2015

Essays on Sub-national Economic Growth: Evidence from A Global Sample

Dachao Ruan
Louisiana State University and Agricultural and Mechanical College, rdcfqz@gmail.com

ESSAYS ON SUB-NATIONAL ECONOMIC GROWTH: EVIDENCE FROM A GLOBAL SAMPLE

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Economics

by
Dachao Ruan
B.S., University of Macau, 2007
M.S., University of Macau, 2010
M.S., Louisiana State University, 2012
August 2015

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Professor Areendam Chanda for his guidance, patience, and friendship during my studies at Louisiana State University. He has been inspiring me and challenging me.

I would like to acknowledge the rest of my committee: Professor W. Douglas McMillin, Dr. Fang Yang, and Dr. Stephen R. Barnes, as well as Professor Carter Hill, Professor Bulent Unel, and Dr. Seunghwa Rho for their invaluable advice.

Finally, I owe a great debt to my wife, Anabela Nogueira do Espirito Santo, who married me when I had nothing. She has always had faith in me and taken care of the family with her love. I thank my mother, Meirong Zhou, for consistently supporting me and cheering me up. I thank my son, Douglas Ruan, for bringing new hope and unrestrained joy in my life. Additionally, I am very grateful for my friends and classmates, particularly Ting Wang, who have supported me unconditionally both in life and research.

Table of Contents

Abstract

In this dissertation, I make three important contributions to the literature on regional economics. In Chapter 2, I construct a measure of early development, urban population density (urban population relative to total land area), that is novel to the growth literature, and apply GIS techniques to define and locate regions/cities and obtain geographic and historical measures across regions and cities. Chapter 3 investigates the persistence in subnational development over the past 150 years. I find that regions that had a relatively higher urban population density in 1850 tend to be relatively more developed today. Geographic and climatic characteristics are significantly correlated with development and explain part of persistence, and human capital and physical capital are potential channels of the persistence. Chapter 4 explores the existence of regional convergence over the past 150 years. I find that regions have been converging at a very slow rate over the past 100 years but the rate of convergence has accelerated over the most recent half century.

Chapter 1. Introduction

Explaining persistent differences in living standards across countries remains a central challenge in economics research. Traditionally, growth theories have emphasized the importance of physical capital, human capital, and technology in explaining cross-country inequalities (Hall and Jones, 1999). However, the bi-directional relationship between these factors and economic growth has posed a significant hurdle. As a consequence, increasingly more attention has been paid to geographic and historical factors as fundamental factors in explaining the failure of most developing countries to catch up with their industrialized counterparts.

At the sub-national level, income disparities across regions are also pervasive and substantial. One would expect such persistence to be less important than cross-country disparities. The movement of goods and people is inherently easier between regions because of lower transport costs, similar national institutions, and fewer political barriers. Despite this, it is often observed that the distribution of economic activity across regions can persist for decades or even hundreds of years. However, research on long-run regional growth is limited by lack of appropriate data. To fill this void, in the first chapter of my dissertation I construct a measure of historical development, urban population density (urban population relative to total land area), that is novel to the growth literature. I apply GIS techniques to define and locate regions and cities and obtain their geographic and historical data. In the second chapter, I 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. Based on strong evidence of persistence found in this chapter, in the last chapter I address the question of whether incomes are converging or diverging across regions within countries. While there are numerous studies on economic growth at the regional level, to the best of my knowledge this is the first to explore this employing a global sample while simultaneously encompassing a period of one hundred and fifty years.

1.1 Use of GIS in Creating Regional Data

The reliance on Geographic Information System (GIS) has become increasingly common in the sphere of applied economics research. In Chapter 2, I demonstrate how to use various ArcGIS tools to process GIS data and to create variables included in this dissertation. This chapter serves the dual purpose of both documenting my work using GIS based data that forms the basis for my econometric work in the subsequent chapters, as well as a providing a brief guide to those who would like to apply these techniques in doing empirical research in international development. In the first part of this chapter, I explain how I use ArcGIS and Stata to construct urban population density for regions based on population estimates of 2,803 historical settlements from Chandler (1987), Bairoch (1998), and Eggimann (1994) for the year 1850. I discuss an econometric problem - spatial correlation - associated with spatial data and approaches to solve this issue. In the second part, I introduce commonly used GIS data such as nighttime luminosity, land suitability, temperatures, altitude, ruggedness, rainfall, etc. with an emphasis on sub-national units.

1.2 Regional Persistence

A 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 (Douglas A. Hibbs and Olsson, 2004; Olsson and Hibbs, 2005; Ashraf and Galor, 2013) as well as early development in technology (Comin, Easterly, and Gong, 2010), state capacity (Bockstette, Chanda, and Putterman, 2002), and agriculture (Galor and Moav, 2007).

Acemoglu, Johnson, and Robinson (2002), on the other hand, is a notable exception and find no such persistence among former European colonies over the past 500 years. They find a reversal of fortune among ex-colonial countries and attribute the pattern to European colonialism - European settlers adopted constructive institutions in colonies that were poor 500 years and extractive institutions in colonies that were more densely populated. They argue that the economic profitability of alternative colonial institutions associated with population density resulted in the reversal. However, some economists, including Rafael et al. (2004), Putterman and Weil (2010), and Chanda, Cook, and Putterman (2014), point out that the human capital which accompanied by the large scale of population movements to the colonies since 1500 should not be neglected. Chanda, Cook, and Putterman (2014) correct Acemoglu, Johnson, and Robinson (2002) early urbanization indicators based on the World Migration Matrix 1500 - 2000 constructed by Putterman and Weil (2010), such that the early development of a country traces civilization attached to the year 1500 ancestry of current population rather than that of the same territory 500 years ago. Given these migration-adjusted early development indicators, they find that all reversal patterns found by Acemoglu, Johnson, and Robinson (2002) become persistence ones. Moreover, they find evidence that human capital is one of the important channels through which early development influences current economic performance.

In Chapter 3, I explore the extent to which contemporary GDP per capita at the sub-national level is correlated with economic development in 1850. Drawing on historical city data, I construct measures of urban population density in 1850 for a sample of 2,058 sub-national regions covering 135 countries. These measures are supplemented with indicators to capture the existence of urban areas within a region, as well as its neighboring regions, urban population densities in neighboring regions, as well as quadratic versions of the density variables to capture non-linearities. I found strong evidence of persistence in regional development. The findings are robust to an extensive range of geographic and spatial controls. This persistence is remarkably robust even for various sub-samples of nations -

grouped by continent, colonization history, current income levels, and also using alternative measures of modern development such as current urbanization, population density, and night-time light density. I also find that past urbanization is associated with contemporary human capital and infrastructure differences across regions. In addition, I look back further in time and find there is persistence of regional development for over 500 hundred years for non-colonized countries.

The results are in line with two theories explaining regional disparities in economic development. 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. Second, the economics of agglomeration postulates that there are advantages to agglomerations 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, Ottaviano, and Temple, 2013). These externalities attract mobile factors from other regions that 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.

1.3 Regional Convergence

Convergence, often called β convergence, is a significant prediction of the neoclassical growth model. The neoclassical growth model suggests that a country converges to its steady state and that its growth rate is negatively correlated with national income at the beginning of the period. Therefore, one should observe that poor economies tend to grow faster than rich countries and eventually catch up with rich countries if convergence exists. Numerous cross-country studies of long run development have examined convergence; for example, see

Barro (1991) ,Mankiw, Romer, and Weil (1992), and Caselli, Esquivel, and Lefort (1996). More particularly, because different countries may have different steady states, economists are interested in conditional β convergence, i.e. convergence conditional on various national characteristics like the savings rate that helps determine the steady-state. The convergence rate, the speed of convergence of an economy to its steady state, is estimated based on regression of the growth rate of GDP per capita on GDP per capita at the beginning of the period holding other variables constant. For evidence of convergence, the coefficient on GDP per capital at the beginning of the period must be negative; the magnitude of the coefficient reflects the convergence rate in percent per year. In addition to the general evidence of β convergence across various studies, another interesting phenomenon emerges - the estimated speed of β convergence is consistently around 2 % per year. For example, Barro (2012) studies convergence across 80 countries with GDP data starting from the 1960s and finds the convergence rate is about 1.7 % per year. For a sample of 34 countries with GDP data starting around the late 19th century, the estimated convergence rate is about 2.4 % per year.

For studies of convergence across regions, many of them are based on regions within a single or several countries (Barro and Sala i Martin, 1991, 1992, 2004; Shioji, 1996; Durlauf and Quah, 1999; Coulombe and Tremblay, 2001; Garofalo and Yamarik, 2002). Gennaioli et al. (2014) use a large sample of regions from over 80 countries since the 1950s and find the convergence rate is around 2 %, essentially the same as the speed of catch up between countries.

However, the robust evidence of persistence in regional inequality over the past 150 years found in the second chapter casts doubt on the existence of regional convergence. The finding of persistence suggests that regions that were more developed in 1850 tend to be richer today. On the other hand, the idea of convergence postulates that poorer regions' economies should grow at faster rates than those of more prosperous areas and should eventually catch up with richer regions. To find out whether regional convergence over the past 150 years exists and the speed at which regions have been converging enable us to better understand regional income disparities. In Chapter 4, I look for convergence and find a robustly slow convergence rate across regions. I look at 827 regions in which cities existed in 1850 across 144 countries and three 50-year intervals between 1850 and 2000. I use urban population density as a measure of regional development and investigate the speed of convergence in development across regions within countries. The estimated within-country convergence rate for urban population density is around 0.35 % per year, conditional on a comprehensive set of geographic and climatic controls as well as country fixed effects. I also find worldwide evidence that the convergence of regional development has mainly occurred within the past 100 years and accelerated over time. Changes in the speed of regional convergence are potentially correlated with increases in overall productivity.

Chapter 2. A Description of the Use of GIS in Creating Regional Data

2.1 Introduction

The use of Geographic Information Systems (GIS) has recently injected new blood into empirical economics. Numerous new data brought in by GIS have helped economists broaden their research frontier and strengthen identification of causal effects. With GIS, one can utilize historical information such as roads in Kenya (Burgess et al., 2013), ethnic group boundaries in Africa (Nunn, 2008; Michalopoulos and Papaioannou, 2014), and colonial railroads in India (Donaldson, 2010) to explore long-run economic impacts of economic activities. In addition, geographic and climatic GIS data such as agriculture suitability, elevation, rainfall, and temperature have been frequently used to construct explanatory variables to facilitate our understanding of the relationship between economic activities and environmental conditions (Nunn and Puga, 2012; Mitton, 2013; Dell, Jones, and Olken, 2012). With the help of remote sensing devices, spatial data has been continuously growing in both breadth and depth. Novel spatial data with higher resolution, such as electromagnetic radiation strength (Burgess et al., 2012), nighttime lights (Henderson, Storeygard, and Weil, 2012), and wind speed and direction in the oceans (Feyrer and Sacerdote, 2009), provides economists detailed and invaluable information on the earth's surface. More importantly, given that geographic and climatic characteristics are seldom affected by socioeconomic activities, GIS data has become an important source for researchers to construct instrumental variables to explore causal effects (Rosenthal and Strange, 2008; Combes et al., 2010; Linden and Rockoff, 2008, for example).

Since I intensively use GIS for entering, analyzing, and mapping spatial data in the next two chapters, this chapter serves to fully document my work with GIS in this dissertation.

Given the importance of GIS in empirical economic studies, however, only a few guides for the use of GIS exist for economists with little GIS background. This chapter is, therefore, a complement to the existing ones.[1]

I start by explaining how I construct a novel measure of early development in the growth literature, urban population density, using historical data. Meanwhile, I introduce various ArcGIS tools, terms, and concepts that help create this measure. In the second part, I introduce commonly used spatial data such as nighttime luminosity, land suitability, temperatures, altitude, ruggedness, rainfall, etc. In the last part, I focus on measures of development constructed at both national level and regional level and explore their correlations by displaying scatterplots.

2.2 Constructing 1850 Urbanization Variables

GIS enables us to associate geospatial references with historical data and other invaluable data so that we can merge them with other socioeconomic variables in studies. It is no exaggeration that GIS has opened up abundant resources of data and reshaped economics studies. Thanks to GIS, in this study, I can utilize data for historical settlements and construct a measure of development that is a novel to the growth literature. I illustrate how I build 1850 urbanization variables in this section and briefly touch on concepts and GIS tools being used.

To examine the long run evolution of regional inequality, 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 19th century, let alone at regional levels. In fact, it is only recently that Gennaioli et al. (2013, 2014) compiled regional GDP per capita for the late 20th century and early 21st century. However, GDP per capita is not the only measure of development. The degree of urbanization, i.e. the

[1] Dell (2009), Lowe (2014), and Kudamatsu's lecture slides (http://economics.mit.edu/files/8945) are good sources to learn how to use GIS.

fraction of the population living in urban areas, is also a strong correlate of development. In addition to urbanization, population densities can also serve as a viable indicator. In fact, as argued by Rappaport and Sachs (2003), population density is preferable to incomes when studying variations across regions within a country.[2] In a similar vein, in the urban economics literature, population trends in urban regions are routinely used to compare relative prosperity. However, in 1850, even population estimates for regions are hard to come by making the construction of both measures, population density and urbanization, difficult for a large sample of countries. At the same time, urban historians, such as Chandler (1987), Bairoch (1998), and Eggimann (1994), drawing upon various sources, have compiled population estimates of urban settlements going back centuries.

