LSU | E. J. Ourso College of Business
Department of Economics

DEPARTMENT OF ECONOMICS WORKING PAPER SERIES

Early Urbanization and the Persistence of Regional Disparities within Countries

Areendam Chanda
Louisiana State University

Dachao Ruan
Capital One Bank

Working Paper 2017-01
http://faculty.bus.lsu.edu/papers/pap17_01.pdf

Department of Economics
Louisiana State University
Baton Rouge, LA 70803-6306
http://www.bus.lsu.edu/economics/

Early Urbanization and the Persistence of Regional Disparities within Countries*

Areendam Chanda[1] and Dachao Ruan[2]

[1]Louisiana State University, Baton Rouge, LA
[2]Capital One Bank, McLean, VA

January 2017

Abstract

We explore the extent to which present day economic development at the sub-national level, captured by GDP per capita, urbanization, and night-time light density, is correlated to regional economic development in 1850. Drawing on historical city data, we construct a measure of urban population density and other features of urbanization in 1850 for as many as 2,054 sub-national regions covering 135 countries. We find strong evidence of persistence in regional development. In our baseline estimates, a one standard deviation increase in 1850 urban density raises 2005 GDP per capita by almost 10%. Further, presence of the largest national city in 1850 confers significant advantages to the region even 150 years later. Though our findings are robust to a large range of geographic and spatial controls, proximity to the coast and rivers continues to have a significant effect. While persistence is generally true, there is also considerable heterogeneity, with it being strongest in Asia and West Europe. Early urbanization is also associated with human capital and infrastructure differences across regions. Finally, for a limited sample of countries that were not subject to European colonization, we find that even 1500CE urbanization is significantly associated with modern development.

Keywords: Regional Economic Development, Persistence, Physical Geography, Urbanization.
JEL Codes: N10, N90, O18, O47, R12.

*Corresponding Author: achanda@lsu.edu. We are grateful to Omer Ozak and David Weil for sharing their city data, Rafael LaPorta for subnational data on GDP, and Feng Chi Hsu, Kytt McManus, Omer Ozak, Mark Roberts, and Adam Storeygard for their advice on using spatial datasets. We are also grateful to David Cuberes, Justin Cook, Carter Hill, Douglas McMillin, Ozkan Eren, Seunghwa Rho, David Weil, and participants at the 2015 World Econometric Congress, DEGIT conference in Nashville, Southern Economic Association Meetings in Atlanta, George Mason, and Louisiana State University. This is revised version of an earlier draft titled "150 Years of Worldwide Regional Income Differences." All errors are our own.

1 Introduction

Research on long run growth has shifted its emphasis from understanding the forces of convergence in the past few decades, to exploring the sources of persistent differences in living standards over centuries, and increasingly, millennia. At the sub-national level, i.e. first level administrative units, one would expect such persistence to be less important. The movement of goods and people is inherently easier between regions because of lower transport costs, similar national institutions, and fewer political barriers and thus convergence should be more rapid than at the cross country levels. Despite this, it is often observed that the uneven distribution of economic activity across regions can persist over decades or even hundreds of years. Economically developed regions also show remarkable resilience to large scale natural disasters. Davis and Weinstein (2002), for example, document that the cities of Hiroshima and Nagasaki in Japan returned to prewar trends of population growth in about 20 years after being substantially damaged by nuclear bombings. Historically capital cities, such as Nanjing in China, and Berlin in Germany, continue to retain their status as an important center of commerce despite repeated mass destruction.[1] In similar vein, Kocornik-Mina et al (2015) use a global sample and show that despite large scale flooding, affected cities tend to return rapidly to their earlier state. On the other hand, there are examples of states like Louisiana in the US, and West Bengal in India, which, while having some of the highest levels of per capita incomes in the mid to late nineteenth century, experienced relative declines since.[2] Given the variety of experiences, in this paper we empirically explore the extent to which regional inequalities persist globally; whether they are driven by geographical differences, whether they vary by continent, and various other such groupings.

More specifically, we examine the relationship between contemporary and 1850 measures of regional economic development drawing on a sample that covers 2,054 sub-national regions from 135 countries. For the year 1850, we circumvent data limitations by relying on historical sources and construct a proxy for development which we refer to as "urban population density" - the urban

[1]The national capital of China has alternated between Beijing and Nanjing over the past 600 years.

[2]Easterlin (1960, pp97) estimates Louisiana's per capita income to have been the second highest in 1840 after Rhode Island. West Bengal which was one of the first states to industrialize under British Rule, but has lagged behind since India's independence in 1947.

population of a subnational region in 1850 relative to total land area as measured by current borders. We draw on various sources of estimates of historical settlements such as Chandler (1987), Bairoch et al (1988), and Eggimann (1994). We supplement this measure with binary indicators to capture the presence of urban settlements within a region, as well as its neighboring regions, the location of largest national cities in 1850, urban population densities in neighboring regions, as well as quadratic versions of the density variables to capture non-linearities. Our results overwhelmingly support worldwide "persistence of fortunes" at the sub-national level during the past 150 years. The existence of sufficiently large urban populations 150 years ago is significantly associated with regional GDP per capita in 2005 as well as other proxies of contemporary economic development such as urbanization rates and night-time light density. We control for country fixed effects and a large range of geographic factors commonly used in the literature. The results are also generally true across different samples of countries grouped by continent, by their colonization history, etc., though there are differences in degree. We also briefly look for mechanisms through which urbanization 150 years ago affects current economic performance at the sub-national level. While not conclusive, we find that both human capital and physical capital, as measured by infrastructure, are more strongly associated with historical urban density than cultural or institutional factors. We also find, unsurprisingly, that regions in the Neo-Europes are exceptions to such persistence.

Our choice of using 1850 as the initial year is dictated largely by data considerations - mainly concerns of accuracy and reasonably exhaustive sample size. As one goes further back in time, measurement error gets worse for at least two reasons. First, the number of cities covered by any source or even a combination of sources is likely to get more and more unreliable. Second, even if a city is recorded, population estimates are likely to be increasingly inaccurate as we go further back in time. Indeed if we go back to 1750 or even 1800, the historical compilations miss population estimates for what were obviously well settled regions (e.g. a number of states in the US North East, the state of Kerala and Orissa in India, Tehran in Iran, to name a few.) It is also true that more developed regions kept longer and more complete historical statistics records. In that case our estimation strategy fails, and any evidence of persistence is really one of persistence of records availability. Finally, while all these reasons are essentially limitations to not going back further, 1850 also remains instructive as a starting point since most non-European countries had only just begun industrializing if at all. This would mean that regions with higher levels of development in a country then either capture a much longer civilization history or some initial advantages related to industrialization and/or colonization. Nevertheless, we also explore the effect of regional urbanization in 1500 on present day GDP per capita. For the set of countries which were not subject to European colonization, we continue to find persistence.

Theories that explain regional disparities in economic development emphasize the role of physical geography and the economics of agglomeration, both of which have implications on the long run persistence of economic activities. There are several channels through which physical geography can lead to persistence. First, permanent characteristics of specific locations, such as temperatures, distance to the coast, and ruggedness of terrain, that determined economic prosperity hundreds or thousands of years ago may still play important roles in contemporary economic development. As indicated earlier, Davis and Weinstein (2002) find that the relative population densities of regions in Japan were only temporarily (though substantially) affected by the Allied bombings during World War II, and emphasize the long run importance of physical geography. Second, geographic characteristics may account for differences in culture and social norms, and local institutional development which persist over time. For instance, historical differences between the arable areas which favored permanent settlement and the pastoral areas led to nomadic culture partly contribute to China's cultural differences (Breinlich et al., 2013). Such geographical advantages can further reinforce the economic advantages of agglomeration which are in turn derived from technological externalities which refer to spillovers of knowledge, ideas, and information, and pecuniary externalities which include bigger labor-market pooling and richer availability of intermediates (Breinlich et al., 2013). These externalities tend to attract mobile factors from other regions which in turn generate higher agglomeration effects until the advantages are offset by higher commuting costs, higher land rents, and other congestion costs. While physical geography might often be a primary determinant, such agglomeration effects might help explain why certain regions sustain their advantages. Bleakley and Lin (2012), for example, study the evolution of economic activity across portage sites built before 1900 to avoid navigational obstacles. They find evidence that there is persistence of relatively high population densities at those sites even though their direct relevance to transport costs has long been obsolete. To ensure that our findings are not simply picking up geographical advantages, we also control for a large number of such variables that are now available at a very granular spatial level (in most cases down to 1 square kilometer). While our measures of early urbanization survive the addition of geographic controls, we also find that some of these variables continue to play an significant role in explaining regional disparities. In particular, proximity to the coast or to rivers, has a significant effect though even here there are variations by region (e.g. access to the coast is more important in Africa but rivers are more important in Asia).

1.1 Related Literature

Our research is inspired by two recent advances in the economic growth literature. First, an increasing availability of sub-national data, beyond industrialized countries, has drawn economists to investigate sources determining within-country differences. Acemoglu and Dell (2010), for example, observe that cross-municipality labor income differences within a country is twice as large as cross-country differences in Latin America. They attribute these to variations in the quality of municipal institutions. Tabellini (2010), on the other hand, suggests that variation in institutions may be important to explain cross-country but not within-country inequality. Gennaioli et al. (2013) use a database of 1,569 regions from 110 countries to look for determinants of regional development. They find a strong association between human capital and regional GDP per capita but little effect of institutions. Their work represents a significant advance in this literature since it is the first paper to examine regional differences with such a comprehensive sample of countries. Acemoglu et al. (2014) claim otherwise. Instrumenting for the current average years of schooling with the share of protestant missionaries per 10,000 people in the early 20th century, they argue that the effect of human capital on income per capita returns to a more reasonable range within former colonized countries. Based on an expanded coverage of regions, Mitton (2016) finds no evidence of a positive effect of institutions on development. When looking at growth rates instead of levels, Genniaoli et al (2014) observe that regional convergence rates are only slightly higher than national rates indicating strong barriers to mobility. More recently, Valero and Van Reenen (2016) use sub-national data along with a world database on universities. They show that doubling the number of universities per capita is associated with 4% higher future GDP per capita.

A separate strand of research on long run development has increasingly found that countries which benefitted from more advantageous conditions hundreds, or even thousands of years ago, tend to be richer today. Such conditions include the importance of geographic factors (Hibbs and Olsson, 2004; Olsson and Hibbs, 2005; Ashraf and Galor, 2013) as well as early development in technology (Comin et al., 2010) and state capacity (Chanda et al 2014, Bockstette et al., 2002). Acemoglu et al. (2002), on the other hand, is a notable exception and find no such persistence among former European colonies over the past 500 years. We revisit the same question in the regional context, albeit for a much shorter time period. The papers most closely related to our line of inquiry Maloney and Valencia Caicedo (2016) and Henderson, Squires, Storeygard and Weil (2016). The former paper examines 18 countries within the Americas, and find that sub-national regions with higher pre-colonial population densities 500 hundred years ago tend to have higher population densities and higher income per capita today. They argue that geographic factors as well as increasing return of population agglomeration are plausible mechanisms of persistence. While similar in nature, our work encompasses more countries though the time scale is shorter. Henderson et al. (2016) examine the distribution of economic activity around the world as captured by detailed satellite light data, and note that more than half of the variation can be explained by a parsimonious set of physical geographic variables. Their unit of analysis is not regions but grids and their focus is not on persistence. Nevertheless our research is very complementary since we look at the relative roles of physical geography and early urbanization in explaining the distribution of sub-national economic activity.

The remainder of the paper is organized as follows. In Section 2 we describe our measure of regional development in 1850 and measures of contemporary development around 2000-2005. In Section 3 we present our empirical strategy and results. In Section 4 we look at potential mechanisms for persistence. In Section 5, we briefly investigate persistence over 500 years. Section 6 concludes.

2 Subnational Data

2.1 Measuring Development at the Regional level in 1850

To examine the long run evolution of economic activity at the regional level, one needs reliable measures of regional development. This is particularly problematic as one goes back in time. GDP per capita does not exist at the national level for most countries in the nineteenth century, let alone at sub-national levels. As a result, research on cross-country differences primarily rely on urbanization and/or population density as indicators of living standards, supplemented by measures such as state capacity (Bockstette et al., 2002) and technological development (Comin et al, 2010). However, in 1850, even population estimates for sub-national regions are hard to come by making the construction of both population density and urbanization difficult for a large sample of countries. At the same time, urban historians, such as Chandler (1987), Bairoch et al (1988), and Eggiman (1994), drawing upon various sources, have compiled population estimates of urban settlements going back centuries. These sources are used extensively to construct national measures of urbanization for cross-country studies of economic growth or the evolution of urbanization. We draw on them to construct our primary indicator of development - the 1850 urban population in a region divided by its total contemporary land area - or what we will refer to as *urban population density.* Urban population density is, by definition, a product of the degree of urbanization and population density since,

$$\frac{\text{Urban Population}}{\text{Land Area}} = \frac{\text{Urban Population}}{\text{Total Population}} \times \frac{\text{Total Population}}{\text{Land Area}}.$$

Hence, while not as precise as the two underlying measures, increases in either or both of them would be reflected in increases in urban population density. Before we can evaluate the usefulness of this measure, we first need to discuss what defines an urban area.

2.1.1 Defining an urban area

Even today, the definition of an urban area varies by country and can depend on the size of the population inhabiting an area or its population density. For our work, we include any location that has a recorded population of 5,000 or more in 1850 from our sources. We follow Acemoglu et al. (2014), Acemoglu et al. (2011), and Cantoni (2015) in this regard. With all the data sources taken together, we identify 2,796 settlements with populations of 5,000 or more in 1850 spanning 128 countries. Mapping these settlements into contemporary regions yields 766 sub-national regions with non-zero urban populations in 1850 and 1288 regions without any urban populations - a total of 2054 sub-national units. In Figure 1, we depict the distribution of urban population in 1850 across the world, aggregated to the regional level. The darker regions are more densely populated. Asia and Europe had many more cities in 1850 as well as higher population per city than other places. Due to data limitations of other variables, we are able to use a subset of 1848 these regions. Of these, 745 regions have non-zero urban populations with a mean urban population density of 26 persons per square kilometer and a standard deviation of 120 persons per square kilometer. We report summary statistics of urban population density in Table 1.[3]

We should note that there is nothing sacrosanct about the threshold value of 5000. There are studies that use other thresholds. For example, when studying cities for the period 800-1800CE, Bosker et al. (2013) only consider those that had at least 10,000 inhabitants. Nunn (2011) constructs national urbanization numbers for the period 1000-1900 using a much higher threshold of 40,000. The advantage of having a larger threshold is that the constructed measure is less likely to be subject to biases from country specific variations in keeping demographic records. The downside is that we would have far fewer regions with non-zero urban populations. In the paper, we examine how our results vary when using higher thresholds of 20,000 as well as 100,000.[4]

2.1.2 Urban population density as a measure of regional development

To what extent does urban population density adequately capture regional differences in 1850? Since we do not have other regional level development measures for 1850 on a worldwide scale, as a starting point, we evaluate the extent to which it is correlated at the country level using GDP per capita data from the Maddison Project (2013). Figure 2 displays the scatterplot for 56 countries.

[3] One might wonder why we don't just use all of the data irrespective of settlement size that is available to us and construct a population density measure. The obvious disadvantage of doing this is that with smaller settlements, the likelihood of missed observations is far greater thus making our measure even more noisy.

[4] The distribution of cities across countries and continents according to different minimum population thresholds is summarized in appendix Table A.7. The listing of all settlements with estimated population are displayed in an online appendix of this paper which can be downloaded at http://www.lsu.edu/achanda/research.

The correlation between the log values is at 0.3 - with some new world countries clearly being the outliers. Another way to evaluate the relationship is to compare urban population density in 2000, and GDP per capita from 2005 at the regional level. Figure 3 displays the logarithmic relationship after controlling for country fixed effects. The correlation for this sample of 1380 regions stands at 0.37. This is despite the fact that by the late 20th century there are many regions in the world that have high urbanization combined with low population densities, or vice versa. Finally, Figure 4 depicts the relationship between year 2000 urban population density and night-light density (also, after removing country fixed effects). This allows for a much larger sample of 1901 regions and delivers a much stronger correlation of 0.80. One might be concerned that most of this correlation is driven by the common denominator in both variabes - the land area. If, instead of densities, we look at total urban populations and total night-lights, the correlation is still high at 0.74 (not pictured). In other words, the perfect correlation in the denominator (area of the region), is not responsible for the overall relationship. To summarize, even though not perfect, urban population density contains meaningful information about the distribution of economic activity across sub-national units.

