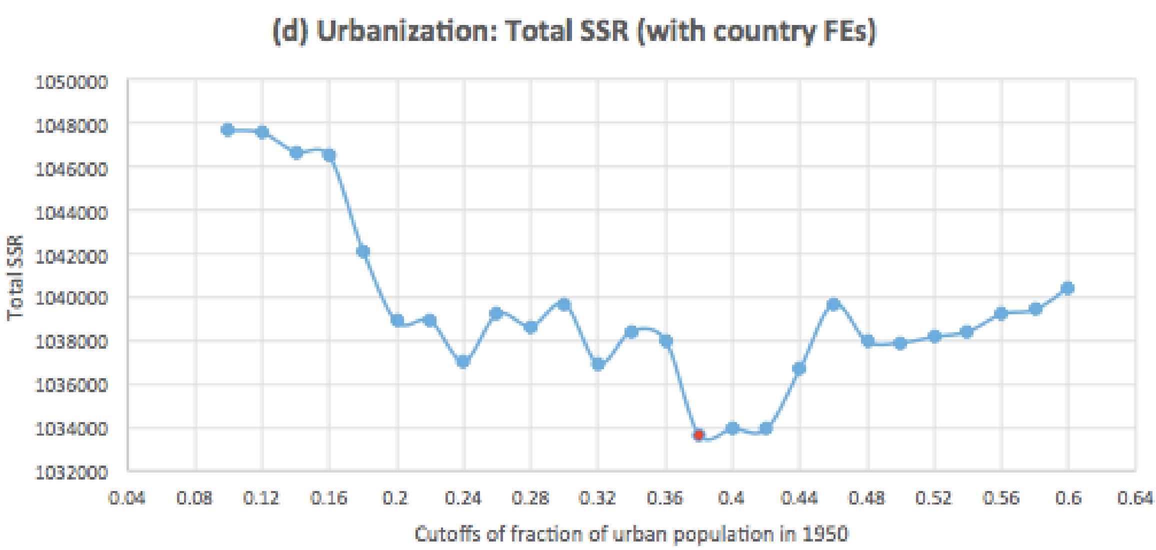
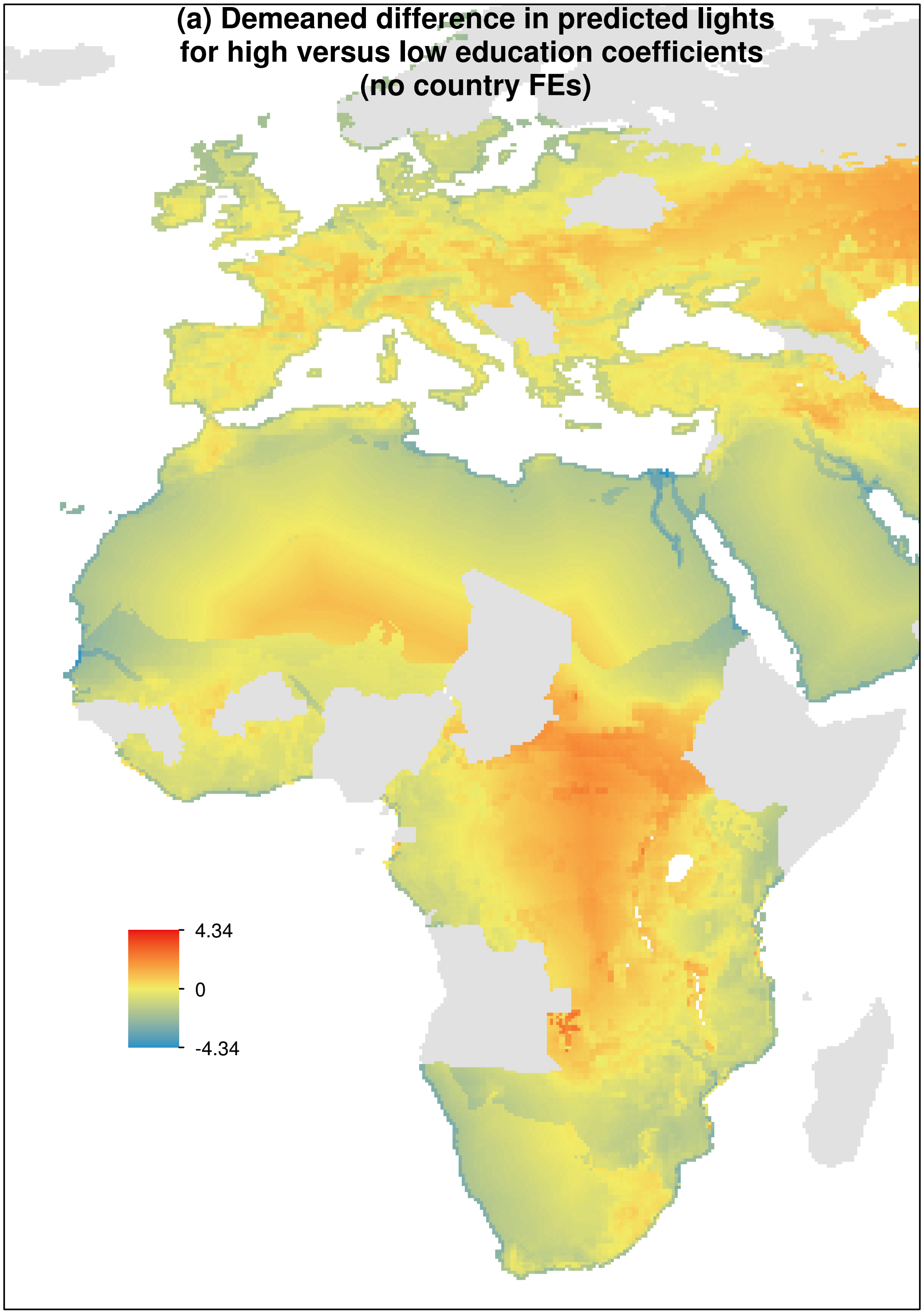

