Agriculture, Transportation and the Timing of Urbanization:
Global Analysis at the Grid Cell Level

Mesbah J. Motamed[a]
Raymond J.G.M. Florax[b]
William A. Masters[c]

[a]   *United States Department of Agriculture, United States of America;*
[b] *Purdue University, United States of America; Faculty of Economics and Business Administration, and VU University Amsterdam, and Tinbergen Institute, The Netherlands;*
[c] *Tufts University, United States of America.*

Agriculture, Transportation and the Timing of Urbanization:
Global Analysis at the Grid Cell Level[*]

Mesbah J. Motamed [a,†], Raymond J.G.M. Florax [b,c,d] and William A. Masters [e]

[a] *Economic Research Service, U.S. Department of Agriculture, Washington D.C., USA*
[b] *Department of Agricultural Economics, Purdue University, West Lafayette, IN, USA*
[c] *Department of Spatial Economics, VU University Amsterdam, Amsterdam, The Netherlands*
[d] *Tinbergen Institute, Amsterdam, The Netherlands*
[e] *Friedman School of Nutrition and Department of Economics, Tufts University, Boston, MA, USA*

November 29, 2013

Abstract

This paper addresses the timing of a location's historical transition from rural to urban activity. We test whether urbanization occurs sooner in places with higher agricultural potential and comparatively lower transport costs, using worldwide data that divide the earth's surface at half-degree intervals into 62,290 cells. From an independent estimate of each cell's rural and urban population history over the last 2,000 years, we identify the date at which each cell achieves various thresholds of urbanization. Controlling for unobserved heterogeneity across countries through fixed effects and using a variety of spatial econometric techniques, we find a robust association between earlier urbanization and agro-climatic suitability for cultivation, having seasonal frosts, better access to the ocean or navigable rivers, and lower elevation. These geographic correlations become smaller in magnitude as urbanization proceeds, and there is some variance in effect sizes across continents. Aggregating cells into countries, we show that an earlier urbanization date is associated with higher per capita income today.

Keywords: economic growth, economic geography, urbanization, agriculture, transportation
JEL-classification: C21, N50, O11, O18, R1

---

1 Introduction

This paper addresses the role of physical geography in the timing and location of urbanization. Our central hypothesis is that, for any given place, higher local agricultural potential and lower physical barriers to transport facilitate the earlier establishment of cities and subsequent economic growth in the associated region. This could help explain late urbanization and the persistence of rural poverty in geographically disadvantaged regions, and help guide intervention to overcome agricultural and infrastructural obstacles to economic transformation.

Our focus on the initial formation and early growth of cities is consistent with new work that links events in the distant past to countries' current state of economic development (see Nunn, 2014). Many of these studies rely on variation in geography and climate to explain economic outcomes, either directly or indirectly. Sokoloff and Engerman (2000) showed that climate and soils favorable to plantation crops led to institutions that concentrated and preserved power among landholding elites. Similarly, Acemoglu et al. (2001) linked differences in disease ecology and the mortality of colonial settlers to the kinds of institutions they established in the New World, Nunn (2008) exploits variations in terrain and elevation to identify the impact of Africa's historic slave trade on modern outcomes, and Nunn and Qian (2011) find that the parts of Europe and Asia naturally more suitable for growing potatoes experienced accelerated rates of population growth and urbanization after potatoes were imported from South America.

Meanwhile, a related stream of work focuses on people rather than places, looking for population traits acquired long ago, their transmission across generations, and their persistent effects on economic development (Spolaore and Wacziarg, 2013). In this literature, the roots of economic growth are traced back to events such as early technology adoption (Comin et al., 2010), ethnolinguistic diversity (Michalopoulos, 2008) and genetic diversity (Ashraf and Galor, 2013), which in turn may be traced back to physical geography.

Jared Diamond's *Guns, Germs, and Steel* is perhaps the most popular exposition of geography's role in shaping human history (Diamond, 1997). Diamond focuses on soil fertility and the presence of edible plants, potentially domesticated wildlife, and infectious diseases, tracing their role in the development of ancient societies around the world. Olsson and Hibbs (2005), in a direct test of Diamond's hypotheses, use continent size, east-west orientation, climate, and availability of potentially domesticated plants and animals to explain the timing of a region's transition into sedentary agriculture. In their empirical results, they find significant evidence in support of Diamond's hypothesis. Putterman (2008), in an extension of Olsson and Hibbs (2005), refines the analysis by observing these variables' effects on population density in the year 1500 as well as on modern-day income, controlling for the transmission of agricultural technologies via migration. A more formal test of the hypothesis of historic populations correlating with modern economic outcomes appears in Putterman and Weil (2010). Additional evidence in support of the geography hypothesis is presented in Ashraf and Galor (2011), who show that the date of a population's first transition into agriculture, as determined by its location's suitability for agriculture, helps explain the level of population density and technological achievement thousands of years later.[1]

---

[1] The papers discussed here all focus on how physical geography might have influenced historical events and institutions. A related but different literature looks at the contemporary effects of geography, including for example the role of access to navigable waterways, frost prevalence, and malarial ecology (Gallup et al., 1999; Masters and McMillan, 2001; Sachs, 2003). Some authors, such as Easterly and Levine (2003), find that time-invariant geographic variables influence modern incomes only through their historical effect on institutions, but climatic conditions could play an ongoing role because of variance over time and the possibility of forecasting future trends (Dell et al., 2013).

Urbanization at a given location is a particularly important aspect of economic history, in part because it is readily identifiable and rarely reversed. One can think of urbanization in terms of a break from farm households' Malthusian dependence on local natural resources to more use of man-made capital, specialization and agglomeration which facilitate the start of economic growth (Lucas, 2000).[2] Variation in the timing of urbanization has been linked to differences in institutional quality, total factor productivity, and geography (Gollin et al., 2002; Ngai, 2004). Economic transition itself entails a variety of contemporaneous changes, including rising productivity in both agriculture and manufacturing; sectoral shifts in employment, output and expenditure from agriculture to manufacturing and services; and a demographic transition in mortality, fertility and the age structure of the population (Kuznets, 1973; Williamson, 1988). We focus only on the most visible, and thus best recorded, aspect of economic transition, which is the shift from rural to urban residence.

Our approach to economic geography utilizes global grid cell data, disaggregating the whole world into regular cells defined only by their latitude and longitude. To our knowledge, the first paper on economic growth to exploit such data directly is Masters and McMillan (2001), who test a variety of possible influences on cell-level population density, and then aggregate cell-level attributes to the country level to test their relationship with national per capita income growth. More recently, Nordhaus (2006) disaggregates all economic activity into 1° by 1° cells, and finds the level of activity to be highly correlated with climate and topography.

Economic data rendered in grid cell units, as opposed to political, administrative, or other geographic regions, offer important advantages for empirical analysis. Such regions are smaller than most countries or even provinces, and thus offer a larger number of distinct observations. The boundaries of each observation are defined arbitrarily in regular intervals of geographic coordinates and not subject to the potential endogeneity of administrative boundaries.[3] The corresponding challenge is to allocate socioeconomic activity across grid cells, with the underlying data originally being collected on the basis of administrative boundaries.

We begin by motivating our model specification in terms of the broader literature on geographic determinants of urbanization. We then demonstrate the need to control for spatial autocorrelation, rejecting the null hypothesis that each cell's urbanization measure is a spatially independent observation. Spatial dependence can be due to spatially correlated omitted variables, or to spatial interactions between cells through common institutions, trade, migration, investment or other

---

[2] Urbanization has been associated with higher productivity and more desirable living standards ever since antiquity. Long before Adam Smith observed his pin factory, the ancient Greeks observed shoe-making. Xenophon's *Cyropedia* (1914) contains this remarkable passage about the value of agglomeration, written in the early fourth century BC:

> In a small city the same man must make beds and chairs and ploughs and tables, and often build houses as well; and indeed he will be only too glad if he can find enough employers in all trades to keep him. Now it is impossible that a single man working at a dozen crafts can do them all well; but in the great cities, owing to the wide demand for each particular thing, a single craft will suffice for a means of livelihood, and often enough even a single department of that; there are shoe-makers who will only make sandals for men and others only for women. Or one artisan will get his living merely by stitching shoes, another by cutting them out, a third by shaping the upper leathers, and a fourth will do nothing but fit the parts together. Necessarily the man who spends all his time and trouble on the smallest task will do that task the best.

[3] For example, if economically successful societies tend to seek control over locations with attractive geographical features (e.g., access to navigable waterways), their institutions would become correlated with geography even if their initial wealth accumulation was caused by something else.

channels (McMillen, 2003). To control for many different types of distance-dependent spatial processes, we employ a very general spatial econometric approach that uses neighbors' information to provide a heteroskedasticity and autocorrelation consistent (HAC) estimator (see Anselin, 2006, p. 931). The specific HAC-approach we use is a non-parametric estimator for spatial data developed by Kelejian and Prucha (2007). Beyond these statistical controls, each cell's agro-climatic suitability for agriculture still correlates closely with its speed of urbanization, with a particularly strong link at early stages of city growth. Access to low-cost transportation also facilitates urbanization, particularly at later stages.

The remainder of this paper is organized as follows. Section 2 motivates the examination of geography's relationship to economic transition. Sections 3–5 introduce the data and the econometric model, and present and discuss the empirical results. Section 6 concludes.

2   Motivation

In the standard literature on economic growth, explanations for the extreme differences in present-day national incomes initially focused on recent growth rates, which were shown to be associated with disparities in national savings rates, technical change, human capital, institutions, and even genetic distances between populations (Solow, 1956; North, 1990; Mankiw et al., 1992; Spolaore and Wacziarg, 2009). More recently, models have focused instead on differences in income levels, which could be primarily driven by the timing of transition from a Malthusian regime of low or zero long-run growth into a modern regime in which higher growth rates are sustained over time (Lucas, 2000). If growth paths within each regime are similar and differences are attributable to the date of transition, then the relevant research question is historical: it is not so much what differences exist today, but rather why some places started down the path of modern economic growth earlier than others.

Figure 1 suggests that both the start date and the pace of modern growth are important. For each region, the development of per capita income has a clear structural break.[4] Western Europe's income began to rise soon after 1800, Latin America's approximately fifty years later, and Asia about fifty years after that. Africa's income appears to break more gradually, though it certainly occurs at the end of the sequence.

Most evidence suggests that per capita incomes grew negligibly prior to industrialization. Over the millennia, human development resulted in considerable population growth but not much improvement in living standards (Birdsall, 1988). Despite constant incomes, however, population growth and the accumulation of knowledge could have changed things so that switching to new technologies eventually became profitable (Galor and Weil, 2000; Hansen and Prescott, 2002; Strulik and Weisdorf, 2008). Such a switch is said to have occurred in eighteenth-century England, perhaps involving changes in property rights (McCloskey, 1972; Clark, 1992; Clark and Clark, 2001), improved soil fertility through animal manure (Chorley, 1981) or nitrogen-fixing legumes (Allen, 2008), or the introduction of New World crops (Nunn and Qian, 2011) raising farm productivity in a way that could have helped sustain industrialization over time.

⟨ Figure 1 about here ⟩

---

[4] The regions are grouped following Maddison et al. (2001), with the exception that we combine east and west Asia into one category from which we exclude Japan.

4

Improvements in agricultural productivity have long been seen as necessary for the sustained growth of industry (Johnston and Mellor, 1961), typically because of nutrition constraints and Engel's Law (Matsuyama, 1992; Kögel and Prskawetz, 2001; Gollin et al., 2002). The need for food implies that, in a closed economy, low farm productivity traps workers in agriculture. Only when agricultural production per worker rises above subsistence needs can resources shift to non-farm activities and sustain urbanization.

A location's successful agriculture does not, however, guarantee industrialization and may not even be necessary if manufacturing involves scale economies and transport costs. Krugman (1991) demonstrated how scale economies interact with transport costs to drive agglomerations of firms in a particular place.[5] Murata (2008) links Engel's Law to the spatial location of firms to show that falling transport costs can raise real wages and drive up food consumption. Later, upon satisfying the food requirement, consumption of manufactured goods begins to rise, and ultimately, this spurs agglomeration. Capital accumulation and the use of intermediate goods can accelerate this process, as in the model of Gallup et al. (1999) where lower transport costs help make capital investment more profitable.

To help explain the location of cities, we test the simple hypothesis that geographic features associated with suitability to grow crops or lower transport costs facilitate earlier urbanization dates, *ceteris paribus*. The use of worldwide grid cell data to test this hypothesis, with a variety of controls for neighborhood effects and country-level institutions, provides a new kind of evidence about economic development.

3   Data

Our focus in this paper is on variation in physical geography and climate, exploiting grid cell data on variables that could be associated with either rural productivity or transportation costs. Since these variables are fixed in time, we can test their link to urbanization in terms of the historical date at which each location's urban population passed various thresholds, driven by some exogenous source of demographic change or knowledge accumulation. Our goal is to explain variance across grid cells in the timing of that urbanization.

The unit of observation is an arbitrarily-defined geographic region, which in our case are grid cells demarcated by lines at intervals of $0.5\,°$ latitude and $0.5\,°$ longitude. Such cells are largest at the equator, where each cell covers approximately 3,000 square kilometers. At the Arctic Circle, grid cells are about one-third that size. The total number of such half-degree cells in our dataset is 62,290.

Empirically, we are testing $T = f\left(G_r, G_u\right)$, where T is the year at which a grid cell acquired an urban population of a specific size, G_r are the grid cell's geographic correlates of rural productivity, and G_u are its geographic correlates of lower transport costs. To permit a causal interpretation of any correlation between T and these geographic features G, we must choose features that are not themselves influenced by urbanization, either directly or indirectly. The strengths and limitations of each regressor are discussed in turn below, after we present the data on urbanization.

---

[5] Numerous variations on this original model have appeared since then, including Fujita and Krugman (1995), Puga (1999) and Fujita et al. (2001).

3.1 Urbanization

Historical data on the location and population of cities have been collected and assigned to grid cells and matched with estimates of the surrounding rural populations by Klein Goldewijk et al. (2010). The authors used a wide range of census figures, historical records, scholarly publications, and interpolations over time and space to construct an internally consistent time series of rural and urban populations for each grid cell, from year 0 to 1600 in century intervals, and from 1700 to 2000 in decade intervals. The details underlying the construction of their data set can be found in Klein Goldewijk et al. (2010) and its supplemental materials.

From those data we compute for each cell a time period T defined as the number of years that have passed since year t at which a cell reached a particular urbanization threshold. Thus, $T = 2000 - t$.[6] Urbanization thresholds can then be defined in either of two ways: as the absolute number of urban people in that grid cell, or as the urban fraction of that cell's total population. Our tests compare results using both definitions, at various levels of the threshold.

First, in terms of the absolute number of urban people, we follow Acemoglu et al. (2002), who built on Bairoch (1998) and judged urbanization to have started once a location could support over 5,000 urban inhabitants. Since grid cells closer to the equator are larger, we normalize by area and define the Bairoch rule in terms of the smallest cell that supported an urban population greater than 5,000 in the year 2000. That cell turns out to have had a size of 881 square kilometers, so our threshold for the absolute number of urban people is actually 5.67 urban inhabitants per square kilometer.[7] This value is then applied to grid cells of other sizes, offering a consistent measure of economic activity for cross-sectional comparison purposes. For comparison, we also use a lower threshold of one person per square kilometer. At the equator, this corresponds to approximately 3,000 urban inhabitants per grid cell.

For robustness, we compare results with the number of urban people against results based on the fraction of people who are urban, for which we follow Williamson (1988) and many others and consider a range of thresholds of 10, 25, and 50 percent. Figure 2 presents the geographic distribution of the duration since transition at the 10 percent level. Maps for the other thresholds have a similar appearance, and formal summary statistics are presented at the end of this section.

The long time period and detailed spatial resolution of the data set raises potential concerns about inaccuracy in the estimated population of each cell, particularly in earlier time periods and for rural people. In practice, however, our results depend only on the urban counts in more recent years. Ninety-nine percent of all cells that achieved 10 percent urbanization as shown in Figure 2 did so in the last 400 years. Among the cells that reached 50 percent urbanization, 99 percent occurred within the past 170 years. In other words, nearly all the historic urbanization addressed in this study occurred in the modern period and has been relatively well measured. The grid cells around ancient civilizations, such as the southern Mediterranean region or Central America, often did not reach even the 10 percent urbanization threshold until recently, and so they are shown lightly shaded in Figure 2.

⟨ Figure 2 about here ⟩

---

[6] We use the year 2000 as the base year, since that is the last year for which data are available.

[7] This smallest of all urbanized grid cells lies above the Arctic Circle, along the northern coast of Norway. A threshold of 5,000 urban people divided by its area of 881 square kilometers yields the 5.67 urban people per square kilometer.

A related concern arises from truncation of our urbanization data, which spans from 0 to 2000 AD. Truncation at 0 AD arises only for the lowest of our five urbanization thresholds, as a few cells actually reached one urban person per square kilometer prior to our earliest observation. Such locations appear as if they urbanized "only" 2000 years ago, whereas they actually passed the threshold earlier. There are very few such grid cells, however, and our regressions use logarithms which limits the influence of errors in the early years. Conversely, locations that have recently or will soon pass each threshold are shown as not yet urbanized, reflecting the historical nature of our analysis. The issue of truncation is also addressed econometrically using Tobit estimates. And finally, a few locations passed back and forth across one of our urbanization thresholds during the 0–2000 time period, for example regions that reached 10 percent urbanization early, then slipped back and may or may not have recovered. These are counted as of their first date to achieve that level of urbanization, and our use of several different thresholds ensures robustness to such cases.

3.2 Cultivation Suitability Index

To capture soil and climate conditions associated with agriculture, we use an index of each grid cell's suitability for crop cultivation prepared by Ramankutty et al. (2002). The index is based on recently constructed global climate and soil maps, from which Ramankutty et al. (2002) calculate four types of agronomic constraints on crop growth involving soil temperature, soil moisture, soil structure and soil chemistry. Each of the four constraints is derived from standard models of plant physiology, and scaled from zero to one. The four variables are then multiplied together to form a single zero-to-one index of agronomic suitability, where zero implies that plant growth is impossible, and one implies that the entire grid cell could potentially be cultivated. Figure 3 presents a map of these data.