I draw on these sources to construct my primary indicator of development - the 1850 urban population in a region divided by the total contemporary land area of the region - or what I call *urban population density.* Urban population density is, by definition, a product of the degree of urbanization and population density since, $\frac{UrbanPopulation}{LandArea} = \frac{UrbanPopulation}{Population} \times \frac{Population}{LandArea}$. Hence, as increases in either or both of them would be reflected in increases in urban population density, urban population density is a valid (and also the only available) measure of early development at the regional level in 1850 in this dissertation.

2.2.1 Defining A Spatial Unit

When using Geographic Information Systems (GIS) data, one must clearly define the spatial unit (e.g. county, city, or state) for which GIS data to be summarized. Therefore, the definition of spatial units and their boundaries are fundamental to spatial studies. In this study, the spatial units I am interested in are mainly the 1,569 sub-national divisions from 110 countries for which data for regional income in 2005 is available from Gennaioli et al. (2013).

[2] The cross-country literature, on the other hand, uses population density as a proxy for development mainly during the pre-industrial era when Malthusian forces were dominant.

After clarifying the spatial unit, one can create data for boundaries of regions. Before that, let me briefly introduce how data is represented in GIS.

Data Types

There are two common formats that GIS uses to store data: raster and vector. The raster format consists of millions or billions of grids (or pixels) in which values of spatial information are assigned. The number of grids relates to the resolution of the raster; data with higher resolution has more grids. Most of the spatial data used in this study have a resolution of 30 arc-second. The 30 arc-second grid spacing equates to about 1 kilometer around the equator.

Vector format assigns values to irregular features, which could be points, polylines, or polygons, and provides coordinate data on the location of these features. For example, data for roads and rivers are polylines, locations of settlements are points, and countries, regions, and lakes are polygons. The data for boundaries of regions that I create are examples of polygons.

Figure 2.1 is an example of polygon data. It is a spatial database of the world's administrative areas, called GADM database of Global Administrative Areas. The data can be downloaded for free at http://gadm.org. This database provides administrative divisions at five regional levels beyond the country level, from provinces, states, and departments, to prefectures, counties, and cities. The numerous lines are the boundaries of the finest sub-division available for all countries, including 218,238 administrative areas.

This is the database I use to derive boundaries of my regions. An "attribute table" associated with GADM database provides useful information for me to accomplish this. Next, I explain what an attribute table is.

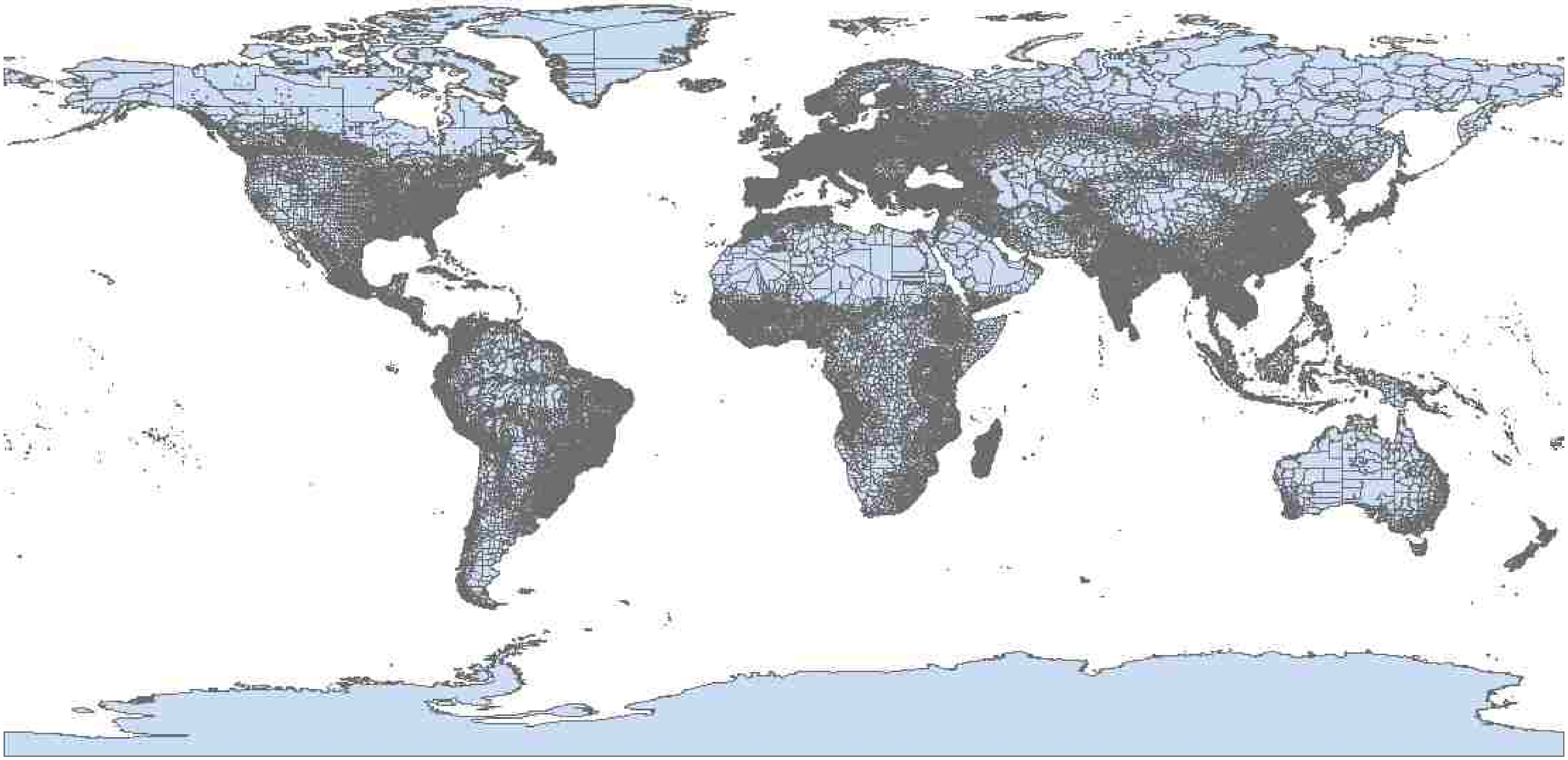


Figure 2.1: Global Administrative Areas (GADM)

Note: This figure is a screen shot of Global Administrative Areas database loaded in ArcGIS 10. It is a spatial database of boundaries of the 218,238 world's administrative divisions. Source:http://gadm.org.

Attribute Table

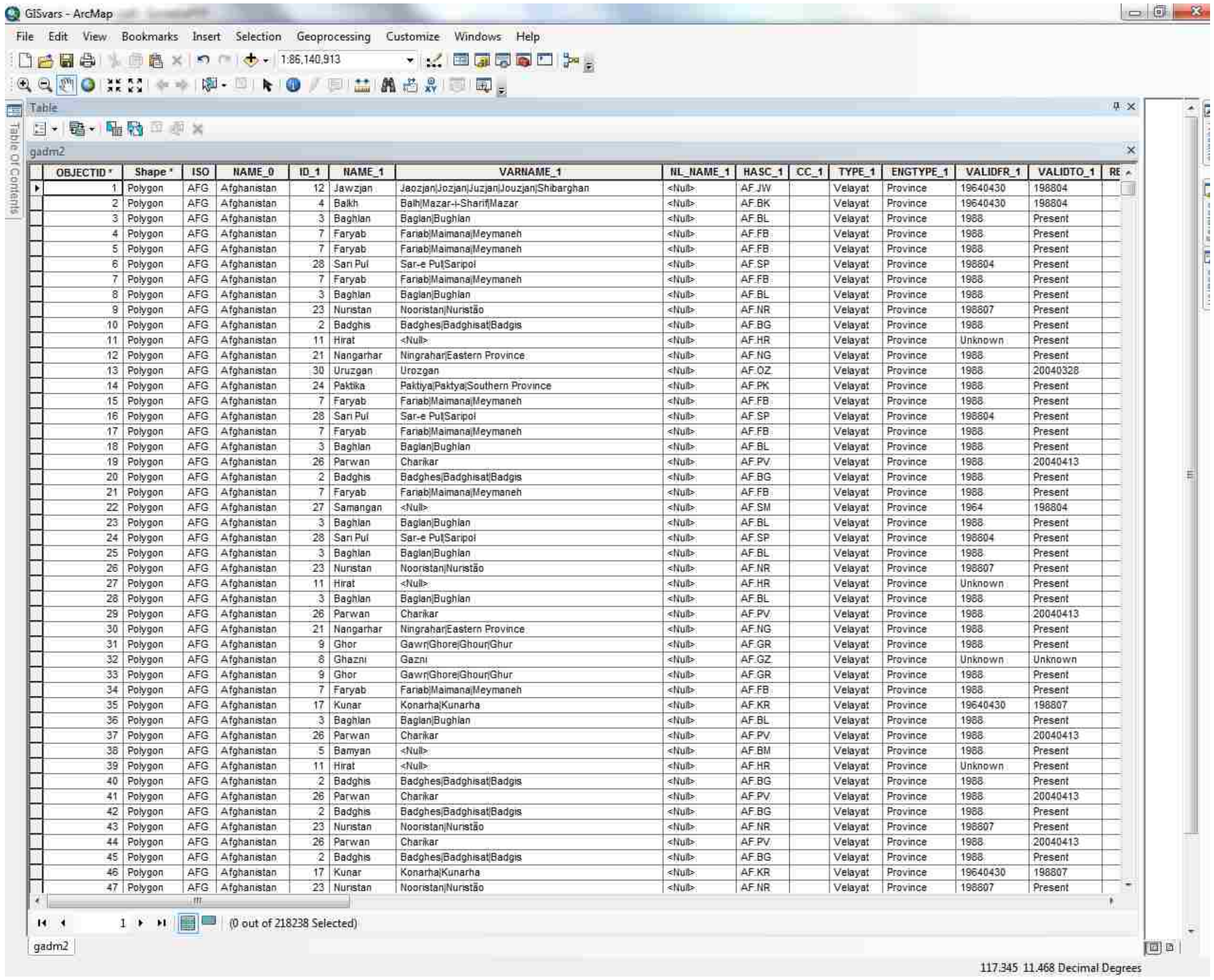
Every vector database is often associated with an attribute table including useful information on each subject such as road, region, a city location. Figure 2.2 shows the attribute table of GADM database, in which each row represents the finest spatial unit containing information such as name, administrative level, names of the higher administrative divisions and so on. In this study, I use this information to identify regions of interest, generate a regional identifier, and create a GIS map for these regions.

An attribute table is a .dbf file, which can be loaded in Stata with Stata's odbc functions. I often use Stata to load attribute tables and manage the attributes information to create desired variables. I need to identify the Gennaioli et al. (2013) regions in GADM database based on their names. This step can be done in Stata. I code Stata to look for the Gennaioli et al. (2013) regions at different GADM sub-national levels and create a unique identifier for the merged regions. In the end, I merge the identifier with the GADM attribute table.

ArcGIS Tool: Dissolve

Once I have customized identifier for interested regions, I can merge several smaller regions under the same identifier number to a bigger region. Because this process dissolves boundaries of the smaller regions, the ArcGIS tool to accomplish this is called Dissolve.

The dissolved output is a GIS dataset of boundaries of new regions. A new attribute table is also created with the number of rows equal to the number of new regions. Figure 2.3 reveals those regions and their boundaries, of which red ones are boundaries of countries; black ones are boundaries of the regions of interest. Gray areas are countries excluded in my study due to a lack of key information. With this data, I am now able to create spatial variables derived from various GIS data.