(d) Urbanization: Total SSR (with country FEs)
Total SSR
1050000
1048000
1046000
1044000
1042000
1040000
1038000
1036000
1034000
1032000
0.04 0.08 0.12 0.16 0.2 0.24 0.28 0.32 0.36 0.4 0.44 0.48 0.52 0.56 0.6 0.64
Cutoffs of fraction of urban population in 1950

Figure 10. Changes from having or being given high education coefficients (over low education ones)



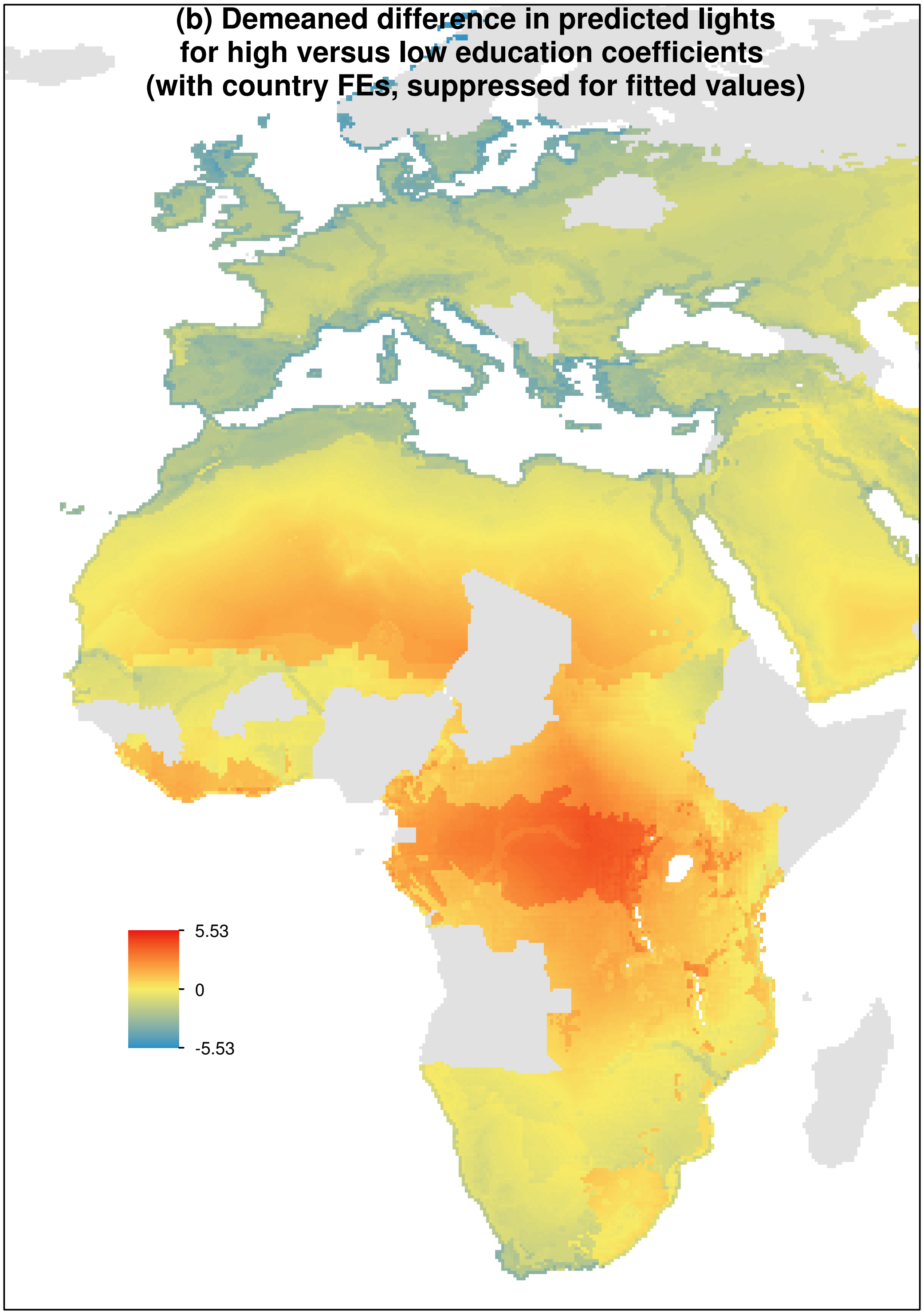

(b) Demeaned difference in predicted lights
for high versus low education coefficients
(with country FEs, suppressed for fitted values)
5.53
0
-5.53