2.2 Measuring Outcomes

We use the logarithm of GDP per capita in 2005 as our main measure of contemporary regional prosperity. The data, which comes from Gennaioli et al, overlaps with 92 countries for which we also have urban population data from 1850. This in turn yields 1,395 regions which forms our baseline sample. Of these, 668 had cities with populations greater than 5,000 in 1850 and the remaining 727 had zero urban populations. If an entire present day country had no settlement recorded in any of our sources, it was completely dropped. We also dropped city-states which comprise of only one region.[5] In Figure 5, we display boundaries of all subdivisions across the world. The areas shaded in dark red are the ones for which we have GDP per capita data. The areas with stripes are ones for which we have no information on settlements but are part of their GDP data set. Thus our baseline analysis is based on the regions which are marked in red and not striped.[6]

[5] However, when examining spatial spillovers, we, include the information from these two groups.

[6] The geographic boundaries for first level administrative units are procured from the Database of Global Administrative Areas Map version 2 (GADMv2) at www.gadm.org. A detailed explanation of regions is provided in Appendix A.1.1.

Relying on GDP per capita alone means that we lose a number of regions for which there is 1850 urban population data. Moreover, it is known that GDP is not accurately measured, especially in developing countries. This problem gets further compounded at the sub-national level since GDP in richer regions may be more accurately reported than in poorer regions. The GDP data itself may not adequately correct for differences in living standards across regions though Genniaoli et al (2014) present evidence that this is not necessarily the case. To ensure that our conclusions are not driven by the limitations of measurement, we use two additional measures of development. These are urbanization (fraction of population living within cities) and the log nighttime light density using satellite data. Both are widely used correlates of development at the country and regional level. Our urbanization data comes from the Global Rural-Urban Mapping Project (GRUMP,Version 1), CIESIN et al (2011). The data is based on census and other officially recorded entries across countries for urban centers. Balk et al. (2006) discuss the various steps involved in creating this data. We should note that the main strength is its reliance on approximately million actual observations rather than having to rely heavily on interpolating and extrapolating data. We aggregate this data to the first level administrative units and then divide it by population data for regions using Gridded Population of the World (GPW version 3, CIESIN (2005)) and is available at a resolution of 30 arc second by 30 arc second (0.86 square km at the equator). The second measure, nighttime lights, sourced from satellite data, has become increasingly popular as a way to circumvent some of the problems related to measurement error in GDP. Henderson, Storeygaard, and Weil (2012) have documented a positive correlation between GDP and nighttime luminosity at the country and regional level, respectively. An increasing number of studies focusing on research questions at the sub-national level also rely on satellite data. In this paper, we use the improved radiance-calibrated data provided by NOAA which is not top-coded compared to some of the earlier research.[7] We use this to construct a density measure- nighttime lights per square kilometer (or luminosity for short). Using these outcomes allows us to expand the coverage up to 135 countries. For the urbanization measure, we use data for the year 2000, which was the last year for which it was constructed. In the case of night-time lights we construct an average of 2000 and 2005.[8]

[7]See Ziskin et al (2010) for an explanation of the underlying methodology.

[8]In the earlier version of our paper, we also considered population density. However, it is strongly correlated with

Table 1 lists summary statistics for these three outcomes. Among them, for the 1,395 regions for which we have GDP per capita data, the mean in 2005 (PPP) is 12,650 US dollars with a standard deviation of 13,380 dollars. The mean value of the logarithm of lights per square kilometer is 0.45 and standard deviation of it is 2.5 which translates to actual pixel values of 1.5 and 12.2 per square kilometer respectively. The maximum value of 6.96 translates to 1053 pixels per square kilometer which underscores the tremendous variation in nighttime light density. Mean urbanization for 2000 is approximately 44% with a standard deviation of 28%. We provide detailed descriptions and the sources of those variables in Appendix Table A.1.

While these two variables offer us a larger coverage compared to GDP per capita, they are not unambiguously better. Nighttime light density, for example, is not very informative in developed countries where income data is already well documented. At the other end of the spectrum, it is known that satellite imagery does not capture nighttime lights for sparsely populated regions and can thus underestimate economic activity. Urbanization, while conceivably an excellent complement to GDP per capita data also has some limitations in terms of the way it is measured. Balk et al (2006) note that its construction is based on three different methods - for developed regions, census data usually is adequate; in other regions it is a combination of light data and census sources. The census source is usually a town or city, and nighttime lights are used to create an urban mask of adjoining areas which allows a more comprehensive measure of the urban population. For some regions in Africa, the absence of sufficient light but the presence of census data implied the adoption of third sources such as aeronautical charts. Keeping these various qualifications in mind, we proceed with displaying their correlations (after controlling for country fixed effects) in Panel A of Table 2. The correlations between GDP per capita and (i) urbanization, and (ii) log light density are 0.52 and 0.41 respectively. The correlation between urbanization and log light density is 0.51. Thus while the three variables are strongly correlated with each other, the correlation is not so strong that any one of them would be redundant. We also look at how these correlations would change if we did not control for country fixed effects. This is highlighted in Panel B. The correlations with GDP per capita are at 0.62 and 0.67 - higher but not so much as to indicate that regional differences are

nighttime light density with a correlation coefficient of 0.86. This was also borne out in the regressions results where the results were very similar. To keep the analysis manageable, we have excluded it from this version.

subordinate. Interestingly, the correlation between urbanization and lights is unaffected by country fixed effects. Henderson et al (2016) also note that country specific factors explain less than 10% of the distribution of economic activity as captured by lights.

3 Empirical Strategy and Results

Our goal is to gauge the strength of association between regional development in 1850, using urban population density as a proxy, and outcomes around the year 2000. To complement this measure we add a number of other indicators of urbanization in 1850. First, more than half the regions in our sample have zero urban population density in 1850. To ensure that our results are not driven by this demarcation, we also use a dummy variable, taking a value of one if urban density is greater than zero i.e. capturing the existence of at least one settlement with population greater than or equal to 5000. In the tables we label this as "Existence of a City in Region, 1850" while in the main text, we will refer to it loosely as the 1850 city dummy. Third, our definition of a region is based on current maps and not those of 1850. In fact, many of the regions were not defined by their current boundaries one and a half centuries ago, that is if they existed at all. Hence, our 1850 density may not be the relevant measure. Even if they did exist, spatial spillovers between adjoining regions needs to be accounted for. To deal with these issues of mis-measurement and spatial correlation, we add two more variables. First, we add another dummy variable identifying whether one or more cities existed within 25 miles geodesic distance of the current region's borders. In the tables, we label this variable as "City in Neighboring Region". In the text we will refer to it more loosely as the 1850 neighboring city dummy.[9] Based on the surrounding cities, we also generate a neighboring urban population density that equals the ratio of aggregated population in neighboring cities to land area of the region being surrounded, hereafter year 1850 neighboring urban population density. 25 miles as a range is clearly an arbitrary choice. We also report results using 100 miles as an alternative range for neighboring areas when we do further robustness tests. Fourth, we also include a dummy variable that equals one if a nation's most populous city in 1850 was located in that region. Our

[9] Geodesic distance refers to the shortest line between two places on the earth's surface, and it does not necessarily mean the shortest path in reality.

inclusion of this is motivated by a large literature on urban bias, and also some of the recent work such as that of Jedwab and Vollrath (2016) highlighting the differences in megacities in developing countries vs developed countries. Here, we explore the possibility that having the largest city located within a region 150 years back, may continue to translate into economic benefits today. Clearly, there is some degree of arbitrariness in this definition - for populous countries there is no reason why we should look at only the largest city. We address this concern later on by seeing how persistence varies based on the size of the largest city within each region.[10]

In light of the discussion so far, we regress measures of contemporary development on urbanization in 1850 using the following specification:

$$\begin{aligned} \mathrm{Y}_{i,t} = \alpha + \beta_1 \text{Urban Dummy}_{i,1850} + \beta_2 \text{UrbPopDensity}_{i,1850} \\ +\beta_3 \text{NeibUrbanDummy}_{i,1850} + \beta_4 \text{NeibUrbPopDensity}_{i,1850} \\ +\beta_5 \text{LargestNatl.City}_{i,1850} + \mu_c + \varepsilon_i \end{aligned} \tag{1}$$

where $\mathrm{Y}_{i,t}$ captures log GDP per capita for region i in year 2005 in our baseline regressions. In additional regressions, it represents log nighttime light density averaged for 2000 and 2005, or the degree of urbanization in 2000. $\text{UrbanDummy}_{i,1850}$ is the year 1850 city dummy of the i^{th} region. $\text{UrbPopDensity}_{i,1850}$ is the year 1850 urban population density of the i^{th} region. $\text{NeibUrbanDummy}_{i,1850}$ is the year 1850 neighboring Urban Dummy of the i^{th} region. $\text{NeibUrbPopDensity}_{i,1850}$ is the year 1850 neighboring urban population density of the i^{th} region. Finally $\text{LargestNatl.City}_{i,1850}$ is a dummy variable to indicate if the region had the largest national city. The term μ_c represents country fixed effects. To mitigate problems of heteroscedasticity, the standard errors are clustered at the country level.[11] We also consider an extended specification which allows for nonlinearity in the relationship between urban population density and income per capita by introducing quadratic terms for both year 1850 urban population density and year 1850 neighboring urban population

[10] One might also wonder why we don't look at the role of 1850 capital cities instead. The simple answer is that for most countries, capital cities did not exist in 1850. According to at least one source, Gordon (2006), there were only forty national capitals as late as 1900.

[11] In the initial scatterplots, we used log urban population density for the year 2000, and yet in the regressions we do not use logarithms. The simple reason is the presence of zero values in 1850. Nevertheless, as we show in the appendix, similar results hold when we restrict the sample to only those regions with positive urban population density. The sample, however, drops by more than half.

density.

$$
\begin{aligned}
\mathrm{Y}_{i,t} = \alpha + \beta_1 \text{Urban Dummy}_{i,1850} + \beta_2 \text{UrbPopDensity}_{i,1850} + \beta_3 \text{UrbPopDensity}^2_{i,1850} \\
+ \beta_4 \text{NeibUrbanDummy}_{i,1850} + \beta_5 \text{NeibUrbPopDensity}_{i,1850} \\
+ \beta_6 \text{NeibUrbPopDensity}^2_{i,1850} + \beta_7 \text{LargestNatl.City}_{i,1850} + \mu_c + \varepsilon_i
\end{aligned} \tag{2}
$$

Finally, in our broadest specification, we introduce a comprehensive set of variables to rule out the possibility that any of the urbanization measures capture advantages related to regional geographic factors. These include average values of temperature, rainfall, altitude, land suitability for agriculture, ruggedness, inverse distance to the coast, and inverse distance to a river, as well as standard deviations of the first four,

$$
\begin{aligned}
\mathrm{Y}_{i,t} = \alpha + \beta_1 \text{Urban Dummy}_{i,1850} + \beta_2 \text{UrbPopDensity}_{i,1850} + \beta_3 \text{UrbPopDensity}^2_{i,1850} \\
+ \beta_4 \text{NeibUrbanDummy}_{i,1850} + \beta_5 \text{NeibUrbPopDensity}_{i,1850} \\
+ \beta_6 \text{NeibUrbPopDensity}^2_{i,1850} + \beta_7 \text{LargestNatl.City}_{i,1850} \\
+ X_i' \delta + \mu_c + \varepsilon_i
\end{aligned} \tag{3}
$$

The vector X_i captures these additional geographic controls. Table 1 lists the summary statistics of all the explanatory variables.

3.1 Baseline Results

In Table 3 we begin with a parsimonious version of 1 by regressing log GDP per capita in 2005 only on 1850 city dummy, the region's 1850 urban population density (which is in 100 persons per sq. km.) and country fixed effects (which we include in every regression in the paper). The coefficient of the dummy variable is 0.086, with a standard error equal to 0.029, while the coefficient of year 1850 urban population density is 0.095 with a standard error of 0.024. These coefficients suggest that regions with at least one city in 1850 were likely to record 9 percent greater GDP per capita in 2005. Furthermore, among the regions that did have cities, a one standard deviation increase in urban residents per square kilometer (approximately 120 persons per sq km.) was associated with an almost 10 percent higher GDP per capita in 2000.

In the second column, we consider the contribution of urbanization of surrounding cities in 1850 to income per capita today. The coefficients of both variables are small in magnitude and insignificant. Coefficients of the regions' own year 1850 city dummy and year 1850 urban population density remain close to their values in column (1). In column (3), we examine the influence of the largest national city in 1850. The coefficient of 0.35 is not only significant but is also large in magnitude. It implies a 35% higher GDP per capita for that region relative to other regions. This is over and above any benefits that the region might have had simply from having an initially high urban density. Furthermore, the 1850 city dummy is no longer significant. One interpretation of this is that the advantage conferred to the region that has the largest 1850 city is so great, that differences in density between regions with zero and non-zero density are no longer relevant.

In column (4), we incorporate quadratic effects for both year 1850 urban population density and year 1850 neighboring urban population density. A negative sign for squared density variables indicates that beyond a point additional increases in urban density may not lead to higher GDP per capita, for example, possibly due to congestion costs. To be sure, such congestion costs are really relevant at the city level. At the regional level urban population density can be high because if the number of cities is high - in which case congestions costs should not be important. Nevertheless, if it is the case, particularly in developing countries, that most regions have really one large city, the non-linear effects may be informative. The results suggest the quadratic effects are indeed present for both the region's own density as well as that for neighboring regions which was insignificant in the earlier columns. If we apply the estimated coefficients to regions with the mean value of 11 persons per sq. km, then it implies a 2.6% higher GDP per capita, a lower effect than the one derived based on only a linear effect. However, this also means that if we considered regions with higher urban population densities, the effects would be larger. Indeed, given the concave shape, one might wonder if there is a maximizing urban population density. The value turns out to be about 700 persons per square kilometer. There are five regions that exceed this value- Basel-City, Beirut, Brussels, London and Vienna. As one can immediately infer, these are all city-regions. Nevertheless, their "excess" density is offset by the fact that they also happened to be the largest cities in their respective countries in 1850. Before we move on to adding geographical controls, we summarize by drawing three main inferences so far - urban population density in 1850 matters for modern regional GDP per capita differences, these effects are non-linear and third, the presence of the largest national city confers significant advantages.

3.2 Geographic Controls

One might be concerned that the association between contemporary income and our early urbanization variables may simply capture the influence of environmental and geographical characteristics, the persistent effects of which, on both national and regional disparities, has been studied extensively. The contribution of our work is tilted more towards the notion of evaluating the importance of early urbanization irrespective of the reason for the existence of a city. This is not to suggest geographical factors are not important determinants. In fact, they may play an important role in the location of newer cities and thus help reduce regional differences.

Physical geography can be captured in many ways, and we include a rich set of variables - altitude, agricultural land suitability, rainfall, ruggedness of terrain, temperature, and proximity to the coast and rivers. We take average values for these as well as the standard deviations in agricultural land suitability, rainfall, and temperature.[12] Many of these variables are obviously correlated to each other. For example, land suitability is partly explained by rainfall and temperatures. Moreover temperature and altitude are negatively correlated. While most of these variables are obvious choices to anyone familiar with the research, we briefly elaborate on a couple of others. We include land suitability for agriculture since Michalopoulos (2012) shows its variance is strongly associated with the emergence of linguistic diversity which in turn has been negatively associated with contemporary economic development. Ruggedness can be expected to adversely affect productivity. Both high elevation and ruggedness means higher costs of economic activities such as construction and transportation. Nunn and Puga (2012) provide evidence of the negative impact of ruggedness on economic development. However, they also note that it protected some African populations from the damaging effects of the slave trade. Finally, we should note that

[12] We do not consider standard deviations in altitude because that is reflected in ruggedness. Since ruggedness itself reflects a standard deviation, we do not consider its standard deviation either.

despite these variables being geographic, exogeneity cannot be assumed. Some of these, such as agricultural suitability and rainfall are based on long averages from the second half of the twentieth century. To some extent, the inclusion of country fixed effects can mitigate this problem though not entirely (Michalopoulous, 2012).[13]

In column (1) of Table 4, we first report impacts of geographic and climatic characteristics on log income per capita in 2005 without including our measures of urban development in 1850. The sample size declines to 1371 because we do not have data on agricultural land suitability for the remaining 24 regions. The results indicate that neither the average, nor the standard deviation in rainfall are significant. The average values for altitude, temperature, and ruggedness have the expected negative effects and are all significant . While the average value of land suitability is not significant, the standard deviation has a negative effect as expected. Interestingly, the standard deviation in temperature has a positive effect on GDP per capita. Finally, proximity to coastlines and rivers confers a distinct advantage for modern outcomes.