OBJECTID *	Shape *	ISO	NAME_0	ID_1	NAME_1	VARNAME_1	NL_NAME_1	HASC_1	CC_1	TYPE_1	ENGTYPE_1	VALIDFR_1	VALIDTO_1	RE
1	Polygon	AFG	Afghanistan	12	Jawzjan	Jaozjan|Jozjan|Juzjan|Jouzjan|Shibarghan	<Null>	AF.JW		Velayat	Province	19640430	198804	
2	Polygon	AFG	Afghanistan	4	Balkh	Balh|Mazar-i-Sharif|Mazar	<Null>	AF.BK		Velayat	Province	19640430	198804	
3	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
4	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
5	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
6	Polygon	AFG	Afghanistan	28	Sari Pul	Sar-e Pul|Saripol	<Null>	AF.SP		Velayat	Province	198804	Present	
7	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
8	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
9	Polygon	AFG	Afghanistan	23	Nuristan	Nooristan|Nuristão	<Null>	AF.NR		Velayat	Province	198807	Present	
10	Polygon	AFG	Afghanistan	2	Badghis	Badghes|Badghisat|Badgis	<Null>	AF.BG		Velayat	Province	1988	Present	
11	Polygon	AFG	Afghanistan	11	Hirat	<Null>	<Null>	AF.HR		Velayat	Province	Unknown	Present	
12	Polygon	AFG	Afghanistan	21	Nangarhar	Ningrahar|Eastern Province	<Null>	AF.NG		Velayat	Province	1988	Present	
13	Polygon	AFG	Afghanistan	30	Uruzgan	Urozgan	<Null>	AF.OZ		Velayat	Province	1988	20040328	
14	Polygon	AFG	Afghanistan	24	Paktika	Paktiya|Paktya|Southern Province	<Null>	AF.PK		Velayat	Province	1988	Present	
15	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
16	Polygon	AFG	Afghanistan	28	Sari Pul	Sar-e Pul|Saripol	<Null>	AF.SP		Velayat	Province	198804	Present	
17	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
18	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
19	Polygon	AFG	Afghanistan	26	Parwan	Charikar	<Null>	AF.PV		Velayat	Province	1988	20040413	
20	Polygon	AFG	Afghanistan	2	Badghis	Badghes|Badghisat|Badgis	<Null>	AF.BG		Velayat	Province	1988	Present	
21	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
22	Polygon	AFG	Afghanistan	27	Samangan	<Null>	<Null>	AF.SM		Velayat	Province	1964	198804	
23	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
24	Polygon	AFG	Afghanistan	28	Sari Pul	Sar-e Pul|Saripol	<Null>	AF.SP		Velayat	Province	198804	Present	
25	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
26	Polygon	AFG	Afghanistan	23	Nuristan	Nooristan|Nuristão	<Null>	AF.NR		Velayat	Province	198807	Present	
27	Polygon	AFG	Afghanistan	11	Hirat	<Null>	<Null>	AF.HR		Velayat	Province	Unknown	Present	
28	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
29	Polygon	AFG	Afghanistan	26	Parwan	Charikar	<Null>	AF.PV		Velayat	Province	1988	20040413	
30	Polygon	AFG	Afghanistan	21	Nangarhar	Ningrahar|Eastern Province	<Null>	AF.NG		Velayat	Province	1988	Present	
31	Polygon	AFG	Afghanistan	9	Ghor	Gawr|Ghore|Ghour|Ghur	<Null>	AF.GR		Velayat	Province	1988	Present	
32	Polygon	AFG	Afghanistan	8	Ghazni	Gazni	<Null>	AF.GZ		Velayat	Province	Unknown	Unknown	
33	Polygon	AFG	Afghanistan	9	Ghor	Gawr|Ghore|Ghour|Ghur	<Null>	AF.GR		Velayat	Province	1988	Present	
34	Polygon	AFG	Afghanistan	7	Faryab	Fariab|Maimana|Meymaneh	<Null>	AF.FB		Velayat	Province	1988	Present	
35	Polygon	AFG	Afghanistan	17	Kunar	Konarha|Kunarha	<Null>	AF.KR		Velayat	Province	19640430	198807	
36	Polygon	AFG	Afghanistan	3	Baghlan	Baglan|Bughlan	<Null>	AF.BL		Velayat	Province	1988	Present	
37	Polygon	AFG	Afghanistan	26	Parwan	Charikar	<Null>	AF.PV		Velayat	Province	1988	20040413	
38	Polygon	AFG	Afghanistan	5	Bamyan	<Null>	<Null>	AF.BM		Velayat	Province	1988	Present	
39	Polygon	AFG	Afghanistan	11	Hirat	<Null>	<Null>	AF.HR		Velayat	Province	Unknown	Present	
40	Polygon	AFG	Afghanistan	2	Badghis	Badghes|Badghisat|Badgis	<Null>	AF.BG		Velayat	Province	1988	Present	
41	Polygon	AFG	Afghanistan	26	Parwan	Charikar	<Null>	AF.PV		Velayat	Province	1988	20040413	
42	Polygon	AFG	Afghanistan	2	Badghis	Badghes|Badghisat|Badgis	<Null>	AF.BG		Velayat	Province	1988	Present	
43	Polygon	AFG	Afghanistan	23	Nuristan	Nooristan|Nuristão	<Null>	AF.NR		Velayat	Province	198807	Present	
44	Polygon	AFG	Afghanistan	26	Parwan	Charikar	<Null>	AF.PV		Velayat	Province	1988	20040413	
45	Polygon	AFG	Afghanistan	2	Badghis	Badghes|Badghisat|Badgis	<Null>	AF.BG		Velayat	Province	1988	Present	
46	Polygon	AFG	Afghanistan	17	Kunar	Konarha|Kunarha	<Null>	AF.KR		Velayat	Province	19640430	198807	
47	Polygon	AFG	Afghanistan	23	Nuristan	Nooristan|Nuristão	<Null>	AF.NR		Velayat	Province	198807	Present	

Figure 2.2: Attribute Table

Note: This figure displays part of attribute table of GADM database of Global Administrative Areas (http://gadm.org) loaded in ArcGIS 10. In the table, each row represents one of the 218,238 indivisible sub-national regions. The field, OBJECTID, is a unique identifier of rows. The field, Shape, indicates data type of each row, and Polygon is the data type of regions. The field, ISO, is the 3-digit International Standard for country codes. NAME_0 gives names of countries in which regions are. The rest of fields in the table provide information of the largest administrative divisions in which regions belong to. More fields behind are not displayed for the sake of convenience.

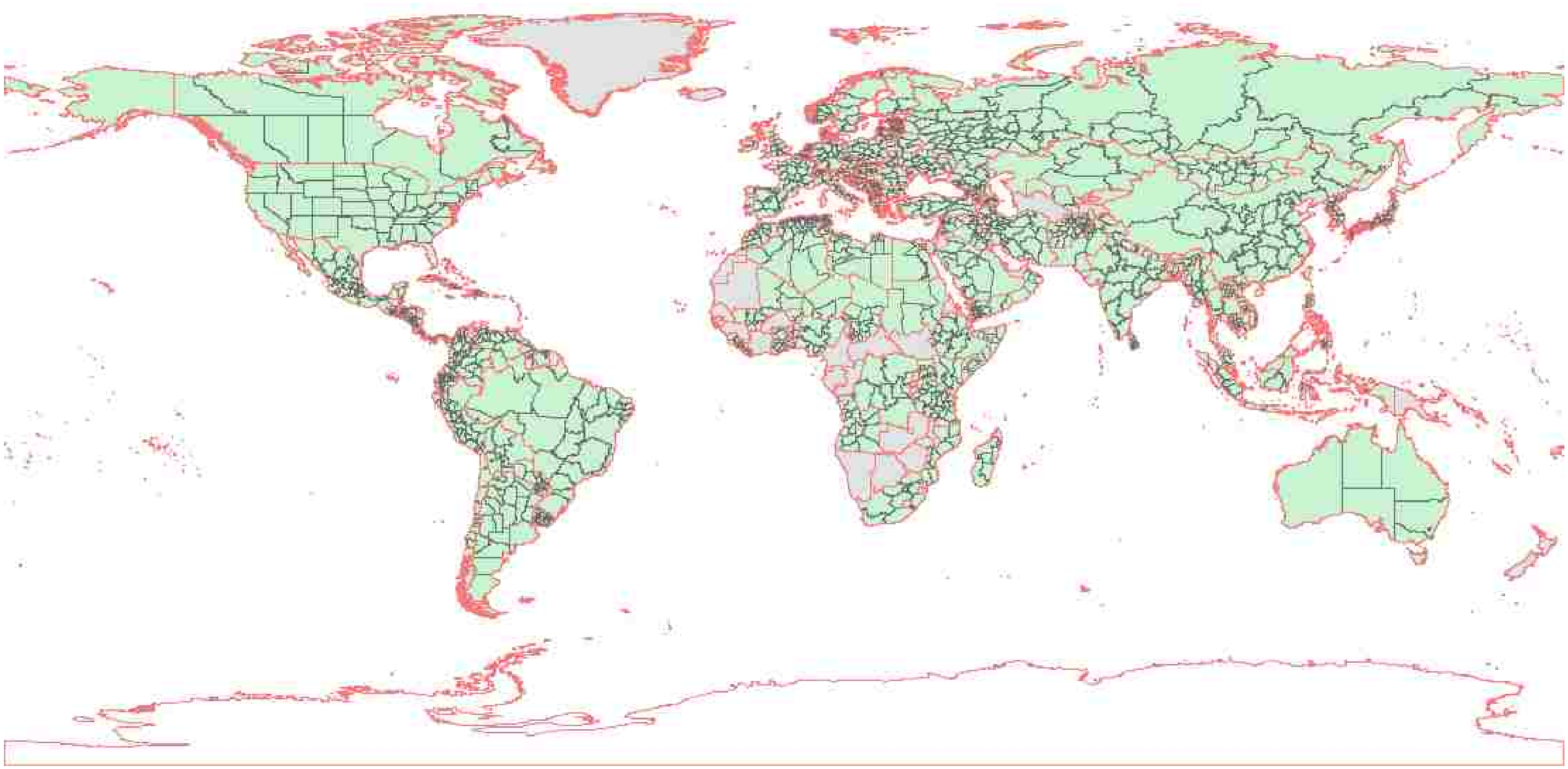


Figure 2.3: Data for Sub-national Regions

Note: This figure shows the data for boundaries of the regions used in this study. The data is derived from GADM database of Global Administrative Areas (http://gadm.org). Red lines divide the world into different countries while the gray ones are boundaries of sub-national units. Green areas consist of countries in which key socioeconomic data is available for this study.

2.2.2 Urban Population Density 1850

In this study, I construct a measure, called urban population density, with the help of GIS. This measure is used as a proxy of historical development at the regional level so that I can study the dynamics of regional development across the world over hundreds of years. To construct this measure, I collect historical settlements and their historical population estimates going back centuries from various sources, such as Chandler (1987). Figure 2.4 displays how the data looks. The page (on the left) lists cities in the Americas in the 19th century. Numbers beside the names of cities are population estimates. There is a citation for how each figure is derived, which is based on either direct quotation from official historical census or indirect derivation from numbers of other events. For example, if a direct population estimate in a city in a year is not available, a number will be determined based on available records regarding the number of churches/temples or any other public facilities, certain goods being traded, or taxes collected in the city within a time period.

The measure, urban population density, is calculated based on equation: $\text{Urban Population Density} = \frac{\text{Aggregate Population living in cities}}{\text{Land Area}}$. I construct urban population density in the following steps: 1) create an XY table for historical settlements; 2) identify cities within boundaries of regions using the ArcGIS Near tool and aggregate population estimates of settlements within regions; 3) measure regional land area; 4) divide aggregate urban population by land area. Below, I outline each of these steps briefly.

Step 1: Creating an XY Table

An XY table is a table of locations associated with longitude (x value) and latitude (y value). To create such a table of historical settlements, I digitize ancient settlements and their population estimates initially recorded in books and look for contemporaneous coordinates (latitudes and longitudes) of these settlements.

I then add the table into ArcGIS, select WGS1984, the most commonly used World Geodetic System, as its spatial reference, and output the data. Now I have a GIS database of historical settlements. Figure 2.5 gives the data I create, in which each spot represents a settlement. In the attribute table associated with this data, population estimates in each settlement are stored.

Step 2: Aggregating Population of Each Region

With the GIS data for settlements and population estimates, I can map settlements into regions created in Figure 2.3 and aggregate urban population within each region. I use the ArcGIS tool of **Generate Near Table** to accomplish this. The tool produces a table indicating the shortest distances of each settlement (point feature) from boundaries of regions (polygon features) and coordinates on regions' boundaries that connects each settlement in the shortest distance. When settlements and a region intersect, or settlements are within a region, distances between the two are zero. To calculate urban population density, one need not know the distances of cities outside regions. However, it is useful for creating other variables that I will explain in the next subsection.

To operate aggregation of population estimates within each region, I load the output table into Stata by using Stata's odbc function, and for each region I sum population estimates of settlements that have zero-distance to that region.

I illustrate the process again in Figure 2.6 which displays an imaginary region with five settlements. The ArcGIS tool of **Generate Near Table** produces a table identifying settlements within the region (two red ones) and distances of other points to the region's boundaries. I load the table in Stata and obtain an urban population of the region by summing population estimates in the two red settlements.