Table A1: Conley Standard errors, kernel cutoff = 40 km

Dependent variable: lrad2010land_ csd

	(1)
ruggedness	-8.70e-06***
	(2.19e-06)
malaria index	-.0348355***
	(.0027154)
tropical moist forest	-.040639
	(.0800286)
tropical dry forest	.9364165***
	(.0994472)
temperate broadleaf	1.771348***
	(.0759276)
temperate conifer	.7689181***
	(.0899183)
boreal forest	-.441965***
	(.0858767)
tropical grassland	-.8617399***
	(.0591735)
temperate grassland	.7103923***
	(.0703476)
montane grassland	.6306202***
	(.0867233)
tundra	-.8745064***
	(.0970188)
Mediterranean forest	.8365996***
	(.0981427)
mangroves	.4176001**
	(.1943355)
temperature	.1737878***
	(.00393)
precipitation	-.0089633***

Continued on next page

Table A1 – *Continued from previous page*

	(1)
	(.000473)
growing days	.0098492***
	(.0003026)
land suitability	2.689555***
	(.059731)
abs(latitude)	.1135415***
	(.0028149)
elevation	.0005011***
	(.0000267)
$\mathbb{1}$ (coast)	.4216031***
	(.0444975)
distance to coast	-.0006645***
	(.0000306)
$\mathbb{1}$ (harbor<25km)	1.62537***
	(.0786706)
$\mathbb{1}$ (river<25km)	.7637808***
	(.0769342)
$\mathbb{1}$ (big lake<25km)	.316677***
	(.0324868)

Notes: *** p<0.01, ** p<0.05, * p<0.1

Table A2: Intensive and extensive margins

	OLS No FEs (1)	LPM No FEs (2)	OLS With FEs (3)	LPM With FEs (4)
ruggedness	-2.53e-05***	9.58e-07***	-3.31e-05***	-6.31e-07**
	(2.14e-06)	(3.05e-07)	(1.94e-06)	(2.70e-07)
malaria index	-0.0377***	-0.00499***	-0.0509***	-0.00640***
	(0.00320)	(0.000451)	(0.00446)	(0.000440)
tropical moist forest	-0.430***	0.0363***	-0.196***	-0.00467
	(0.0571)	(0.0109)	(0.0551)	(0.00988)
tropical dry forest	-0.117*	0.172***	-0.124**	0.0581***
	(0.0613)	(0.0126)	(0.0625)	(0.0112)
temperate broadleaf	0.735***	0.174***	0.561***	0.154***
	(0.0546)	(0.00987)	(0.0519)	(0.00921)
temperate conifer	0.384***	0.0985***	0.110*	0.0457***
	(0.0631)	(0.0120)	(0.0610)	(0.0117)
boreal forest	-0.242***	-0.0777***	-0.370***	-0.176***
	(0.0681)	(0.0118)	(0.0689)	(0.0121)
tropical grassland	-0.917***	-0.0948***	-0.273***	0.0149**
	(0.0580)	(0.00878)	(0.0586)	(0.00757)
temperate grassland	-0.0446	0.148***	0.0762	0.183***
	(0.0524)	(0.00984)	(0.0506)	(0.00883)
montane grassland	0.0293	0.0776***	0.396***	0.0835***
	(0.0683)	(0.0133)	(0.0669)	(0.0122)
tundra	-0.261***	-0.177***	-0.532***	-0.230***
	(0.0972)	(0.0131)	(0.106)	(0.0134)
Mediterranean forest	0.247***	0.0725***	0.385***	0.214***
	(0.0634)	(0.0123)	(0.0685)	(0.0129)
mangroves	-0.144	0.0676***	-0.0447	-0.0162
	(0.122)	(0.0230)	(0.113)	(0.0200)
temperature	0.0828***	0.0262***	0.0160***	0.0236***
	(0.00346)	(0.000508)	(0.00452)	(0.000601)