Next, we reinstate our measures of urban development in 1850 in column (2). The variables capturing urban density in 1850 continue to be significant and the magnitude of their coefficients are now only marginally lower in value. As a result the net effect is slightly lower (for the mean urban population density, the net effect is now 2.1% rather than 2.6%). The effects of neighboring regions are no longer significant. Finally the presence of the largest city continues to be significant with a coefficient only marginally lower than the previous estimate. Among the geography variables, other than ruggedness, all the variables retain the significance and there is hardly any change in their size of their coefficients. The only other notable change is that average land suitability is now negative and significant. This is not what other papers find. One possible explanation is that this may reflect that after controlling for various aspects of early urbanization as well as other geographical determinants, it may simply be picking up the failure to successfully transition to industry for some agrarian regions. The conclusion so far seems to be that while geography is important and continues to play a role in explaining regional differences, regions that had larger urban populations

[13] We also considered absolute value of latitude in earlier versions. However, we have not included it here given that the current set of geographic variables capture more detail, and the fact that mean temperature has a 0.9 correlation with latitude.

have benefited more over the long run.

While the results so far provide initial evidence supporting the association between urban development in 1850 and current outcomes, one may still be rightly concerned that the results might be driven by our particular choices for dependent and independent variables. To investigate further, we test along two broad dimensions- using alternative outcome variables, and making modifications to the 1850 urbanization measures. First, we investigate robustness to additional dependent variables that have already alluded to - nighttime luminosity as captured by satellites and urbanization. Second, we investigate robustness of the results based on alternative constructs of the 1850 variables. More specifically, we examine a) how the results change if we raise the threshold population value in defining a city and b) if we expand the definition of neighboring areas by increasing the distance from 25 to 100 miles, c) restrict the sample only to regions with positive urban density, and d) consider log urban density instead of the quadratic form.

3.3 Alternative Measures of Modern Regional Development

In this section, we examine the effect of 1850 urbanization on nighttime Light Density and urbanization. Our results using these alternative dependent variables are displayed in Table 5. For each of the dependent variables, we list two regressions, with and without geographic controls. Irrespective of the dependent variable, and whether or not we add geographic controls, we see that at least four of our 1850 variables- the existence of a city, urban density (and its squared value), and the largest national city play a significant role. For example, from column (2) we can infer that the presence of a city in 1850 meant that year 2000 urbanization would be 6.7 percentage points higher. The presence of the largest national city would have an additional effect of 13 percentage points. Finally, at the mean level of urban population density, the marginal effect is a 1 percentage point increase in urbanization with the effects increasing as urban density increases. In the case of urbanization, the maximizing 1850 urban density turns out to be coincidentally at around 700 persons per square km - just as with GDP per capita. As far as geographic variables are concerned, ruggedness and the standard deviation of land suitability, proximity to the coast, and proximity to the rivers continue to have significant effects with the expected signs.

If, we instead focus on the log of light density, the same four 1850 urbanization measures are significant. We also see that the presence of a city in the neighboring region is quite significant and this is robust to geographic controls. With respect to the geographic controls, proximity to coasts and rivers continue to be significant. Average agricultural suitability now has a positive and significant effect. This is not due to the fact that we have a larger sample. In fact, as a check on the consistency of all the results in this table, we reran the regressions with the sample restricted to that of column (2) in Table 4 and found them to be unchanged. One reason the average agricultural suitability maybe positive is that light density captures lights in agricultural areas that would normally not be a part of an urbanization measure.

To summarize our findings here, urbanization in 1850 continues to have a significant effect on urbanization and light density even though these two variables are not perfectly correlated with GDP per capita. The geography variables are less consistent, though proximity to rivers and coasts continue to have significant effects.

3.4 Redefining Urban Density

While we have earlier explained our choice for using 5000 as a threshold, one might be worried that the number is so small that many settlements in 1850 with population slightly greater than 5,000 may be missing historical records leading to a measurement error of urban population density in 1850. Indeed, we do find that for some continents or countries, the city data seems to only record settlements with populations far greater than 5000. For example, most settlements in 1850 in Africa and Asia in our data have a population size higher than 15,000. Thus even if we raised the threshold from 5,000 to 15,000, we would find that for these continents, the number of sub-national regions with positive urban populations, and the total urban population, would remain largely unchanged. On the other hand, the numbers drop substantially for Europe and the Americas.[14] To investigate the effect of using larger population thresholds, we reconstruct variables measuring development in 1850 by using minimum population thresholds of 20,000 and 100,000. We start with a threshold of

[14] If it were the case that settlements within each country were recorded uniformly based on a consistent population threshold, but different countries used different thresholds, then country fixed effects could mitigate the impact of losing of small cities. However, for 1850, that would be wishful thinking.

20,000 for the reason that Chandler's (1987) work, one of the most influential source of historical cities and the benchmark for a lot of the related research, is based on the same threshold. When the threshold is raised to 20,000 the number of regions with positive urban populations declines from 668 to 431 (out of the 1395 regions for which we have 2005 GDP per capita and 1850 settlement data). In Asia and Africa, the decline is muted - from 165 to 145, and 24 to 20 respectively. Americas and Europe experience the bulk of the decline - from 157 to 82 and from 321 to 183 respectively. When the threshold is raised to 100,000, the number of regions drops precipitously to 87 with most of them concentrated in Western Europe, India, China, and the United States.

We report our results in Tables 6. To keep the analysis manageable, moving forward we only list results when geographical controls are added. We can summarize our main findings in this table - first, the switch from 5000 to 20000 does not seem to have major repercussions for our 1850 urbanization variables or the geography controls. Extending the threshold to 100,000 leads to some of the 1850 measures being no longer significant. However, this is to be expected. A 100,000 threshold is high for cities in 1850 and removes a lot of the variation that one would see in city data. A more interesting result is that, as we increase the threshold, irrespective of the dependent variable, the coefficient of the largest 1850 city dummy *increases.* This is particularly stark in the case of log light density. There are possibly two factors at play here. First, in countries which had more than one city with population greater than 100,000, having a higher threshold may tease out the difference, if there is one, between the largest city and other large cities. Second, the largest city dummy now captures the advantage the corresponding regions had, if any, in countries with less populated largest cities. To see if either, or both of these two effects were at play, we also looked at the differential effects based on two sub-samples: countries with cities greater than or equal to 100,000 and less. Interestingly, the effect of the largest city survives in both samples, indicating both factors might be at work. The effect of the largest city was smaller for countries with cities that had populations of 100,000 or more but this varied by our measure of development. It was lower for GDP per capita and light density but not for urbanization. The fact that this effect survives might be taken to be supportive of a superstar effect. Consider for example, the US which had six cities with populations greater than a hundred thousand in 1850- New York, Baltimore, Boston, Philadelphia, New Orleans and Cincinnati (in that order). By 2005, only New York and Philadelphia remained, in the top ten cities with New York continuing to retain its position as the largest city. Thus, while difficult to assert outright, controlling for other large cities may highlight the superstar effect of places like New York. One should keep in mind though that there were only 32 countries that had cities with populations greater than 100,000. With respect to the remaining variables, we can observe that variations in urban density is no longer significant when we look at GDP per capita and we use the 100,000 population. Finally, the geographical factors remain generally consistent in their significance irrespective of the threshold used. Like in the earlier tables, proximity to coasts and rivers continues to be consistently significant across all outcomes.

In addition to changing the size of a city, we also explored the robustness of the persistence results along additional dimensions - changing the extent of surrounding regions, restricting the sample to only regions with non-zero urban populations and examining a logarithmic rather than the quadratic specification. All of the results are listed in the appendix and we go over them briefly here. First, we considered the effect of increasing the size of the surrounding area. Instead of considering 25 miles, we extended the area to 100 miles (160 km) to check if we were under-emphasizing the extent of spatial spillovers. The results are listed in appendix Table A.2. Comparing the coefficients to their counterparts in Tables (4) and (5), we find little effect of changing the extent of the surrounding region. At best, with nightime light density, we can observe a slight increase in the estimated coefficient of neighboring urban density, and also a noticeable decline in the size of the effect of the largest city. Next, we look at what happens if we consider only regions with positive urban populations. These results are listed in appendix Table A.3. While the linear effect of urban population density is slightly higher and the effect of the largest national city is slightly lower, the overall pattern of the results are largely unchanged. We also consider what happens if we abandon the quadratic specification and instead consider the logarithm of urban population density (own and neighboring regions). To preserve the number of observations, all regions with zero urban populations are assigned the lowest observed non-zero urban population density of 0.01 persons per square kilometer (and similarly for the neighboring region density). Given that logarithmic specification itself assumes non-linearity, we now drop the quadratic terms. As the results in Table

A.4 indicate, we continue to see a significant effect of urban population density in 1850.

3.4.1 Robustness to Extreme Values

The quadratic specification implies that the net effects are largest for regions with the highest urban density, albeit at a decreasing rate. Also, a quick review of the summary statistics from Table 1 indicates that there is a huge variation in our measure of urban density even within the regions with non zero urban populations. Furthermore, we just saw that the log specification also fits the data very well. All of these taken together raises the question whether the finding of persistence with our specification so far is being really driven by a small group of regions with extreme values. To further explore this, we pursue two variations of our exercise. First, since urban populations are incorrectly measured in 1850, instead of placing emphasis on the actual density, we group the data by creating dummy variables based on range of values. We then test if the effects are driven by those regions within the highest range of values. Since the distribution, even for non zero values, is highly skewed, we divided the sample into five groups based on the logarithm of the density - 2 standard deviations below the log mean, between 1 and 2 standard deviations below the log mean, between 1 standard deviation below and 1 standard deviation above the log mean, between 1 and 2 standard deviations above the log mean, and above 2 standard deviations above the log mean. These groups are summarized in Table 7. Table 8 displays the regression results. The base group comprises regions with zero urban density. From all the three columns it is clear that higher densities in 1850 lead to higher outcomes. The effects are clearest for nighttime lights, but we also see the dominating effect of the regions within the two highest intervals of urban population density for the other dependent variables as well. The lowest two groups exhibit less clear patterns. In other words, regions with the highest values might indeed be driving the results. As a subsequent exercise, we re-estimated our main specification but now drop the 103 regions that make up groups 4 and 5. Once we drop these regions, the mean urban population density for areas with positive urban populations declines dramatically to 3 persons per square kilometer and the standard deviation declines to 4.5 persons per sq. km. compared to 26 and 120 respectively (see Table 1 for the latter two). The results based on the truncated sample are presented in Table 9. Urban Population Density and the largest national city are still significant. However, the quadratic effects are no longer present which is not surprising given that some of the largest values have been dropped. The effect of the largest national city is lower compared to what we have seen in Tables 4 and 5. This is because of the 103 regions that we dropped, about half had their countries' largest cities. Another observation one can make is the dramatic increase in the size of the coefficients for urban population density compared to earlier tables. However, when we evaluate the effects of a one-standard deviation increase, the effects are much the same. For example, a one standard deviation increase in urban population density still leads to an approximately 10% increase in GDP per capita. To further check the robustness of the result, we excluded all regions with zero urban populations. We include this in the appendix Table A.5. They are surprisingly robust.

While the previous exercise focused on the role played by regions with extremely large values, we now consider a second problem - the fact that there might be regions where despite the presence of cities, urban population density might be an underestimate of living standards because of large areas of uninhabitable land (glaciers, deserts, etc). To address this, we create an adjusted measure of urbanization - urpan population divided by "geographically predicted" total population. The latter is a variable we obtain by regressing the logarithm of observed population in the year 2000 on all the geography controls, country fixed effects and logarithm of the land area. The idea here is to get a sense of what might be a sustainable population based on physical geography factors. This predicted population is used to construct what we call an "Adjusted Urbanization" in 1850. We are under no illusion that these are very hypothetical numbers- after all the relationship between geography and population is known to change over time and hence what might be appropriate for 2000 need not be relevant for 1850. In fact, based on geography, some sub-national units have predicted populations in 2000 substantially lower than their actual population. For some city states, our adjusted urbanization measure is greater than 1 - highlighting the fact that much of the residual population growth is probably due to path dependence and other factors driving agglomeration and not physical geography. Despite this anomaly, the adjusted urbanization has an advantage in terms of creating a new variable that has a substantially reduced variation compared to urban population density. In column 1 of appendix Table A.6 we present the coefficients of the regression that is used to estimate the predicted population. In the remaining columns we use the implied Adjusted Urbanization as our independent variable. Despite the reduced variation, we continue to find significant evidence of persistence.

3.5 Evidence in Subsamples

Having dwelled on persistence at the global level, we next investigate whether this also holds for sub-groups of countries based on geographical or historical circumstances. In our first set of regressions, we look at major continent groupings to see if there is heterogeneity along those lines. It should not be surprising to find that the extent of persistence to be different between, say, Western Europe and Africa. We also look at what happens when we restrict our attention only to Neo-Europes. Beyond clear geographical demarcations, we also examine the extent to which the results might be different when we group countries based on whether they were colonized. To save space, we do not include all the geography coefficients in the tables, but will discuss any interesting findings as we proceed below. We begin by reporting regressions for regions in different continent groups in Tables 10a to 10d. Like before, we show the results for each of the three outcome variables. This is particularly important in sub-groups, since the lack of GDP per capita data is more pronounced. For example, we have only 13 countries with GDP per capita data in Africa. However this number almost doubles (and the number of regions is almost tripled) once we use the other two dependent variables. We present the results for Africa, Americas, Asia and Europe respectively. Starting with Africa, we see that for GDP per capita, none of the urban development measures in 1850 are significant. This may not be surprising given the small sample. However, even when we substitute the other dependent variables, the results are generally inconsistent. For example urban population density has a non-linear effect when using urbanization but not when we use luminosity. More consistent is the observation that the largest national city seems no longer important irrespective of our choice of dependent variable. Our own expectation was the opposite given that unbalanced regional development is more of a concern in third world countries. It is quite possible that the data itself on largest city size may be more suspect than elsewhere or that large cities were not so large to begin with. Among the geography variables, the distance to coast is uniformly significant, and the standard deviation in agricultural suitability also has a significant effect on both urbanization and light density. The latter finding was not true for the worldwide sample but is supportive of Michalopoulos (2012)

Moving on to Americas, in stark contrast to Africa, it is clear that the largest national city in 1850 is strongly associated with modern development irrespective of the dependent variable. Beyond that, we can also see that urban population density in 1850 has a significant effect though the quadratic specification is not useful in looking at GDP per capita differences. Thus for Americas, with some generalization, the results are largely in keeping with that of the global sample. On the other hand, for geographic variables (not shown here), the Americas do not follow the global sample. First, average rainfall now seems to have a significant negative effect and proximity to the river seems not to have played an important role. Proximity to the coast is also not uniformly significant. We also repeated the specifications after dropping US and Canada given their obviously very different trajectories compared to the rest of the continent. The results were largely the same.