⟨ Figure 3 about here ⟩

The four constraints in the Ramankutty et al. (2002) index are: (1) the number of growing degree-days (GDD) during a standard growing season, (2) the level of soil moisture available (α) to plants in a generic crop growth model, (3) the total carbon density (C_{soil}) in the top 30 cm of soils, and (4) the acidity (pH_{soil}) of that soil. Each cell's average annual level of GDD and α is computed from underlying daily temperature, precipitation, and many other variables, while C_{soil} and pH_{soil} estimates are computed from soil maps. Each variable is then transformed into a zero-to-one index of suitability using sigmoidal or stepwise linear functional forms, whose parameters are calibrated to fit around the fraction of land in each grid cell that was actually observed to have been under cultivation in 1992.

The Ramankutty et al. (2002) procedure uses a kind of envelope function, capturing the degree to which any of the four agronomic constraints limits crop growth. The resulting index classifies as agronomically suitable many areas that are not actually cultivated, presumably because other constraints are binding. For the world as a whole, Ramankutty et al. (2002) find that in 1992 only 44 percent of agronomically suitable land was under crops. Most of this suitable but uncultivated area was in the remote inland tropical woodlands of Africa and Latin America, where cultivation is limited by factors other than plant growth potential. The authors also point out that South Asia had much more cultivated area than its natural endowment of suitable land, presumably due to widespread irrigation, fertilization and other investments that improve soils beyond their naturally-occurring levels of moisture, carbon and pH.

7

Before using the Ramankutty et al. (2002) index to test whether suitability for crop cultivation facilitated urbanization, we must consider two possible kinds of endogeneity bias in that relationship. First, the index might capture aspects of the environment, such as disease transmission, that affect urbanization independently from crop cultivation. Second, the index might have been calibrated to aspects of the environment that capture modern but not historical agriculture, and thereby reflect a consequence rather than a cause of urbanization. Both omitted-variable and reverse-causality biases could mistakenly attribute to agriculture a causal effect on urbanization that is actually due to something else. Such biases cannot be ruled out by statistical tests, and our identification strategy therefore rests on the data generation process.

In our model, identification of a causal effect for agriculture relies on the way in which Ramankutty et al. (2002) fit their cultivation suitability index to natural constraints on the growth of plants. These constraints differ from natural constraints on the growth of animals, insects, or other life forms, and they differ from constraints that are specific to the growth of modern crops. Much of the underlying data that are used to compute the Ramankutty et al. (2002) index would also appear in measures of other natural phenomena, but they would then take very different functional forms. For example, the index for transmissibility of malaria computed by Kiszewski et al. (2004) also uses temperature and moisture, but mosquitoes and parasites respond to heat very differently from plants, resulting in a very different spatial distribution than the map of cultivation suitability in Figure 3. An index of suitability for modern agriculture in particular would also produce a very different-looking map, if only because modern crops have been bred to succeed in temperate zones whereas plants in general grow freely in the tropics.

Given that urbanization requires many factors other than plant growth, we might expect little or no correlation between cultivation suitability and the location of cities at any given time. The bivariate relationship between them is illustrated in Figure 4, plotting the mean Klein Goldewijk et al. (2010) urbanization rate at each level of the Ramankutty et al. (2002) cultivation index in selected years. Means are computed using a local polynomial regression to compute confidence intervals at each level of cultivation suitability over all cells that were actually inhabited in that year.[8] The resulting bivariate relationships do not control for transport costs or other factors, and omit all of the uninhabited cells where peoples' subsistence needs are not met.

⟨ Figure 4 about here ⟩

Figure 4 reveals that there was indeed very little correlation between cultivation suitability and urbanization in 1800, but that the bivariate correlation became steeper as urbanization rates rose over time. Whatever was driving increased urbanization may have done so faster in the cells more suitable for agriculture. It is this difference in upward trend that we exploit in our cross-sectional tests, regressing the number of years since each individual cell crossed various thresholds on that cell's own cultivation suitability index and other factors described below.

3.3 Transportation cost

In open economies, urbanization could arise independently of local farm productivity, perhaps using low transport costs to support agglomeration. To capture this channel we use a grid cell's distance from a navigable waterway, for access to trade on the ocean or by river.

---

[8] The 95-percent confidence intervals around each line correspond to fitting a zero-degree polynomial with bandwidths of 0.12 for the year 1800, 0.1 for the year 1900, and 0.09 for the remaining years. Bandwidths are selected by rule of thumb using Stata's *lpolyci* command.

Coastal access is measured by distance in kilometers from the centroid of each half-degree grid cell to the nearest unbounded sea or ocean. Coastal transport is presumably most important for long-distance trade, to enlarge the market for locally produced goods and to bring in goods from elsewhere.

River access is measured by the highest Strahler order of streams in each grid cell (Vörösmatry et al., 2000). The Strahler index is a categorical variable ranging from 1 to 6, with six representing the largest river into which all tributaries flow and one representing a small stream with no tributaries; zero signals the absence of rivers and streams. We interpret larger order values to reflect greater inherent navigability of a river. The location and Strahler order of rivers and streams are observed in the 1990s, but these attributes rarely change and using them avoids the potential endogeneity associated with measures of stream flow or actual navigability. Water flow and navigability are more mutable, and could be influenced by human interventions such as water use, dredging, locks, and the design of suitable boats.

3.4   Seasonal frost and elevation

In addition to cultivation suitability and access to the coast or a navigable river, we also control for seasonal frost and elevation. These capture additional dimensions of the physical environment that could matter for urbanization, perhaps through disease transmission or other factors associated with seasonality and altitude.

For frost, we follow Masters and McMillan (2001) and use a dummy variable to capture whether or not a cell receives frost in winter, after a frost-free summer. This measure is intended to capture a specific aspect of seasonality, by which annual frosts help drive insects, micro-organisms and many animals into a period of dormancy until the next warm season. This could help people control pests, pathogens and disease vectors more easily than in places where the lack of frost allows more diverse organisms to thrive year-round. The prevalence of ground frost is computed from IPCC (2002) as a dummy variable equal to one for cells that receive a significant level of frost in winter, conditional on having no frost in summer. For this purpose, winter is defined as December-February in the northern hemisphere and June-August in the southern hemisphere, and vice-versa for summer. The cutoff number of frost days in winter we use is 2.11 frost days per month, corresponding to the threshold estimated econometrically by Funke and Zuo (2003). All data correspond to averages over the 1961–1990 period; the frequency of frost will have varied over the centuries due to climate change, but its spatial pattern is more stable and is accurately estimated only for the modern period.

For elevation, we use the average altitude of each grid cell from Hijmans et al. (2005). This variable may be related to disease transmission and also the ease of surface transportation. More complex measures of topography were considered for this paper, including a grid cell's standard deviation of elevation, its average gradient, and its ruggedness, but these are all highly correlated with average elevation. For hypothesis testing we chose the simplest possible measure, so as to avoid any concern about endogeneity due to ex-post selection from among a large pool of candidate variables.

Including average elevation and winter frost alongside each cell's cultivation-suitability index and access to a coast or navigable river gives us several distinct dimensions of physical geography that might influence a city's initial location and subsequent growth. To keep the focus entirely on physical geography our regressions control for country or continent fixed effects, and are subject to a variety of tests for spatially correlated errors and heteroskedasticity.

3.5 Summary statistics

To test for links from physical geography to urbanization, we exclude cells that were entirely uninhabited over the entire time frame. Also, to permit controls for spatial effects we exclude isolated cells that have no immediate neighbors. The resulting sample consists of 52,143 grid cell observations, summarized in Table 1. The dependent variables of interest are the number of years since various thresholds of urbanization were reached. Their maximum values reveal that, at the start of our data about 2000 years ago, some cells had already reached the first threshold of one urban inhabitant per square kilometer. About 200 years later the first cells reached the threshold of having 10 percent of their population in urban areas. The mean date of urbanization, however, was much more recent and many cells have not yet passed any of our five thresholds.

⟨ Table 1 about here ⟩

4 Econometric Model

Our basic specification is to regress a grid cell's time since its date of urbanization, as measured by one of the five thresholds listed above, on our geographic variables plus country fixed effects:

$$\ln T_{ij} = \beta_0 + \beta_1 \ln cultiv_{ij} + \beta_2 \ln coast_{ij} + \beta_3 river_{ij} + \beta_4 frost_{ij} + \beta_5 \ln elevation_{ij} + \delta_j + \varepsilon_{ij}, \ (1)$$

where T_{ij} represents the number of years elapsed since grid cell i in country j passed the given level of urbanization, *cultiv* is the cultivation suitability index, *coast* is the distance from a grid cell's centroid to the nearest unbounded sea or ocean coast, *river* indicates the highest Strahler order of streams of a river present in the grid cell, *frost* indicates the presence of frost in winter after a frost-free summer, and *elevation* is the average elevation of the cell's land area. All of the continuous variables are in logarithmic form, so coefficients can be interpreted as elasticities.[9]

Country fixed effects are captured by the parameters δ_j, using a year-2000 map of country borders to absorb social or institutional variation at the country level in a consistent manner. Our data set comprises 181 countries. Grid cells that span more than one country are assigned to the country in which most of their area is situated. The bulk of observed urbanization occurred under contemporary borders, but even locations that urbanized earlier and then changed nationalities are well served by using the most recent country definitions, because over time the number of countries increased, effectively leading to a less restrictive specification. As a result of the fixed effects specification, our estimated coefficients refer only to within-country variation. Robustness checks include regressions without these controls as well as regressions with artificial fixed effects based on creating rectangular "super-grids" across the globe. Given the geographic location of landmasses, we define 169 artificial fixed effects for the grid cells in our sample.[10] The artificial fixed effects are comparable in number to the original set of country fixed effects, but they do not suffer from the potential endogeneity that could be associated with country borders.

---

[9] In applying the logarithmic transformation to the urbanization variables, log(0) is defined as zero. The log of distances shorter than 1 km is also fixed at zero, and negative elevations are set to the smallest positive elevation in the data set.

[10] The borders of the super-grids are defined as the 5-percentile points of the entire longitude and latitude span of the globe, which leads to 19×19 (= 361) super-grids. The spatial distribution of inhabited grid cells is such that they are located in 169 of these super-grids.

As described earlier, cells that have not reached a particular urbanization threshold receive zero values for T. In fact, across the different urbanization thresholds, more than two-thirds of the observations have not experienced any urbanization; concurrently, there are very few grid cells that were already urbanized over 2000 years ago.[11] This outcome can be interpreted in two ways. First, the presently-observed sample may reflect a long-run equilibrium outcome; thus the rural or urban character of a cell is no longer expected to change. Or, alternatively, the outcomes could be understood to reflect a snapshot of a process that is not yet complete. In the future, some of these zero-valued cells may reach some urban threshold, implying a presently negative value of T instead of the observed zero that is reported here. If T is censored at zero (and at 2000), naïve estimation of equation (1) could involve substantial attenuation bias, which we can address using Tobit regressions as detailed below.

Our regression specification is highly stylized, omitting a wide range of variables, which could lead to bias in estimating equation (1) by ordinary least squares (OLS). We include country fixed effects to mitigate this. However, modern country boundaries typically refer to smaller regions sharing unobservable factors, which may cause spatial autocorrelation. In addition, we use only the cell's own values in the model, although one could expect there to be at least some interactions between cells that influence the timing of each others' urbanization. As illustrated in Figure 2, urbanization outcomes appear clustered in certain parts of the world, including Europe, the eastern seaboard of the United States, Central America, and Japan. To confirm this, Moran's I statistics, which measure the extent of a variable's spatial autocorrelation, are reported in Table 1 for each of the urbanization threshold variables.

To account for the spatial autocorrelation due to shared unobserved characteristics and/or unmodeled interaction across grid cells we use a non-parametric heteroskedasticity and autocorrelation consistent (HAC) estimator of the variance-covariance matrix for spatial data, developed by Kelejian and Prucha (2007). This spatial variance-covariance matrix is estimated as:

$$\hat{\Omega}_{\text{HAC}} = n^{-1} \sum_i \sum_j x_i x_j \hat{\varepsilon}_i \hat{\varepsilon}_j K(d_{ij}/d), \tag{2}$$

where n refers to sample size, i and j now both refer to grid cells, x corresponds to the explanatory variables given in equation (1), $\hat{\varepsilon}$ are the estimated OLS residuals from equation (1), and K is a kernel function to ensure that the estimated variance-covariance matrix is positive-definite. In the kernel estimation, grid cells i and j separated by a distance d_{ij} smaller than the suitable cutoff-distance d, are utilized to estimate the (co)variances. In the empirical analysis we use a Gaussian kernel based on arc-distances between the geographical midpoints of the grid cells, with a maximum of 48 neighbors (comparable to the contiguity matrix described in footnote b of Table 1).

Alternatively, one can use a parametric approach based on explicitly modeled spatial processes, which are typically introduced through the inclusion of a spatial lag of the dependent variable and/or a spatially autoregressive error component. Lagrange Multiplier (LM) tests can guide the choice of specification (Anselin et al., 1996). This approach is further detailed below.

In the following section we present and discuss the OLS/HAC results. We also report on various robustness checks, including Tobit estimation, parametric spatial process models, and the introduc-

---

[11] The large probability mass at zero makes the distribution of the urbanization threshold variables very right-skewed and leptokurtic (peaked). For instance, the 10-percent urbanization variable has a skewness of 56 and a kurtosis of 51. The logarithmic transformation results in a distribution that is more similar to the normal distribution, with skewness (0.76) and kurtosis (–0.99) indicators being much closer to zero.

tion of various forms of spatial heterogeneity using fixed effects and spatially varying parameters.

5 Empirical results and discussion

5.1 Global effects

Table 2 presents the spatial HAC results with and without country-level fixed effects, which clearly reject the hypothesis that urbanization's timing is uncorrelated with our primary geographic variables, except for elevation in the case of the model without country fixed effects. The effects are "global" in the sense that the coefficients are restricted to be constant across space, an assumption that will be relaxed below. The use of HAC standard errors is clearly warranted, given the results for the misspecification tests. The Moran's I test on the residuals reveals a pattern of modest spatial correlation, with values of the quasi-correlation coefficient ranging from 0.15 to 0.23, and strong evidence that the null hypothesis of spatial randomness should be rejected. As is often the case in spatial data analysis, the null hypothesis of homoskedasticity is rejected as well.

⟨ Tables 2 and 3 about here ⟩

As a robustness check for the spatial HAC results reported above we also performed regressions based on a parametric spatial perspective. For all urbanization thresholds, using the third-order queen contiguity weights, the Lagrange Multiplier (LM) test procedure outlined in Anselin et al. (1996) revealed the spatial autoregressive error model as the most likely alternative specification.[12] Since spatially correlated errors do not invalidate the OLS estimator in terms of unbiasedness and consistency, but only affect the efficiency of OLS, the impact on the magnitude of the estimated coefficients is negligible. The estimation results for these parametric spatial models do therefore not meaningfully alter the conclusions drawn utilizing the non-parametric spatial HAC estimator.

Table 3 represents another robustness check and takes into account that the dependent variables for the different urbanization thresholds are censored from below, and in one case also from above. The table documents the estimated coefficients using the Tobit estimator as well as the associated average marginal effects. The magnitude of the marginal effects of the Tobit estimator and the HAC estimator and their statistical significance are virtually identical; even although the Tobit estimator is implemented with clustered standard errors by country, rather than with spatially autocorrelated errors as in the HAC estimator.

In order to provide intuition with respect to the estimation results, we discuss the results in terms of marginal effects. Consider, for example, the results for the 10-percent urbanization model with country fixed effects from Table 2, noting that marginal effects at the sample mean, in years, depend on the type of right-hand side variable (discrete or continuous) and whether a variable enters the regression equation in logarithmic form.[13] Based on the coefficient for cultivation

---

[12] A spatial lag specification, comprising the spatially lagged dependent variable, is not well suited to these economic data since it would assume that spillovers are global in nature. We also experimented with the addition of spatial crossregressive variables, which measure the interaction with immediate neighbors and hence imply local spillovers, but these estimated effects are difficult to interpret. Detailed results are not reported here for reasons of space, but they are available from the authors upon request.

[13] The estimated coefficients for cultivation suitability, distance to the coast, and elevation are elasticities, and their marginal effect at the sample mean, in years, therefore equals $\partial y/\partial x = \beta \bar{y}/\bar{x}$. The estimated coefficient for order of streams is a rate of change, with $\partial y/\partial x = \beta \bar{y}$ as the corresponding marginal effect. Frost is a dummy variable and we define its "marginal" effect as $\partial y/\partial x = (\exp^{\beta} - 1)\bar{y}_0$, where $\bar{y}_0$ is the sample mean of y in grid cells without frost (see

suitability in the 10 percent urbanization model, a 0.01-increase in the cultivation suitability index introduces transition approximately 0.3 years, or nearly four months earlier. Increasing distance from the coast by 100 km delays transition by 1.3 years. Increasing elevation by 100 meters delays transition by almost 0.2 years, or about 2 months. Raising a cell's order of streams by one unit and the occurrence of seasonal frost introduce urbanization slightly more than 8 and 18 years earlier, respectively. Across all specifications, we find strong statistical significance for all of our variables, with the exception of elevation in certain specifications and at certain thresholds.

An interesting feature of our results is that most geographic characteristics decline in importance at higher urbanization levels. Specifically, cultivation suitability, river order, and frost exhibit smaller absolute values of the coefficients going across the first two columns as well as across the last three columns of the models including fixed country effects in Table 2.[14] These aspects of physical geography are closely linked to the initial formation of urban settlements, but have significant but lower correlation with further urbanization beyond that. This is also visible in the Tobit marginal effects in Table 3, where the same downward trend appears for distance to the coast.