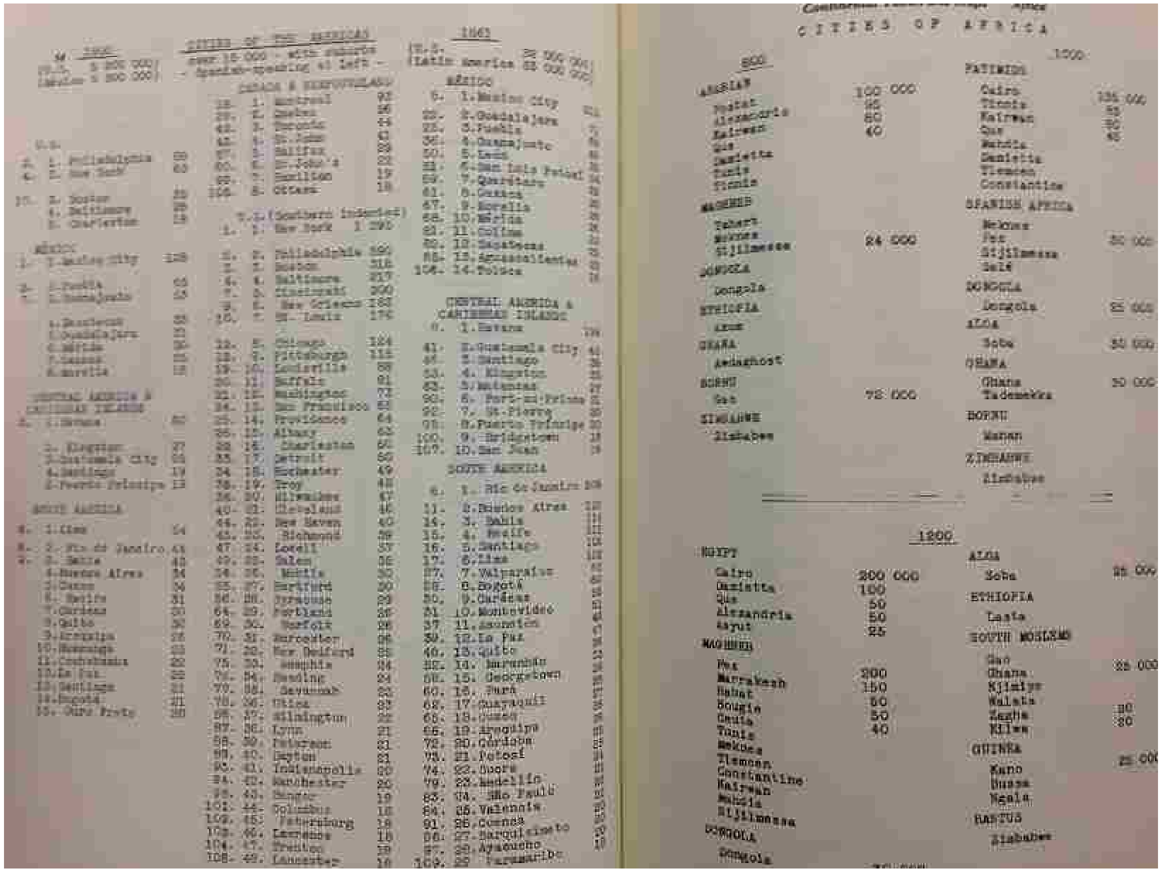


Figure 2.4: Data for Historical Settlements with Population Estimates, 1850

Note: This figure shows an example of the data for historical cities used in this study. The data is printed in “Four Thousand Years of Urban Growth” by Chandler (1987). The page (on the left) lists cities in the Americas and their corresponding population estimates in the 19th century.

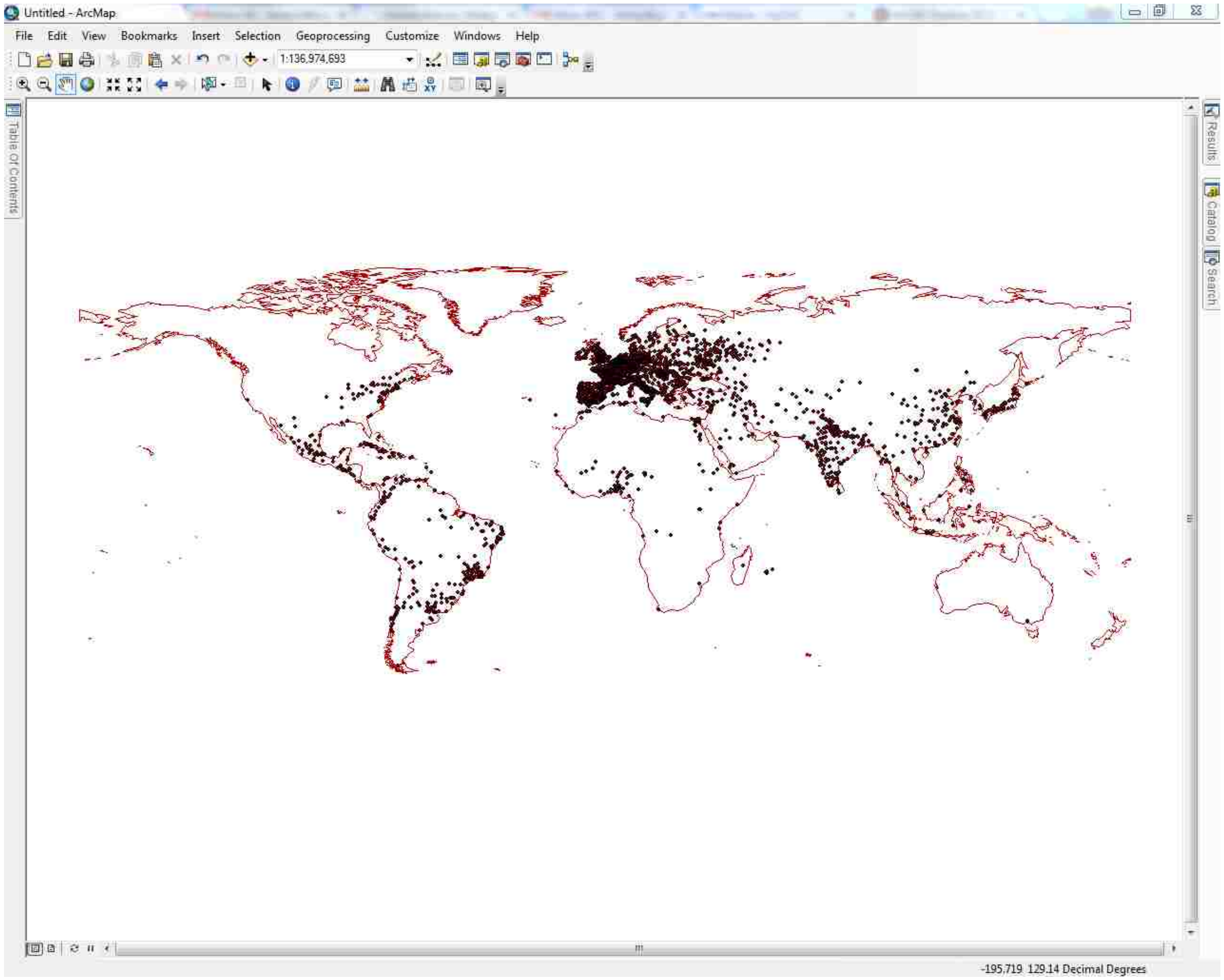


Figure 2.5: Spatial Data for Settlements, 1850

Note: This figure displays GIS data for historical settlements in 1850. Spots in black are the locations of the settlements in 1850. The settlements in 1850 collected from Chandler (1987), Bairoch (1998), and Eggimann (1994).

Step 3: Measuring Land Area

The geographical data for land area is in raster format with a resolution of 30arc second. The data is available at the Center for International Earth Science Information Network (CIESIN) and can be downloaded at http://sedac.ciesin.columbia.edu. Figure 2.7 provides a glimpse of the data in ArcGIS.

Figure 2.6: Aggregating Population Estimates at Regions
Note: This figure includes an imaginary region in yellow, two cities within the region in red, and three cities outside the region in black.

To process raster data, one usually use the ArcGIS tool of **Zonal Statistics as Table**, which summarizes values of a raster within spatial units of another dataset and outputs results to a table. To calculate land area within the spatial units, one need to specify the statistics type option of the tool in SUM so that values of all grids within regions are aggregated.

Step 4: Creating Maps of Urban Population Density

Now, I have data on regional urban population and land area, and can merge the two according to identifiers of regions and obtain urban population density by dividing the former by the later.

One can further make a map of urban population density in ArcGIS - just right click on properties and click on the "Symbology" tab. Figure 2.8 shows the distribution of urban population density in 1850 I generate using ArcGIS. The darker regions are more densely populated.

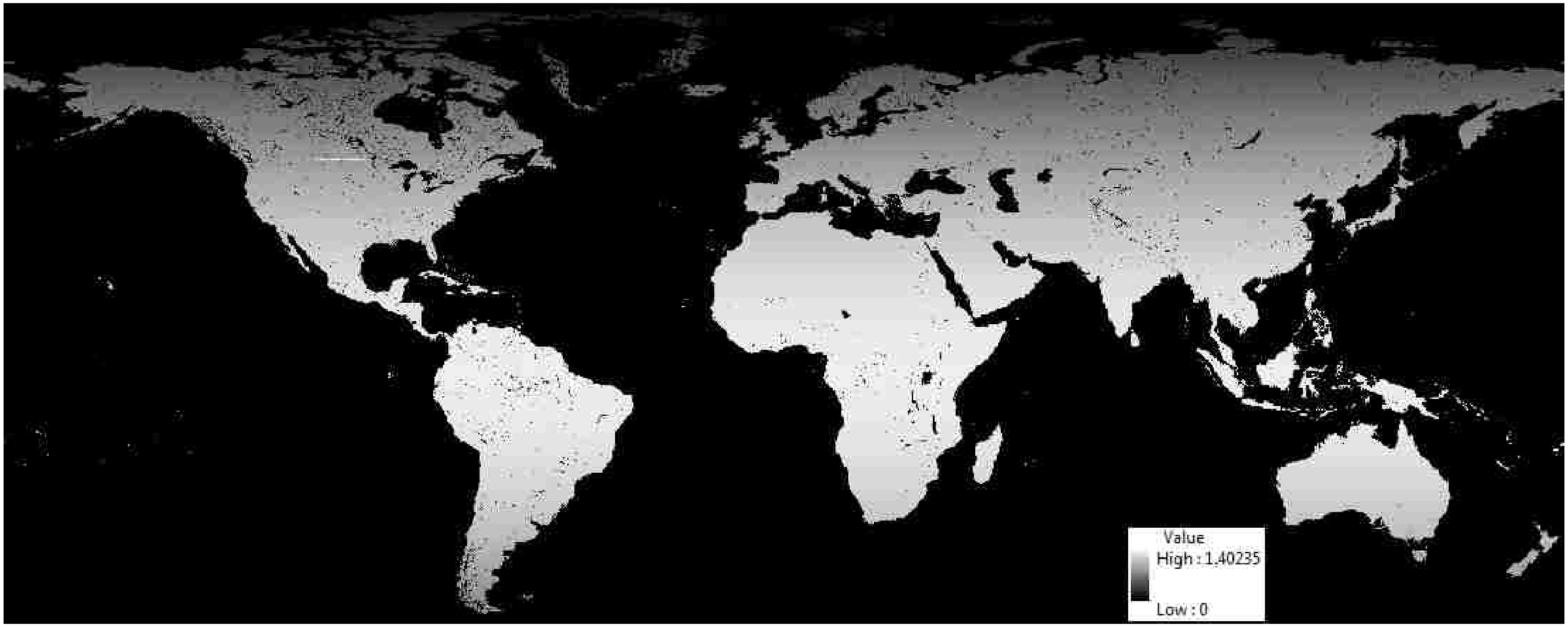


Figure 2.7: Land and Geographic Unit Area Grids, 2000

Note: This is a 30 arc-second raster of global land area in 2000. The data measures land areas in square kilometers. Source: 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): Land and Geographic Unit Area Grids. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4K935FC. Accessed DAY MONTH YEAR.