For Asia, we get a consistent set of results for the quadratic effects of urban population density and the largest national city. For urbanization and nighttime lights, we also find that existence of a city in 1850, as well as density of the neighboring regions play a significant role. Thus, so far, persistence is most convincing here. As far as geographic determinants are concerned, here too we do not see any convincing effect of proximity to the coast (only for GDP per capita). In fact, for Asia, proximity to the rivers is significant for urbanization and night-time light density

For Europe, in Table 10d, we see a set of interesting results. First, the presence of city in 1850 turns out to be important for all three outcomes. This might be due to a combination of two factors. Smaller cities were better documented in Europe, and by 1850, areas in Europe which were already inhabitable had been settled. In other words, regions that did not already have cities were not as economically viable. Second, urban population density is not significant in explaining GDP per capita. Given that GDP per capita is likely to be better measured for European regions compared to elsewhere this comes as a surprise. However, upon further investigation it turns out that this result is actually due to East European countries. If the sample is restricted only to West Europe as we do in columns 4-6, the quadratic effects of urban population density manifest themselves again.

Third, the presence of the largest national city continues to play an important role despite the fact that Europe was certainly the most urbanized continent. Finally, as far as geography variables are concerned, like Asia, in Europe too we find that proximity to rivers tends to be more significant than proximity to the coast.

Having examined selected continent groupings, we now classify countries based on some other factors. First, we examine the strength of persistence by examining only the Neo-Europes. Compared to the rest of the world, the Neo-Europes (Australia, Canada, New Zealand, and USA) have experienced remarkable movements of labor and capital over the past 150 years. For example, the US states such as California and Texas that were underdeveloped 150 years ago have been growing rapidly. For Australia, the city data indicates Melbourne as the only settlement in 1875 with a population greater than 5000 and New Zealand is entirely missing from our data.[15] In Table 10e we present our estimates for the three remaining countries. As expected, we do not see any consistent evidence of persistence in regional inequalities over the 150 years. Urban population density is only significant at 10% with the urbanization outcome, and the largest national city, if anything, has a negative effect on GDP per capita also at 10%.

To conclude this discussion, even when we consider various subsets of countries, persistence in urbanization seems to be the norm though unsurprisingly, there are exceptions, and the strength of the relationship varies. In particular it is strong for Europe, Asia and the Americas but is not supported as strongly in the data for the Neo-Europes, and also for Africa but possibly for different reasons. Finally, there is also heterogeneity in geographic factors. Unlike the global sample, the role of distance to the coast seems more varied. It matters for Africa, not the Americas and takes a backseat to rivers in Europe and Asia.

3.5.1 Colonization

Next, we look to see if patterns of persistence were any different for colonized vs non-colonized countries. Maloney and Valencia Caicedo (2016) document such persistence at the subnational level over the past 500 years for the Americas despite its tumultuous history. We have also seen this to

[15] According to the New Zealand Government's statistics, the 1858 population estimate was 115000, of which 59000 were Maoris.

hold in our shorter time span of 150 years. In Table 11, we investigate this for the worldwide sample of colonized countries using two strategies. First, we continue our standard fixed effect estimates (Panel A). Second, instead of fixed effects we use a random effects estimator and control for the country level log of population density in 1500, the variable used by Acemoglu et al to document reversals (Panel B). From all six columns, we can clearly see continuing evidence of persistence. The last three columns provide some support that national level population density has negative effect on regional outcomes, though it goes away when we use night light density as the outcome variable. This is not entirely surprising, since Acemoglu et al note that the reversal had occurred by 1850.[16] The magnitude of the 1850 urban population density is much larger in size compared to those in Tables (4) and (5) while the effect of the largest national city is lower. In Panel C, we also include results for non-colonized countries. The coefficients for these countries are much closer to that of the global sample.

4 Potential Mechanisms

So far we have found evidence of persistence in long run development at the regional level over the past 150 years. An equally interesting but harder question is through what channels is early urban development linked to income today at the regional level. Is it through institutions, human capital, or physical capital? To provide a concrete answer one would need to find proper instruments given the feedback effects of income on these variables. Here, instead we carry out an exercise similar to that of Putterman and Weil (2010) and Genniaoli et al (2013) to look for the potential channels in Table 12. We begin by looking at the relationship between urbanization in 1850 and years of education in 2005 in column (1). The coefficient of year 1850 city dummy, the non-linear effects of urban population density, as well as neighboring population density, and the largest national city are all significant. Among the geography variables, as in many of the earlier regressions, access to the coast and rivers is significant. We also see that higher temperatures has a negative effect. In Column (2) we consider an indicator of culture, trust in others. Despite the large number of variables, there is no strong message that can be inferred. In columns (3) - (5) we regress three measures of regional

[16] See Figure IVa in their paper

institutions - informal payments, access to financing and log days without electricity. The estimates do not indicate any strong association with 1850 measures of urban development and in fact, the presence of cities seems to actually increase days without electricity. [17]

The remaining two columns capture the effect of urbanization in 1850 on infrastructure measured by log power line density in column (6) and log travel time in column (7). Both columns show that regions with a city in 1850 tend to have larger and more efficient infrastructure. However, it is clear that our set of 1850 variables have a significant effect on lowering travel time. Moreover a range of geography variables including proximity to coasts and rivers have a significant effect on both these variables. Overall, while the enter exercise is at best suggestive, it is clear that there is a stronger association between our 1850 measures and the human capital and infrastructure variables.

5 Going Back Further

Given the extensive literature on long run persistence in development across countries, one might be curious as to whether contemporary regional disparities might originate even earlier than 1850. As a final exercise, we extend our urbanization measures back to 1500. Needless to mention, we are skeptical about the accuracy of the data. In our estimates, we control for both 1500 and 1850 urbanization variables. However, we drop the quadratic versions and focus only on log GDP per capita as the dependent variable to keep the discussion brief. In addition to estimating the relationship for all countries, given the time-span, we also separate them based on their (European) colonization experience. As the results in Table 13 indicate, for the global sample there are no meaningful inferences apart from the fact that the largest national city in 1850 continues to have a significant effect. Once we break the sample into its two parts based on colonization, we see some more variation in the results. For colonized countries, both variations in 1500 and 1850 density matter while 1500 cities has a negative effect. All of these are not necessarily contradictory - it is possible that many of the important cities are recorded only in the 1850 data. As a result the variation in density is important in 1500 for the ones that are recorded, but the ones that show up

[17] The data for all these variables are taken from Gennaioli et al (2013) and so are the classification as human capital, culture, infrastructure, etc.

in 1850 are more important (and hence the negative sign for existence of cities).

In the case of non-colonized countries (column 3) the existence of 1500 and 1850 cities are both important but it turns out that variations in density in 1500 is significant. This might reflect the fact that in these countries, cities from 1500 are better documented. Thus while the importance of cities by 1850 are important, the true variations come from an even longer run persistence of urbanization. Finally, for both colonized and non-colonized countries, the largest national city has a sifnificant effect with the size of the coefficient being remarkably similar across all three columns and also that in column (2) of Table 4.

6 Conclusion

The debate regarding sources of economic prosperity has attracted economists' attention to historical and geographic factors. Existing studies have documented cross-country evidence that economic activities hundreds or thousands of years ago play an important role in shaping the distribution of the world economy today. For inequality of economic development at the sub-national level, however, most of studies are restricted to a single country or to a single continent. In this paper, using a global sample, we show that regions with larger urban populations in 1850 continue to be better off today. Further, regions that had the largest cities benefited even more.

We should conclude with some limitations of our work which can guide future research. First, we have used only two years, 1850 and 2000 (or 2005) for our analysis. An obvious step forward would be to collect information in the intervening time period to see whether there were strong breaks in the pattern. An obvious candidate would be approximately around the middle of the 20th century. A second limitation is that persistence does not imply lack of convergence. The problem is that the data for initial and terminal years use different definitions of urbanization. Therefore, even though the correlation is strong, it is harder to make inferences regarding convergence. Third, while the 1850 data was constructed using well known historical compilations, a number of years have passed since their publication. It should be possible in the future to create a more detailed dataset for settlements (particularly in developing economies) based on newer archival work.[18] Finally, while

[18] Reba et al (2016) have made important progress in this direction.

we document persistence in the past 150 years, some of the geographical advantages may not hold in the future. With increasing concerns about climate change and vulnerability along the coast, it might well be that proximity to the coast flips signs in the next 100 years.

References

[1] Acemoglu, Daron, Simon Johnson, and James Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." Quarterly Journal of Economics 107: 1231-1294.

[2] Acemoglu, Daron, and Melissa Dell. 2010. "Productivity Differences Between and Within Countries." American Economic Journal - Macroeconomics 2 (1):169-88.

[3] Acemoglu, Daron, Davide Cantoni, Simon Johnson, and James Robinson. 2011. "The Consequences of Radical Reform: The French Revolution." American Economic Review 101 (7): 3286-3307.

[4] Acemoglu, Daron, Francisco A. Gallego, and James Robinson. 2014. "Institutions, Human Capital and Development." Annual Review of Economics, 6:875–912.

[5] Ashraf, Quamrul, and Oded Galor. 2013. "The "Out of Africa" Hypothesis, Human Genetic Diversity, and Comparative Economic Development." American Economic Review 103 (1): 1 - 46.

[6] Balk, D.L., U. Deichmann, G. Yetman, F. Pozzi, S. I. Hay, and A. Nelson. 2006. "Determining Global Population Distribution: Methods, Applications and Data." Advances in Parasitology 62:119-156. http://dx.doi.org/10.1016/S0065-308X(05)62004-0.

[7] Bairoch, Paul, Jean Batou, and Pierre Chevre. 1998. *La Population des villes Europeenes de 800 a 1850: Banque de Donnees et Analyse Sommaire des Resultats.* Geneva: Centre d'histoire economique Internationale de l'Universite de Gneve, Libraire Droz.

[8] Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence." Quarterly Journal of Economics 127 (2): 587-644.

[9] Bockstette, Valerie, Areendam Chanda, and Louis Putterman. 2002. "States and Markets: The Advantage of an Early Start." Journal of Economic Growth 7 (4): 347 - 369.

[10] Bosker, Maarten, Eltjo Buringh, and Jan Luiten 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.

[11] Breinlich, Holger, Gianmarco I.P. Ottaviano, and Jonathan R.W. Temple. 2013. "Regional Growth and Regional Decline." Handbook of Economic Growth Vol.2. Amsterdam: North Holland.

[12] Chandler, Tertius. 1987. *Four Thousand Years of Urban Growth: An Historical Census.* Lewiston, NY: St. David's University Press.

[13] Cantoni, Davide. 2015. "The Economic Effects of The Protestant Reformation: Testing The Weber Hypothesis in The German Lands." Journal of the European Economic Association, Vol. 13, N. 4 (August 2015): pp. 561–598

[14] Center for International Earth Science Information Network - CIESIN - Columbia University, and Centro Internacional de Agricultura Tropical - CIAT. 2005. Gridded Population of the World, Version 3 (GPWv3): Population Density Grid. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4XK8CG2. Accessed May 11, 2016

[15] Center for International Earth Science Information Network - CIESIN - Columbia University, International Food Policy Research Institute - IFPRI, The World Bank, and Centro Internacional de Agricultura Tropical - CIAT. 2011. Global Rural-Urban Mapping Project, Version 1 (GRUMPv1): Settlement Points. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4M906KR. Accessed May 11, 2016.

[16] Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Gridded Population of the World, Version 4 (GPWv4): Land and Water Area. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H45M63M9. Accessed May 11, 2016.

[17] Comin, Diego, William Easterly, and Erick Gong. 2010. "Was the Wealth of Nations Determined in 1000 BC?" American Economic Journal -Macroeconomics 2 (3): 65-97.

[18] Davis, Donald R., and David E. Weinstein. 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity." American Economic Review 92 (5): 1269-1289.

[19] Easterlin, Richard. 1960. "Interregional Differences in Per Capita Income, Population, and Total Income, 1840-1950" in NBER, Trends in the American Economy in the Nineteenth Century, Studies in Income and Wealth 24, Princeton, NJ: Princeton University Press. pp 73-142.

[20] Eggimann, Gilbert. 1999. *La Population des villes des Tiers-Mondes, 1500 - 1950.* Geneva: Centre d'histoire economique Internationale de l'Universite de Geneve, Libraire Droz.

[21] Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2013. "Human Capital and Regional Development." The Quarterly Journal of Economics 128 (1): 105-164.

[22] Gennaioli, Nicola, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2014. "Growth in Regions." Journal of Economic growth 19: 259 - 309.

[23] Jedwab, Remi, and Dietz Vollrath, 2016, The Urban Mortality Transition and the Rise of Poor Mega-Cities, Manuscript, University of Houston.

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

[25] Henderson, J. Vernon, Adam Storeygard, Tim L. Squires, and David N. Weil, "The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade," NBER Working Paper No. 22145, April 2016.

[26] Hibbs, Douglas A. Jr., and Ola Olsson. 2004. "Geography, Biogeography, and Why Some Countries Are Rich and Others Are Poor." Proceedings of the National Academy of Sciences 101: 3715-3720.

[27] Hodler, Roland, and Paul A. Raschky. 2014. "Regional Favoritism." The Quarterly Journal of Economics: 995 - 1033. doi:10.1093/qje/qju004.

[28] Kocornik-Mina, Adriana , Thomas K.J. McDermott, Guy Michaels, and Ferdinand Rauch, 2015, "Flooded Cities," Working Paper No. 221, Grantham Research Institute on Climate Change and the Environment, London School of Economics.

[29] Maddison Project, 2013, http://www.ggdc.net/maddison/maddison-project/home.htm.

[30] Maloney, William F., and Felipe Valencia Caicedo. 2016. "The Persistence of (Subnational) Fortune." Available at SSRN: http://ssrn.com/abstract=2397299, forthcoming in *The Economic Journal*.

[31] Michalopoulos, Stelios, 2012, "The Origins of Ethnolinguistic Diversity," American Economic Review, 2012, 102(4): 1508-1539

[32] Mitton, Todd. 2016. "The Wealth of Subnations: Geography, Institutions and Within-country Development." Journal of Development Economics, Volume 118, Pages 88-111, 2016.

[33] Nunn, Nathan, and Nancy Qian, 2011, The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment, The Quarterly Journal of Economics, 2011, 126, 593–650

[34] Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." The Review of Economics and Statistics 94 (1): 20-36.

[35] Olsson, Ola, and Douglas A. Hibbs, Jr. 2005. "Biogeography and Long-Run Economic Development." European Economic Review 49: 909-938.

[36] Putterman, Louis with David Weil. 2010. "Post-1500 Population Flows and the Long-Run Determinants of Economic Growth and Inequality." Quarterly Journal of Economics 125 (4): 1627-1682.

[37] 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): 377–92.

[38] Reba, Meredith, Femke Reitsma and Karen C. Seto, 2016, "Spatializing 6,000 years of global urbanization from 3700 BC to AD 2000," Science Data 3:160034 doi: 10.1038/sdata.2016.34.

[39] Tabellini, Guido. 2010. "Culture and Institutions: Economic Development in The Regions of Europe." Journal of the European Economic Association 8 (4): 677-716.

[40] Valero, Anna and John Van Reenen, "The Economic Impact of Universities: Evidence from Across the Globe," NBER Working Paper No. 22501, August 2016.

[41] Ziskin, D., K. Baugh, F. C. Hsu, T. Ghosh, C. Elvidge, 2010, "Methods Used For the 2006 Radiance Lights," Proceedings of the Asia-Pacific Advanced Network, v.30, p.131-142.

Table 1: Summary Statistics of Key Variables

Variable	N	Mean	Std. Dev.	Min.	Max.
1850 Urbanization Measures:					
Existence of a City, 1850*	1848	.4	.49	0	1
Urban Population Density, 1850	1848	.1	.78	0	15.71
—Regions with Urb. Popn.>0	745	.26	1.20	.0001	15.71
Existence of City in Neighboring Regions, 1850*	1848	.47	.5	0	1
Urban Popn. Dens. in Neighboring Regions, 1850	1848	.64	7.28	0	197.58
—Regions with Urb. Popn.>0	745	.40	4.31	0	103.47
Largest National City in 1850*	1848	.06	.25	0	1
Dependent Variables:					
GDP per capita in 1,000 USD, const. 2005 PPP	1395	12.65	13.38	.07	143.48
Urbanization, 2000	1848	.44	.28	0	1
Ln(Nighttime Light density), 2000 & 2005 Avg.	1848	.45	2.59	-9.46	6.96
Geographic Controls:					
Mean Temperature (Cel.)	1848	15.93	8.44	-14.4	29.59
Mean Land Suitability	1848	.43	.28	0	1
Mean Altitude (100m)	1848	5.85	6.51	-.13	48.8
Mean Rainfall (meters)	1848	1.04	.74	0	5.4
Mean Ruggedness (100m)	1848	1.39	1.39	.01	10.13
Std. Devn. Temperature	1848	16.55	15.43	0	103.17
Std. Devn. Land Suitability	1848	.09	.07	0	.38
Std. Devn. Rainfall	1848	168.06	202.07	0	2175.21
Inverse Distance to Coast	1848	.82	.16	.33	1
Inverse Distance to River	1848	.86	.14	.34	1

Note: Unit of population density is 100 persons per square kilometers. Variables with * take values of 0 or 1. A detailed explanation on these variables and their sources is in Appendix Table A.1.