Controlling for country fixed effects, the results reported above correspond to within-country differences. Those country effects are themselves potentially of interest, since they capture impacts of social, cultural, institutional and policy factors beyond the explicitly modeled geographic factors and simple distance-dependent effects absorbed in the spatially correlated errors. Table 2 shows that the effects of geography for the within-country results are generally somewhat smaller than for the base model; in terms of statistical significance, the results are very similar (except for the elevation variable). We report adjusted R^2-measures for three different specifications: the base model that only contains the geography variables, the model with country fixed effects in addition to the geography variables, and a model including country fixed effects and the unique cross-products of the geography variables. The latter can be seen as a specification that maximizes the "impact" of the geography variables, but we do not report detailed estimation results here because the interpretation of marginal effects of such a specification is cumbersome. Although we do not formally address the geography–culture–institution debate (Acemoglu and Robinson, 2012), the measures of fit demonstrate that the amount of variance explained by our geography factors is generally bigger than the variance explained by "institutional" differences modeled by means of country fixed effects, at all levels of urbanization except the highest 50-percent threshold. The least restrictive specification, which also includes interactions among the geography factors, increases the variance explained by the model even more, but only marginally so.[15]

There is of course the possibility that country borders as such are endogenous, and incorporating country fixed effects would then invalidate our identification approach. Given the similarity

---

Halvorsen and Palmquist, 1980). To provide intuition about the magnitude of the effects we report "marginal effects" as the number of years earlier a cell reaches the 10% urbanization threshold, based on a 0.01-change in cultivation suitability, a 100 km increase in coastal distance, a 100 m increase in elevation, an increase in the Strahler index by 1 unit, and a unity switch for the frost variable. In judging these "marginal effects" one should keep in mind that for large parts of the world, urbanization is a relatively recent phenomenon, and comparing the effects against the entire development span of 2000 years would not be appropriate.

[14] Although not a formal statistical test, it is easy to see that 99-percent confidence intervals for the estimated coefficients across the models with different urbanization thresholds are frequently non-overlapping.

[15] An important question for future research is to investigate the interaction effects among geographical endowments in more detail. Our regressions containing interaction terms generally show that access to transportation is a complement rather than a substitute to agricultural suitability. However, interaction effects are easier to interpret for variables without logarithmic transformations, and an interaction model should preferably also include threshold effects and non-linearities in addition to coefficients that are allowed to vary over space.

of the OLS/HAC results and the within-results this is rather unlikely, but we further investigated this possibility by introducing fixed effects derived by means of exogenously determined rectangular super-grid cells rather than countries (see footnote 10 for details). For reasons of space the estimation results are not reported here, but they do not differ qualitatively from the results with fixed effects defined by country borders in the year 2000.

For illustration purposes, the effects of geography and (unobserved) country characteristics on the predicted number of years since meeting the 10-percent urbanization threshold are presented in Figures 5 and 6. These maps show the effect on the timing of urbanization of differential geography characteristics for the "average" country, and of differences in country characteristics for grid cells with "average" geographical endowments.[16]

<center>⟨ Figures 5 and 6 about here ⟩</center>

The geographic variables alone predict urbanization to occur earliest in a handful of unsurprising locations: the northern coast of Europe, including Amsterdam, the area around Venice in southern Europe, the Mediterranean coast, eastern China around Shanghai, and the area around Dhaka in South Asia. In the Western Hemisphere, this includes the deltas of the Mississippi River around New Orleans, as well as the Rio de la Plata which terminates in Buenos Aires, Argentina. Interestingly, the model predicts a large unrealized spatial concentration of early urbanizers around the Caspian Sea, which is classed as a seacoast but had no outlet to the open ocean until completion of the Volga-Don canal in 1952. Africa falls nearly entirely in the later half of the predicted urbanization period. The only grid cell predicted to urbanize early corresponds to the city of Accra in Ghana. Similarly, the interiors of Latin America and South Asia are predicted to lag in urbanization, relative to Europe and the Asian and North American coasts. Locations predicted to experience urbanization last are typically surrounded by deserts, mountains, or permafrost. Population densities in these parts of the world are so sparse that they may never urbanize.

Figure 5 shows that the variation in the number of years ago since urbanization attributable solely to our five geographic variables ranges from zero to 26 years. These values are small relative to the full span of our data, but given that grid cells meeting the 10-percent urbanization threshold did so on average approximately 100 years ago, our five variables can account for substantial

---

[16] The predictions are calculated using the OLS coefficients in column (3) of the second panel of Table 2, and the associated unreported coefficients for fixed effects. Since the prediction is in years rather than in the logarithm of years, two subtleties are important. First, assuming independence of the errors and the explanatory variables, it can be shown that (Cameron and Trivedi, 2010):

$$\mathrm{E}(T|X) = \mathrm{E}(\exp(\varepsilon))\,\mathrm{E}(\exp(\ln T)).$$

We can therefore calculate the predicted effect on the timing of urbanization in years as:

$$\widehat{T} = \alpha_0 \exp(\widehat{\ln T_G} + \widehat{\ln T_C}),$$

where the subscripts refer to the geography and the country fixed effects parts of the specification, and α_0 is the expected value of $\exp(\varepsilon)$. A consistent estimate of α_0 can be obtained from the auxiliary regression:

$$T = \alpha_0 \exp(\widehat{\ln T}) + \mu.$$

In our case, the OLS estimate for α_0 equals 1.45. Second, in the maps we show $\widehat{T_G}$ and $\widehat{T_C}$ at the mean sample value for $\widehat{\ln T_C}$ and $\widehat{\ln T_G}$, respectively. Our results should therefore be interpreted as showing the global distribution of the number of years since a grid cell belonging to an "average" country reaches the 10-percent urbanization threshold due to its geography. Similarly, we show the global distribution of the number of years since a country with "average" geographical endowments reaches the 10-percent urbanization threshold due to its country characteristics.

<center>14</center>

differences in the timing of urbanization over the relevant time period. Clearly some of the variation is systematically absorbed by country dummies. Figure 6 shows earlier-than-average urbanization primarily in Northwest Europe, but also the rest of Europe, the Near East, plus Argentina, parts of Mexico and Uruguay, a few coastal countries in Africa, Mongolia and New Zealand.

5.2 Continent-specific effects

The results presented above summarize the global effects of selected geographic features on a location's date of urbanization. A concern arises, however, that these global effects might be driven by one continent (in particular, Europe) and that the selected geographic variables might not similarly explain outcomes in other continents. To address this question, we describe the variation in urbanization outcomes and geography at the continent level, and subsequently subject our model to a variety of restrictions on the continental coefficients to test whether the geographic variables behave similarly across different continents.

Table 4 presents continent-specific average values for each of the urbanization thresholds and the selected geographic features. Certain broad patterns emerge from the table. On average, Europe experiences urbanization, across all thresholds, much earlier than any other continent. Meanwhile, Europe's average grid cell has the highest cultivation suitability, the lowest distance to the coast, the second-highest order of streams (just slightly behind South America), the highest occurrence of winter frost, and the second lowest elevation (after Australia). In contrast, Africa is among the latest urbanizers, rivaling only sparsely-populated Australia. Africa's geography is unfavorable, with its average grid cell possessing the lowest cultivation suitability, the second-highest distance to the coast, the second-lowest order of streams, the second-lowest occurrence of winter frost, and the second-highest elevation.

⟨ Tables 4 and 5 about here ⟩

The gap in urbanization outcomes and geographic traits between Europe and Africa raise the question whether the global patterns detected in the earlier section also operate within continents. Using the 10-percent threshold model as our test case, we introduce continent-level interactions with each of the geographic variables. Table 5 presents the results using the spatial HAC estimator for the base model including continent fixed effects at the 10-percent urbanization threshold. Cultivation suitability, distance to the coast, and order of streams appear statistically significant in all six continents. Perhaps surprisingly, frost and elevation do not exert a significant effect on the timing of urbanization in Europe, and they are statistically less significant in Africa and Asia. Apart from the robustness across continents, we test the restriction of parameter equality across all continents using a spatially-explicit Chow test (Anselin, 1988). Results from the test on the model parameters reveal systematic variation across continents, for each variable separately as well as overall.

The estimated coefficients for the continuous explanatory variables are elasticities, describing how sensitive the onset of urbanization is to various factors. The cultivation suitability elasticity of the onset of urbanization is highest in Europe and South America. Concurrently, Europe has the lowest coastal distance elasticity. The elevation elasticity is practically zero in Europe, but interestingly positive in Australia and negative in North America. The effect of the order of streams is relatively similar across continents, except for Australia, and the impact of frost is relatively high in the Americas and, to a lesser extent, in Australia.

The effects described above might be more intuitively interpreted if we recast them in terms of their original units of measurement and calculated continent-specific marginal effects at the sample

means of each continent. The bottom of Table 5 presents the number of years earlier a location reaches the 10% urbanization threshold based on the parameters presented in the top half of the table, and the mean sample values from Table 4. The results show that large differences appear across the continents. Except for frost and elevation, marginal effects at the sample mean are invariably highest in Europe. Although many of the effects, measured in years, appear small, one should keep in mind that the vast majority of cells that experienced urbanization at the 10% level, did so less than 200 years ago. From that perspective, the effects of distance to the coast and order of streams in Europe, and of order of streams in the Americas and Australia are quite substantial.

It is interesting to compare Europe to Africa, which frequently has the lowest marginal effects. Africa's suitability index must be six times higher than Europe's to establish the same effect on the onset of urbanization. Its order of streams must rise by more than 14 times. Even with relatively modest differences in the estimated coefficients across continents, the difference in marginal effects at the sample mean are substantial between, for instance, Africa and Europe because of the large difference in geographical endowments.

5.3 Transition and modern national income

We have shown that physical geography is closely linked to the date at which populations urbanize. Recall Figure 1 in which incomes are plotted over time for different regions. A similar pattern holds with urbanization for different geographies. For example, Figure 7 shows the time path of average urbanization rates for the universe of all grid cells that eventually achieved at least 10% urbanization by the year 2000, divided into two samples: one line traces average urbanization for those cells whose cultivation suitability index is above the median value, 0.41, and the other traces the average of all grid cells whose cultivation suitability index is below that value. Figure 7 illustrates the idea that locations with higher agricultural land quality may have, on average, urbanized earlier than areas with lower agricultural land quality.

⟨ Figures 7 and 8 about here ⟩

Since urbanization plausibly links geography to modern income, we ask whether the historic date of a country's urbanization matters to income today. Following Olsson and Hibbs (2005) and Putterman (2008), the implicit hypothesis is that the earlier a country was able to urbanize, the higher its present-day income can be. Figure 8 illustrates this possibility, plotting a country's real per capita income against the number of years since any of its grid cells first reached 10 percent urbanization.[17] There is a strong positive correlation between current income and the time spell since urbanization began across countries, but it is also evident that there are three distinctive time periods. On the right hand side of the graph we find China and India, representing old civilizations with a substantial degree of urbanization already 1,500 and more years ago. The plurality of countries are "late" urbanizers, with urbanization at the 10% level occurring less than 400 or 500 years ago. In the middle part of the graph we find many European countries, including the older city-states of the Ancient Greeks and Romans, as well as Japan.

We test the relationship between national income and cell-level urbanization more formally in a regression framework, again accounting for spatial dependence using the non-parametric spatial HAC estimator. The dependent variable here is per capita national income in PPP prices in 2000

---

[17] We use the country's first cell to urbanize, rather than the average date of a country's cells, to capture the start of a country's transition wherever it may occur.

16

taken from the Penn World Table (Heston et al., 2006).[18] We regress the logarithm of per capita income on the number of years since at least one of a country's grid cells passed each of the urbanization thresholds.

$$\ln(income_i) = \beta_0 + \beta_1 trans_i + \epsilon_i, \tag{3}$$

where $income_i$ is country i's per capita income in the year 2000 and the variable $trans_i$ represents country i's duration since transition of at least one grid cell to the various urbanization thresholds. Table 6 presents the results of the estimation and shows, as before with the grid cell analysis, clear evidence of spatial correlation as measured by Moran's I applied to the regression residuals. In reference to Figure 8 we provide results utilizing the spatial HAC estimator for a linear and a quadratic version of the hypothesized relationship between per capita income and time since first urbanization, as well as for the linear model using a subsample of countries which first urbanized during the last 400 years.[19] Additional robustness tests, which are not reported in terms of numerical results here, include regressions based on a parametric spatial approach as well as specifications including continent fixed effects.

⟨ Table 6 about here ⟩

Across all urbanization criteria, the historical transition date is a statistically significant correlate of incomes today, when controlling for spatial dependence among neighbors and using a linear or quadratic specification. From Table 6, reaching the 10 percent urbanization threshold one year earlier is associated with nearly 0.1 percent higher per capita income in 2000, and reaching the 50 percent urbanization one year earlier is associated with a 0.4 percent higher per capita income, assuming a simple linear model. If we restrict the sample to countries that have first met an urbanization threshold during the last 400 years, the effects are slightly smaller, but clearly still significantly different from zero. Alternatively, using the quadratic specification to accommodate the distinct pattern among "early" and "late" urbanizers shown in Figure 8, we find increasing impacts of early transition dates on GDP per capita for extended periods of time. Moreover, all transition durations in the sample are associated with positive impacts, and at the sample mean of the transition variable the effect of early transition has an even higher impact than in the linear model.

As robustness checks, in addition to the controls for spatial autocorrelation and heteroskedasticity described above, we also subjected our model to additional continent-level controls. The results were robust to these added variables, with estimated coefficients that closely matched the results from the base model in both linear and quadratic forms.

In sum, for regions where sizeable levels of urbanization have yet to be achieved, the transition effect can account for economically important differences in present-day incomes. The results presented here confirm that a country's income is correlated with having had an earlier start of urbanization, demonstrating the importance of our finding that geographic variation in the

---

[18] The Penn World Table Version 6.2 has 188 countries. In order to facilitate the spatial analysis and correct for missing data we excluded small island economies (Antigua, Bahrain, Barbados, Bermuda, Dominica, Grenada, Kiribati, Macao, Maldives, Malta, Micronesia, Netherlands Antilles, Palau, Seychelles, St. Kitts & Nevis, St. Lucia, St. Vincent & Grenadines, Tonga) as well as Serbia and Montenegro, leading to a sample of 169 countries.

[19] The latter regression basically only takes into account the left hand side of the scatterplot in Figure 8. In the case of subsampling we used the OLS rather than the spatial HAC estimator, because subsampling creates artificial spatial "holes" in the data set.

start of urbanization is correlated with a grid cell's agro-ecological characteristics and transport opportunities.

6   Conclusion

A growing body of research is concerned with the role of physical geography in shaping historical events associated with modern economic outcomes. This paper focuses on one of the most important and highly visible of such events, namely the formation and growth of cities. We motivate our work from the idea that the extent of urban population (or employment) might depend on crop production closely around that city, and transport costs between the city and elsewhere.

We test this theory with independent data from Klein Goldewijk et al. (2010) on urban and rural populations around the world over the past 2000 years, from which we derive the number of years that have elapsed since each of the world's 62,290 half-degree grid cells passed various thresholds of urbanization. Most of the world's grid cells have never urbanized by any measure, but over time an increasing fraction of cells passes our thresholds of 1 and then above 5 urban people per square kilometer, or the thresholds of 10, 25 and then 50 percent of the population living in urban areas. Our central finding is that, controlling for country fixed effects and spatially correlated omitted variables, cells urbanize earlier when their land is more suitable for cultivation or is exposed to seasonal frost, using the Ramankutty et al. (2002) index of cultivation suitability (based on growing degree days, soil moisture, soil carbon density and soil pH) and the IPCC (2002) data on frost in winter (after a frost-free summer). We also find that cells urbanize earlier when they are closer to the seacoast, or have larger rivers. Our results for the influence of elevation are much more mixed, although the influence of elevation can not be ignored in certain parts of the world. The major effects are of similar magnitude within each continent, so our results are not driven by European history or other distinctive regional experiences. We go on to show that countries with earlier urbanization dates now have higher incomes, again controlling for spatially correlated omitted characteristics.

An additional finding is that most geographic factors, specifically cultivation suitability, distance to the coast, frost and rivers, were more closely linked to urbanization in the early stages of city formation and initial growth, rather than its later stages. At this junction we do want to stress that our results have been derived exploiting cross-sectional variation among grid cells of the world. An interesting and necessary next step would be to use spatial panel data to identify the dynamics of urbanization processes, for instance, along the lines of Bosker et al. (2013).

Our results are historical in nature, aiming to help explain why cities arose where they did. Geography is not destiny, of course: new technologies and new institutional arrangements can offer man-made substitutes for geographic amenities observed elsewhere. Identifying which factors mattered historically is helpful to focus these efforts on factors such as disease control, transport and access to agricultural products, but other kinds of evidence would be needed to estimate likely impacts and economic rates of return from specific investments. The data analyzed here show only that over the long span of history, urbanization and economic development have occurred sooner in places more suitable for crop production as well as closer to oceans and rivers.

Appendix

This appendix brings together more background documentation on the variables and the regression

results, in particular the robustness analyses alluded to in the main text.

A   Maps

Figures A.1-A.10 provide quantile choropleth maps of the time since urbanization for the different urbanization thresholds, the explanatory variables except for the fixed effects and the interaction variables, and the predicted effects of geography and fixed effects. These figures pertain to the grid cell analyses, and maps already included in the main text (in Figures 2, 3, 5 and 6) are omitted from the Appendix. Figures A.11-A.16 are concerned with the country-level analyses of per capita income of the year 2000, and the explanatory variables of the linear model indicating when at least one of a country's grid cells meets a specific urbanization threshold.

⟨ Figures A.1–A.16 about here ⟩

B   Timing of urbanization

Tables A.1 and A.2 provide estimation results for a less restrictive specification where we include our base variables as well as country fixed effects and unique cross-product terms of the explanatory variables. This specification maximizes the impact of the geography variables, although at the expense of interpretability because marginal effects are now dependent on the explanatory variables. The tables show, however, that the primary effects remain intact and are statistically significant, except for cultivation suitability in the models utilizing the percent urbanization thresholds. Further research would be required to investigate the degree to which cultivation suitability and the transportation variables, distance to the coast and order of streams, are complements or substitutes (see also footnote 15).

In Table A.3 we show estimation results related to the investigation of the importance of the potential endogeneity of country borders. The intention is to show that the conclusions based on a specification with country fixed effects are unaffected if a strictly exogenous definition of fixed effects based on super-grid cells is used. In a qualitative sense, the Table shows exactly that. The predicted effects of geography and super-grid fixed effects are visually represented in Figures A.9 and A.10. The corresponding results for the specification utilizing country fixed effects are given in Figures 5 and 6.