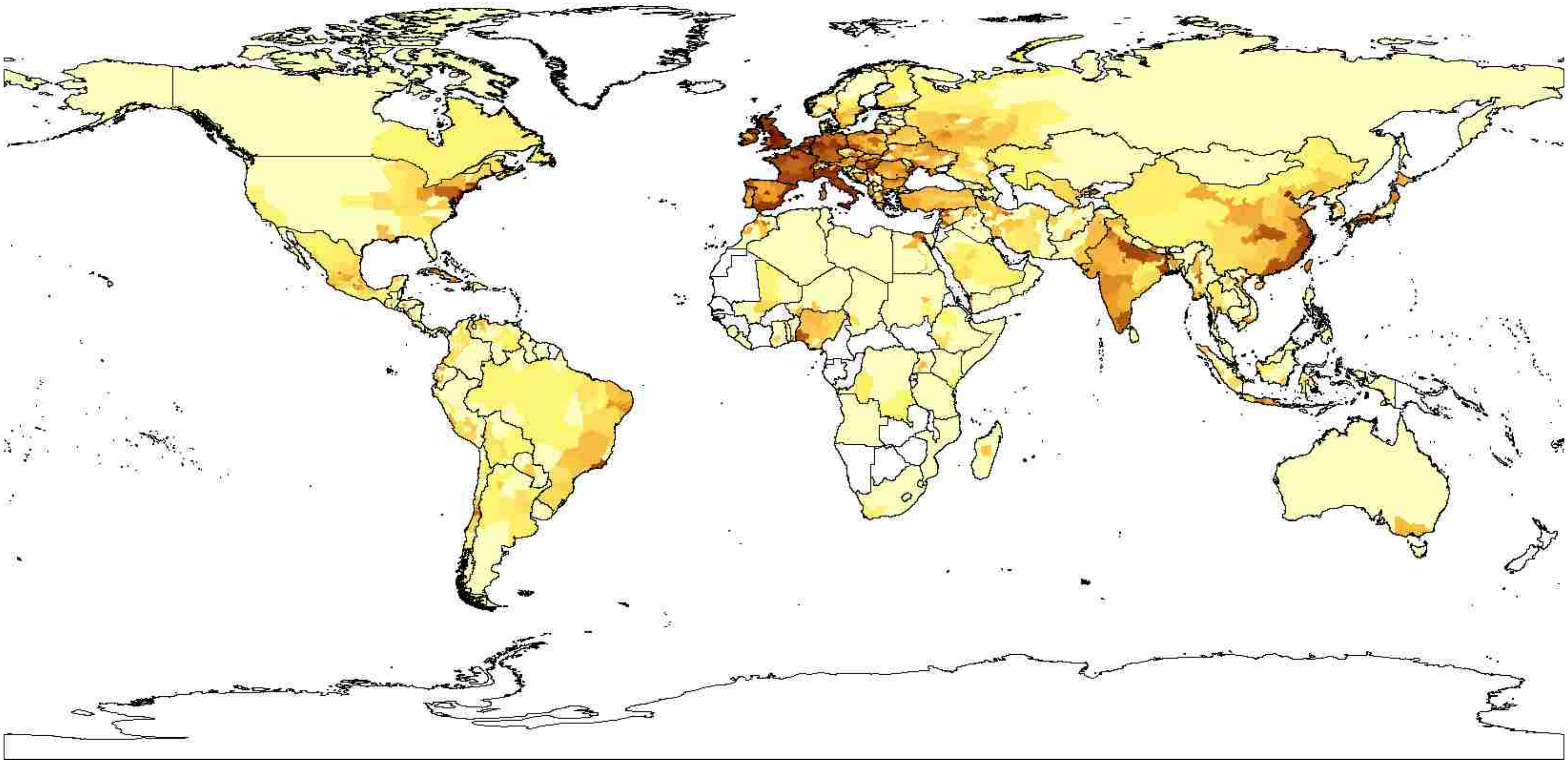


Figure 2.8: Distribution of Urban Population in 1850

Note: Empty areas consist of countries that do not appear in my data for cities in 1850. Shaded areas consist of regions with the urban population in 1850, and the darker regions is more densely populated in 1850. Unit of urban population density is 100 persons per square kilometer.

2.2.3 Controlling for Spatial Correlation

For studies on small spatial units, such as sub-national regions in this study, one cannot ignore interregional inflows of goods and production factors. Failure to include spatial interaction effects will lead to misspecification problems. There are three traditional ways to model spatial correlation: 1) incorporate a spatial lag of the dependent variable; 2) a spatial lag of the independent variable; and 3) a spatial lag of the error term. The strategy I use to control for the spatial correlation is to include urban population density in neighboring areas, which is an example of the second approach.

I divide the process of creating spatial variables while controlling for spatial correlation into four steps. First, I use GIS to help estimate the shortest geodesic distances between settlements and regions' boundaries. Second, I use 25 miles from regions' neighboring areas and aggregate population estimates of cities in these areas. Third, I calculate regional land area. Lastly, I obtain urban population density in neighboring areas by dividing the aggregate population estimates of cities by regional land area.

The first step is done in Section 2.2.2 by using the ArcGIS tool of Generate Near Table, through which I obtain shortest distances of settlements from regions' boundaries. The regional area in step three is the same as the one in Section 2.2.2. Therefore, the only thing I need to do is aggregate population estimates of settlements conditional on the settlements within 25 miles from boundaries of each region and divide the total numbers by the land area. This can be done in Stata.

Alternatively, one can identify cities in neighboring areas by using the ArcGIS tool of **Buffer**. This tool helps define neighborhoods around regions by specifying a 25-miles buffer around existing regions.

I illustrate the process again in Figure 2.9. For an imaginary region colored in yellow, I use 25 miles from its boundaries as the region's neighborhood (presented by green area). Instead of creating a 25-miles buffer around the region and looking at settlements within it, I apply the ArcGIS tool of **Generate Near Table** to output a table identifies settlements within 25 miles from the region's boundaries (two red points). I load the table in Stata and obtain an urban population of the region by summing population estimates in the two red settlements. At last, I divide the urban population in neighboring areas by land area of the region.

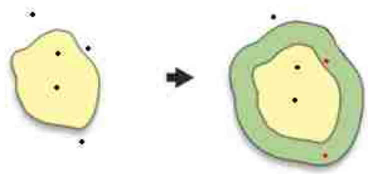


Figure 2.9: Defining Neighboring Areas

Note: This figure gives an example of how to define neighboring areas of an imaginary region in yellow. Areas in green are the neighboring areas of the region using 25 miles from the region's boundary. For the five locations of cities, the two in red are considered cities in the neighboring areas.

Projection and Distance Calculation

When calculating distances in meters and areas in square meters, one needs to project the spherical surface of the Earth onto a plane. There are numerous ways to do that; however, each type of projection introduces some specific distortions in calculation. Therefore, one should use the one that creates the least distortions. There are three points to note when we project coordinate systems. First, we need to ensure that all layers in performing the calculation are in the same coordinate system. 2) When spatial units have small areas, such as US states or Japanese prefectures, the Universal Transverse Mercator (UTM) projection is usually a good option bringing the least distortion in any dimension. 3) Maps associated with different projections have different looks. However, in ArcGIS, the visual output follows the projection of the first layer. Therefore, one should always check the projection of each layer by looking at its properties.

The best way that I am aware of to calculate geodesic distance is to use "globdist" .ado directly utilizing coordinates of locations. In my study, for example, I often need to calculate the shortest distances between points and polygons. ArcGIS calculates distance first looking for coordinates of a location of a polygon that is closest to a point and then measure the distance between the two. As I have mentioned above, distortions will only be introduced in the second stage when the projection is involved. By using "globdist," observations are treated as being on the surface of a perfect spherical planet with the world-radius slightly adjusted with locations of observations. The distance is computed in kilometers based on great-circle distance formulas and is claimed to be accurate to 0.1 kilometers.

2.3 Other Spatial Data

In this section, I briefly introduce other GIS data included in this study and how I process them with ArcGIS.

Nighttime Lights

Data on nighttime lights, as displayed in Figure 2.10 are measured by satellites from outer space. The data are available at the National Centers for Environmental Information and can be downloaded at http://www.ngdc.noaa.gov. Each data file is a 30 arc-second raster. Therefore, I use the ArcGIS tool of **Zonal Statistics as Table** with the statistics type option specified in SUM. I further divide it by the land area and use it as a measure of contemporaneous development.

Population in 2000

Outcome variables of regional population density 2000, urban population density 2000, and urbanization are derived from GIS data for the Population Density grids in Figure 2.11 and Settlement Points in Figure 2.12. Both are available at the CIESIN and can be downloaded at http://sedac.ciesin.columbia.edu.

The population density grids measure population per square kilometer. It is a 30 arc-second raster. Therefore, I use the ArcGIS tool of **Zonal Statistics as Table** with the statistics type option specified in MEAN to calculate regional population density.

The Settlement Points consists of all urban and rural settlements. I use the ArcGIS tool of **Generate Near Table** to aggregate both urban and total populations for each sub-national unit. I obtain regional urban population density by dividing the urban population by regional land area, and urbanization by dividing the urban population by total population.

Agriculture Suitability

Data on agriculture suitability, as displayed in Figure 2.13 is introduced by Ramankutty et al. (2002). The measure is constructed with consideration of cultivability of land, climatic conditions, and soil properties in growing agricultural products. The data are available in the Atlas of the Biosphere: Mapping the Bioshpere and can be downloaded at http://www.sage.wisc.edu. The data file is a 5-degree raster. I use the ArcGIS tool of **Zonal Statistics as Table** with the statistics type option specified in MEAN to calculate the average land suitability for agriculture in a region.

Latitude

Data on latitude, as displayed in Figure 2.14, is derived from data for sub-national regions in Figure 2.3. The ArcGIS tool of **Median Center** is used to output the average latitudes and longitudes for all regions' centroids. I further take the absolute values of latitudes and include the variable in this study.

Coast Lines

I derive the proximity to the coast from data for Boundary between Land and Ocean displayed in Figure 2.15. The data are available at the National Centers for Environmental Information and can be downloaded at http://www.ngdc.noaa.gov. The original file is a polygon vector, therefore, I first use the ArcGIS tool of **Feature To Line** to converting polygon boundaries to lines. I then use the ArcGIS tool of **Generate Near Table** to calculate the shortest geodesic distances between the lines to regional centroids in Figure 2.14. Finally, I calculate the proximity to the coast as the reciprocal of one plus the distance in 1,000 kilometers for each region.

Rivers

I derive the proximity to rivers from data for Rivers displayed in Figure 2.16. The data are available at the National Centers for Environmental Information and can be downloaded at http://www.ngdc.noaa.gov. The data file is a polyline vector. Therefore, I use the ArcGIS tool of **Generate Near Table** to calculate the shortest geodesic distances between the river lines to regional centroids in Figure 2.14. Finally, I calculate the proximity to rivers as the reciprocal of one plus the distance in 1,000 kilometers for each region.

National Capitals

The proximity to the national capital is derived from data for Capital Cities displayed in Figure 2.17. The data are available at the World Urbanization Prospects, the 2011 Revision, and can be downloaded at http://esa.un.org/unpd/wup. The data file is a points vector. What I want to create first is the shortest distance between a region's centroid to the capital city of the country where the region exists. Therefore, I use the ArcGIS tool of **Generate Near Table** to calculate the shortest geodesic distances between the national capital to regional centroids in Figure 2.14 within a country, and I loop over all the countries. At last, I calculate the proximity to the national capital as the reciprocal of one plus the distance in 1,000 kilometers for each region.

Diamond Mines

Dummy variable indicating whether a diamond mine exists in a region or not is derived from the Diamond Dataset, as displayed in Figure 2.18. The data are available at the Peace Research Institute Oslo (PRIO) and can be downloaded at https://www.prio.org. The data file is a points vector. Therefore, I use the ArcGIS tool of **Generate Near Table** to calculate the shortest geodesic distances between the points to regions' boundaries in Figure 2.3. I obtain the diamond mine dummy by checking whether a region intersects at least a diamond mine location.

2.4 Correlations between Various Development Measures

I have mentioned above that urbanization and population density are measures of historical development commonly used in the literature. To what extent is urban population density actually correlated with urbanization and population density? Since I do not have data for the latter two measures at the regional level in 1850, as a starting point, I evaluate the extent to which the former captures these measures by examining correlations using country level data. Figures 2.19(a) and 2.19(b) are scatterplots of the two variables against the logarithm of urban population density.[3] Both plots indicate that log urban population density is strongly associated with log population density and log urbanization. For the 251 sovereign economies for which I could gather data, the simple correlations are 0.88 and 0.45, respectively.

Moreover, GDP per capita, nighttime light intensity, urbanization, and population density are also valid measures of contemporaneous development of regions. As these measures at the regional level are available in contemporaneous years such as 2000 and 2005, I investigate to what extent urban population density captures these measures by plotting

[3] Data for total population in 1850 at the country level is taken of world population history (1978). I calculate urbanization in 1850 as the ratio of urban population to the total population.

their correlations. Nighttime light intensity, urbanization, and population density in 2000 are constructed with using ArcGIS through ways I have discussed above. Data for GDP per capita in 2005 is available for 1380 regions from Gennaioli et al. (2013). As shown in Figure 2.20, for measures constructed at the regional level in 2000 and 2005, urban population density is positively correlated with GDP per capita, urbanization, population density, and nighttime light intensity. In addition, in Figure 2.21, I provide further evidence that these four commonly used measures of regional development are significantly correlated with each other. Given all these, therefore, urban population density is a valid measure of development.