Table 2: Correlations Between Present Day Outcomes

Panel A: Partial Correlations After Removing Country Fixed Effects

	Ln. GDP pc	Urbanization	Ln. Light Density
Ln. GDP pc.	1 (n=1395)		
Urbanization	0.50 (n=1395)	1 (n=2054)	
Ln. Light Density	0.42 (n=1395)	0.50 (n=2054)	1 (n=2055)

Panel B: Partial Correlations before Removing Country Fixed Effects

Ln GDP pc	1 (n=1395)		
Urbanization	0.62 (n=1395)	1 (n=2054)	
Ln. Light Density	0.67 (n=1395)	0.49 (n=2054)	1 (n=2055)

Table 3: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	(1)	(2)	(3)	(4)
	Ln. GDP pc 2005			
Existence of a City (1850)	0.086[b]	0.096[a]	0.024	0.015
	(0.029)	(0.026)	(0.023)	(0.023)
Urban Population Density 1850	0.097[a]	0.092[a]	0.056[b]	0.240[a]
	(0.025)	(0.025)	(0.022)	(0.065)
Sq Urban Pop. Den. 1850				-0.017[b]
				(0.006)
City in Neighboring Regions (1850)		-0.050	-0.043	-0.042
		(0.049)	(0.046)	(0.044)
Urb. Pop. Den. in Neib. 1850		0.007	0.009	0.067[b]
		(0.006)	(0.008)	(0.028)
Squ. Urb. Pop. Den. in Neib. 1850				-0.002[b]
				(0.001)
Largest National City in 1850			0.354[a]	0.304[a]
			(0.052)	(0.054)
Countries	92	92	92	92
Observations	1395	1395	1395	1395
Adjusted R^2	0.04	0.04	0.08	0.11

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Fixed-effects estimates include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 4: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850 and Geographic Controls

	(1)	(2)
	Ln. GDP pc 2005	
Existence of a City (1850)		0.025
		(0.030)
Urban Population Density 1850		0.193^b
		(0.059)
Sq Urban Pop. Den. 1850		-0.013^b
		(0.005)
City in Neighboring Regions (1850)		-0.029
		(0.032)
Urb. Pop. Den. in Neib. 1850		0.069
		(0.042)
Squ. Urb. Pop. Den. in Neib. 1850		-0.003
		(0.002)
Largest National City in 1850		0.274^a
		(0.056)
Mean Temperature (Cel.)	-0.034^b	-0.033^b
	(0.010)	(0.011)
Mean Land Suitability	-0.075	-0.125^c
	(0.064)	(0.070)
Mean Altitude (100m)	-0.018^b	-0.020^b
	(0.006)	(0.007)
Mean Rainfall (meters)	-0.072	-0.059
	(0.048)	(0.048)
Mean Ruggedness (100m)	-0.056^b	-0.037
	(0.022)	(0.022)
SD Temperature	0.003^c	0.003^b
	(0.001)	(0.001)
SD Land Suitability	-0.770^a	-0.564^b
	(0.169)	(0.176)
SD Rainfall	-0.000	-0.000
	(0.000)	(0.000)
Proximity to the Coast	0.893^a	0.771^a
	(0.183)	(0.179)
Proximity to Rivers	0.481^b	0.404^b
	(0.182)	(0.188)
Constant	8.532^a	8.620^a
	(0.261)	(0.278)
Countries	92	92
Observations	1371	1371
Adjusted R^2	0.15	0.22

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 5: Urbanization and Nighttime Light Density as Dependent Variables

	(1)	(2)	(3)	(4)
	Urbanization, 2000		Ln. Light Density, 2001-05	
Existence of a City (1850)	0.074[a]	0.068[a]	0.752[a]	0.559[a]
	(0.016)	(0.017)	(0.106)	(0.075)
Urban Population Density 1850	0.136[a]	0.109[b]	1.250[a]	1.024[a]
	(0.037)	(0.035)	(0.212)	(0.180)
Sq Urban Pop. Den. 1850	-0.010[b]	-0.008[b]	-0.089[a]	-0.074[a]
	(0.004)	(0.003)	(0.020)	(0.016)
City in Neighboring Regions (1850)	-0.010	-0.018	0.754[a]	0.457[a]
	(0.014)	(0.014)	(0.094)	(0.088)
Urb. Pop. Den. in Neib. 1850	-0.005	-0.006	0.019	0.014
	(0.005)	(0.004)	(0.019)	(0.013)
Squ. Urb. Pop. Den. in Neib. 1850	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Largest National City in 1850	0.141[a]	0.135[a]	0.957[a]	0.863[a]
	(0.026)	(0.026)	(0.163)	(0.151)
Mean Temperature (Cel.)		-0.001		0.035
		(0.002)		(0.028)
Mean Land Suitability		0.023		1.527[a]
		(0.035)		(0.272)
Mean Altitude (100m)		-0.001		0.010
		(0.002)		(0.023)
Mean Rainfall (meters)		-0.035		-0.350[c]
		(0.024)		(0.202)
Mean Ruggedness (100m)		-0.019[c]		-0.008
		(0.010)		(0.065)
SD Temperature		-0.000		-0.006
		(0.001)		(0.007)
SD Land Suitability		-0.288[b]		-0.742
		(0.089)		(0.729)
SD Rainfall		-0.000		-0.000
		(0.000)		(0.000)
Proximity to the Coast		0.217[b]		4.685[a]
		(0.069)		(0.771)
Proximity to Rivers		0.150[b]		2.579[a]
		(0.064)		(0.722)
Countries	127	127	127	127
Observations	1848	1848	1848	1848
Adjusted R^2	0.14	0.17	0.20	0.33

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 6: Using Higher Population Thresholds in Defining an 1850 City.
Thresholds of 20,000 and 100,000.

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln. GDP pc 2005		Urbanization, 2000		Ln. Light Density, 2000-05	
Existence of a City (1850)	0.015	0.030	0.059^b	0.047^c	0.457^a	0.092
	(0.032)	(0.078)	(0.018)	(0.024)	(0.089)	(0.116)
Urban Population Density 1850	0.200^b	0.122	0.090^b	0.059^b	0.916^a	0.599^b
	(0.061)	(0.083)	(0.031)	(0.023)	(0.186)	(0.220)
Sq Urban Pop. Den. 1850	-0.014^b	-0.009	-0.006^b	-0.004^b	-0.068^a	-0.045^b
	(0.005)	(0.006)	(0.003)	(0.002)	(0.018)	(0.017)
City in Neighboring Regions (1850)	0.005	-0.014	-0.021	-0.044^c	0.300^b	0.156
	(0.027)	(0.060)	(0.014)	(0.022)	(0.096)	(0.117)
Urb. Pop. Den. in Neib. 1850	0.111	0.019	0.006	0.007	0.146^b	0.186^c
	(0.070)	(0.062)	(0.013)	(0.015)	(0.069)	(0.104)
Squ. Urb. Pop. Den. in Neib. 1850	-0.008	-0.002	-0.000	-0.001^c	-0.002^b	-0.004^c
	(0.006)	(0.005)	(0.000)	(0.000)	(0.001)	(0.002)
Largest National City in 1850	0.282^a	0.328^a	0.143^a	0.196^a	0.981^a	1.456^a
	(0.056)	(0.051)	(0.027)	(0.028)	(0.170)	(0.177)
Mean Temperature (Cel.)	-0.034^b	-0.034^b	-0.000	-0.001	0.046	0.049
	(0.011)	(0.011)	(0.002)	(0.002)	(0.030)	(0.031)
Mean Land Suitability	-0.136^c	-0.118^c	0.024	0.028	1.655^a	1.804^a
	(0.070)	(0.062)	(0.035)	(0.033)	(0.291)	(0.295)
Mean Altitude (100m)	-0.020^b	-0.019^b	-0.000	0.000	0.018	0.022
	(0.007)	(0.006)	(0.002)	(0.002)	(0.024)	(0.025)
Mean Rainfall (meters)	-0.057	-0.062	-0.035	-0.036	-0.342	-0.378^c
	(0.050)	(0.050)	(0.025)	(0.025)	(0.221)	(0.225)
Mean Ruggedness (100m)	-0.036	-0.044^c	-0.021^b	-0.025^b	-0.013	-0.049
	(0.022)	(0.022)	(0.010)	(0.010)	(0.065)	(0.067)
SD Temperature	0.003^b	0.003^c	-0.000	-0.000	-0.006	-0.008
	(0.001)	(0.001)	(0.001)	(0.001)	(0.008)	(0.008)
SD Land Suitability	-0.574^b	-0.665^a	-0.287^b	-0.281^b	-0.758	-0.601
	(0.187)	(0.164)	(0.089)	(0.086)	(0.747)	(0.745)
SD Rainfall	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Proximity to the Coast	0.758^a	0.796^a	0.234^b	0.253^a	4.930^a	5.259^a
	(0.182)	(0.176)	(0.071)	(0.070)	(0.794)	(0.801)
Proximity to Rivers	0.386^b	0.410^b	0.154^b	0.152^b	2.589^a	2.710^a
	(0.188)	(0.187)	(0.064)	(0.066)	(0.740)	(0.754)
Countries	92	92	128	128	128	128
Observations	1371	1371	1857	1857	1857	1857
Adjusted R^2	0.22	0.21	0.17	0.15	0.31	0.28

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 7: Summary of Urban Density Groups

	N	Mean	Std. Dev.
Group 1	10	.0003	.0001
Group 2	105	.0016	.0008
Group 3	548	.0434	.0467
Group 4	74	.6309	.382
Group 5	29	6.34	6.82

Note: Regions with non-zero urban populations are divided into five groups based on the mean and standard deviation of logarithmic values. Group 1 includes regions with density less than two standard deviations below the mean of log density, Group 2 between 1 and 2 standard deviations, Group 3 between 1 std. deviation below and 1 std. deviation above the mean, Group 4 between 1 and 2 std. deviations above the mean, and Group 5 is 2 std. deviations above the mean. Mean and Standard Deviations are in 100 persons per sq. km.

Table 8: Urban Density in 1850 as a Discontinuous Variable

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2001-05
Urban Pop Den (1850), Group 1	0.305[c]	0.078	-0.310
	(0.162)	(0.051)	(0.294)
Urban Pop Den (1850), Group 2	-0.040	0.018	0.258[b]
	(0.057)	(0.016)	(0.123)
Urban Pop Den (1850), Group 3	0.041	0.081[a]	0.691[a]
	(0.030)	(0.018)	(0.088)
Urban Pop Den (1850), Group 4	0.262[a]	0.305[a]	1.964[a]
	(0.071)	(0.045)	(0.277)
Urban Pop Den (1850), Group 5	0.465[a]	0.287[a]	2.650[a]
	(0.136)	(0.076)	(0.395)
City in Neighboring Regions (1850)	-0.029	-0.018	0.441[a]
	(0.031)	(0.014)	(0.086)
Urb. Pop. Den. in Neib. 1850	0.055	-0.005	0.019
	(0.044)	(0.005)	(0.015)
Squ. Urb. Pop. Den. in Neib. 1850	-0.002	0.000	-0.000
	(0.002)	(0.000)	(0.000)
Largest National City in 1850	0.247[a]	0.100[a]	0.713[a]
	(0.054)	(0.025)	(0.162)
Countries	91	127	127
Observations	1362	1848	1848
Adjusted R^2	0.23	0.19	0.34

Notes: For explanation of group variables see notes to Table 7. The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$..

Table 9: Urban Density in 1850
Dropping Regions with Highest Values

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05
Existence of a City (1850)	-0.028	0.025	0.364^a
	(0.040)	(0.017)	(0.080)
Urban Population Density 1850	3.055^b	1.951^b	6.446^c
	(1.260)	(0.629)	(3.876)
Sq Urban Pop. Den. 1850	-8.907	-5.539^c	-9.272
	(6.702)	(3.078)	(18.915)
City in Neighboring Regions (1850)	-0.041	-0.021	0.471^a
	(0.028)	(0.014)	(0.085)
Urb. Pop. Den. in Neib. 1850	0.144^c	-0.001	0.042^c
	(0.080)	(0.005)	(0.025)
Squ. Urb. Pop. Den. in Neib. 1850	-0.011^c	-0.000	-0.000^c
	(0.006)	(0.000)	(0.000)
Largest National City in 1850	0.215^a	0.078^b	0.668^a
	(0.059)	(0.032)	(0.192)
Countries	91	127	127
Observations	1288	1759	1759
Adjusted R^2	0.17	0.09	0.26

Notes: Regions that belong to groups 4 and 5 from Table 7 have been dropped. The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 10a: Investigating Persistence by Continent: Africa

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05
Existence of a City (1850)	-0.084	0.018	0.340
	(0.210)	(0.065)	(0.304)
Urban Population Density 1850	0.476	0.385^b	2.122^c
	(4.839)	(0.166)	(1.168)
Sq Urban Pop. Den. 1850	0.971	-0.064^c	-0.376
	(14.666)	(0.037)	(0.233)
City in Neighboring Regions (1850)	0.076	-0.090^c	-0.220
	(0.164)	(0.044)	(0.227)
Urb. Pop. Den. in Neib. 1850	-15.246^c	0.273	2.020
	(7.646)	(0.249)	(1.443)
Squ. Urb. Pop. Den. in Neib. 1850	91.945	-0.127	-0.606
	(54.654)	(0.131)	(0.695)
Largest National City in 1850	0.207	0.076	0.623
	(0.210)	(0.088)	(0.462)
Countries	13	25	25
Observations	122	331	331
Adjusted R^2	0.25	0.14	0.39

Table 10b: Investigating Persistence by Continent: Americas

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2001-05
Existence of a City (1850)	0.007	0.036	0.589^a
	(0.047)	(0.026)	(0.111)
Urban Population Density 1850	0.503^b	0.624^a	2.920^a
	(0.217)	(0.099)	(0.614)
Sq Urban Pop. Den. 1850	-0.050	-0.179^a	-0.535^b
	(0.054)	(0.023)	(0.163)
City in Neighboring Regions (1850)	-0.067	-0.001	0.435^b
	(0.058)	(0.022)	(0.121)
Urb. Pop. Den. in Neib. 1850	0.354	-0.068	2.079^b
	(0.236)	(0.128)	(0.970)
Squ. Urb. Pop. Den. in Neib. 1850	-0.045	0.059^c	-0.247
	(0.064)	(0.033)	(0.251)
Largest National City in 1850	0.265^b	0.113^c	1.007^b
	(0.097)	(0.058)	(0.340)
Countries	20	26	26
Observations	383	454	454
Adjusted R^2	0.27	0.21	0.46

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 10c: Investigating Persistence by Continent: Asia

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2001-05
Existence of a City (1850)	-0.022	0.108^a	0.564^b
	(0.041)	(0.026)	(0.164)
Urban Population Density 1850	0.395^b	0.143^c	1.527^b
	(0.165)	(0.079)	(0.513)
Sq Urban Pop. Den. 1850	-0.038^b	-0.014^c	-0.146^b
	(0.016)	(0.008)	(0.052)
City in Neighboring Regions (1850)	0.073	-0.040^b	0.226
	(0.053)	(0.020)	(0.159)
Urb. Pop. Den. in Neib. 1850	0.161	0.154^b	0.977^b
	(0.211)	(0.047)	(0.313)
Squ. Urb. Pop. Den. in Neib. 1850	-0.012	-0.011^b	-0.057^b
	(0.016)	(0.004)	(0.023)
Largest National City in 1850	0.255^b	0.137^b	1.111^a
	(0.116)	(0.049)	(0.309)
Countries	24	36	36
Observations	366	514	514
Adjusted R^2	0.21	0.27	0.43