Table A.4 reports on a parametric spatial specification for the base model and the model including country fixed effects. The model specification used is the spatially autoregressive error model in which $y = X\beta + \varepsilon$ and the errors are correlated following a spatial autoregressive pattern, $\varepsilon = \lambda W \varepsilon + \mu$, based on adjacency or proximity specified through the spatial weights matrix W. The spatial error rather than the spatial lag model was selected on the basis of the Lagrange Multiplier test decision rule suggested in Anselin et al. (1996). The estimator used is the general methods of moments estimator (GMM) for the spatial error model allowing for heteroskedasticity (Kelejian and Prucha, 2010), and the weights matrix is defined using the third-order queen principle (see Table 1 for details). Since OLS is an unbiased and consistent estimator for the spatial error model, the results in Table A.4 are not notably different from our original results.

⟨ Tables A.1–A.4 about here ⟩

C Modern-day income

We provide two robustness checks for the modern-day country level per capita income regressions. One adds continent fixed effects to the linear and quadratic specification of (the logarithm) of per capita income of the year 2000 for countries as a function of the date of first urbanization. The other maintains the continent fixed effects, but adds a parametric spatial autoregressive component.

Table A.5 shows that the within-continent correlation between present-day per capita income and the first date of urbanization is comparable to the previous results without continent fixed effects reported in Table 6, although the correlations are slightly smaller.

Table A.6 reports on a parametric spatial version of the modern-day income regressions using a maximum likelihood (ML) estimator for the spatial autoregressive lag model (Anselin, 2006). The spatial lag model rather than the spatial error model was again selected on the basis of the Lagrange Multiplier test decision procedure suggested in Anselin et al. (1996). The ML estimator does not correct for heteroskedasticity, so the estimated standard errors may be biased downwards. In the spatial lag model estimated coefficients do not represent marginal effects, because the lag model $y = \rho W y + X\beta + \mu$ reads in reduced form as $y = (I - \rho W)^{-1}(X\beta + \mu)$. Marginal effects are thus a function of β, ρ and the a priori specified weights matrix W. Asymptotically, marginal effects are equal to $(1-\rho)^{-1}\beta$, which makes the marginal effects of the spatial lag model comparable in magnitude to the results of the original non-parametric spatial HAC results provided in Table 6. There is, however, a crucial distinction in interpretation: the spatial lag model incorporates a distinctive second-nature geography process of spatial interactions (or spillovers) between countries that decays with distance.

⟨ Tables A.5 and A.6 about here ⟩

References

Acemoglu, D., S. Johnson, and J. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Acemoglu, D., S. Johnson, and J. Robinson (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics 117*(4), 1231–1294.

Acemoglu, D. and J. Robinson (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business.

Allen, R. C. (2008). The nitrogen hypothesis and the English agricultural revolution: A biological analysis. *The Journal of Economic History 68*(1), 182–210.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer.

Anselin, L. (2006). Spatial econometrics. In T. Mills and K. Patterson (Eds.), *The Palgrave Handbook of Econometrics, Volume 1 Econometric Theory*, Chapter 26, pp. 901–969. New York: Oxford University Press.

Anselin, L., A. K. Bera, R. Florax, and M. J. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics 26*(1), 77–104.

Ashraf, Q. and O. Galor (2011). Dynamics and stagnation in the Malthusian epoch. *The American Economic Review 101*(5), 2003–2041.

Ashraf, Q. and O. Galor (2013). The 'Out of Africa' hypothesis, human genetic diversity, and comparative economic development. *American Economic Review 103*(1), 1–46.

Bairoch, P. (1998). *Cities and Economic Development: From the Dawn of History to the Present.* The University of Chicago Press.

Birdsall, N. (1988). Economic approaches to population growth. In H. Chenery and T. Srinivasan (Eds.), *Handbook of Development Economics*, Volume 1, Chapter 12, pp. 477–542. Elsevier.

Bosker, E., E. Buringh, and J. van Zanden (2013). From Baghdad to London: Unraveling urban development in Europe, the Middle East, and North Africa, 800–1800. *The Review of Economics and Statistics 95*(4), 1418–1437.

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics Using Stata.* Stata Press.

Chorley, G. P. H. (1981). The agricultural revolution in northern Europe, 1750–1880: Nitrogen, legumes, and crop productivity. *The Economic History Review 34*(1), 71–93.

Clark, G. (1992). The economics of exhaustion, the Postan thesis, and the agricultural revolution. *The Journal of Economic History 52*(1), 61–84.

Clark, G. and A. Clark (2001). Common rights to land in England, 1475–1839. *The Journal of Economic History 61*(4), 1009–1036.

Comin, D., W. Easterly, and E. Gong (2010). Was the wealth of nations determined in 1000 B.C.? *American Economic Journal: Macroeconomics 2*(3), 65–97.

Dell, M., B. Jones, and B. Olken (2013). What do we learn from the weather? The new climate-economy literature. *Journal of Economic Literature* (forthcoming).

Diamond, J. M. (1997). *Guns, Germs and Steel: the Fates of Human Societies.* Jonathan Cape.

Easterly, W. and R. Levine (2003). Tropics, germs, and crops: How endowments influence economic development. *Journal of Monetary Economics 50*(1), 3–39.

Fujita, M. and P. Krugman (1995). When is the economy monocentric?: von Thünen and Chamberlin unified. *Regional Science and Urban Economics 25*(4), 505–528.

Fujita, M., P. Krugman, and A. J. Venables (2001). *The Spatial Economy: Cities, Regions, and International Trade.* The MIT Press.

Funke, M. and J. Zuo (2003). Annual hard frosts, scale effects and economic development: A case not closed. Quantitative Macroeconomics Working Papers 20308, Hamburg University, Department of Economics.

Gallup, J. L., J. D. Sachs, and A. D. Mellinger (1999). Geography and economic development. *International Regional Science Review 22*(2), 179–232.

Galor, O. and D. N. Weil (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *The American Economic Review 90*(4), 806–828.

Gollin, D., S. Parente, and R. Rogerson (2002). The role of agriculture in development. *The American Economic Review 92*(2), 160–164.

Halvorsen, R. and R. Palmquist (1980). The interpretation of dummy variables in semilogarithmic equations. *The American Economic Review 70*(3), 474–475.

Hansen, G. D. and E. C. Prescott (2002). Malthus to Solow. *The American Economic Review 92*(4), 1205–1217.

Heston, A., R. Summers, and B. Aten (2006). Penn World Table Version 6.2. Technical report, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology 25*(15), 1965–1978.

IPCC (2002). *The IPCC Data Distribution Centre: Downloading Scenarios and Climate Data from the DDC.*

Johnston, B. F. and J. W. Mellor (1961). The role of agriculture in economic development. *The American Economic Review 51*(4), 566–593.

Kelejian, H. H. and I. R. Prucha (2007). HAC estimation in a spatial framework. *Journal of Econometrics 140*(1), 131–154.

Kelejian, H. H. and I. R. Prucha (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics 157*(1), 53–67.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. Sachs, and J. Sachs (2004). A global index representing the stability of malaria transmission. *The American Journal of Tropical Medicine and Hygiene 70*(5), 486–498.

Klein Goldewijk, K., A. Beusen, and P. Janssen (2010). Long-term dynamic modeling of global population and built-up area in a spatially explicit way: Hyde 3.1. *The Holocene 20*(4), 565–573.

Kögel, T. and A. Prskawetz (2001). Agricultural productivity growth and escape from the Malthusian trap. *Journal of Economic Growth 6*(4), 337–357.

Krugman, P. (1991). Increasing returns and economic geography. *The Journal of Political Economy 99*(3), 483–499.

Kuznets, S. (1973). Modern economic growth: Findings and reflections. *The American Economic Review 63*(3), 247–258.

Lucas, R. E. (2000). Some macroeconomics for the 21st century. *The Journal of Economic Perspectives 14*(1), 159–168.

Maddison, A., O. for Economic Co-operation, and Development. (2001). *The World Economy: A Millennial Perspective.* Development Centre of the Organisation for Economic Co-operation and Development.

Mankiw, N. G., D. Romer, and D. N. Weil (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics 107*(2), 407–437.

Masters, W. A. and M. S. McMillan (2001). Climate and scale in economic growth. *Journal of Economic Growth 6*(3), 167–186.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory 58*(2), 317–334.

McCloskey, D. N. (1972). The enclosure of open fields: Preface to a study of its impact on the efficiency of English agriculture in the eighteenth century. *The Journal of Economic History 32*(1), 15–35.

McMillen, D. (2003). Spatial autocorrelation or model misspecification? *International Regional Science Review 26*(2), 208–217.

Michalopoulos, S. (2008). The origins of ethnolinguistic diversity: Theory and evidence. *American Economic Review 102*(4), 1508–1539.

Murata, Y. (2008, August). Engel's law, Petty's law, and agglomeration. *Journal of Development Economics 87*(1), 161–177.

Ngai, L. R. (2004). Barriers and the transition to modern growth. *Journal of Monetary Economics 51*(7), 1353–1383.

Nordhaus, W. D. (2006). Geography and macroeconomics: New data and new findings. *Proceedings of the National Academy of Sciences of the United States of America 103*(10), 3510–3517.

North, D. C. (1990). *Institutions, Institutional Change, and Economic Performance.* Cambridge University Press.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *Quarterly Journal of Economics 123*(1), 139–176.

Nunn, N. (2014). Historical development. In S. Durlauf and P. Aghion (Eds.), *Handbook of Economic Growth, Volume 2a*, Chapter 7. Elsevier.

Nunn, N. and N. Qian (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics 126*(2), 593–650.

Olsson, O. and D. A. Hibbs (2005). Biogeography and long-run economic development. *European Economic Review 49*(4), 909–938.

Puga, D. (1999). The rise and fall of regional inequalities. *European Economic Review 43*(2), 303–334.

Putterman, L. (2008). Agriculture, diffusion and development: Ripple effects of the neolithic revolution. *Economica 75*(300), 729–748.

Putterman, L. and D. Weil (2010). Post–1500 population flows and the long-run determinants of economic growth and inequality. *The Quarterly Journal of Economics 125*(4), 1627–1682.

Ramankutty, N., J. A. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography 11*(5), 377–392.

Sachs, J. D. (2003). Institutions don't rule: Direct effects of geography on per capita income. Working Paper 9490, National Bureau of Economic Research.

Sokoloff, K. L. and S. L. Engerman (2000). History lessons: Institutions, factors endowments, and paths of development in the new world. *The Journal of Economic Perspectives 14*(3), 217–232.

Solow, R. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics 70*(1), 65–94.

Spolaore, E. and R. Wacziarg (2009). The diffusion of development. *Quarterly Journal of Economics 124*(2), 469–529.

Spolaore, E. and R. Wacziarg (2013). How deep are the roots of economic development. *Journal of Economic Literature 51*(2), 1–45.

Strulik, H. and J. Weisdorf (2008). Population, food, and knowledge: A simple unified growth theory. *Journal of Economic Growth 13*(3), 195–216.

Vörösmatry, C. J., B. M. Fekete, M. Meybeck, and R. Lammers (2000). Global system of rivers: Its role in organizing continental land mass and defining land-to-ocean linkages. *Global Biogeochemical Cycles 14*(2), 599–622.

Williamson, J. G. (1988). Migration and urbanization. In H. Chenery and T. Srinivasan (Eds.), *Handbook of Development Economics*, Volume 1, Chapter 11, pp. 425–465. Elsevier.

Xenophon (1914). *Cyropaedia: The Education of Cyrus*. J. M. Dent and Sons. Translated by H.G. Dakyns, revised by F. M. Stawell.

Table 1: Summary statistics for all variables.[a]

Urbanization threshold	Mean	Std. Dev.	Min.	Max.	Moran's I[b]
1 inhabitant per km^2	35.30	118.98	0	2000	0.31
5.67 inhabitants per km^2	18.83	67.25	0	1600	0.32
10% urbanization	40.64	102.38	0	1800	0.23
25% urbanization	23.51	57.84	0	1100	0.26
50% urbanization	11.42	33.50	0	700	0.23
cultivation suitability index	0.31	0.32	0.0001	1.00	
distance to coast (km)	566.49	525.71	0	2514.70	
Strahler order of streams	1.04	1.03	0	6	
seasonal frost	0.41	0.49	0	1	
land elevation (m)	630.84	821.49	−29.33	5717.11	

Source: Authors' calculations from sources detailed in the text.

[a] The number of observations equals 52,143.

[b] Moran's I measures the degree of spatial correlation in the time of urbanization between a grid cell and its surrounding grid cells. Neighboring grid cells are defined on the basis of third-order queen contiguity, implying that a grid cell and its immediate neighbors share a common border (including a single point), and neighbors of the neighbors as well as their neighbors are also included. Each cell has a maximum of 48 neighbors, which corresponds to an area of 385×385 km at the Equator and 196×196 km at the Arctic Circle. The null hypothesis of spatial randomness is rejected at conventional significance levels in all cases. Results are similar, with slightly higher correlation coefficients, for a first-order contiguity matrix, which only includes direct neighbors.

Table 2: Timing of urbanization, spatial HAC results.[a]

	(1) Base model				
Urbanization threshold	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
ln(cultivation suitability)	0.312***	0.225***	0.216***	0.149***	0.071***
	(0.007)	(0.006)	(0.009)	(0.008)	(0.007)
ln(distance to coast)	−0.250***	−0.239***	−0.243***	−0.228***	−0.176***
	(0.012)	(0.011)	(0.014)	(0.012)	(0.010)
order of streams	0.249***	0.218***	0.274***	0.229***	0.156***
	(0.013)	(0.011)	(0.014)	(0.013)	(0.010)
frost	0.693***	0.667***	0.524***	0.529***	0.394***
	(0.039)	(0.033)	(0.040)	(0.036)	(0.027)
ln(elevation)	0.026	0.020	−0.016	−0.011	−0.015
	(0.016)	(0.014)	(0.016)	(0.014)	(0.011)
constant	2.763***	2.137***	3.070***	2.486***	1.686***
	(0.088)	(0.080)	(0.093)	(0.084)	(0.071)
	(2) Base model with country fixed effects				
ln(cultivation suitability)	0.285***	0.197***	0.219***	0.165***	0.092***
	(0.008)	(0.006)	(0.011)	(0.010)	(0.008)
ln(distance to coast)	−0.147***	−0.148***	−0.164***	−0.164***	−0.144***
	(0.012)	(0.010)	(0.016)	(0.014)	(0.011)
order of streams	0.186***	0.161***	0.201***	0.164***	0.119***
	(0.012)	(0.010)	(0.013)	(0.011)	(0.009)
frost	0.552***	0.419***	0.472***	0.405***	0.297***
	(0.038)	(0.031)	(0.043)	(0.037)	(0.029)
ln(elevation)	−0.019*	−0.030**	−0.030*	−0.001	0.013
	(0.015)	(0.013)	(0.016)	(0.014)	(0.011)
constant	2.414***	1.767***	2.345***	1.489***	0.757***
	(0.183)	(0.148)	(0.165)	(0.124)	(0.088)
Goodness of fit[b] and misspecification tests[c]					
R^2-adjusted					
Model (1)	0.24	0.22	0.14	0.12	0.08
Model (2)	0.37	0.37	0.26	0.25	0.20
Model (3)[d]	0.39	0.39	0.27	0.26	0.21
Moran's I	0.21***	0.18***	0.23***	0.20***	0.15***
Breusch-Pagan	7657***	9396***	3377***	4435***	6110***

[a] The number of observations equals 52,143. HAC standard errors are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of the time since a grid cell first met a certain urbanization threshold.

[b] Likelihood Ratio tests for jointly adding fixed country effects and/or interaction terms are invariably statistically significant.

[c] The misspecification tests are performed for Model (2). The result for the Breusch-Pagan test is the χ^2-value with random coefficients as the alternative hypothesis. The Koenker-Bassett version of the test is reported if the null hypothesis of normally distributed errors is rejected according to the Jarque-Bera test.

[d] Model (3) represents the base model specification augmented with 181 country fixed effects and interaction terms among the above reported geography variables. Results are not presented in detail here, but are discussed in the main text.

Table 3: Timing of urbanization, Tobit results.[a]

	Tobit coefficients including country fixed effects				
Urbanization threshold	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
ln(cultivation suitability)	1.227***	1.277***	0.661***	0.621***	0.555***
	(0.115)	(0.130)	(0.136)	(0.136)	(0.166)
ln(distance to coast)	−0.414***	−0.588***	−0.416***	−0.492***	−0.616***
	(0.136)	(0.161)	(0.128)	(0.123)	(0.116)
order of streams	0.458***	0.587***	0.449***	0.459***	0.511***
	(0.094)	(0.095)	(0.065)	(0.060)	(0.052)
frost	1.343***	1.602***	0.958**	0.915*	0.939*
	(0.337)	(0.395)	(0.444)	(0.474)	(0.495)
ln(elevation)	−0.001	−0.018	−0.058	−0.014	−0.074
	(0.178)	(0.191)	(0.152)	(0.148)	(0.151)
constant	2.849***	1.623***	2.046***	−0.137	−4.965***
	(0.513)	(0.546)	(0.562)	(0.378)	(1.236)
Pseudo-R^{2b}	0.16	0.19	0.09	0.09	0.10
Left-censored observations	35,387	40,507	31,280	35,017	40,912
Uncensored observations	16,749	11,636	20,863	17,126	11,231
Right-censored observations	7	0	0	0	0
	Tobit marginal effects[c]				
ln(cultivation suitability)	0.435***	0.309***	0.291***	0.222***	0.127***
	(0.039)	(0.029)	(0.061)	(0.050)	(0.039)
ln(distance to coast)	−0.147***	−0.142***	−0.182***	−0.176***	−0.142***
	(0.048)	(0.038)	(0.056)	(0.044)	(0.026)
order of streams	0.162***	0.142***	0.198***	0.164***	0.117***
	(0.033)	(0.022)	(0.029)	(0.022)	(0.012)
frost	0.476***	0.387***	0.422**	0.328*	0.216*
	(0.119)	(0.093)	(0.193)	(0.168)	(0.113)
ln(elevation)	−0.0003	−0.004	−0.025	−0.005	−0.017
	(0.063)	(0.046)	(0.067)	(0.053)	(0.035)

[a] The number of observations equals 52,143, and all regressions include fixed effects for 181 countries. Standard errors adjusted for 181 country clusters are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05 and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of the time since a grid cell first met a certain urbanization threshold.

[b] The reported R^2-value is the squared correlation between the predicted and the actual time since meeting a specific urbanization threshold.