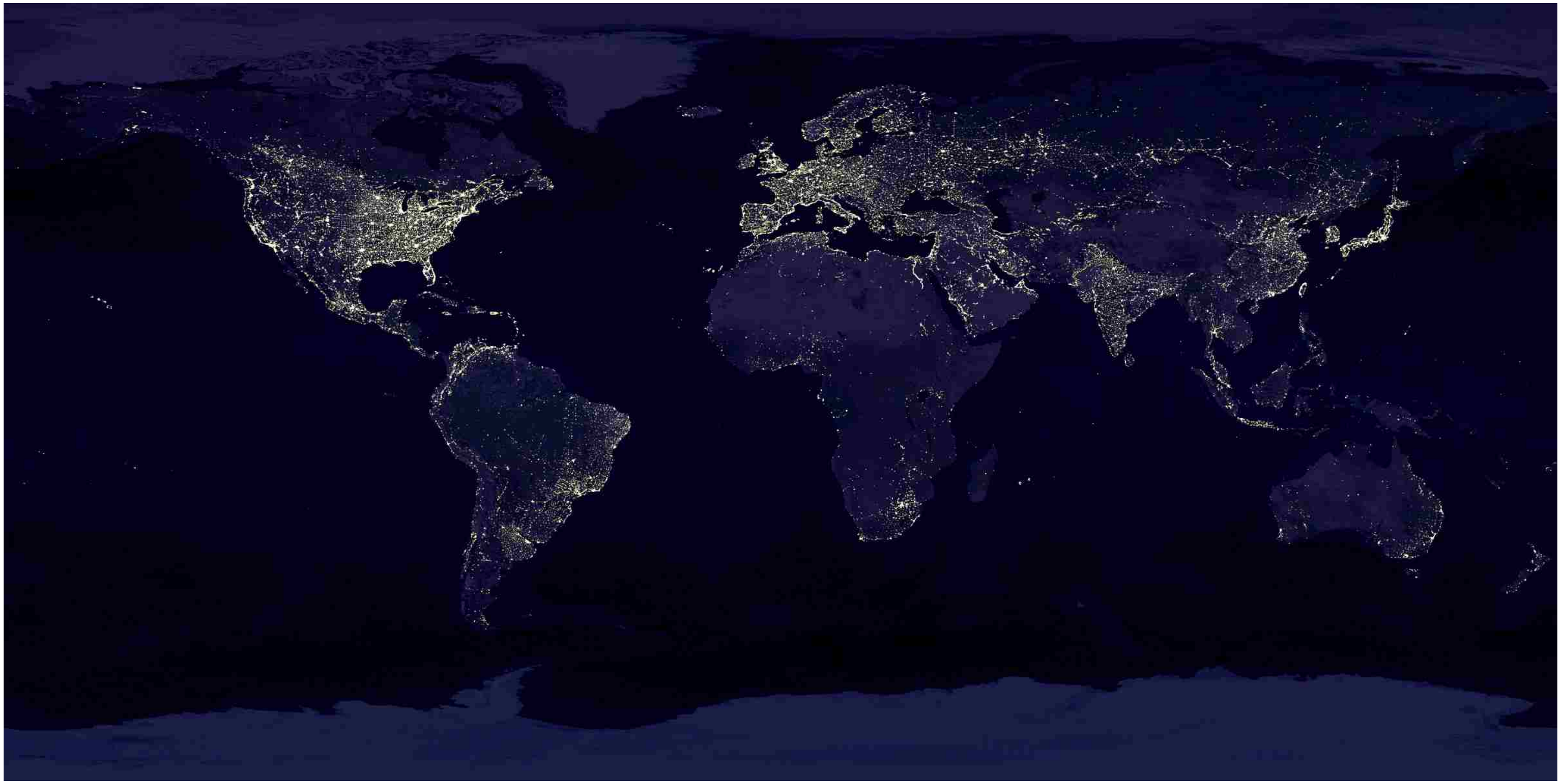


Figure 2.10: Data for Nighttime Light Intensity

Note: This image shows the nighttime light of the Earth during October 1, 1994 - March 31, 1995, created by the Defense Meteorological Satellite Program (DMSP) Operational Linescan System (OLS). Source: http://visibleearth.nasa.gov/view.php?id=55167.

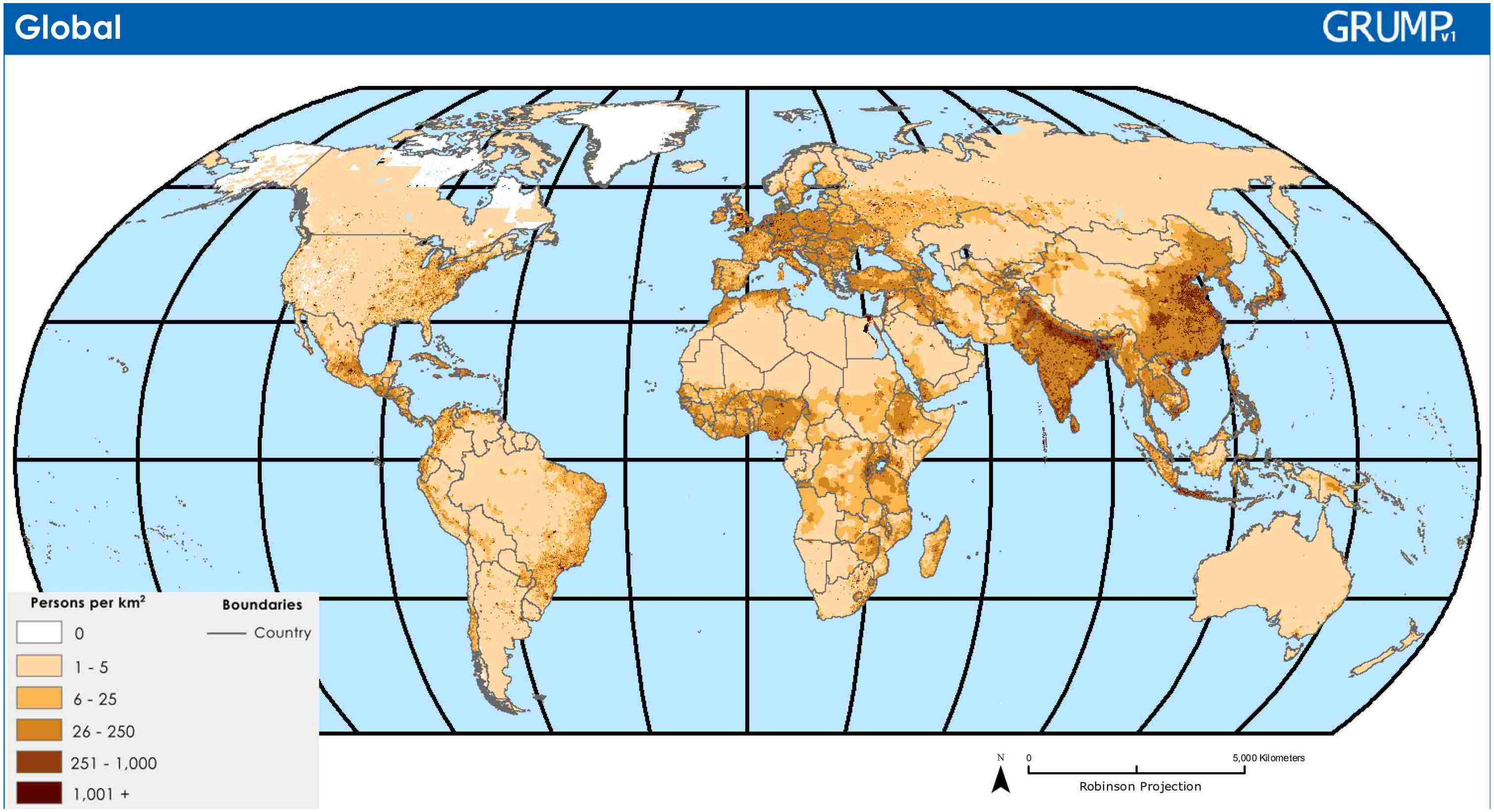


Figure 2.11: Data for Population Density, 2000

Note: This imagine shows the world population density in 2000. Source: Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute (IFPRI), the World Bank, and Centro International de Agricultura Traopical (CIAT). Global Rural-Urban Mapping Project (GRUMP), Population Density. Available at http://sedac.ciesin.columbia.edu/gpw.

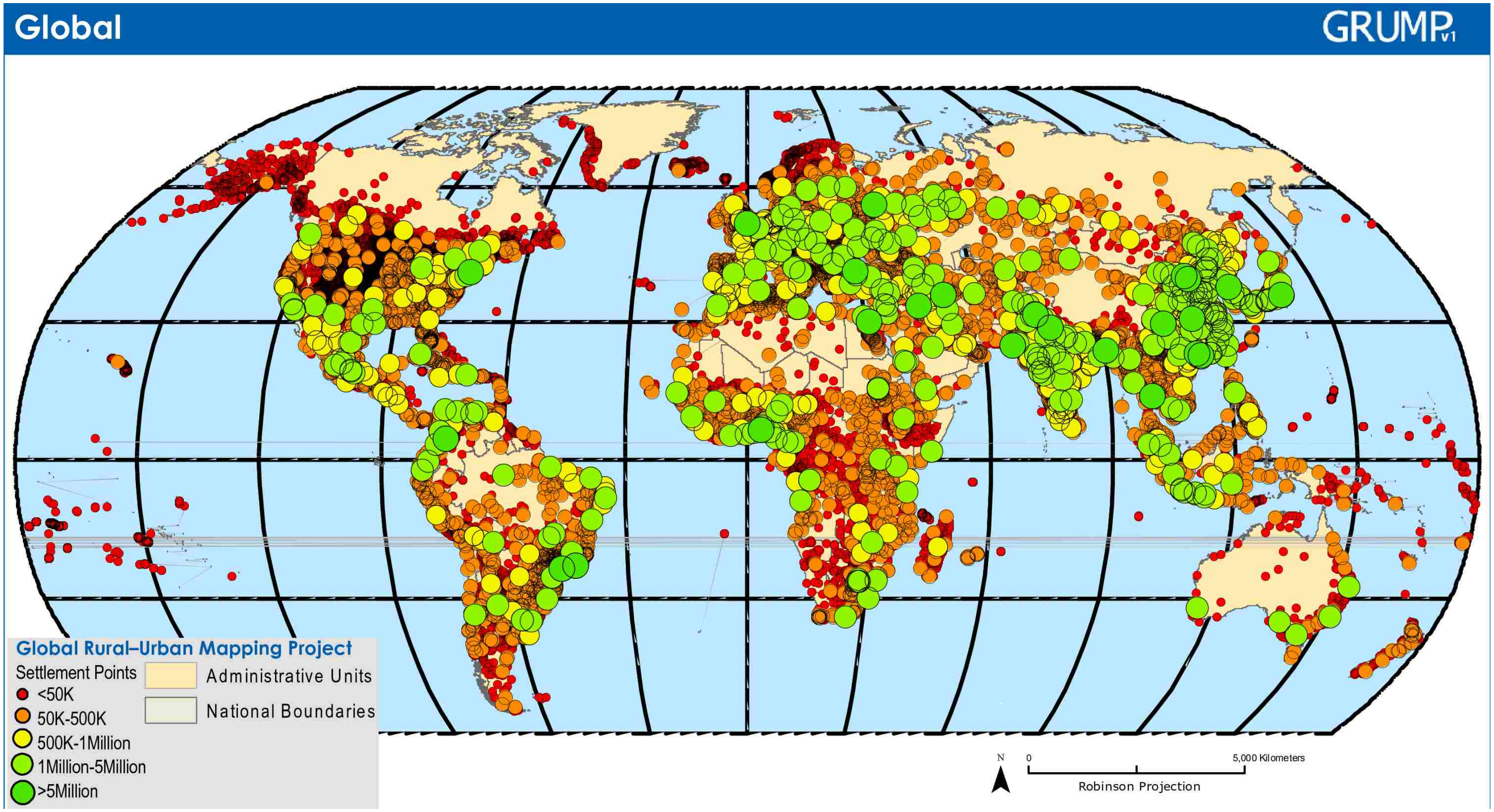


Figure 2.12: Data for Settlement Points, 2000

Note: This imagine shows the world settlement points in 2000. Source: Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute (IFPRI), the World Bank, and Centro International de Agricultura Traopical (CIAT). Global Rural-Urban Mapping Project (GRUMP), Population Density. Available at http://sedac.ciesin.columbia.edu/gpw.