Continued on next page

Table A2 – *Continued from previous page*

	OLS No FEs (1)	LPM No FEs (2)	OLS With FEs (3)	LPM With FEs (4)
precipitation	-0.00327***	-0.00116***	-0.00519***	-0.00157***
	(0.000369)	(6.50e-05)	(0.000419)	(7.20e-05)
growing days	0.00418***	0.00132***	0.00500***	0.00110***
	(0.000217)	(3.92e-05)	(0.000231)	(4.08e-05)
land suitability	0.697***	0.417***	0.736***	0.348***
	(0.0402)	(0.00784)	(0.0414)	(0.00780)
abs(latitude)	0.0479***	0.0175***	0.00240	0.00935***
	(0.00212)	(0.000373)	(0.00327)	(0.000502)
elevation	-1.36e-05	8.78e-05***	-0.000310***	4.67e-05***
	(2.54e-05)	(3.84e-06)	(2.75e-05)	(4.08e-06)
$\mathbb{1}$ (coast)	0.916***	0.0102*	0.898***	0.0107**
	(0.0352)	(0.00528)	(0.0322)	(0.00452)
distance to coast	-0.000389***	-9.85e-05***	-0.000309***	-9.88e-05***
	(2.92e-05)	(4.56e-06)	(3.31e-05)	(5.21e-06)
$\mathbb{1}$ (harbor<25km)	0.492***	0.152***	0.568***	0.138***
	(0.0449)	(0.00824)	(0.0403)	(0.00752)
$\mathbb{1}$ (river<25km)	0.320***	0.118***	0.306***	0.106***
	(0.0548)	(0.00895)	(0.0485)	(0.00843)
$\mathbb{1}$ (big lake<25km)	0.246***	0.0443***	0.182***	0.0259***
	(0.0273)	(0.00393)	(0.0257)	(0.00361)
Observations	98,941	243,985	98,940	243,974
R-squared	0.262	0.385	0.359	0.475

Notes: Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3: List of high/ low education vs. high/ low urbanization countries

Country name	Education level	Urbanization level	High/ low, educ		High/ low, urban	
			No FEs	With FEs	No FEs	With FEs
New Zealand	9.19	0.725	high	high	high	high
Switzerland	8.84	0.444	high	high	high	high
United States of America	8.4	0.642	high	high	high	high
Slovakia	8.13	0.300	high	high	low	low
Czech Republic	8.10	0.542	high	high	high	high
Australia	8.04	0.770	high	high	high	high
Canada	7.60	0.609	high	high	high	high
Norway	7.40	0.505	high	high	high	high
Israel	7.30	0.710	high	high	high	high
Belize	7.23	0.553	high	high	high	high
Armenia	7.22	0.403	high	high	low	high
Hungary	7.13	0.530	high	high	high	high
Germany	6.80	0.681	high	high	high	high
Belgium	6.75	0.915	high	high	high	high
Sweden	6.75	0.657	high	high	high	high
Japan	6.73	0.534	high	high	high	high
United Kingdom	6.39	0.790	high	high	high	high
Ireland	6.23	0.401	high	high	low	high
Estonia	6.13	0.497	high	high	high	high
Netherlands	6.08	0.561	high	high	high	high
Austria	5.97	0.636	high	high	high	high
Slovenia	5.86	0.199	high	high	low	low
Iceland	5.70	0.728	high	high	high	high
Croatia	5.66	0.223	high	high	low	low
Denmark	5.51	0.680	high	high	high	high
Poland	5.40	0.383	high	high	low	high
Trinidad and Tobago	5.00	0.214	high	high	low	low
Argentina	4.85	0.653	high	high	high	high
Chile	4.81	0.584	high	high	high	high
Republic of Korea	4.50	0.214	high	high	low	low
Romania	4.38	0.256	high	high	low	low
Ukraine	4.37	0.355	high	high	low	low
China, Hong Kong SAR	4.36	0.852	high	high	high	high
Uruguay	4.34	0.779	high	high	high	high
France	4.33	0.552	high	high	high	high
Guyana	4.24	0.280	high	high	low	low
Italy	4.21	0.541	high	high	high	high
Greece	4.14	0.522	high	high	high	high
Tajikistan	4.13	0.294	high	high	low	low
Kyrgyzstan	4.04	0.265	high	high	low	low
South Africa	4.03	0.422	high	high	low	high
Finland	3.86	0.430	high	high	low	high

Continued on next page

Table A3 – *Continued from previous page*

Country name	Education level	Urbanization level	High/ low, educ		High/ low, urban	
			No FEs	With FEs	No FEs	With FEs
Latvia	3.84	0.464	high	high	high	high
Russian Federation	3.83	0.441	high	high	high	high
Spain	3.83	0.519	high	high	high	high
Bulgaria	3.82	0.276	high	high	low	low
Panama	3.76	0.358	high	high	low	low
Lithuania	3.71	0.288	high	high	low	low
Fiji	3.62	0.244	high	high	low	low
Jamaica	3.59	0.241	low	high	low	low
Cyprus	3.56	0.284	low	high	low	low
Costa Rica	3.55	0.335	low	high	low	low
Cuba	3.49	0.565	low	high	high	high
Sri Lanka	3.40	0.153	low	high	low	low
Luxembourg	3.39	0.672	low	high	high	high
Republic of Moldova	3.28	0.185	low	high	low	low
Taiwan	3.03	.	low	high	high	high
Réunion	2.85	0.235	low	low	low	low
Peru	2.83	0.410	low	low	low	high
Singapore	2.71	0.994	low	low	high	high
Paraguay	2.69	0.346	low	low	low	low
Albania	2.60	0.205	low	low	low	low
Kazakhstan	2.59	0.364	low	low	low	low
Ecuador	2.55	0.283	low	low	low	low
Dominican Republic	2.52	0.237	low	low	low	low
Mauritius	2.51	0.293	low	low	low	low
Viet Nam	2.47	0.116	low	low	low	low
Lesotho	2.47	0.018	low	low	low	low
Namibia	2.40	0.134	low	low	low	low
Colombia	2.33	0.327	low	low	low	low
Bolivia (Plurinational State of)	2.32	0.338	low	low	low	low
Saudi Arabia	2.31	0.213	low	low	low	low
Philippines	2.21	0.271	low	low	low	low
Mexico	2.19	0.427	low	low	low	high
Malaysia	2.08	0.204	low	low	low	low
Brazil	2.08	0.362	low	low	low	low
Thailand	2.04	0.165	low	low	low	low
Brunei Darussalam	2.02	0.268	low	low	low	low
Portugal	1.90	0.312	low	low	low	low
Zambia	1.77	0.115	low	low	low	low
Senegal	1.76	0.172	low	low	low	low
Honduras	1.64	0.176	low	low	low	low
Venezuela (Bolivarian Republic of)	1.63	0.473	low	low	high	high