[c] Average marginal effects for the expected value of the latent variable, conditional on the observed dependent variable being greater than zero.

Table 4: Means of continent-specific variables.

Urbanization threshold	Europe	Africa	Asia	North America	South America	Australia
1 inhabitant per km^2	171	20	31	26	20	8
5.67 inhabitants per km^2	99	7	17	14	10	4
10% urbanization	150	26	31	45	39	16
25% urbanization	87	10	16	33	27	12
50% urbanization	41	3	6	19	15	8
cultivation suitability index	0.47	0.27	0.28	0.32	0.40	0.29
distance to coast (km)	171	664	685	431	514	284
Strahler order of streams	1.23	0.83	1.14	0.97	1.28	0.28
seasonal frost	0.71	0.35	0.46	0.41	0.12	0.59
land elevation (m)	323	637	743	596	575	275
Number of observations	3,401	9,353	21,978	8,865	6,268	2,278

Source: Authors' calculations from sources detailed in the text.

Table 5: Timing of 10% urbanization, coefficients and marginal effects by continent, spatial HAC results.[a]

	Europe	Africa	Asia	North America	South America	Australia	Chow test[b]
Coefficients							
ln(cultivation suitability)	0.386***	0.211***	0.185***	0.212***	0.388***	0.185***	50.03***
	(0.045)	(0.019)	(0.011)	(0.030)	(0.033)	(0.042)	
ln(distance to coast)	−0.152***	−0.206***	−0.074***	−0.192***	−0.355***	−0.441***	91.42***
	(0.040)	(0.030)	(0.020)	(0.037)	(0.035)	(0.044)	
order of streams	0.251***	0.101***	0.199***	0.229***	0.220***	0.541***	20.81***
	(0.045)	(0.032)	(0.020)	(0.036)	(0.034)	(0.113)	
frost	0.055	0.151**	0.147**	1.093***	0.829***	0.595***	77.51***
	(0.225)	(0.075)	(0.058)	(0.117)	(0.165)	(0.096)	
ln(elevation)	−0.022	−0.119**	−0.041**	−0.323***	0.270***	0.442***	150.17***
	(0.063)	(0.046)	(0.018)	(0.047)	(0.040)	(0.065)	
continent fixed effects	5.003***	3.579***	2.221***	4.514***	2.560***	0.574*	177.44***
	(0.378)	(0.255)	(0.109)	(0.234)	(0.279)	(0.296)	
Marginal effects[c]							
cultivation suitability	1.23	0.20	0.20	0.30	0.38	0.10	
distance to coast	−13.33	−0.81	−0.33	−2.00	−2.69	−2.48	
order of streams	37.65	2.63	6.17	10.31	8.58	8.66	
frost	5.60	4.24	3.96	63.46	47.77	14.63	
elevation	−1.02	−0.49	−0.17	−2.44	1.83	2.57	

[a] The number of observations equals 52,143. HAC standard errors are in parentheses. Coefficients are tagged with ***, **, and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of the time since a grid cell first met the 10-percent urbanization threshold.

[a] The χ^2-value for the overall Chow test is 1,591.62, with a p-value smaller than 0.01.

[c] The marginal effects are determined as the number of years earlier a cell reaches the 10% urbanization threshold, based on a 0.01-change in cultivation suitability, a 100 km increase in coastal distance, a 100 m increase in elevation, an increase in the Strahler index by 1 unit, and a unity switch in the frost variable, at continental sample mean values.

Table 6: Income and date of first urbanization.[a]

Urbanization threshold	$1/km^2$	$5.67/km^2$	10%	25%	50%
Entire sample, linear, HAC results					
transition	0.001**	0.002**	0.001***	0.003***	0.004***
	(0.0003)	(0.001)	(0.0004)	(0.001)	(0.001)
constant	8.184***	8.131***	8.079***	7.928***	7.993***
	(0.209)	(0.210)	(0.210)	(0.199)	(0.202)
Entire sample, quadratic, HAC results					
transition	0.003***	0.006***	0.005***	0.006***	0.009***
	(0.0008)	(0.001)	(0.001)	(0.001)	(0.002)
transition^{2b}	−0.001***	−0.004***	−0.003***	−0.004***	−0.011***
	(0.0005)	(0.0006)	(0.0007)	(0.001)	(0.002)
constant	7.739***	7.525***	7.480***	7.624***	7.767***
	(0.203)	(0.210)	(0.222)	(0.202)	(0.185)
Top (in years)	1,166	815	948	745	399
Sample mean transition	368	199	278	188	110
Restricted sample, transition $\leq$ 400 years, OLS results					
transition	0.001	0.005***	0.003***	0.004***	0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
constant	8.045***	7.577***	7.659***	7.701***	7.877***
	(0.259)	(0.171)	(0.221)	(0.163)	(0.133)
Number of observations	143	157	154	162	168
R^2-adjusted	0.05	0.09	0.08	0.14	0.13
Moran's I^c	0.43***	0.40***	0.40***	0.35***	0.36***
Breusch-Pagan^d	0.04	13.62***	2.05	0.46	7.95***

[a] The number of observations in the entire sample equals 169. Spatial HAC and OLS standard errors are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of per capita income in 2000. Measures of fit and misspecification tests pertain to the linear model using the entire sample.

[b] For ease of exposition the coefficients and standard errors are multiplied by 1,000.

[c] The weights matrix used for testing is a (row-standardized) inverse distance matrix containing distances between the geographical midpoints of countries, with a cutoff point of 2,876 km. The latter ensures sparsity of the weights matrix as well as each country having at least one neighbor. The average number of neighbors per country is 25, which is also the number implemented in the kernel estimation for the spatial HAC results.

[d] The result for the Breusch-Pagan test is the χ^2-value with random coefficients as the alternative hypothesis. The Koenker-Bassett version of the test is reported if the null hypothesis of normally distributed errors is rejected according to the Jarque-Bera test.

Table A.1: Timing of urbanization, spatial HAC results with interactions.[a]

Urbanization threshold	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
ln(suitability)	0.375***	0.323***	0.047	−0.020	−0.080***
	(0.027)	(0.024)	(0.037)	(0.034)	(0.030)
ln(distance to coast)	−0.165***	−0.195***	−0.104***	−0.123***	−0.108***
	(0.031)	(0.027)	(0.036)	(0.032)	(0.027)
order of streams	0.595***	0.513***	0.607***	0.521***	0.378***
	(0.057)	(0.051)	(0.063)	(0.058)	(0.050)
frost	1.876***	1.655***	1.892***	1.959***	1.58***
	(0.201)	(0.180)	(0.195)	(0.180)	(0.151)
ln(elevation)	0.105***	0.075***	0.067*	−0.073**	0.009
	(0.040)	(0.035)	(0.041)	(0.036)	(0.029)
ln(suitability) × ln(distance)	−0.022***	−0.037***	−0.014**	−0.010*	0.017***
	(0.004)	(0.004)	(0.006)	(0.005)	(0.005)
ln(suitability) × order	0.065***	0.058***	0.043***	−0.028***	0.016***
	(0.006)	(0.005)	(0.007)	(0.006)	(0.005)
ln(suitability) × frost	0.171***	0.158***	−0.238***	−0.226***	0.157***
	(0.015)	(0.013)	(0.016)	(0.014)	(0.011)
ln(suitability) × ln(elevation)	−0.011**	−0.003	−0.003	−0.003	0.0003***
	(0.006)	(0.005)	(0.006)	(0.005)	(0.005)
ln(distance) × order	−0.059***	−0.058***	−0.043***	0.035***	0.021***
	(0.009)	(0.008)	(0.010)	(0.009)	(0.007)
ln(distance) × frost	−0.092***	−0.108***	−0.091***	−0.097***	−0.107***
	(0.025)	(0.021)	(0.026)	(0.023)	(0.018)
ln(distance) × ln(elevation)	−0.014***	−0.012***	−0.006	−0.002	0.007
	(0.005)	(0.004)	(0.006)	(0.005)	(0.004)
order × frost	0.261***	0.247***	0.167***	0.140***	0.101***
	(0.030)	(0.025)	(0.030)	(0.027)	(0.021)
order × ln(elevation)	0.004	0.012	−0.017	−0.021**	−0.025***
	(0.010)	(0.008)	(0.011)	(0.010)	(0.007)
frost × ln(elevation)	−0.103***	−0.055*	−0.094***	−0.098***	−0.063***
	(0.035)	(0.030)	(0.033)	(0.029)	(0.023)
constant	2.122***	1.792***	1.839***	1.331***	0.865***
	(0.176)	(0.159)	(0.093)	(0.163)	(0.139)
R^2-adjusted	0.27	0.26	0.16	0.14	0.10

[a] The number of observations equals 52,143. HAC standard errors are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of the time since a grid cell first met a certain urbanization threshold.

Table A.2: Timing of urbanization, spatial HAC results with interactions and country fixed effects.[a]

Urbanization threshold	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
ln(suitability)	0.338***	0.240***	0.062	0.043	−0.003
	(0.025)	(0.021)	(0.041)	(0.038)	(0.033)
ln(distance to coast)	−0.240***	−0.172***	−0.205***	−0.260***	−0.246***
	(0.032)	(0.027)	(0.036)	(0.032)	(0.027)
order of streams	0.359***	0.302***	0.340***	0.317***	0.264***
	(0.049)	(0.043)	(0.054)	(0.049)	(0.042)
frost	1.326***	1.042***	1.299***	1.035***	0.747***
	(0.168)	(0.143)	(0.172)	(0.152)	(0.127)
ln(elevation)	0.039	0.080***	−0.059	−0.079**	−0.111***
	(0.035)	(0.030)	(0.038)	(0.033)	(0.026)
ln(suitability) × distance	−0.021***	−0.030***	0.019***	0.012*	0.014**
	(0.004)	(0.003)	(0.007)	(0.006)	(0.006)
ln(suitability) × order	0.031***	0.031***	0.020***	0.019***	0.015***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.004)
ln(suitability) × frost	0.184***	0.153***	0.223***	0.168***	0.095***
	(0.016)	(0.012)	(0.017)	(0.015)	(0.012)
ln(suitability) × ln(elevation)	−0.006	0.005	−0.006	−0.003	−0.004
	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)
ln(distance) × order	−0.035***	−0.034***	−0.021**	−0.019**	−0.010
	(0.007)	(0.006)	(0.009)	(0.008)	(0.007)
ln(distance) × frost	0.130***	0.057***	0.103***	0.070***	0.004
	(0.024)	(0.021)	(0.026)	(0.023)	(0.018)
ln(distance) × ln(elevation)	0.004	−0.011***	0.019***	0.028***	0.032***
	(0.005)	(0.004)	(0.006)	(0.005)	(0.004)
order × frost	0.183***	0.177***	0.101***	0.085***	0.070***
	(0.027)	(0.022)	(0.027)	(0.024)	(0.019)
order × ln(elevation)	0.005***	0.010	−0.003	−0.005	−0.013*
	(0.009)	(0.007)	(0.009)	(0.008)	(0.007)
frost × ln(elevation)	−0.221***	−0.133***	−0.176***	−0.126***	−0.054***
	(0.032)	(0.026)	(0.031)	(0.026)	(0.020)
constant	2.405***	1.589***	1.940***	1.363***	0.804***
	(0.213)	(0.182)	(0.205)	(0.177)	(0.147)
R^2-adjusted	0.39	0.38	0.16	0.14	0.10

[a] The number of observations equals 52,143. HAC standard errors are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of the time since a grid cell first met a certain urbanization threshold.

Table A.3: Timing of urbanization, spatial HAC results with country fixed effects or artificial super-grid effects.[a]

Urbanization threshold	$1/km^2$	$5.67/km^2$	10%	25%	50%
(1) Base model with country fixed effects					
ln(cultivation suitability)	0.285***	0.197***	0.219***	0.165***	0.092***
	(0.008)	(0.006)	(0.011)	(0.010)	(0.008)
ln(distance to coast)	−0.147***	−0.148***	−0.164***	−0.164***	−0.144***
	(0.012)	(0.010)	(0.016)	(0.014)	(0.011)
order of streams	0.186***	0.161***	0.201***	0.164***	0.119***
	(0.012)	(0.010)	(0.013)	(0.011)	(0.009)
frost	0.552***	0.419***	0.472***	0.405***	0.297***
	(0.038)	(0.031)	(0.043)	(0.037)	(0.029)
ln(elevation)	−0.019*	−0.030**	−0.030*	−0.001	0.013
	(0.015)	(0.013)	(0.016)	(0.014)	(0.011)
constant	2.414***	1.767***	2.345***	1.489***	0.757***
	(0.183)	(0.148)	(0.165)	(0.124)	(0.088)
(2) Base model with artificial super-grid effects					
ln(cultivation suitability)	0.259***	0.184***	0.241***	0.182***	0.106***
	(0.009)	(0.008)	(0.010)	(0.009)	(0.007)
ln(distance to coast)	−0.178***	−0.169***	−0.168***	−0.173***	−0.146***
	(0.012)	(0.011)	(0.013)	(0.012)	(0.010)
order of streams	0.191***	0.159***	0.203***	0.176***	0.130***
	(0.011)	(0.009)	(0.012)	(0.011)	(0.009)
frost	0.175***	0.156***	0.152***	0.142***	0.120***
	(0.036)	(0.029)	(0.041)	(0.035)	(0.027)
ln(elevation)	0.004	−0.004	−0.012	−0.008	−0.004
	(0.017)	(0.015)	(0.016)	(0.013)	(0.010)
constant	2.173***	1.580***	2.104***	1.498***	0.875***
	(0.111)	(0.095)	(0.114)	(0.097)	(0.078)
R^2-adjusted	0.40	0.39	0.30	0.29	0.23

[a] The number of observations equals 52,143. HAC standard errors are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of the time since a grid cell first met a certain urbanization threshold. Details for the 181 country fixed effects and 169 super-grid effects are not reported; see Figures 6 and A.10 for the spatial distribution of the estimated fixed effects.

Table A.4: Timing of urbanization, spatial autoregressive error model with heteroskedasticity.[a]

	(1) Base model				
Urbanization threshold	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
ln(cultivation suitability)	0.234***	0.172***	0.225***	0.182***	0.112***
	(0.009)	(0.007)	(0.010)	(0.009)	(0.007)
ln(distance to coast)	−0.074***	−0.077***	−0.105***	−0.141***	−0.146***
	(0.012)	(0.011)	(0.013)	(0.012)	(0.010)
order of streams	0.140***	0.114***	0.183***	0.165***	0.132***
	(0.008)	(0.007)	(0.009)	(0.009)	(0.008)
frost	0.153***	0.160***	0.140***	0.167***	0.164***
	(0.034)	(0.029)	(0.036)	(0.032)	(0.027)
ln(elevation)	−0.158***	−0.151***	−0.152***	−0.097***	−0.059***
	(0.014)	(0.013)	(0.015)	(0.014)	(0.011)
λ	0.897***	0.890***	0.877***	0.854***	0.814***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
constant	2.972***	2.406***	3.337***	2.777***	1.983***
	(0.100)	(0.088)	(0.100)	(0.088)	(0.071)
Pseudo-R^2	0.21	0.18	0.12	0.10	0.07
	(2) Base model with country fixed effects				
ln(cultivation suitability)	0.248***	0.181***	0.238***	0.194***	0.121***
	(0.009)	(0.007)	(0.010)	(0.008)	(0.007)
ln(distance to coast)	−0.059***	−0.073***	−0.085***	−0.122***	−0.133***
	(0.012)	(0.010)	(0.013)	(0.012)	(0.010)
order of streams	0.137***	0.115***	0.179***	0.160***	0.128***
	(0.008)	(0.007)	(0.009)	(0.009)	(0.008)
frost	0.166***	0.153***	0.144***	0.153***	0.140***
	(0.034)	(0.028)	(0.035)	(0.032)	(0.027)
ln(elevation)	−0.153***	−0.140***	−0.151***	−0.092***	0.049***
	(0.014)	(0.012)	(0.015)	(0.013)	(0.011)
λ	0.853***	0.825***	0.837***	0.799***	0.736***
	(0.007)	(0.009)	(0.007)	(0.007)	(0.009)
constant	2.909***	2.325***	2.841***	2.052***	1.286***
	(0.292)	(0.216)	(0.288)	(0.214)	(0.115)
Pseudo-R^2	0.35	0.36	0.25	0.24	0.19

[a] The number of observations equals 52,143. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively, using a general method of moments estimator that allows for heteroskedasticity (Kelejian and Prucha, 2010). The dependent variable is the logarithm of the time since a grid cell first met a certain urbanization threshold.

Table A.5: Income and date of first urbanization, spatial HAC results with continent fixed effects.[a]

	(1) Base model, linear				
Urbanization threshold	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
transition	0.0004***	0.0006	0.0006**	0.0014***	0.0021***
	(0.0002)	(0.0004)	(0.0003)	(0.0004)	(0.0009)
Asia	0.961***	0.954***	0.966***	0.944***	0.973***
	(0.225)	(0.224)	(0.224)	(0.227)	(0.230)
Australia	1.721***	1.674***	1.696***	1.650***	1.554***
	(0.365)	(0.360)	(0.361)	(0.337)	(0.319)
Europe	1.962***	1.916***	1.914***	1.837***	1.877***
	(0.252)	(0.270)	(0.258)	(0.246)	(0.241)
North America	1.441***	1.403***	1.414***	1.364***	1.328***
	(0.224)	(0.223)	(0.224)	(0.218)	(0.213)
South America	1.298***	1.267***	1.252***	1.145***	1.067***
	(0.177)	(0.179)	(0.181)	(0.189)	(0.210)
constant	7.366***	7.398***	7.356***	7.300***	7.325***
	(0.152)	(0.148)	(0.154)	(0.150)	(0.151)
R^2-adjusted	0.39	0.39	0.40	0.41	0.41
	(2) Base model, quadratic				
transition	0.001*	0.003***	0.003**	0.003**	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
transition^{2b}	−0.001*	−0.002***	−0.001**	−0.002*	−0.006***
	(0.0004)	(0.001)	(0.001)	(0.001)	(0.002)
Asia	0.943***	0.846***	0.949***	0.925***	0.975***
	(0.223)	(0.228)	(0.221)	(0.230)	(0.230)
Australia	1.781***	1.647***	1.729***	1.628***	1.506***
	(0.347)	(0.307)	(0.325)	(0.314)	(0.295)
Europe	1.871***	1.634***	1.776***	1.780***	1.807***
	(0.245)	(0.260)	(0.252)	(0.247)	(0.241)
North America	1.453***	1.329***	1.385***	1.319***	1.282***
	(0.227)	(0.220)	(0.223)	(0.219)	(0.213)
South America	1.270***	1.149***	1.140***	1.053***	0.987***
	(0.181)	(0.185)	(0.194)	(0.210)	(0.214)
constant	7.148***	7.137***	7.051***	7.172***	7.208***
	(0.235)	(0.182)	(0.244)	(0.176)	(0.162)
R^2-adjusted	0.40	0.42	0.41	0.41	0.42

[a] The number of observations in the entire sample equals 169. Spatial HAC standard errors are in parentheses. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of per capita income in 2000.
[b] For ease of exposition the coefficients and standard errors are multiplied by 1,000.