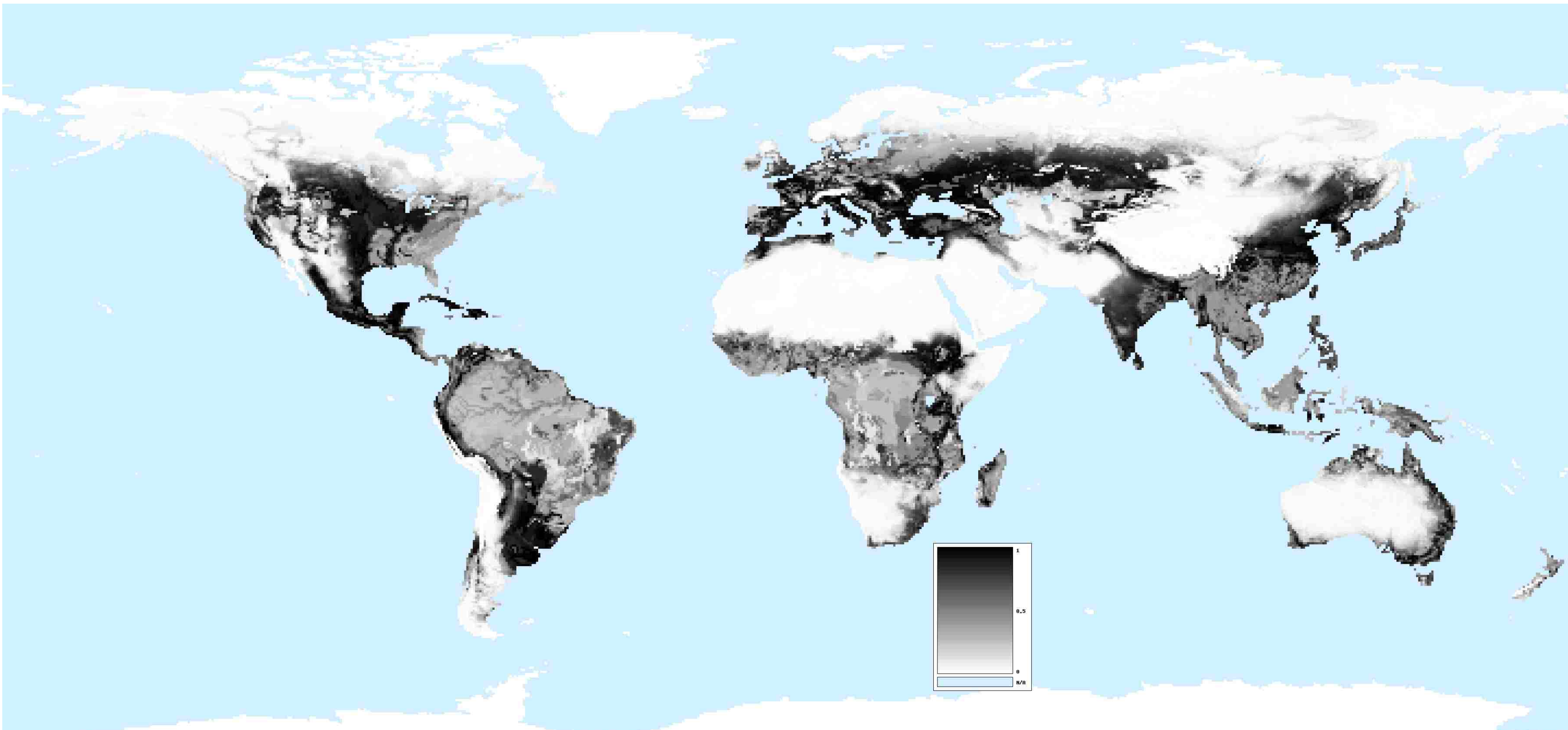


Figure 2.13: Agriculture Suitability, 2000

Note: The figure shows the data for land suitability for agriculture constructed by Ramankutty et al. (2002). Source: Atlas of the Biosphere (http://atlas.sage.wisc.edu).

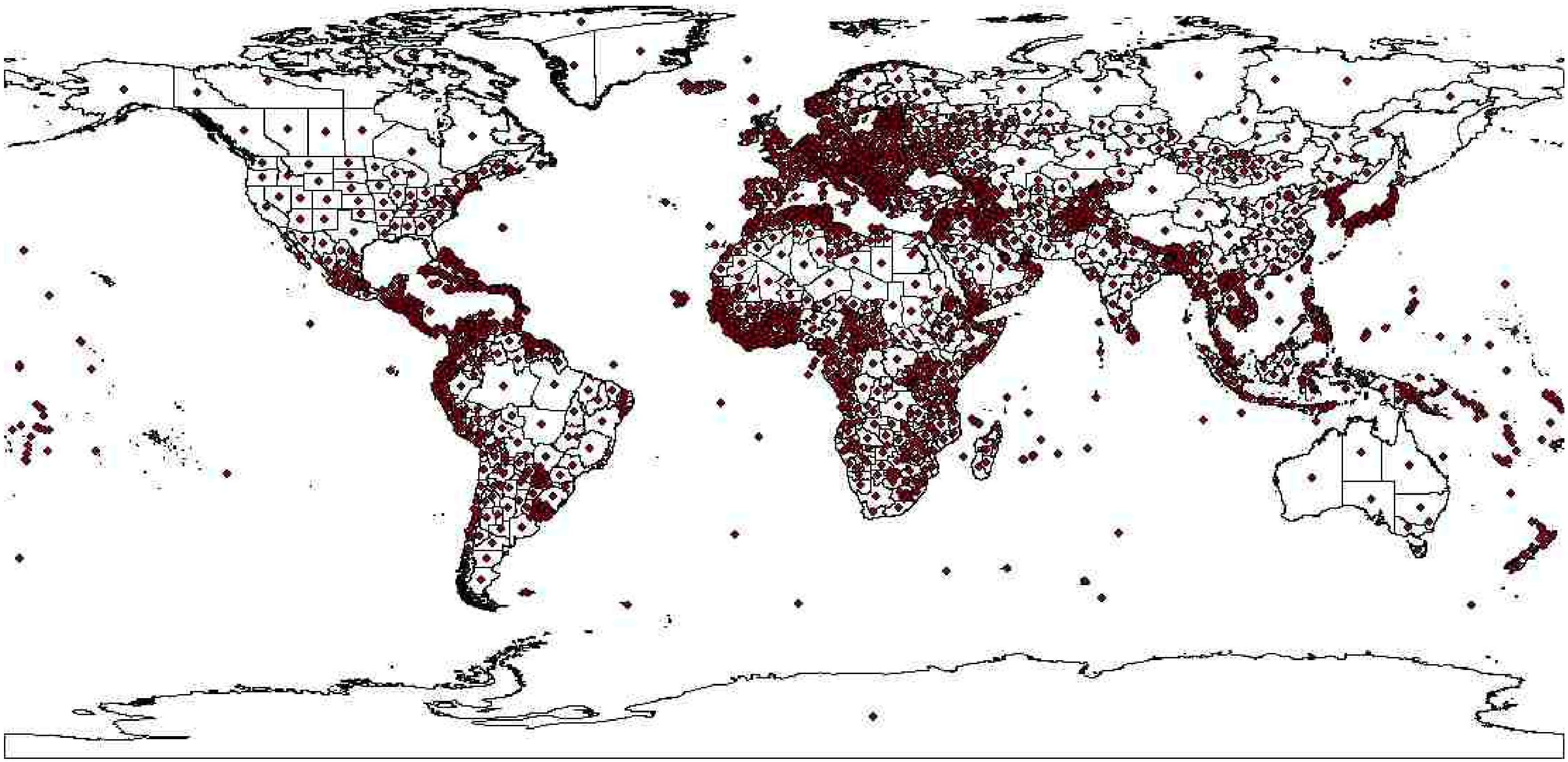


Figure 2.14: Centroids of Regions

Note: This figure shows both regions' boundaries and centroids in this study. A centroid is the median center of a region.

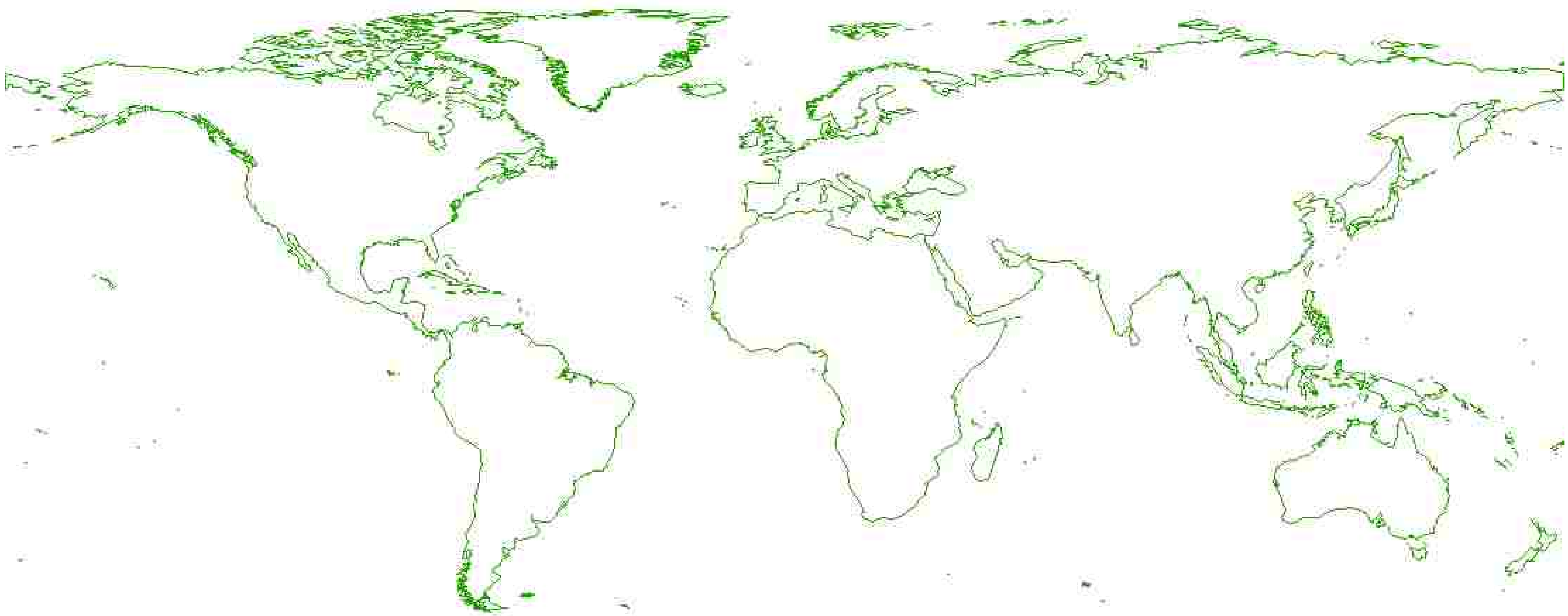


Figure 2.15: Coast

Note: The graph displays the data for boundary lines between land and ocean. The original data is from the National Centers for Environmental Information at http://www.ngdc.noaa.gov.

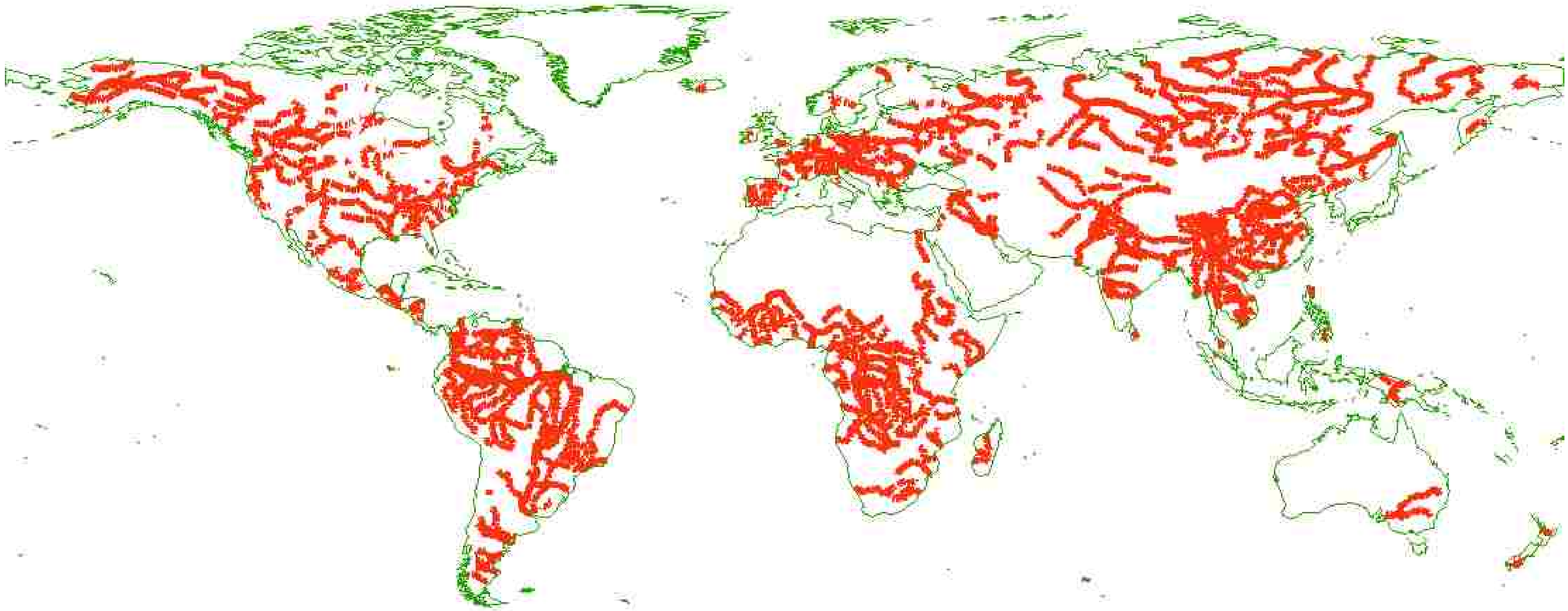


Figure 2.16: Rivers

Note: The graph displays the data for rivers. Red lines are the major rivers in the world. The original data is from the National Centers for Environmental Information at http://www.ngdc.noaa.gov.