Continued on next page

Table A3 – *Continued from previous page*

Country name	Education level	Urbanization level	High/ low, educ		High/ low, urban	
			No FEs	With FEs	No FEs	With FEs
Qatar	1.63	0.805	low	low	high	high
Mongolia	1.61	0.200	low	low	low	low
Zimbabwe	1.58	0.106	low	low	low	low
China	1.58	0.118	low	low	low	low
Nicaragua	1.55	0.352	low	low	low	low
El Salvador	1.53	0.365	low	low	low	low
Kuwait	1.48	0.615	low	low	high	high
Botswana	1.38	0.027	low	low	low	low
Jordan	1.33	0.370	low	low	low	low
Guatemala	1.31	0.251	low	low	low	low
Mauritania	1.27	0.031	low	low	low	low
Lao People's Democratic Republic	1.25	0.072	low	low	low	low
Swaziland	1.23	0.020	low	low	low	low
United Republic of Tanzania	1.21	0.035	low	low	low	low
Kenya	1.16	0.056	low	low	low	low
Myanmar	1.15	0.162	low	low	low	low
Turkey	1.11	0.248	low	low	low	low
Indonesia	1.09	0.124	low	low	low	low
Bahrain	1.00	0.644	low	low	high	high
Pakistan	0.99	0.175	low	low	low	low
India	0.99	0.170	low	low	low	low
Malawi	0.96	0.035	low	low	low	low
Bangladesh	0.93	0.043	low	low	low	low
Uganda	0.89	0.028	low	low	low	low
Algeria	0.85	0.222	low	low	low	low
Syrian Arab Republic	0.85	0.327	low	low	low	low
Côte d'Ivoire	0.84	0.100	low	low	low	low
United Arab Emirates	0.79	0.545	low	low	high	high
Congo	0.79	0.249	low	low	low	low
Cameroon	0.70	0.093	low	low	low	low
Ghana	0.68	0.154	low	low	low	low
Tunisia	0.65	0.323	low	low	low	low
Haiti	0.59	0.122	low	low	low	low
Democratic Republic of the Congo	0.58	0.191	low	low	low	low
Liberia	0.57	0.130	low	low	low	low
Iran (Islamic Republic of)	0.54	0.275	low	low	low	low
Egypt	0.52	0.319	low	low	low	low
Papua New Guinea	0.51	0.017	low	low	low	low
Mozambique	0.49	0.035	low	low	low	low
Gabon	0.47	0.114	low	low	low	low

Continued on next page

Table A3 – *Continued from previous page*

Country name	Education level	Urbanization level	High/ low, educ		High/ low, urban	
			No FEs	With FEs	No FEs	With FEs
Libya	0.44	0.195	low	low	low	low
Benin	0.44	0.050	low	low	low	low
Cambodia	0.42	0.102	low	low	low	low
Burundi	0.42	0.017	low	low	low	low
Sierra Leone	0.41	0.126	low	low	low	low
Gambia	0.40	0.103	low	low	low	low
Central African Republic	0.39	0.144	low	low	low	low
Togo	0.34	0.044	low	low	low	low
Rwanda	0.32	0.021	low	low	low	low
Sudan	0.32	0.068	low	low	low	low
Niger	0.32	0.049	low	low	low	low
Morocco	0.28	0.262	low	low	low	low
Afghanistan	0.27	0.058	low	low	low	low
Iraq	0.24	0.351	low	low	low	low
Mali	0.15	0.085	low	low	low	low
Nepal	0.11	0.027	low	low	low	low
Yemen	0.02	0.058	low	low	low	low
Gibraltar	.	1.000	.	.	high	high
Monaco	.	1.000	.	.	high	high
French Guiana	.	0.537	.	.	high	high
Isle of Man	.	0.529	.	.	high	high
Bahamas	.	0.521	.	.	high	high
Falkland Islands (Malvinas)	.	0.510	.	.	high	high
Greenland	.	0.490	.	.	high	high
Suriname	.	0.469	.	.	high	high
Azerbaijan	.	0.457	.	.	high	high
Turkmenistan	.	0.450	.	.	high	high
Puerto Rico	.	0.406	.	.	low	high
Djibouti	.	0.398	.	.	low	high
Andorra	.	0.388	.	.	low	high
Georgia	.	0.369	.	.	low	low
Guadeloupe	.	0.358	.	.	low	low
Lebanon	.	0.320	.	.	low	low
Dem. People's Republic of Korea	.	0.310	.	.	low	low
Uzbekistan	.	0.289	.	.	low	low
Belarus	.	0.262	.	.	low	low
New Caledonia	.	0.246	.	.	low	low
Montserrat	.	0.158	.	.	low	low
Equatorial Guinea	.	0.155	.	.	low	low
Cabo Verde	.	0.142	.	.	low	low
Bosnia and Herzegovina	.	0.137	.	.	low	low