Table A.6: Income and date of first urbanization, spatial autoregressive lag model with continent fixed effects.[a]

Urbanization threshold	(1) Spatial lag model, linear				
	$1/\text{km}^2$	$5.67/\text{km}^2$	10%	25%	50%
lag ln(GDP per capita)	0.406***	0.426***	0.419***	0.396***	0.400**
	(0.130)	(0.128)	(0.129)	(0.132)	(0.132)
transition	(0.0003)	0.001*	0.001*	0.001***	0.002**
	(0.0002)	(0.0003)	(0.0003)	(0.0005)	(0.001)
Asia	0.618***	0.582***	0.602***	0.604***	0.625***
	(0.221)	(0.223)	(0.221)	(0.220)	(0.219)
Australia	1.345***	1.288***	1.315***	1.294***	1.204***
	(0.494)	(0.491)	(0.490)	(0.485)	(0.486)
Europe	1.267***	1.173***	1.188***	1.161***	1.190***
	(0.290)	(0.293)	(0.292)	(0.295)	(0.291)
North America	0.911***	0.851***	0.871***	0.856***	0.818***
	(0.301)	(0.298)	(0.299)	(0.299)	(0.300)
South America	0.836***	0.783**	0.779**	0.710**	0.635***
	(0.316)	(0.314)	(0.315)	(0.317)	(0.322)
constant	4.318***	4.176***	4.197***	4.320***	4.310***
	(1.000)	(0.987)	(0.995)	(1.017)	(1.014)
R^2-adjusted	0.44	0.45	0.45	0.45	0.45

	(2) Spatial lag model, quadratic				
lag ln(GDP per capita)	0.364***	0.338**	0.352***	0.369***	0.358***
	(0.135)	(0.138)	(0.137)	(0.136)	(0.137)
transition	0.001*	0.003***	0.002**	0.002**	0.004**
	(0.007)	(0.001)	(0.001)	(0.001)	(0.001)
transition^{2b}	−0.0004	−0.002***	−0.001*	−0.001	−0.005
	(0.0004)	(0.0007)	(0.0006)	(0.001)	(0.003)
Asia	0.641***	0.569**	0.647***	0.613***	0.664***
	(0.221)	(0.222)	(0.221)	(0.220)	(0.219)
Australia	1.426***	1.345***	1.400***	1.303***	1.206**
	(0.492)	(0.477)	(0.483)	(0.481)	(0.482)
Europe	1.275***	1.090***	1.198***	1.166***	1.210***
	(0.296)	(0.307)	(0.301)	(0.299)	(0.296)
North America	0.974***	0.903***	0.935***	0.858***	0.838***
	(0.304)	(0.300)	(0.301)	(0.301)	(0.301)
South America	0.864***	0.784**	0.769**	0.673**	0.621*
	(0.317)	(0.315)	(0.317)	(0.322)	(0.324)
constant	4.481***	4.621***	4.465***	4.431***	4.544***
	(1.032)	(1.053)	(1.046)	(1.041)	(1.048)
R^2-adjusted	0.44	0.46	0.45	0.46	0.46

[a] The number of observations in the entire sample equals 169. Coefficients are tagged with ***, ** and * for the 0.01, 0.05, and 0.1 statistical significance level, respectively. The dependent variable is the logarithm of per capita income in 2000. The model is estimated with the maximum likelihood estimator for the spatial lag model, without adjustments for heteroskedasticity (Anselin, 2006). The weights matrix is defined as the (row-standardized) inverse distance matrix containing distances between the geographical midpoints of countries, with a cutoff point of 2,876 km.
[b] For ease of exposition the coefficients and standard errors are multiplied by 1,000.
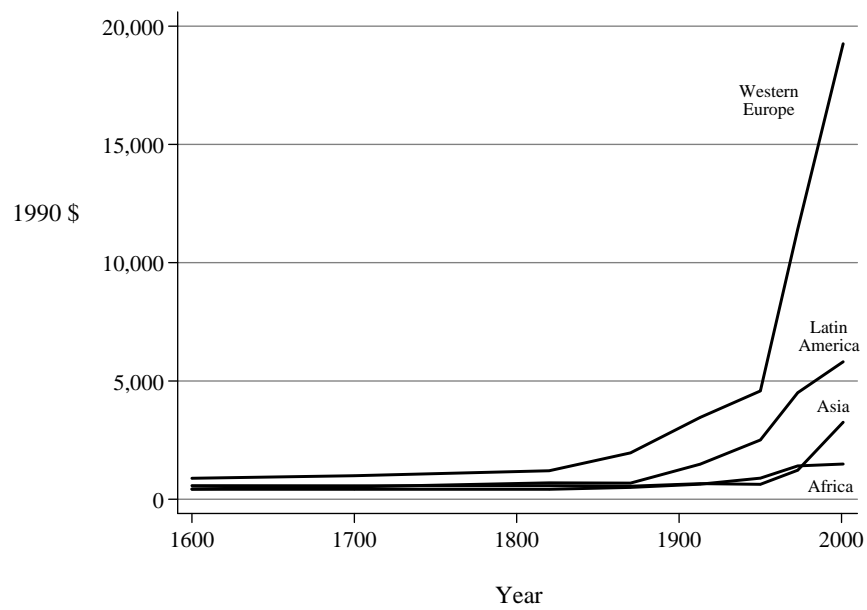
Figure 1: Regional per capita GDP, 1600–2000, in 1990 international dollars, with Asia category excluding Japan.
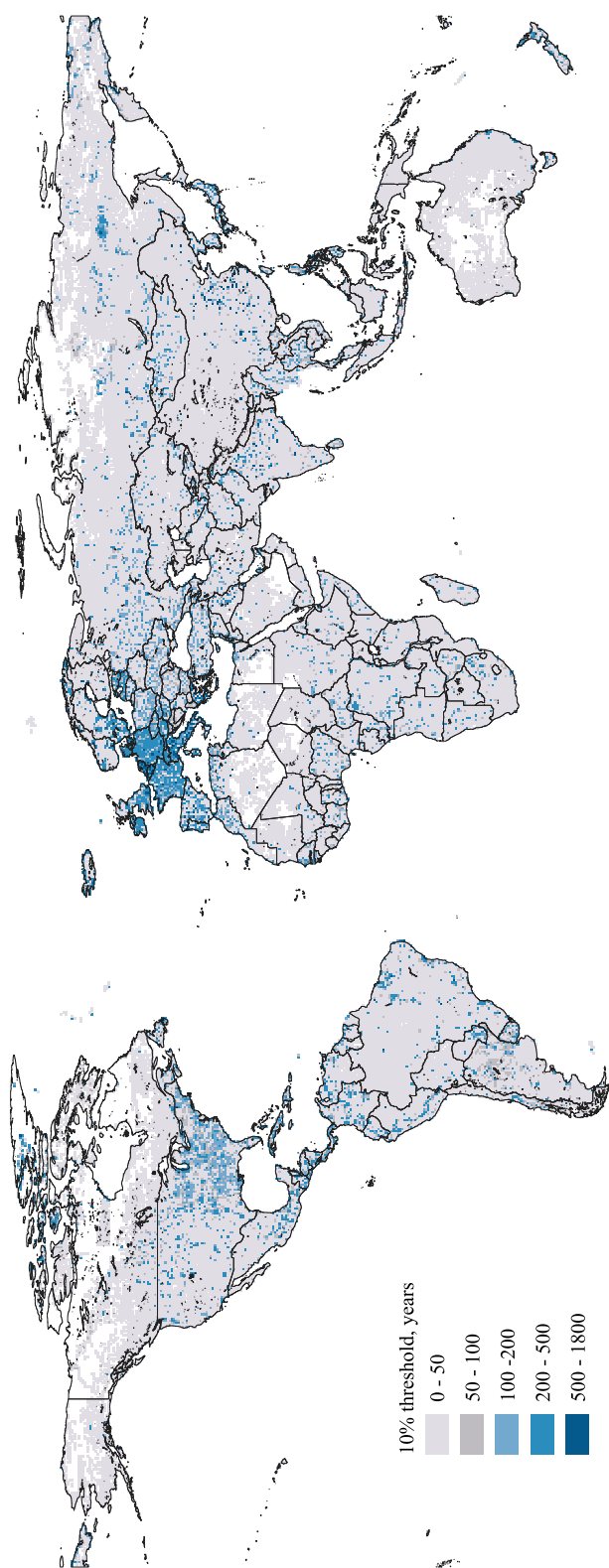Source: Maddison et al. (2001)

Figure 2: Number of years since transition to a 10 percent urban population or greater.
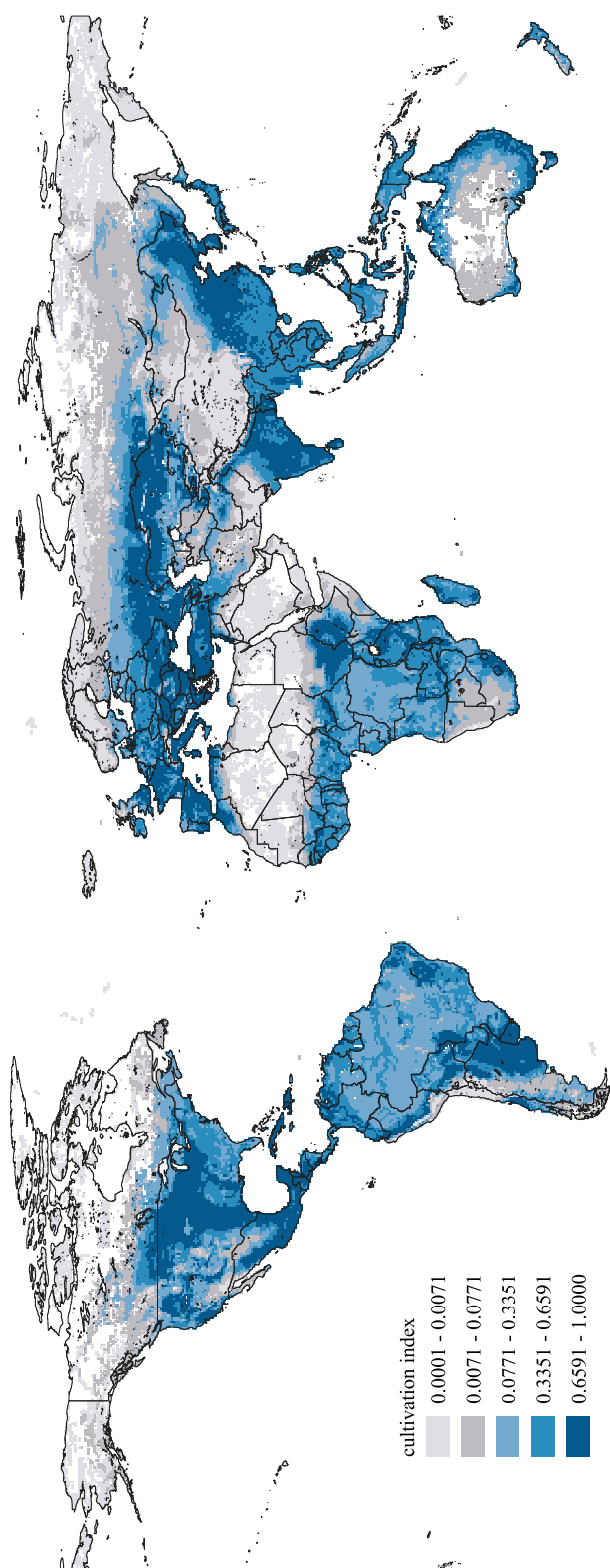Source: Authors' calculations based on Klein Goldewijk et al. (2010); see Table 1 for details.

Figure 3: Worldwide distribution of land suitability for cultivation (quantile map). Source: Ramankutty et al. (2002); see Table 1 for details.
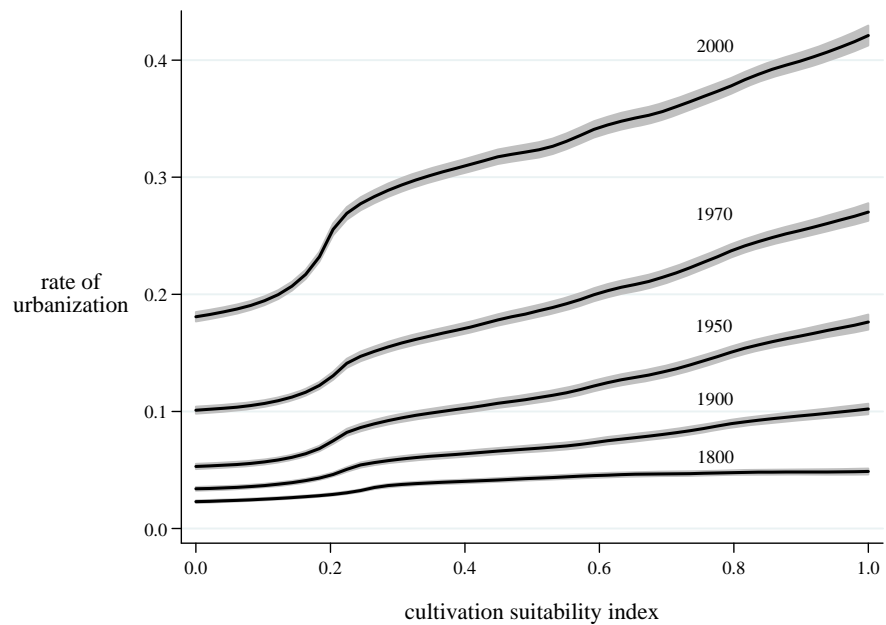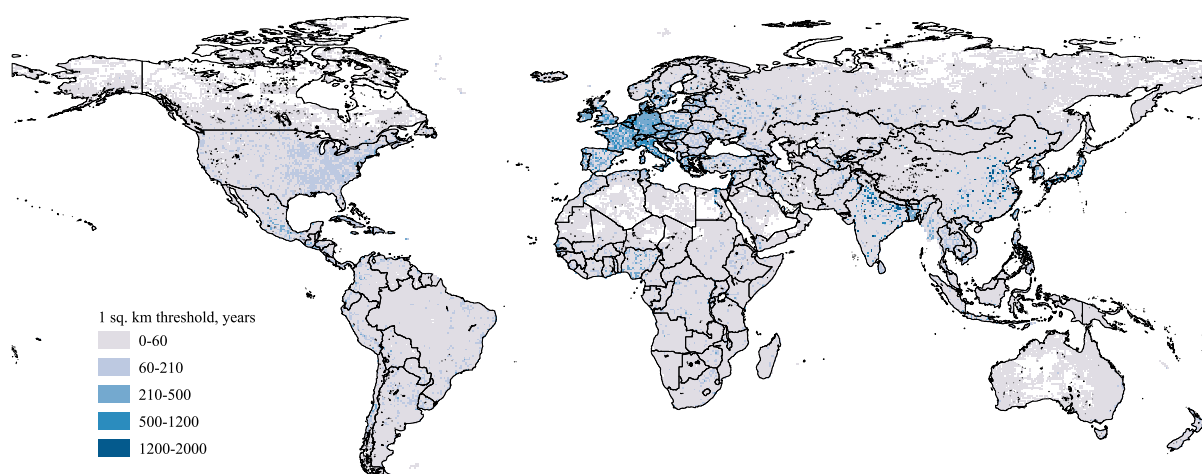
cultivation index

0.0001 - 0.0071
0.0071 - 0.0771
0.0771 - 0.3351
0.3351 - 0.6591
0.6591 - 1.0000

Figure 4: Local polynomial regressions of the rate of urbanization on cultivation suitability with shaded 95 percent confidence intervals around each line.
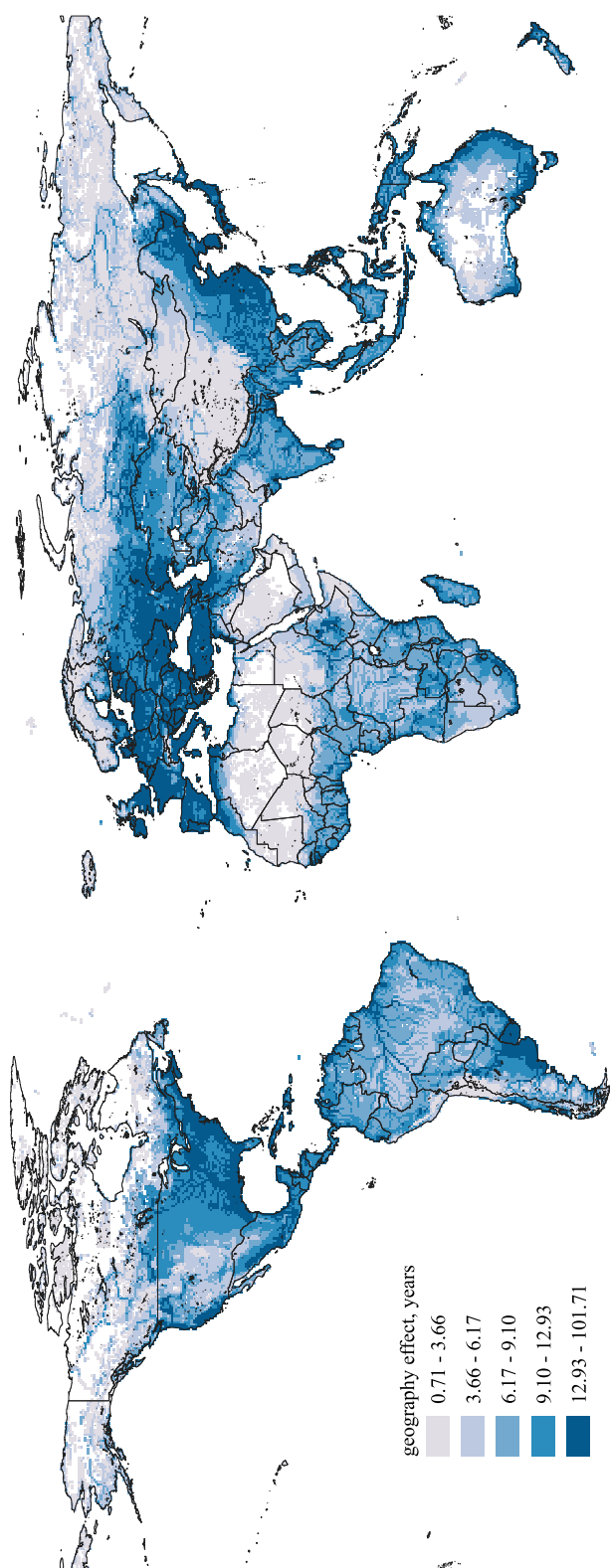
Figure 5: Predicted effect of geography for grid cells belonging to an "average" country, in years, for the 10-percent urbanization threshold (quantile map).