Table 10d: Investigating Persistence by Continent: All of Europe and Western Europe

	(1)	(2)	(3)	(4)	(5)	(6)
		All Europe			Western Europe	
	Ln. GDP pc 2005	Urbanization 2000	Ln. Light Density, 2000-05	Ln. GDP pc 2005	Urbanization 2000	Ln. Light Density, 2000-05
Existence of a City (1850)	0.160^a	0.076^b	0.203^b	0.110^c	0.126^a	0.256^b
	(0.041)	(0.024)	(0.073)	(0.052)	(0.022)	(0.111)
Urban Population Density (1850)	0.097	0.124^a	0.861^a	0.066^b	0.105^a	0.582^a
	(0.075)	(0.034)	(0.227)	(0.029)	(0.022)	(0.097)
Sq Urban Pop. Den. 1850	-0.006	-0.009^b	-0.057^b	-0.004^c	-0.006^a	-0.034^a
	(0.005)	(0.003)	(0.018)	(0.002)	(0.001)	(0.007)
City in Neighboring Regions (1850)	-0.118^b	-0.017	0.243^b	-0.052	-0.007	0.462^a
	(0.048)	(0.033)	(0.074)	(0.042)	(0.072)	(0.106)
Urb. Pop. Den. in Neib. 1850	0.067^b	-0.008^b	0.004	0.074^b	-0.018^a	-0.016
	(0.023)	(0.004)	(0.004)	(0.025)	(0.004)	(0.011)
Squ. Urb. Pop. Den. in Neib. 1850	-0.002^b	0.000	-0.000	-0.002^c	0.000^b	0.000
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Largest National City in 1850	0.316^b	0.147^a	0.679^a	0.227^c	0.137^b	0.767^b
	(0.097)	(0.032)	(0.146)	(0.109)	(0.043)	(0.205)
Countries	34	39	39	16	17	17
Observations	492	541	541	202	208	208
Adjusted R^2	0.42	0.30	0.57	0.35	0.50	0.65

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 10e: Investigating Persistence in Selected Regions: Neo-Europes (Australia, Canada and USA)

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2001-05
Existence of a City (1850)	0.042	0.063^c	0.446
	(0.029)	(0.018)	(0.210)
Urban Population Density 1850	0.643	1.203^c	3.181
	(0.233)	(0.361)	(2.408)
Sq Urban Pop. Den. 1850	-0.064	-0.332	-0.576
	(0.072)	(0.118)	(0.790)
City in Neighboring Regions (1850)	0.030^b	-0.004	0.192^c
	(0.003)	(0.006)	(0.064)
Urb. Pop. Den. in Neib. 1850	0.854	0.533	5.869
	(0.381)	(0.186)	(4.936)
Squ. Urb. Pop. Den. in Neib. 1850	-0.289	-0.838^c	-5.533
	(0.369)	(0.208)	(5.402)
Largest National City in 1850	-0.227^c	-0.235	0.322
	(0.065)	(0.103)	(0.649)
Countries	3	3	3
Observations	70	71	71
Adjusted R^2	0.56	0.56	0.64

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 11: Persistence within Colonized and Non Colonized Countries

	Colonized Countries					
	Panel A Fixed-effects			Panel B Random Effects		
	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05	(4) Ln. GDP pc 2005	(5) Urbanization 2000	(6) Ln. Light Density, 2000-05
Existence of a City (1850)	-0.037	0.041^c	0.617^a	-0.028	0.054^b	0.701^a
	(0.036)	(0.022)	(0.092)	(0.037)	(0.022)	(0.093)
Urban Population Density (1850)	0.650^b	0.453^a	3.456^a	0.681^b	0.462^a	3.564^a
	(0.215)	(0.107)	(0.480)	(0.213)	(0.103)	(0.484)
Sq Urban Pop. Den. 1850	-0.103^c	-0.107^b	-0.692^a	-0.112^b	-0.107^b	-0.710^a
	(0.054)	(0.042)	(0.111)	(0.054)	(0.041)	(0.113)
City in Neighboring Regions (1850)	-0.047	-0.016	0.347^b	-0.038	-0.010	0.421^a
	(0.044)	(0.017)	(0.111)	(0.044)	(0.016)	(0.107)
Urb. Pop. Den. in Neib. (1850)	0.386	-0.068	1.483^c	0.377	-0.075	1.548^b
	(0.232)	(0.078)	(0.745)	(0.232)	(0.080)	(0.730)
Squ. Urb. Pop. Den. in Neib. 1850	-0.038	0.067^b	-0.101	-0.034	0.069^b	-0.103
	(0.069)	(0.026)	(0.240)	(0.070)	(0.027)	(0.236)
Largest National City in 1850	0.220^b	0.095^b	0.635^b	0.204^b	0.081^b	0.544^b
	(0.069)	(0.040)	(0.239)	(0.069)	(0.041)	(0.231)
Log Population Density, 1500				-0.301^a	-0.054^a	0.135
				(0.078)	(0.013)	(0.141)
Countries	42	63	63	42	63	63
Observations	639	970	970	639	970	970
Adjusted R^2	0.19	0.15	0.40	.	.	.

	Panel C: Non-Colonized countries		
	(1) Ln. GDP pc 2005	(2) Urbanization 2000	(3) Ln. Light Density 2000-05
Existence of a City (1850)	0.099^b	0.092^a	0.254^b
	(0.039)	(0.025)	(0.099)
Urban Population Density (1850)	0.097	0.093^b	0.824^a
	(0.058)	(0.039)	(0.198)
Sq Urban Pop. Den. (1850)	-0.006	-0.007^c	-0.058^a
	(0.004)	(0.003)	(0.017)
City in Neighboring Regions (1850)	-0.035	-0.031	0.222^b
	(0.042)	(0.021)	(0.087)
Urb. Pop. Den. in Neib. 0 (1850)	0.032	-0.006	0.012
	(0.021)	(0.004)	(0.010)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001	0.000	-0.000
	(0.001)	(0.000)	(0.000)
Largest National City in 1850	0.305^a	0.164^a	1.047^a
	(0.082)	(0.032)	(0.178)
Countries	49	64	64
Observations	723	878	878
Adjusted R^2	0.35	0.25	0.36

Note: Colonized countries here refer to those subject to European colonization. The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in mm, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100meters), ruggedness, proximity to the coast, and proximity to a river. Fixed-effects estimates include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 12: Impact of Urbanization in 1850 on Contemporary Education, Culture, Institutions, and Infrastructure

	(1) Yrs of Educ.	(2) Trust	(3) Inf. Payments	(4) Acc.to Finance	(5) Electricity.	(6) Power Lines	(7) Travel time
Existence of a City (1850)	0.134[c]	0.006	0.105	0.044	0.276[b]	0.184[a]	-0.247[a]
	(0.070)	(0.010)	(0.152)	(0.029)	(0.135)	(0.037)	(0.049)
Urban Population Density 1850	0.346[b]	-0.016	0.071	-0.020	-0.101	0.058	-0.655[a]
	(0.133)	(0.015)	(0.125)	(0.023)	(0.128)	(0.129)	(0.114)
Sq Urban Pop. Den. 1850	-0.025[b]	0.001	0.004	0.002	0.003	-0.013	0.042[a]
	(0.009)	(0.001)	(0.013)	(0.002)	(0.013)	(0.010)	(0.010)
City in Neighboring Regions (1850)	-0.038	-0.018[b]	0.115	-0.021	-0.122	0.205[a]	-0.219[a]
	(0.053)	(0.009)	(0.115)	(0.023)	(0.139)	(0.053)	(0.046)
Urb. Pop. Den. in Neib. 1850	0.403[b]	0.001	-1.530[b]	0.198	-0.141	0.106	-0.251[b]
	(0.149)	(0.009)	(0.556)	(0.139)	(0.483)	(0.114)	(0.101)
Squ. Urb. Pop. Den. in Neib. 1850	-0.021[b]	-0.000	0.476[b]	-0.034	0.017	-0.005	0.012[b]
	(0.008)	(0.000)	(0.145)	(0.043)	(0.125)	(0.005)	(0.004)
Largest National City in 1850	0.801[a]	-0.003	-0.011	0.008	0.038	0.144	-0.076
	(0.113)	(0.012)	(0.088)	(0.021)	(0.100)	(0.094)	(0.082)
Mean Temperature (Cel.)	-0.035[b]	-0.002	-0.030	0.002	0.035	0.002	-0.047[b]
	(0.015)	(0.003)	(0.023)	(0.004)	(0.025)	(0.013)	(0.017)
Mean Land Suitability	-0.080	0.007	0.470	-0.019	0.268	0.580[a]	-0.778[a]
	(0.218)	(0.026)	(0.546)	(0.046)	(0.293)	(0.115)	(0.121)
Mean Altitude (100m)	-0.025	-0.000	0.023	0.002	0.027[c]	0.014[c]	-0.027[b]
	(0.016)	(0.002)	(0.019)	(0.003)	(0.016)	(0.008)	(0.013)
Mean Rainfall (meters)	-0.111	-0.014	-0.059	0.047[c]	0.104	-0.020	0.395[a]
	(0.112)	(0.015)	(0.199)	(0.026)	(0.176)	(0.083)	(0.096)
Mean Ruggedness (100m)	0.022	-0.003	-0.072	0.016	-0.065	-0.004	-0.001
	(0.038)	(0.006)	(0.071)	(0.011)	(0.078)	(0.031)	(0.031)
SD Temperature	0.002	0.000	-0.004	-0.000	-0.000	-0.008[b]	0.016[a]
	(0.003)	(0.001)	(0.006)	(0.001)	(0.005)	(0.003)	(0.003)
SD Land Suitability	-0.761	0.129[c]	-0.000	-0.095	-0.269	-0.287	-0.076
	(0.558)	(0.071)	(0.921)	(0.129)	(0.526)	(0.285)	(0.246)
SD Rainfall	-0.000	-0.000	0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Proximity to the Coast	1.118[b]	0.073	0.137	-0.049	0.775[c]	1.078[a]	-1.607[a]
	(0.416)	(0.056)	(0.904)	(0.132)	(0.432)	(0.270)	(0.352)
Proximity to Rivers	0.869[b]	0.082	1.543[c]	-0.031	0.221	0.711[b]	-0.964[b]
	(0.425)	(0.060)	(0.795)	(0.116)	(0.551)	(0.300)	(0.370)
Countries	90	61	65	68	63	92	92
Observations	1336	654	330	371	202	1371	1371
Adjusted R^2	0.21	0.00	0.01	0.01	0.03	0.17	0.51

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. All regressions include country fixed-effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table 13: Exploring Effects of 1500 Urbanization

	(1) All Countries	(2) Colonized Countries	(3) Non-Colonized Countries
Existence of a City (1500)	0.005	-0.218[a]	0.084[b]
	(0.048)	(0.052)	(0.040)
Existence of a City (1850)	0.030	-0.016	0.088[b]
	(0.029)	(0.034)	(0.037)
Urban Population Density 1500	0.182	1.280[a]	0.211[c]
	(0.130)	(0.221)	(0.108)
Urban Population Density 1850	0.015	0.299[a]	-0.001
	(0.020)	(0.076)	(0.012)
City in Neighboring Regions (1500)	-0.016	-0.148	0.008
	(0.043)	(0.104)	(0.035)
City in Neighboring Regions (1850)	-0.031	-0.007	-0.045
	(0.036)	(0.052)	(0.042)
Urb. Pop. Den. in Neib. 1500	-0.123	-0.341	-0.059
	(0.102)	(0.247)	(0.062)
Urb. Pop. Den. in Neib. 1850	0.052	0.237[a]	0.029
	(0.043)	(0.048)	(0.029)
Largest National City in 1500	0.112	0.165	0.077
	(0.070)	(0.102)	(0.083)
Largest National City in 1850	0.297[a]	0.238[b]	0.293[a]
	(0.055)	(0.070)	(0.082)
Countries	92	43	49
Observations	1371	648	723
Adjusted R^2	0.22	0.21	0.36

Note: Colonized here refers only to European colonization. The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