Continued on next page

Table A3 – *Continued from previous page*

Country name	Education level	Urbanization level	High/ low, educ		High/ low, urban	
			No FEs	With FEs	No FEs	With FEs
Sao Tome and Principe	.	0.135	.	.	low	low
Samoa	.	0.129	.	.	low	low
Somalia	.	0.127	.	.	low	low
Guinea-Bissau	.	0.100	.	.	low	low
Timor-Leste	.	0.099	.	.	low	low
Vanuatu	.	0.088	.	.	low	low
Oman	.	0.086	.	.	low	low
Nigeria	.	0.078	.	.	low	low
Madagascar	.	0.078	.	.	low	low
Angola	.	0.076	.	.	low	low
Eritrea	.	0.071	.	.	low	low
Guinea	.	0.067	.	.	low	low
Comoros	.	0.066	.	.	low	low
Ethiopia	.	0.046	.	.	low	low
Chad	.	0.045	.	.	low	low
Burkina Faso	.	0.038	.	.	low	low
Solomon Islands	.	0.038	.	.	low	low
Bhutan	.	0.021	.	.	low	low

Notes: Education cutoffs are 3.6 (no FEs) and 3 (with FEs), and urbanization cutoffs are 0.44 (no FEs) and 0.38 (with FEs).

Table A4: High education differentials

Dependent variable: lrad2010land_ csd

	(1) Variable	(2) Interaction	(3) Variable	(4) Interaction
constant		-2.790***		-1.371
		(0.330)		(395.7)
ruggedness	-1.39e-05***	4.50e-06	-1.69e-05***	5.55e-06
	(3.10e-06)	(4.20e-06)	(2.79e-06)	(3.62e-06)
malaria index	-0.0419***	0.00893	-0.0274***	-0.0610***
	(0.00282)	(0.0127)	(0.00280)	(0.00875)
tropical moist forest	-0.0213	-0.947***	-0.0737	0.870***
	(0.0934)	(0.250)	(0.0812)	(0.246)
tropical dry forest	1.037***	0.758	0.350***	0.341
	(0.102)	(0.503)	(0.0922)	(0.345)
temperate broadleaf	1.595***	-0.0942	0.961***	0.301**
	(0.0922)	(0.147)	(0.0913)	(0.136)
temperate conifer	0.610***	0.0487	0.316***	-0.131
	(0.119)	(0.171)	(0.118)	(0.161)
boreal forest	-0.416***	-0.114	-0.0320	-0.964***
	(0.142)	(0.185)	(0.136)	(0.174)
tropical grassland	-0.481***	-0.797***	-0.0120	-0.356***
	(0.0729)	(0.124)	(0.0669)	(0.107)
temperate grassland	0.267***	0.199	0.363***	0.471***
	(0.0895)	(0.130)	(0.0878)	(0.125)
montane grassland	0.745***	-0.0139	0.297***	1.248***
	(0.110)	(0.183)	(0.101)	(0.169)
tundra		-0.509***		-0.685***
		(0.128)		(0.115)
Mediterranean forest	0.886***	-0.232	1.915***	-1.034***
	(0.109)	(0.169)	(0.127)	(0.164)
mangroves	-0.684***	1.261*	-0.733***	0.703
	(0.193)	(0.711)	(0.178)	(0.467)
temperature	0.154***	0.0148*	0.123***	-0.0132
	(0.00631)	(0.00781)	(0.00794)	(0.00911)