geography effect, years

- 0.71 - 3.66
- 3.66 - 6.17
- 6.17 - 9.10
- 9.10 - 12.93
- 12.93 - 101.71
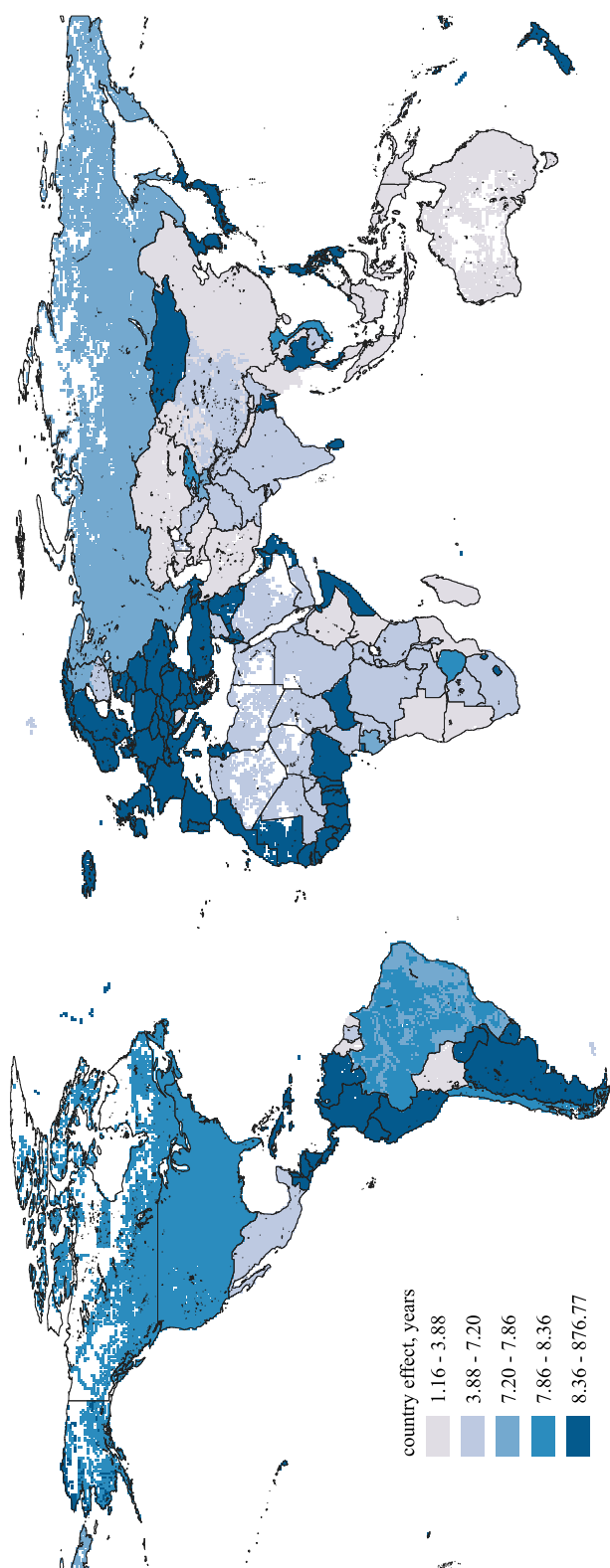
Figure 6: Predicted effect of country characteristics for countries with "average" geographical endowments, in years, for the 10-percent urbanization threshold (quantile map).

country effect, years

1.16 - 3.88
3.88 - 7.20
7.20 - 7.86
7.86 - 8.36
8.36 - 876.77
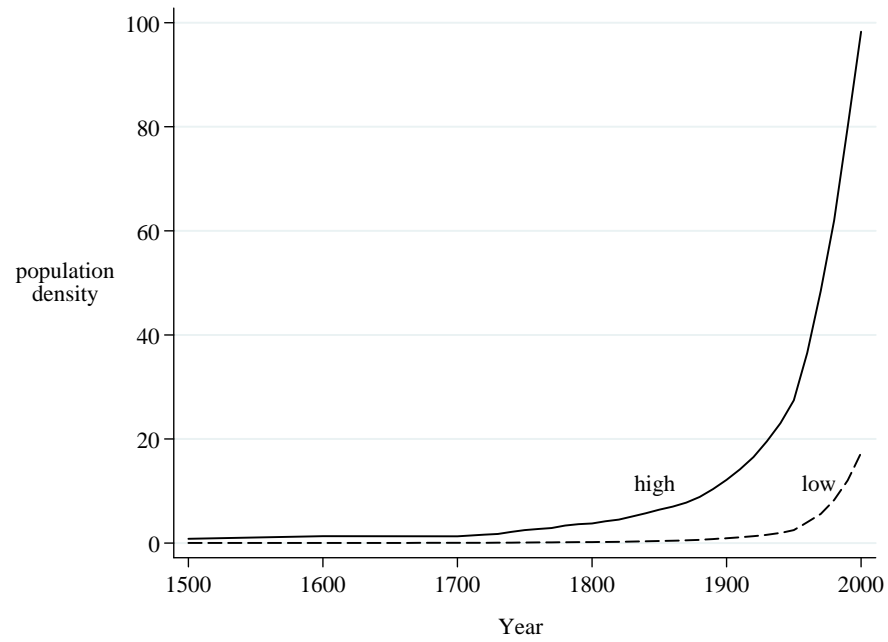
Figure 7: Average urban population density across grid cells with above and below median cultivation suitability, 1500–2000.

Figure 8: Log per capita income in the year 2000 and the number of years since at least one of a country's grid cells reached 10 percent urbanization.