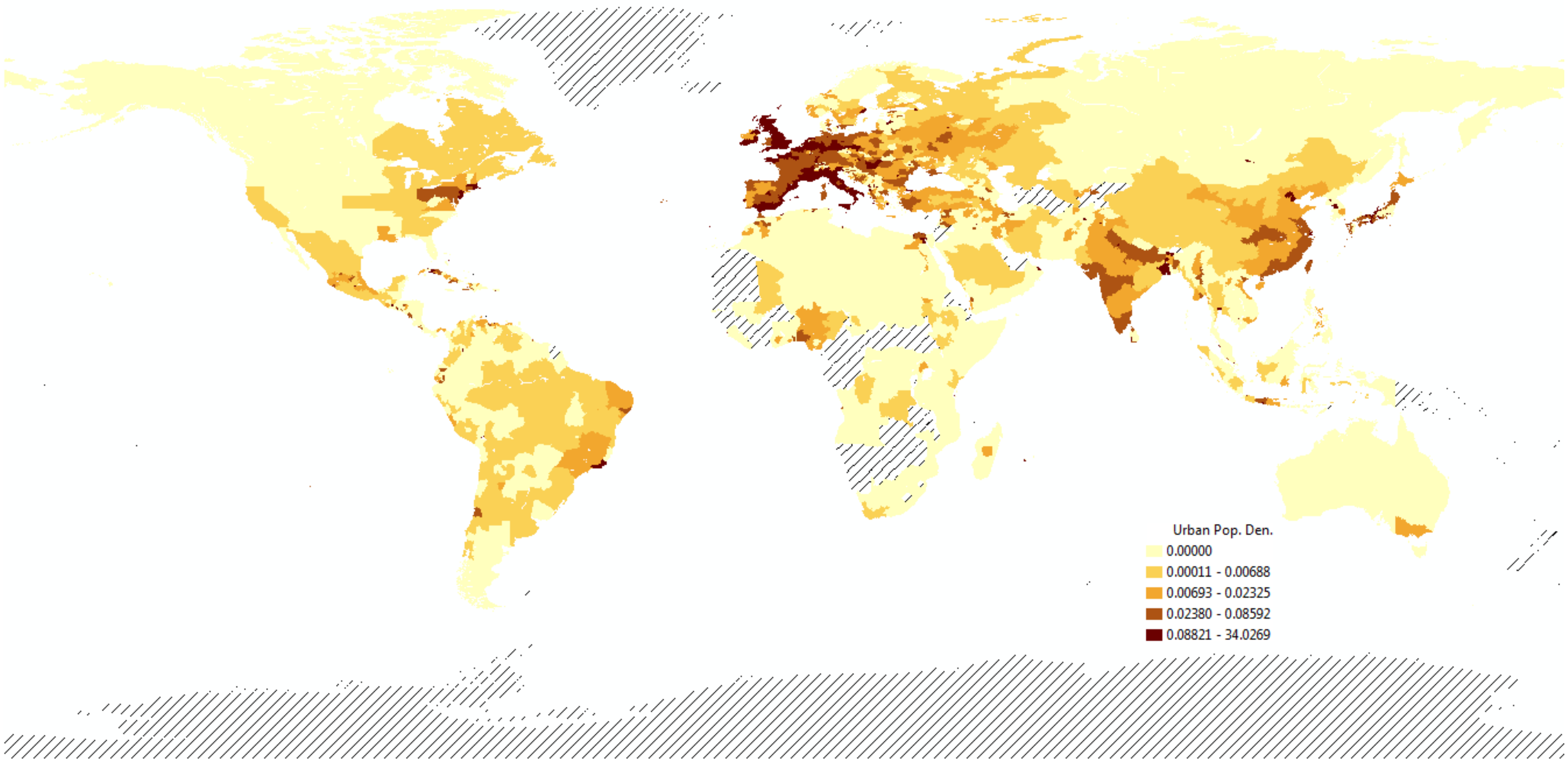

Figure 1: Urban Population Density, 1850

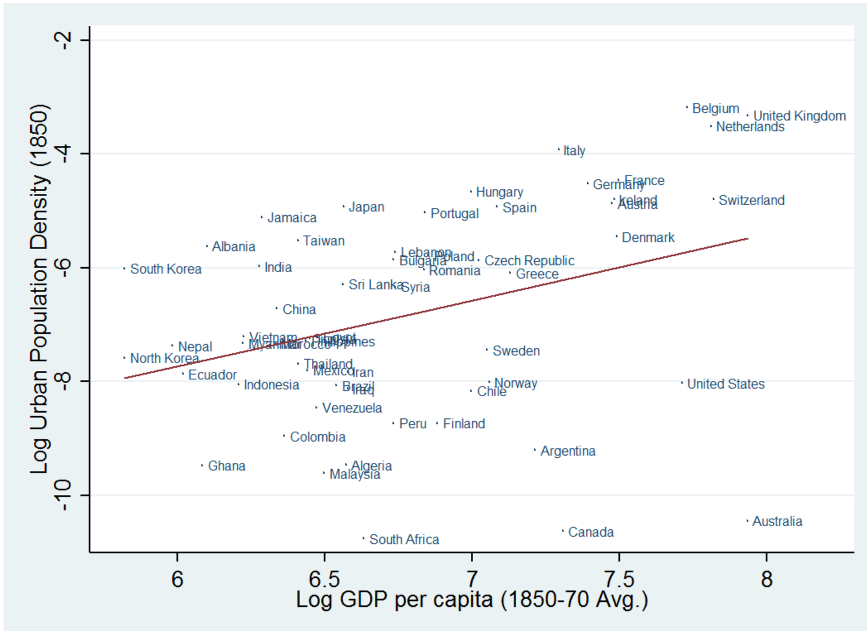

Figure 2: Urban Population Density vs GDP pc, c1850
n=56, r=0.3

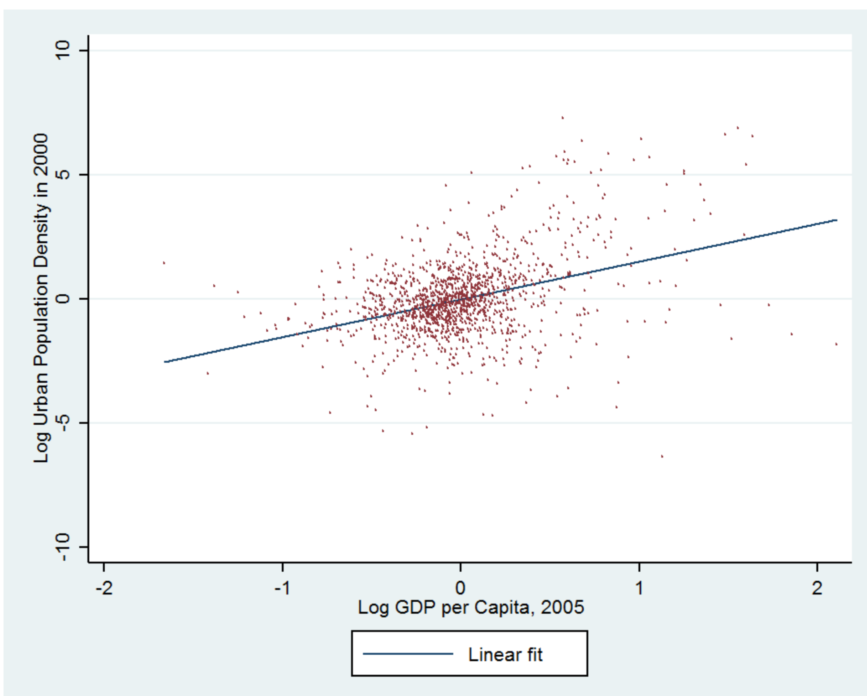

Figure 3: Regional Urban Population Density (2000) vs GDP pc. (2005)
n=1380, r=0.37

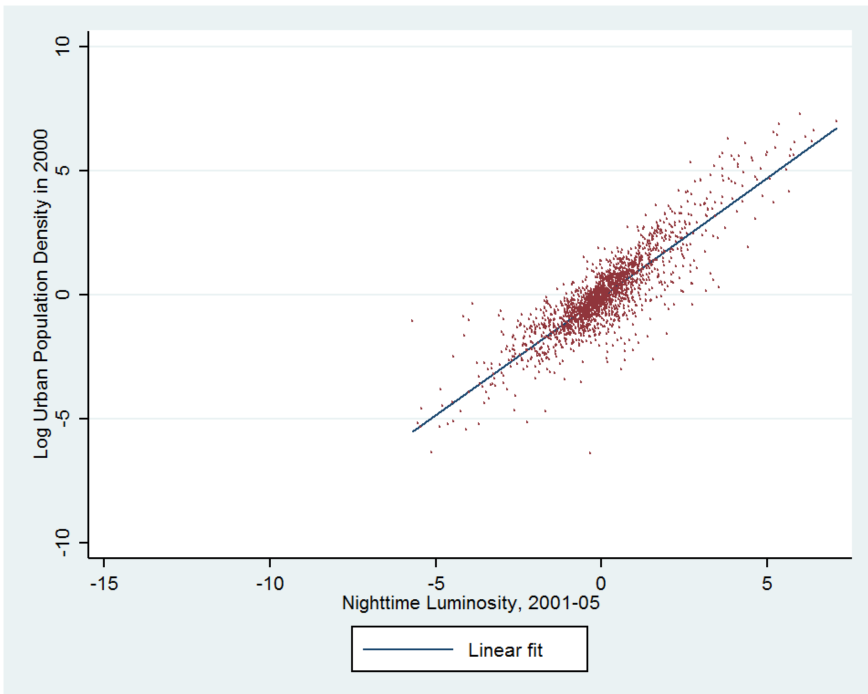

Figure 4: Regional Urban Population Density (2000) vs Nighttime Light Density (2000-2005)
n=1901, r=0.8

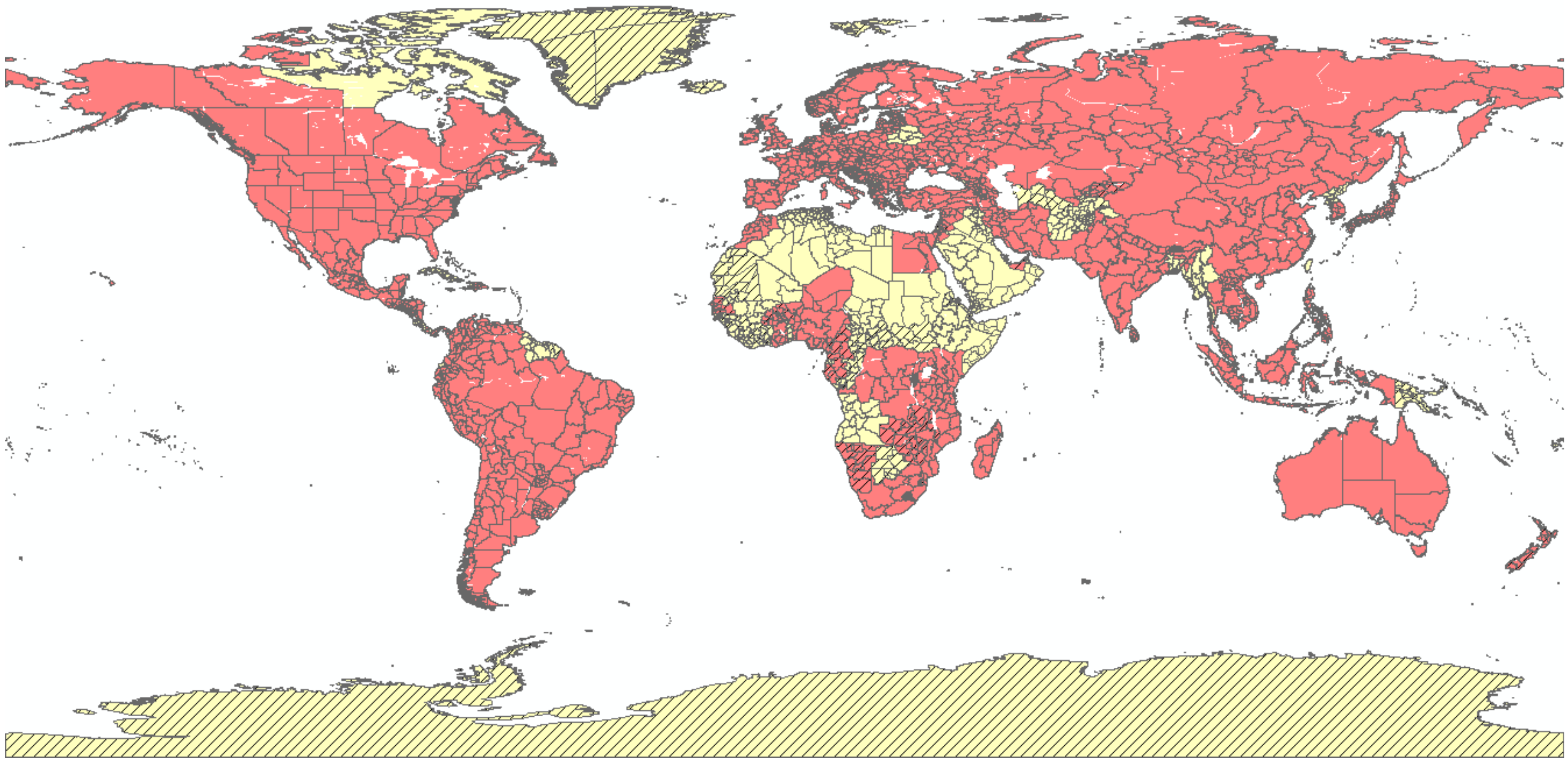

Figure 5: Subdivisions across The World

Note: Shaded areas present regions whose income per capita in 2005 is available. Striped areas consist of countries that do not appear in our data for cities in 1850.

A Appendix

A.1 Variable Descriptions

Table A.1: Variables, Descriptions, and Sources

Variable	Description and Sources
	1850 Urbanization Measures:
Existence of a City (1850)	A dummy indicating regions in which at least a locality with population greater than 5,000 existed in 1850. To generate this variable, we load coordinates of the localities in 1850 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. We code 1 for regions contain at least one of these coordinates; 0, otherwise.
Urban Population Density in 1850	Defined as 100 urban inhabitants per square kilometer of the total land area of the region. To generate this variable, we load localities in 1850 with population greater than 5,000 and the worldwide regions' digital map derived from the Database of Global Administrative Areas.
Existence of a City in Neighboring Regions (1850)	A dummy identifying one or more year 1850 cities existed within 25 miles geodesic distance away from the regions. To generate this variable, we load coordinates of the localities in 1850 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. We code 1 for regions if outside the regions within 25 miles away from the regions' boundaries there exists at least one of these coordinates; 0, otherwise. In TableA.2, we extend the distance to 100 miles.
Urban Pop. Den. in Neighboring Regions, 1850	Defined as 100 urban inhabitants within 25 miles of surrounding area divided by the land area of the region (and not the surrounding region). To generate this variable, we load localities in 1850 with population greater than 5,000 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. In Table 6 the population of neighboring cities is raised to 20,000 and 100,000.
	Sources: Urban settlement data from Chandler (1987), Bairoch (1988), Eggimann (1994), and Rozenblat. Land Area from CIESIN (2016), subnational maps from GADM database of Global Administrative Areas (v2).
	Dependent Variables:
GDP per capita	Regional income per capita in PPP constant 2005 international dollars in 2005. Source: Gennaioli et al. (2013).
Urbaniza- tion (2000)	Urban Population Divided by Total Population of Each Region (2000).
	Source : Balk (2006) and CIESIN et al (2011) for urban settlement points and CIESIN and CIAT (2005) for population.

Continued on next page...

Table A.1 – Continued

Ln(Avg. Nighttime Light Density), 2000 & 2005 avg	The logarithm of average nighttime light intensity yearly averaged for the years 2000 and 2005-06. To produce the regional numbers, we average the radiance calibrated night lights data for 2000 (after applying the inter-satellite calibration multiplier) and 2005/6. We take the ratio of total light intensity in each region to the land area of the region (GPWv4). Sources: Light Data: https://www.ngdc.noaa.gov/eog/dmsp/download_radcal.html
	Baseline Geographic controls:
Temperature	The raw data is Average Temperature during 1950 - 2000 in Celsius. To produce the regional numbers, we load the global temperature grid and take weighted averages based on the land area of each grid (GPWv4). Standard deviation is a simple calculation based on cells within a region. Source: Global Climate Data (Worldclim.org).
Land Suitability	An index of the suitability for agriculture based on temperature and soil quality measurements. We take the weighted average of the index within regions. Weights are based on land area of each cell. Standard deviation is a simple calculation based on cells within a region. Due to the coarse nature of the data, we resampled the data to match the resolution for land area. Source: Ramankutty et al (2002)
Altitude	Weighted (by land area) Average of Altitude in regions (in 100 meters). Source: Global Climate Data (Worldclim.org).
Ruggedness	Weighted (by land area) Average terrain ruggedness in regions (in 100 meters). Source: Nunn and Puga (2012).
Rainfall	Weighted (by land area) Average precipitation in regions during 1950 - 2000 in meter. Standard deviation is a simple calculation based on cells within a region. Source: Global Climate Data (Worldclim.org).
Proximity to the Coast	The reciprocal of 1 plus the distance of regions' centroid to the nearest coastlines in 1,000 kilometers. To produce the numbers, we load the world coastline grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. We generate regions' median centroid and keep coordinates of them. We calculate the distance of the centroid to nearest coastlines. Source: National Geophysical Data Center (NOAA), Global Self-consistent Hierarchical High-resolution Geography v2.3.
Proximity to a River	The reciprocal of 1 plus the distance of regions' centroid to the nearest rivers in 1,000 kilometers. To produce the numbers, we load the world river grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. We generate regions' median centroid and keep coordinates of them. We calculate the distance of the centroid to nearest rivers. Source: National Geophysical Data Center (NOAA), Global Self-consistent Hierarchical High-resolution Geography v2.3.

Continued on next page...

Table A.1 – Continued

Variables Capturing Potential Mechanisms

Years of Education	Average years of schooling beyond primary school for those who are 15 years old and older.
Trust in Others	Percent of respondents who think most people can be trusted.
Informal Payments	Percent of sales goes as informal payments to public officials for activities such as customs, taxes, licenses, etc, averaged across all respondents within regions.
Access to Financing	Percent of respondents think that access to financing is at least a moderate obstacle to business.
Ln(Days without Electricity)	The logarithm of 1 plus the regional average of days with no electricity in the past year reported by respondents.
Ln(Power Line Density)	The logarithm of 1 plus the length in kilometers of power lines per 10 square kilometers in 2007.
Ln(Travel Time)	The logarithm of the regional average of estimated travel time in minutes to the neatest city with population greater than 50,000 in 2000.
	Source: Gennaioli et al. (2013).

A.1.1 Definition of a Region

In this subsection we briefly review how we define regions. We match Gennaioli et al.'s (2013) regions with the Database of Global Administrative Areas Map version 2 (GADMv2). For regions that are not included in Gennaioli et al. (2013), subdivisions at the largest disaggregated level provided in GADMv2 are used. Most of Gennaioli et al.'s (2013) regions are the first-level administrative divisions, and other regions require combining two or more such subdivisions according to at what aggregate level a variable is available. We find those regions' boundaries in the GADMv2. Among Gennaioli et al.'s 1,537 regions, there are 17 regions whose boundaries are not available at the most disaggregated level of the GADMv2. We aggregated the 17 regions into 8 bigger ones that can be found in the GADMv2. The 8 regions (with regions being aggregated displayed after colon) are Copenhagen: Copenhagen and Frederiksberg and Copenhagen county, Daugavpils: Daugavpils city and Daugavpils district, Jelgava: Jelgava city and Jelgava district, Liepaja: Liepaja city and Liepaja district, Rezekne: Rezekne city and Rezekne district, Riga: Riga city, Jurmala city, and Riga district, Ventspils: Ventspils city and Ventspils district, and Selangor: Selangor and Wilayah Persekutuan. Data for the 8 aggregated regions are calculated as the population-weighted average of the regions being combined. Finally, we exclude regions in countries that do not have a single region with settlement data.