Continued on next page

Table A4 – *Continued from previous page*

	(1) Variable	(2) Interaction	(3) Variable	(4) Interaction
precipitation	-0.00823***	0.00185**	-0.00986***	0.00125
	(0.000563)	(0.000940)	(0.000651)	(0.000851)
growing days	0.00919***	0.00288***	0.00708***	0.00142**
	(0.000395)	(0.000591)	(0.000413)	(0.000563)
land suitability	2.040***	1.383***	2.106***	-0.123
	(0.0735)	(0.111)	(0.0788)	(0.111)
abs(latitude)	0.121***	0.00698	0.0908***	-0.0887***
	(0.00410)	(0.00559)	(0.00573)	(0.00697)
elevation	0.000444***	6.30e-05	0.000233***	-0.000615***
	(3.92e-05)	(5.69e-05)	(4.11e-05)	(5.82e-05)
$\mathbb{1}$ (coast)	1.237***	-1.041***	1.291***	-1.225***
	(0.0780)	(0.0897)	(0.0732)	(0.0809)
distance to coast	-0.00156***	0.00182***	-0.00145***	0.00155***
	(3.96e-05)	(5.87e-05)	(4.83e-05)	(6.61e-05)
$\mathbb{1}$ (harbor<25km)	1.697***	0.0416	1.687***	-0.304**
	(0.116)	(0.145)	(0.108)	(0.128)
$\mathbb{1}$ (river<25km)	1.203***	-0.900***	1.094***	-0.639***
	(0.112)	(0.132)	(0.108)	(0.125)
$\mathbb{1}$ (big lake<25km)	0.375***	-0.0135	0.425***	-0.341***
	(0.0519)	(0.0603)	(0.0484)	(0.0557)
Country FEs	No		Yes	
Observations	228,690		228,690	
R-squared	0.492		0.581	

Notes: Columns (1) and (3) report the coefficient on the variables listed, and columns (2) and (4) report the coefficient on the variable's interaction with the high-education dummy. High education equals 1 if average years of education is greater than or equal to 3.6 (no country FEs version) or 3 (with country FEs version). Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A5: High urbanization differentials

Dependent variable: lrad2010land_ csd

	(1)	(2)	(3)	(4)
	Variable	Interaction	Variable	Interaction
constant		-1.206***		-1.756
		(0.307)		
ruggedness	-1.69e-05***	1.12e-05***	-1.87e-05***	9.10e-07
	(2.82e-06)	(3.94e-06)	(2.53e-06)	(3.37e-06)
malaria index	-0.0393***	-0.0262***	-0.0271***	-0.122***
	(0.00271)	(0.00982)	(0.00251)	(0.0100)
tropical moist forest	0.282***	-1.860***	0.272***	-2.090***
	(0.0882)	(0.193)	(0.0778)	(0.195)
tropical dry forest	1.181***	-0.865***	0.535***	-0.640***
	(0.103)	(0.259)	(0.0931)	(0.192)
temperate broadleaf	1.814***	-0.506***	1.178***	0.111
	(0.0848)	(0.144)	(0.0853)	(0.131)
temperate conifer	1.155***	-0.687***	0.671***	-0.587***
	(0.112)	(0.167)	(0.112)	(0.157)
boreal forest	0.719***	-1.409***	0.121	-1.053***
	(0.192)	(0.226)	(0.135)	(0.173)
tropical grassland	-0.254***	-1.385***	0.163***	-0.547***
	(0.0679)	(0.121)	(0.0606)	(0.104)
temperate grassland	0.705***	-0.440***	0.623***	0.158
	(0.0863)	(0.129)	(0.0821)	(0.122)
montane grassland	0.885***	-0.505***	0.515***	0.877***
	(0.0940)	(0.195)	(0.0917)	(0.158)
tundra	-2.955***	2.348***	-0.698***	
	(0.519)	(0.534)	(0.115)	
Mediterranean forest	1.234***	-0.801***	2.014***	-1.224***
	(0.105)	(0.170)	(0.124)	(0.162)
mangroves	0.205	-1.241***	-0.0140	-1.302***
	(0.198)	(0.479)	(0.178)	(0.359)
temperature	0.164***	-0.0163**	0.138***	-0.0328***
	(0.00583)	(0.00731)	(0.00761)	(0.00881)

Continued on next page

Table A5 – *Continued from previous page*

	(1) Variable	(2) Interaction	(3) Variable	(4) Interaction
precipitation	-0.00753***	0.00109	-0.00925***	-1.17e-05
	(0.000539)	(0.000881)	(0.000625)	(0.000824)
growing days	0.00807***	0.00412***	0.00770***	0.000552
	(0.000371)	(0.000572)	(0.000384)	(0.000539)
land suitability	1.952***	1.580***	1.853***	0.415***
	(0.0656)	(0.110)	(0.0720)	(0.106)
abs(latitude)	0.123***	-0.0139***	0.103***	-0.104***
	(0.00355)	(0.00517)	(0.00551)	(0.00675)
elevation	0.000454***	-4.84e-05	0.000220***	-0.000361***
	(3.55e-05)	(5.30e-05)	(3.93e-05)	(5.49e-05)
$\mathbb{1}$ (coast)	0.974***	-0.765***	1.127***	-0.951***
	(0.0710)	(0.0832)	(0.0682)	(0.0760)
distance to coast	-0.00138***	0.00165***	-0.00136***	0.00128***
	(3.61e-05)	(5.70e-05)	(4.65e-05)	(6.46e-05)
$\mathbb{1}$ (harbor<25km)	1.522***	0.226*	1.462***	-0.0359
	(0.106)	(0.136)	(0.0967)	(0.119)
$\mathbb{1}$ (river<25km)	1.173***	-0.787***	0.915***	-0.390***
	(0.106)	(0.127)	(0.107)	(0.125)
$\mathbb{1}$ (big lake<25km)	0.375***	-0.0479	0.359***	-0.238***
	(0.0469)	(0.0560)	(0.0458)	(0.0535)
Country FEs	No		Yes	
Observations	243,661		243,661	
R-squared	0.483		0.579	