Figure A.1: Number of years since transition to a 1 person per km^2 urban population or greater. Source: Authors' calculations based on Klein Goldewijk et al. (2010); see Table 1 for details.
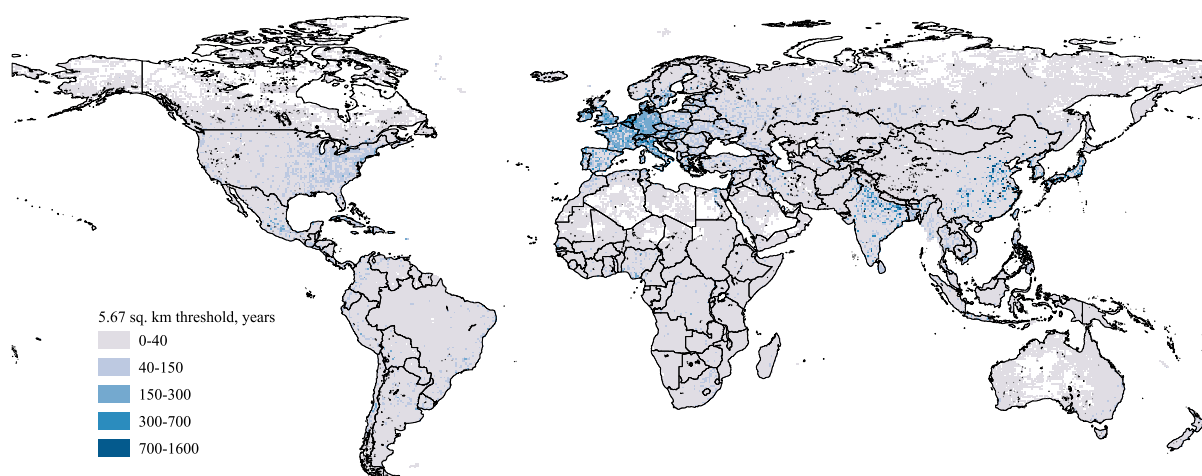
Figure A.2: Number of years since transition to a 5.67 person per km^2 urban population or greater. Source: Authors' calculations based on Klein Goldewijk et al. (2010); see Table 1 for details.
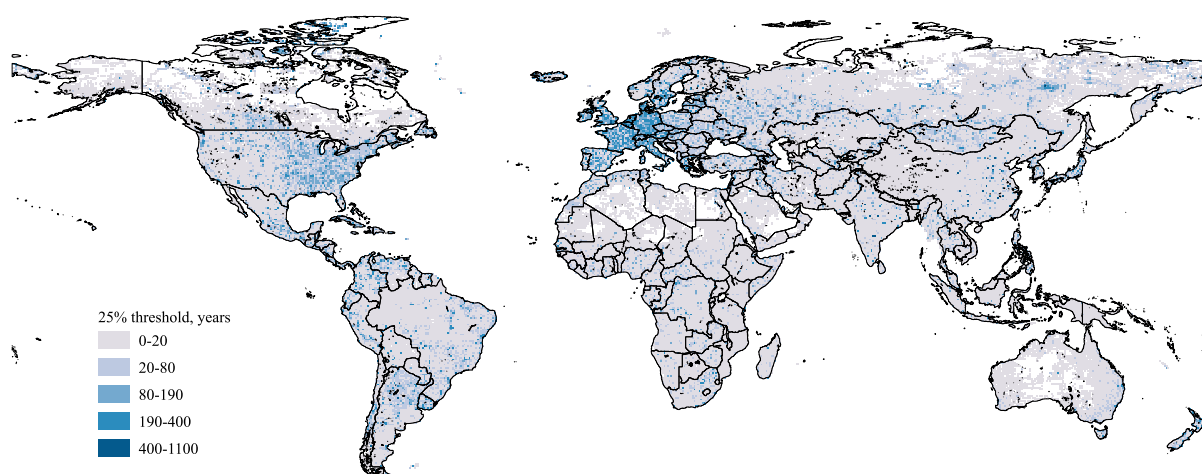
Figure A.3: Number of years since transition to a 25 percent urban population or greater.
Source: Authors' calculations based on Klein Goldewijk et al. (2010); see Table 1 for details.
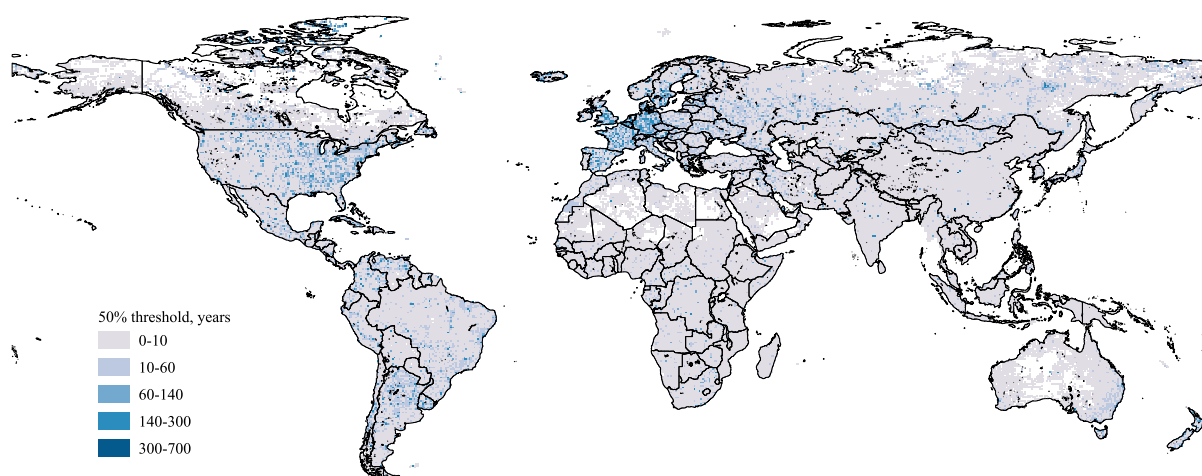
Figure A.4: Number of years since transition to a 50 percent urban population or greater.
Source: Authors' calculations based on Klein Goldewijk et al. (2010); see Table 1 for details.
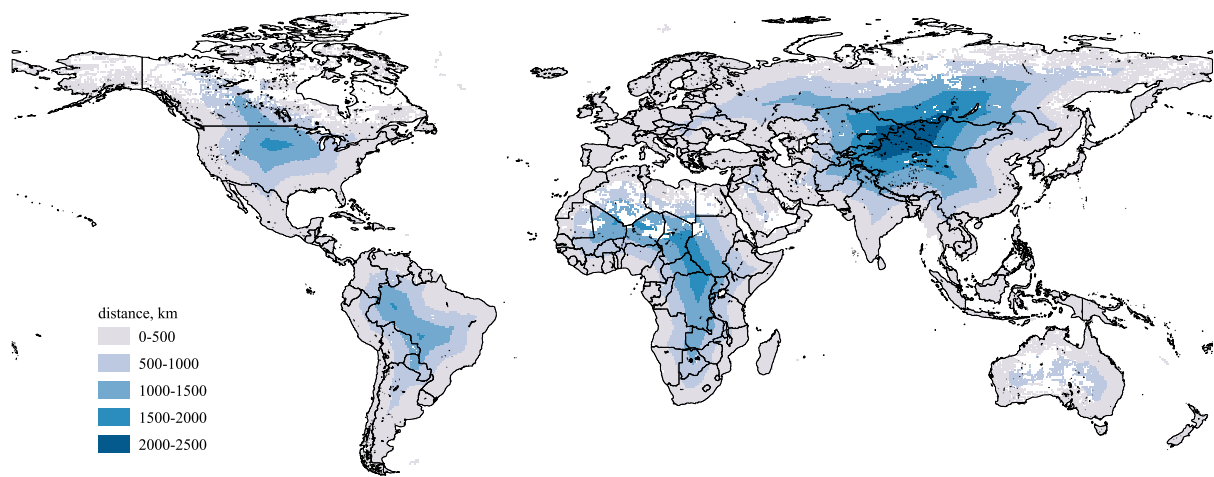
Figure A.5: Worldwide distance to the coast (quantile map).
Source: Authors' calculations; see Table 1 for details.
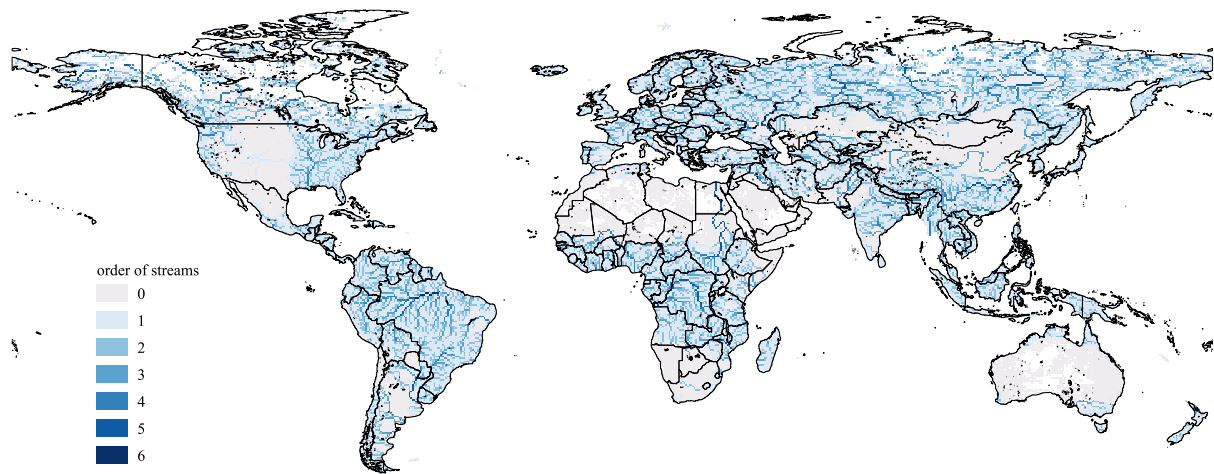
Figure A.6: Worldwide order of streams index (quantile map).
Source: Vörösmatry et al. (2000); see Table 1 for details.
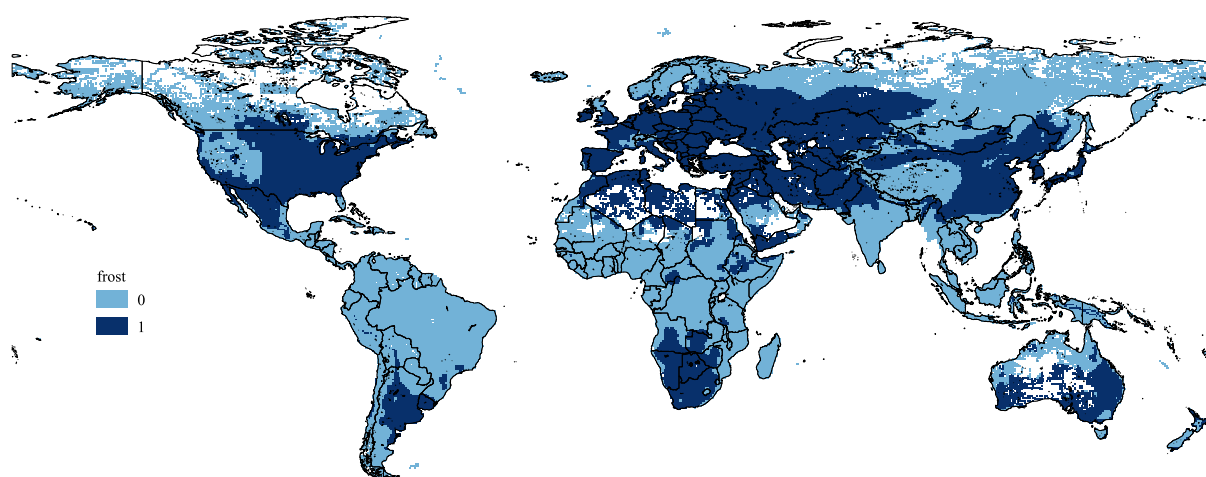
Figure A.7: Worldwide distribution of frost occurrence (quantile map).
Source: IPCC (2002); see Table 1 for details.
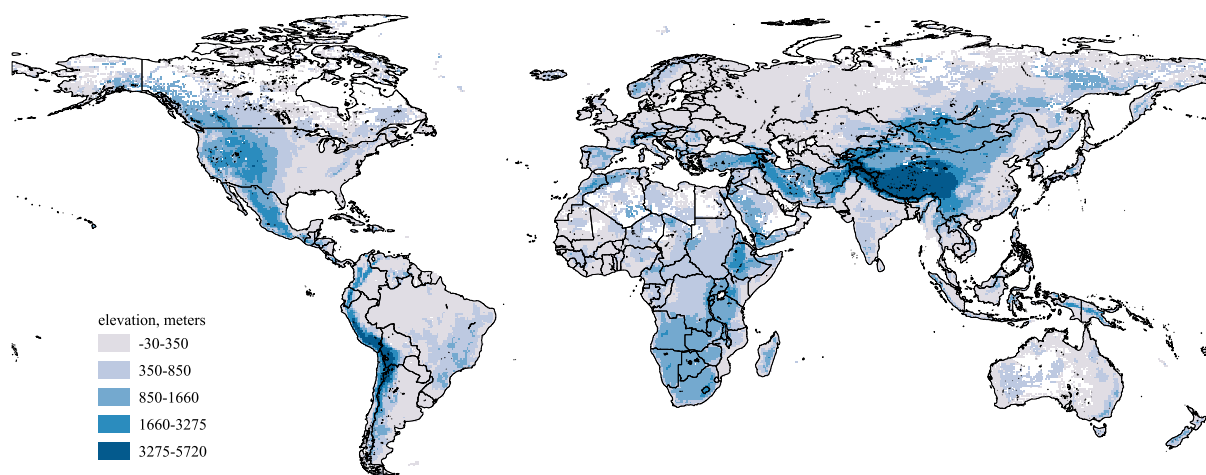
Figure A.8: Worldwide elevation levels (quantile map).
Source: Hijmans et al. (2005); see Table 1 for details.
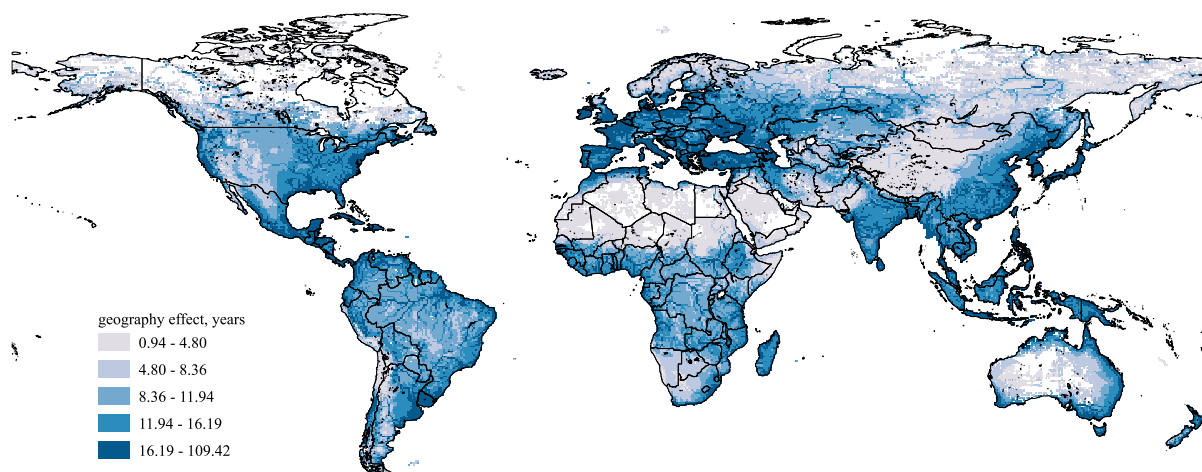
Figure A.9: Predicted effect of geography for grid cells belonging to an "average super-grid" cell, in years, for the 10-percent urbanization threshold (quantile map).
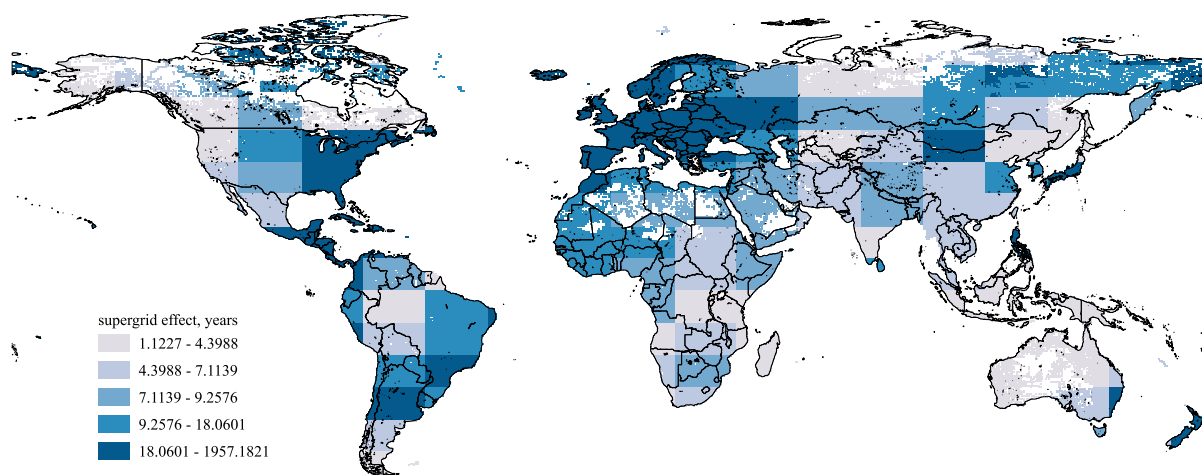
Figure A.10: Predicted effect of "super-grid" cell characteristics for "super-grid" cells with "average" geographical endowments, in years, for the 10-percent urbanization threshold (quantile map).
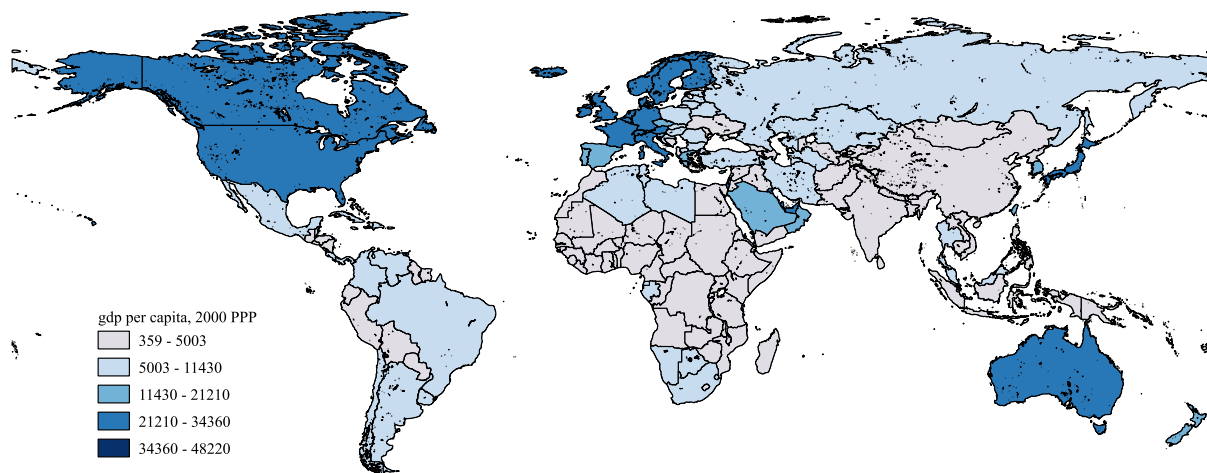
Figure A.11: Per capita GDP in 2000 for 169 countries of the world, in PPP-units.
Source: Penn World Table (Heston et al., 2006)
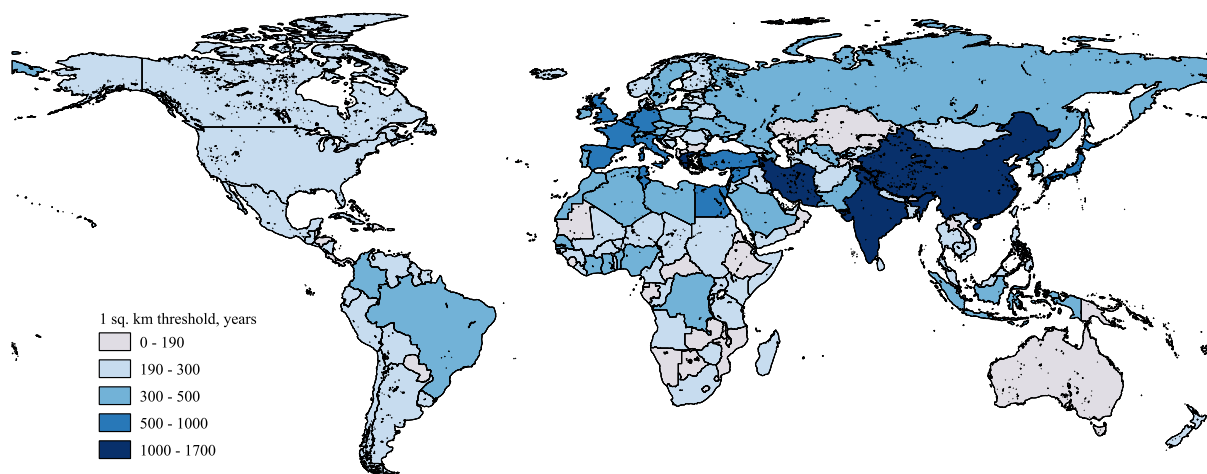
Figure A.12: Number of years since at least one of a country's grid cells reaches the 1 person per km² urbanization threshold.
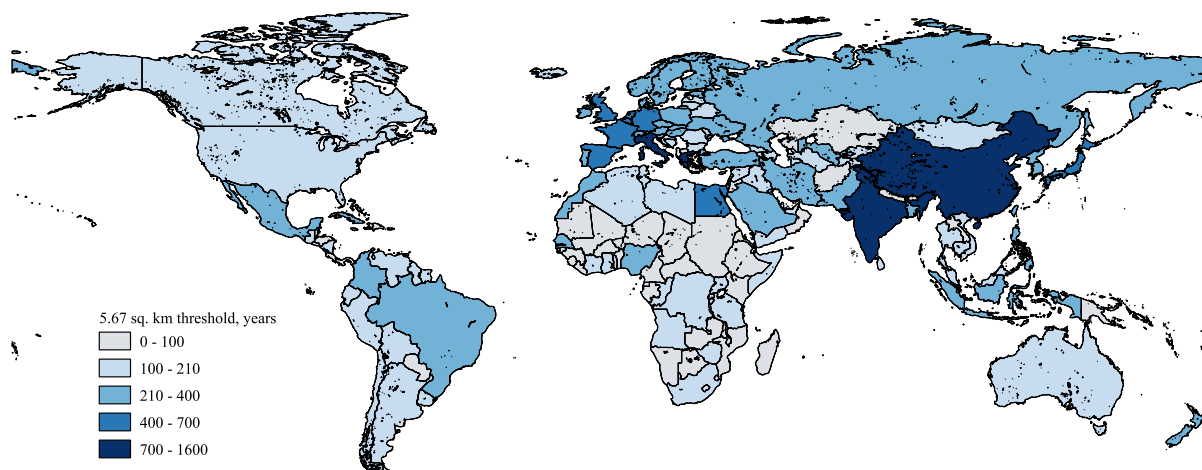Source: Authors' calculations based on Klein Goldewijk et al. (2010)

Figure A.13: Number of years since at least one of a country's grid cells reaches the 5.67 person per km^2 urbanization threshold.
Source: Authors' calculations based on Klein Goldewijk et al. (2010)
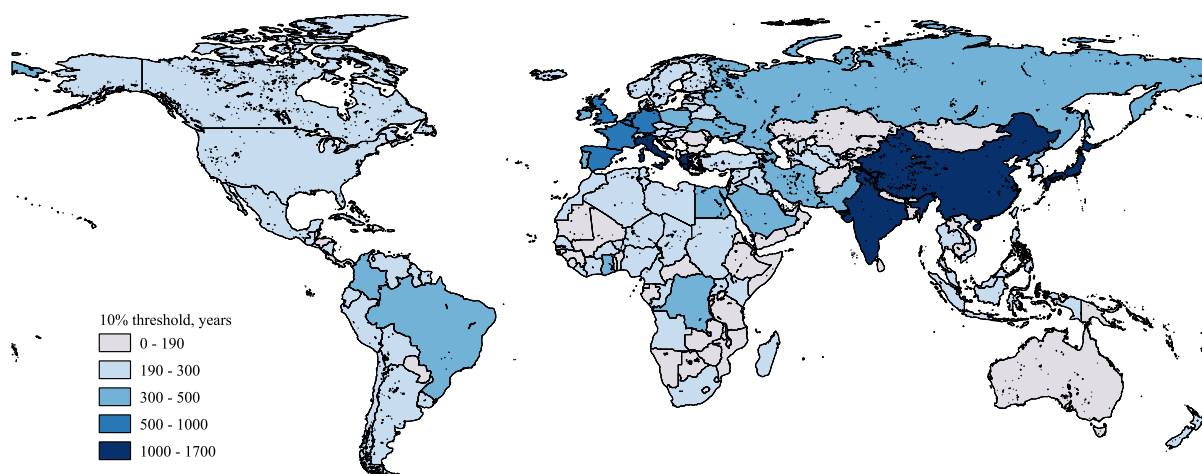

Figure A.14: Number of years since at least one of a country's grid cells reaches the 10 percent urbanization threshold.
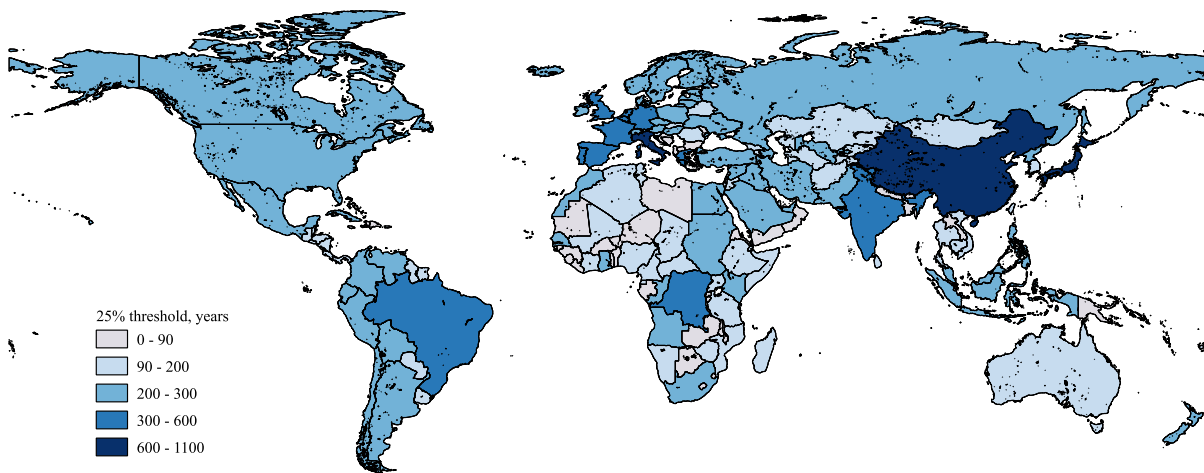Source: Authors' calculations based on Klein Goldewijk et al. (2010)

Figure A.15: Number of years since at least one of a country's grid cells reaches the 25 percent urbanization threshold.
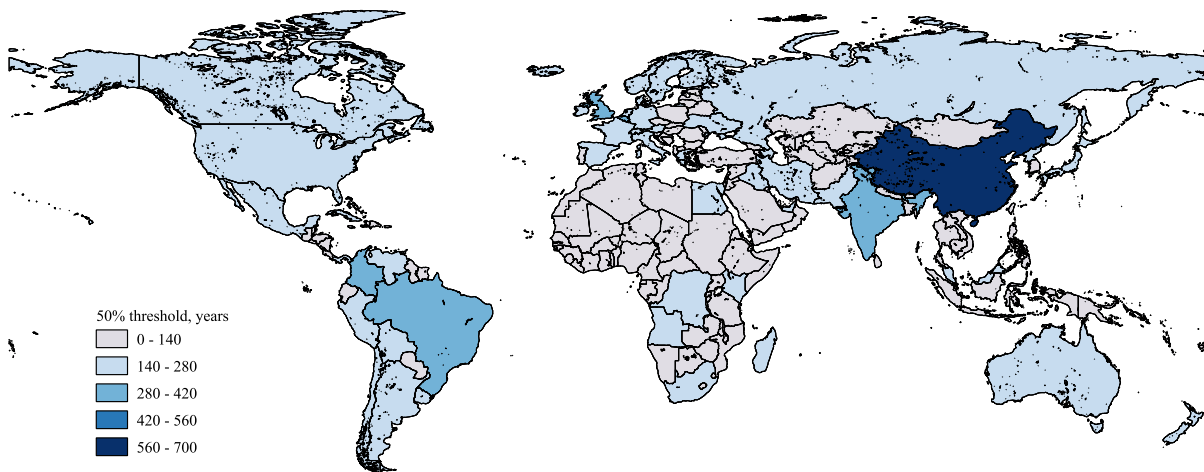Source: Authors' calculations based on Klein Goldewijk et al. (2010)

Figure A.16: Number of years since at least one of a country's grid cells reaches the 50 percent urbanization threshold.
Source: Authors' calculations based on Klein Goldewijk et al. (2010)