A.1.2 Urban Population

We include any location that has a recorded population of 5,000 or more in 1850 from our sources. In an effort to enlarge our sample, we also include locations with records from 1825 and 1875 but absent in 1850. Only Melbourne is therefore considered as a city in 1850 though its estimated population, 222,000 according to Rozenblat's estimates, is only available in 1875. When all of our data sources taken together, we have 3,044 settlements spanning 141 contemporary countries in 1850, of which 2,832 are with a population of 5,000 or greater. However, a city is considered identified only if we are able to confirm in which region the city is located. There are another 29 settlements in 1850 that fit the definition of city but are excluded because their locations are unidentified. These 2,803 settlements are from 772 regions. Among these regions, there are six city states - Gibraltar, Guernsey, Hong Kong, Macao, Malta, and Singapore, which are dropped in our study. We end up with 766 regions from 128 countries in our whole sample had urban areas in 1850.

A.2 Additional Regressions

Table A.2: Robustness to Variation in Distance to Neighboring Cities
100 Miles

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05
Existence of a City (1850)	0.024	0.071^{a}	0.555^{a}
	(0.030)	(0.016)	(0.077)
Urban Population Density 1850	0.191^{b}	0.110^{b}	0.964^{a}
	(0.063)	(0.033)	(0.163)
Sq Urban Pop. Den. 1850	-0.013^{b}	-0.008^{b}	-0.069^{a}
	(0.005)	(0.003)	(0.014)
City in Neighboring Regions (1850)	-0.044	-0.028	0.480^{b}
	(0.062)	(0.024)	(0.191)
Urb. Pop. Den. in Neib. 1850	0.005	-0.001	-0.001
	(0.008)	(0.001)	(0.001)
Squ. Urb. Pop. Den. in Neib. 1850	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)
Largest National City in 1850	0.263^{a}	0.132^{a}	0.879^{a}
	(0.057)	(0.026)	(0.152)
Countries	91	127	127
Observations	1362	1848	1848
Adjusted R^2	0.22	0.17	0.34

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in mm, Temperatur in Celsius, Agricultural Land suitability, and Average values of altitude (100meters), ruggedness, proximity to the coast, and proximity to a river. Fixed-effects estimates include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table A.3: Regressions on Urbanization in 1850
Only For Regions with Positive Urban Population

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05
Urban Population Density 1850	0.225^b	0.130^a	1.270^a
	(0.069)	(0.032)	(0.185)
Sq Urban Pop. Den. 1850	-0.015^b	-0.008^b	-0.083^a
	(0.005)	(0.002)	(0.016)
City in Neighboring Regions (1850)	0.023	0.016	0.462^a
	(0.038)	(0.017)	(0.116)
Urb. Pop. Den. in Neib. 1850	0.034	-0.016	-0.089
	(0.021)	(0.012)	(0.058)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001	0.000	0.001
	(0.001)	(0.000)	(0.001)
Largest National City in 1850	0.248^a	0.131^a	0.753^a
	(0.064)	(0.026)	(0.129)
Countries	87	117	117
Observations	661	745	745
Adjusted R^2	0.27	0.31	0.52

Table A.4: Log Urban Density

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05
Ln(Urban Pop. Den.) 1850	0.021^a	0.023^a	0.178^a
	(0.005)	(0.004)	(0.021)
Ln(Neighb. Urb Pop. Den.) 1850	-0.000	-0.002	0.090^a
	(0.006)	(0.003)	(0.017)
Largest National City in 1850	0.278^a	0.112^a	0.818^a
	(0.053)	(0.025)	(0.156)
Countries	91	127	127
Observations	1362	1848	1848
Adjusted R^2	0.21	0.17	0.34

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table A.5: Urban Density in 1850
Excluding Regions with (a)Zero Urban Populations and (b) Highest Values.

	(1) Ln. GDP pc 2005	(2) Urbanization, 2000	(3) Ln. Light Density, 2000-05
Urban Population Density 1850	2.693^{c}	2.384^{a}	18.311^{a}
	(1.397)	(0.659)	(3.073)
Sq Urban Pop. Den. 1850	-7.262	-6.844^{b}	-58.982^{a}
	(6.779)	(3.051)	(15.528)
City in Neighboring Regions (1850)	0.016	-0.002	0.343^{a}
	(0.030)	(0.017)	(0.080)
Urb. Pop. Den. in Neib. 1850	0.061	-0.065	0.273^{b}
	(0.059)	(0.058)	(0.101)
Squ. Urb. Pop. Den. in Neib. 1850	-0.003	0.008	-0.042^{b}
	(0.008)	(0.007)	(0.014)
Largest National City in 1850	0.184^{b}	0.075^{b}	0.436^{b}
	(0.069)	(0.034)	(0.136)
Countries	85	108	108
Observations	587	656	656
Adjusted R^2	0.13	0.16	0.40

Notes: Regions with no recorded urban settlements of 5000 or more have been dropped. Regions that belong to groups 4 and 5 from Table 7 have also been dropped. Robust standard errors clustered at the country level are shown in parentheses. Geographic controls included are Average and Standard Deviations of : Rainfall in meters, Temperature in Celsius, Agricultural Land suitability, and Average values of altitude (100 meters), ruggedness (100 meters), proximity to the coast, and proximity to a river. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

Table A.6: Adjusted Urbanization and Persistence

	(1) Log. Population 2000	(2) Ln GDP pc 2005	(3) Urbanization 2000	(4) Ln Light Density 2000-05
Existence of a City (1850)		0.015	0.068^{a}	0.589^{a}
		(0.027)	(0.018)	(0.078)
Adjusted Urbanization, 1850		0.230^{b}	0.186^{a}	1.533^{a}
		(0.106)	(0.034)	(0.284)
Largest National City in 1850		0.312^{a}	0.150^{a}	0.954^{a}
		(0.057)	(0.027)	(0.152)
Mean Temperature (Cel.)	0.042^{b}	-0.034^{b}	-0.001	0.036
	(0.017)	(0.011)	(0.003)	(0.029)
Mean Land Suitability	1.285^{a}	-0.112^{c}	0.019	1.750^{a}
	(0.192)	(0.064)	(0.033)	(0.274)
Mean Altitude (100m)	0.013	-0.020^{b}	-0.000	0.015
	(0.011)	(0.007)	(0.002)	(0.024)
Mean Rainfall (meters)	-0.356^{b}	-0.062	-0.031	-0.362^{c}
	(0.149)	(0.050)	(0.025)	(0.201)
Mean Ruggedness (100m)	-0.097^{b}	-0.043^{c}	-0.021^{b}	-0.029
	(0.041)	(0.023)	(0.010)	(0.066)
SD Temperature	0.006^{c}	0.003^{b}	-0.000	-0.008
	(0.003)	(0.001)	(0.001)	(0.007)
SD Land Suitability	2.215^{a}	-0.660^{a}	-0.308^{a}	-0.832
	(0.367)	(0.160)	(0.088)	(0.739)
SD Rainfall	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Proximity to the Coast	2.353^{a}	0.798^{a}	0.219^{b}	5.052^{a}
	(0.510)	(0.189)	(0.071)	(0.756)
Proximity to Rivers	1.781^{b}	0.416^{b}	0.143^{b}	2.707^{a}
	(0.570)	(0.190)	(0.064)	(0.732)
Log Land Area (sq. km.)	0.259^{a}			
	(0.063)			
Countries	127	91	127	127
Observations	1848	1362	1848	1848
Adjusted R^2	0.22	0.21	0.16	0.31

Note: Adjusted Urbanization is urban population in 1850 divided by predicted population of 2000. The latter variable is obtained from the fitted values in column (1). Robust standard errors clustered at the country level are shown in parentheses. All regressions include country fixed effects. a $p < 0.01$, b $p < 0.05$, c $p < 0.10$.

A.3 Countries and Regions

Table A.7: Number of Regions by Country

Code	Country	Sample Regions	No. of Regions year 1850 localities inhabitants over (5,000)	Population in 1,000 year 1850 localities inhabitants over (5,000)	No. of Regions year 1850 localities inhabitants over (20,000)	(50,000)	(100,000)
CHN	China	32	31	11243	31	24	15
GBR	United Kingdom	12	12	8674	12	12	6
IND	India	35	22	7909	16	13	10
FRA	France	22	22	6314	22	12	5
ITA	Italy	20	20	5848	13	10	8
DEU	Germany	16	16	3840	15	8	3
ESP	Spain	19	16	3633	13	5	3
USA	United States	51	26	2981	25	9	6
JPN	Japan	47	32	2670	32	10	4
BRA	Brazil	27	19	2628	15	3	3
RUS	Russia	80	47	2537	22	5	2
TUR	Turkey	12	12	1807	12	4	2
BEL	Belgium	11	11	1264	8	6	2
NGA	Nigeria	7	5	1188	5	3	0
UKR	Ukraine	27	22	1064	11	3	0
NLD	Netherlands	14	12	1029	8	2	2
POL	Poland	16	16	945	9	3	2
HUN	Hungary	7	7	867	3	1	1
MEX	Mexico	32	27	795	14	3	1
EGY	Egypt	4	3	715	3	1	1
IRN	Iran	30	15	642	14	4	0
AUT	Austria	9	9	630	3	2	1
IDN	Indonesia	33	12	601	8	3	1
PRT	Portugal	7	7	594	3	2	1
IRL	Ireland	2	2	565	1	1	1
ROU	Romania	8	8	564	8	2	0
CUB	Cuba	15	10	496	7	2	1
MMR	Myanmar	14	7	436	5	3	1
PAK	Pakistan	8	4	375	3	2	0
SYR	Syria	14	4	330	4	2	1
BGR	Bulgaria	6	6	318	6	0	0
CHE	Switzerland	26	14	318	4	0	0
GRC	Greece	14	9	295	3	1	0
ARG	Argentina	24	13	276	3	1	0
MAR	Morocco	15	5	270	5	3	0
UZB	Uzbekistan	5	5	247	5	3	0
SWE	Sweden	8	6	246	2	1	0
KOR	South Korea	7	2	241	2	1	1
VNM	Vietnam	8	3	240	3	3	0
THA	Thailand	7	3	234	3	1	1
AUS	Australia	11	1	222	1	1	1
CAN	Canada	13	5	221	5	0	0
CZE	Czech Republic	8	5	221	2	1	1
CHL	Chile	13	8	210	2	1	0
PER	Peru	25	10	207	4	1	0
PHL	Philippines	17	4	200	2	1	1

Continued on next page...

Table A.7 – Continued

SAU	Saudi Arabia	13	6	193	5	0	0
VEN	Venezuela	24	12	193	4	0	0
DNK	Denmark	14	5	180	1	1	1
DZA	Algeria	48	7	179	3	1	0
SRB	Serbia	19	3	170	1	0	0
AFG	Afghanistan	32	5	164	4	1	0
SVK	Slovakia	8	7	154	1	0	0
BGD	Bangladesh	6	3	153	3	1	0
BLR	Belarus	6	6	146	2	0	0
COL	Colombia	33	13	145	1	0	0
TWN	Taiwan	4	2	145	2	1	0
IRQ	Iraq	18	4	130	3	1	0
YEM	Yemen	21	4	130	4	0	0
LKA	Sri Lanka	9	3	120	3	1	0
BOL	Bolivia	9	6	116	3	0	0
TUN	Tunisia	24	1	110	1	1	1
NOR	Norway	19	7	103	2	0	0
ALB	Albania	12	8	102	2	0	0
ECU	Ecuador	22	4	97	3	0	0
LVA	Latvia	26	3	94	2	1	0
NPL	Nepal	5	1	90	1	0	0
MDA	Moldova	5	3	86	1	1	0
BIH	Bosnia - Herzegovina	3	2	85	1	1	0
MLI	Mali	9	3	84	2	0	0
LTU	Lithuania	10	2	71	1	1	0
JAM	Jamaica	14	1	66	1	0	0
NER	Niger	8	2	66	2	0	0
PRK	North Korea	14	1	62	1	1	0
COD	Dem. Rep. Congo	11	2	60	2	0	0
MNG	Mongolia	22	1	60	1	0	0
OMN	Oman	8	1	60	1	1	0
TZA	Tanzania	26	1	60	1	1	0
NIC	Nicaragua	18	3	57	1	0	0
HRV	Croatia	20	5	56	0	0	0
SLV	El Salvador	14	3	56	1	0	0
GTM	Guatemala	8	3	54	1	0	0
FIN	Finland	5	3	50	1	0	0
MDG	Madagascar	6	1	50	1	1	0
MUS	Mauritius	12	1	49	1	0	0
REU	Reunion	4	3	48	0	0	0
HTI	Haiti	10	4	45	1	0	0
UGA	Uganda	6	1	45	1	0	0
PRY	Paraguay	18	1	44	1	0	0
GEO	Georgia	12	2	42	1	0	0
EST	Estonia	16	3	40	1	0	0
ETH	Ethiopia	11	2	39	1	0	0
AZE	Azerbaijan	11	2	36	1	0	0
BRN	Brunei	4	1	36	1	0	0
PRI	Puerto Rico	79	2	35	1	0	0
LBN	Lebanon	6	2	34	1	0	0
BEN	Benin	12	2	33	0	0	0
ARM	Armenia	12	1	30	1	0	0
KHM	Cambodia	15	1	30	1	0	0
KWT	Kuwait	5	1	30	1	0	0

Continued on next page...

Table A.7 – Continued

MAC	Macao	1	1	29	1	0	0
MTQ	Martinique	4	2	29	1	0	0
SDN	Sudan	6	2	29	1	0	0
HND	Honduras	18	2	26	0	0	0
MKD	Macedonia	8	2	26	1	0	0
ZAF	South Africa	10	1	26	1	0	0
DOM	Dominican Republic	9	2	24	0	0	0
SVN	Slovenia	12	2	24	0	0	0
LUX	Luxembourg	3	1	22	1	0	0
MYS	Malaysia	13	2	22	0	0	0
BRB	Barbados	11	1	20	1	0	0
TCD	Chad	18	1	20	1	0	0
CRI	Costa Rica	7	1	20	1	0	0
LBY	Libya	32	1	20	1	0	0
SUR	Suriname	10	1	20	1	0	0
KO-	Kosovo	7	1	19	0	0	0
GHA	Ghana	10	1	18	0	0	0
TTO	Trinidad - Tobago	14	1	18	0	0	0
GUY	Guyana	10	1	17	0	0	0
LAO	Laos	18	1	15	0	0	0
AGO	Angola	18	1	14	0	0	0
PAN	Panama	12	1	12	0	0	0
SLE	Sierra Leone	4	1	11	0	0	0
KAZ	Kazakhstan	6	1	10	0	0	0
BHS	Bahamas	32	1	8	0	0	0
SMR	San Marino	9	1	7	0	0	0
KEN	Kenya	8	1	6	0	0	0
TGO	Togo	5	1	6	0	0	0
BHR	Bahrain	5	1	5	0	0	0
BLZ	Belize	6	0	0	0	0	0
GMB	Gambia	6	0	0	0	0	0
LBR	Liberia	15	0	0	0	0	0
MNE	Montenegro	21	0	0	0	0	0
MOZ	Mozambique	10	0	0	0	0	0
SOM	Somalia	18	0	0	0	0	0
URY	Uruguay	19	0	0	0	0	0
Africa (28 countries):		363	49	3149	32	11	2
Americas (31 countries):		601	183	8921	97	20	11
Asia (36 countries):		528	205	28878	178	81	37
Europe (39 countries):		555	334	41352	187	83	39
Oceania (1 country):		11	1	222	1	1	1
World Total (135 countries):		2058	772	82523	495	196	90

Note: The 3rd column shows the total number of regions included in each country. The 4th and 5th columns are the number of regions that had settlements in 1850 and total population across these settlements, respectively, focusing on settlements with population higher than 5,000. The rest of columns display the number of settlements based on a minimum population of 20,000, 50,000, and 100,000, respectively. The last several rows show the information at the continent level. Countries are sorted according to the year 1850 population in settlements with population greater than 5,000.