Notes: Columns (1) and (3) report the coefficient on the variables listed, and columns (2) and (4) report the coefficient on the variable's interaction with the high-urbanization dummy. High urbanization equals 1 if the fraction of urban population is greater than or equal to 0.44 (no country FEs version) or 0.38 (with country FEs version). Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A6: Spatial autoregressive model results

	Full sample		Lit sample	
	Rook	Queen	Rook	Queen
	(1)	(2)	(3)	(4)
Neighbors' lights (rook), avg	0.990***		0.658***	
	(0.00135)		(0.00260)	
Neighbors' lights (queen), avg		1.001***		0.642***
		(0.00151)		(0.00276)
ruggedness	-8.21e-06***	-9.97e-06***	-1.92e-05***	-2.12e-05***
	(5.88e-07)	(6.78e-07)	(1.08e-06)	(1.16e-06)
malaria index	-0.00144***	-0.00143**	-0.00612***	-0.00687***
	(0.000445)	(0.000563)	(0.00145)	(0.00153)
tropical moist forest	-0.0478***	-0.0570***	-0.202***	-0.215***
	(0.0137)	(0.0166)	(0.0240)	(0.0262)
tropical dry forest	-0.00282	-0.0157	-0.182***	-0.193***
	(0.0173)	(0.0211)	(0.0249)	(0.0273)
temperate broadleaf	-0.0226*	-0.0536***	0.0338	0.0198
	(0.0125)	(0.0155)	(0.0233)	(0.0256)
temperate conifer	-0.00362	-0.0226	0.0534*	0.0532*
	(0.0170)	(0.0209)	(0.0284)	(0.0313)
boreal forest	-0.0414***	-0.0402**	0.0123	0.0266
	(0.0142)	(0.0178)	(0.0307)	(0.0339)
tropical grassland	-0.0394***	-0.0394***	-0.0463*	-0.0560**
	(0.00927)	(0.0115)	(0.0249)	(0.0270)
temperate grassland	-0.00187	-0.00811	-0.184***	-0.208***
	(0.0112)	(0.0141)	(0.0226)	(0.0249)
montane grassland	0.0654***	0.0621***	-0.0228	-0.0227
	(0.0189)	(0.0230)	(0.0338)	(0.0370)
tundra	-0.0339**	-0.0250	0.270***	0.338***
	(0.0158)	(0.0197)	(0.0462)	(0.0514)
Mediterranean forest	0.00543	-0.000963	0.0434*	0.0377
	(0.0156)	(0.0193)	(0.0256)	(0.0280)

Continued on next page

Table A6 – *Continued from previous page*

	Full sample		Lit sample	
	Rook	Queen	Rook	Queen
	(1)	(2)	(3)	(4)
mangroves	-0.212***	-0.227***	-0.0793	-0.0903
	(0.0653)	(0.0729)	(0.0669)	(0.0718)
temperature	-0.00406***	-0.00747***	-0.0160***	-0.0195***
	(0.000601)	(0.000751)	(0.00154)	(0.00170)
precipitation	-0.000412***	-0.000469***	-0.000804***	-0.000933***
	(7.55e-05)	(9.31e-05)	(0.000154)	(0.000169)
growing days	0.000158***	6.45e-05	0.000264***	0.000188**
	(4.85e-05)	(5.99e-05)	(8.69e-05)	(9.51e-05)
land suitability	0.0601***	0.0428***	-0.152***	-0.177***
	(0.0103)	(0.0128)	(0.0170)	(0.0187)
abs(latitude)	-0.00292***	-0.00550***	-0.0114***	-0.0138***
	(0.000433)	(0.000539)	(0.000943)	(0.00104)
elevation	-3.83e-05***	-5.37e-05***	-0.000126***	-0.000146***
	(4.99e-06)	(6.14e-06)	(1.15e-05)	(1.26e-05)
$\mathbb{1}$ (coast)	0.0745***	0.106***	0.537***	0.595***
	(0.00991)	(0.0118)	(0.0171)	(0.0186)
distance to coast	1.90e-05***	3.50e-05***	3.21e-05**	4.89e-05***
	(4.67e-06)	(6.05e-06)	(1.27e-05)	(1.41e-05)
$\mathbb{1}$ (harbor<25km)	0.361***	0.491***	0.233***	0.295***
	(0.0217)	(0.0256)	(0.0223)	(0.0245)
$\mathbb{1}$ (river<25km)	0.155***	0.222***	0.173***	0.229***
	(0.0157)	(0.0197)	(0.0228)	(0.0251)
$\mathbb{1}$ (big lake<25km)	0.0261***	0.0376***	0.0871***	0.0986***
	(0.00649)	(0.00794)	(0.0129)	(0.0142)
Observations	243,970	243,973	98,934	98,935
R-squared	0.823	0.803	0.631	0.600

Notes: Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1