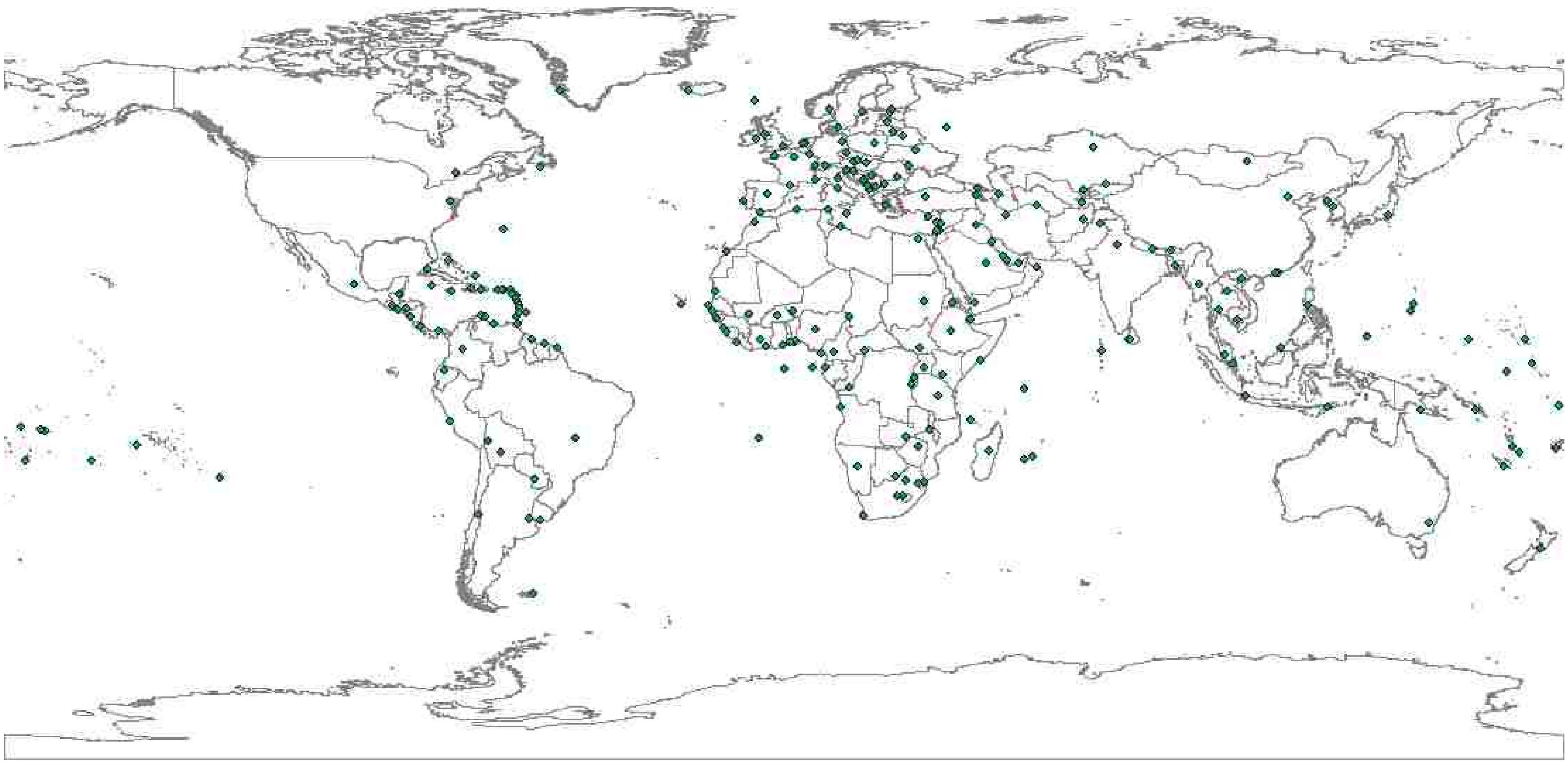


Figure 2.17: National Capital

Note: The figure displays the data for the world national capital cities. Spots in green are the locations of the national capital cities. The data is available at the World Urbanization Prospects, the 2011 Revision, and can be downloaded at http://esa.un.org/unpd/wup.

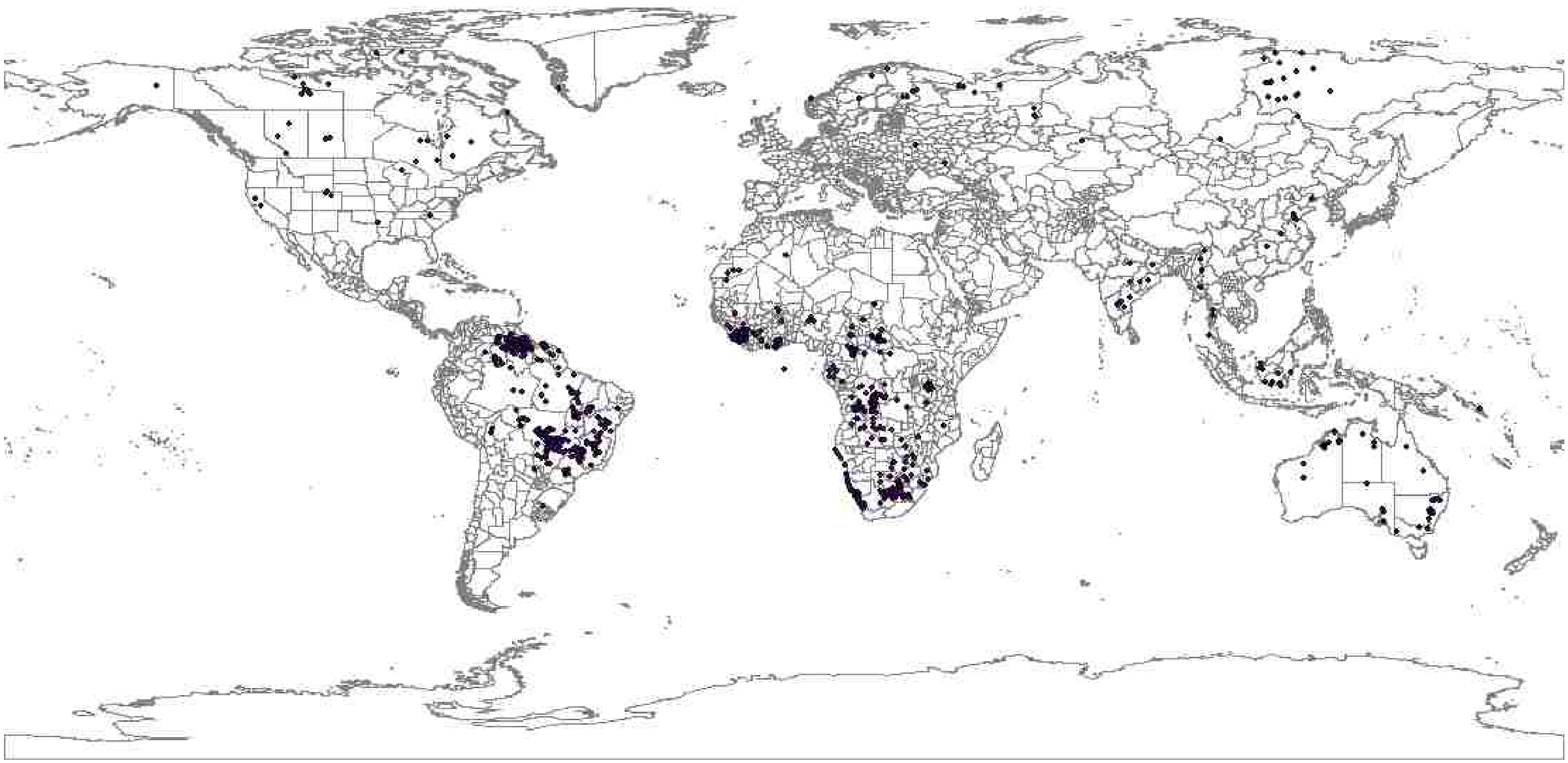
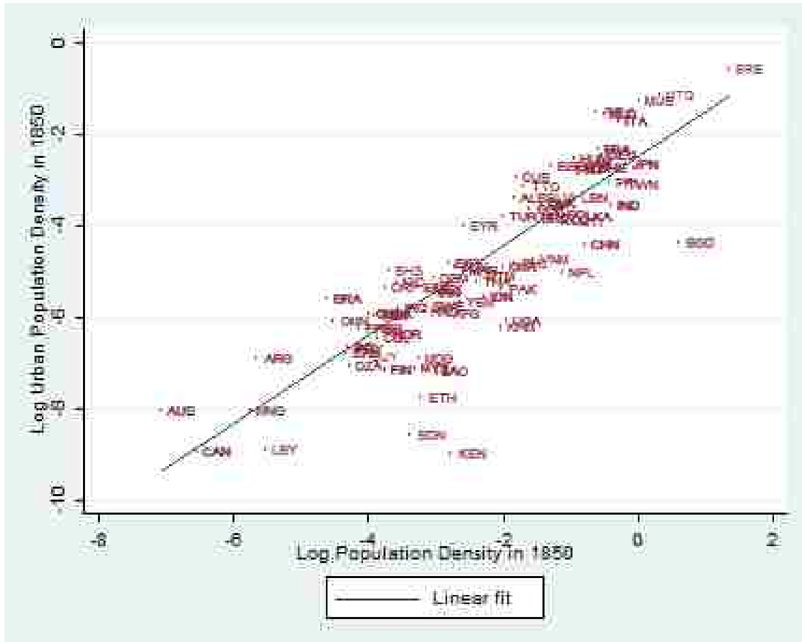
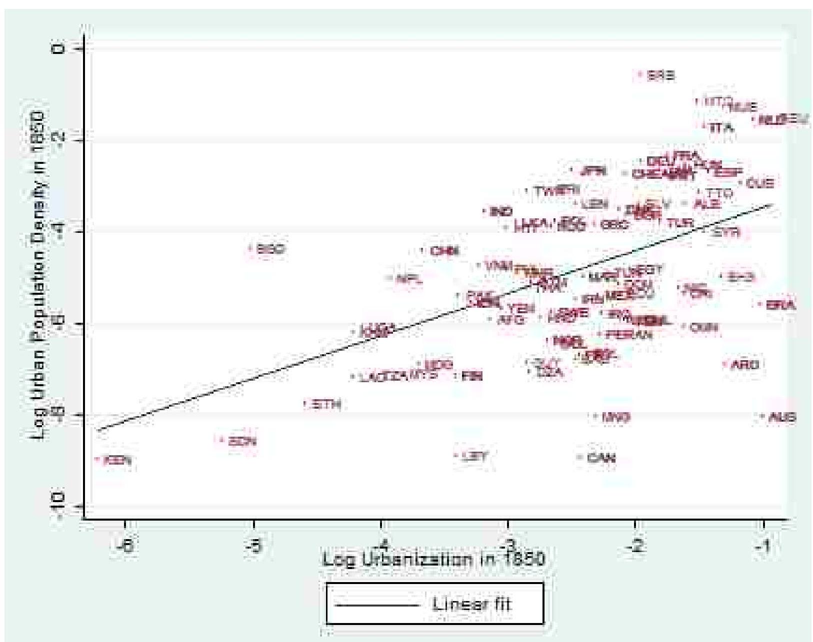


Figure 2.18: Diamond Mines

Note: This imagine shows the data for the world diamond locations. The data is available at the Peace Research Institute Oslo (PRIO) and can be downloaded at https://www.prio.org.



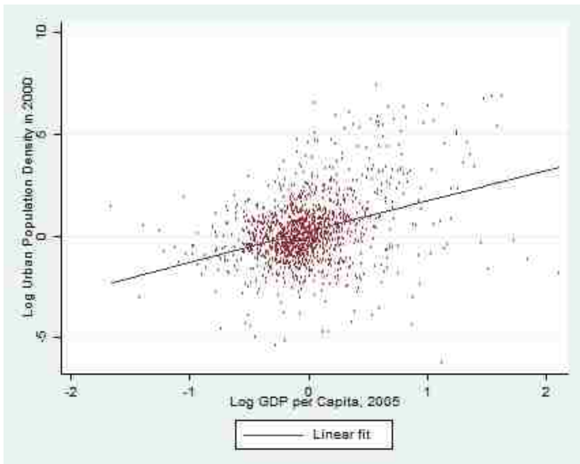
(a) Ln(Urban Population Density) & Ln(Population Density), in 1850
Correlation: 0.8761, P-value: 0.0000



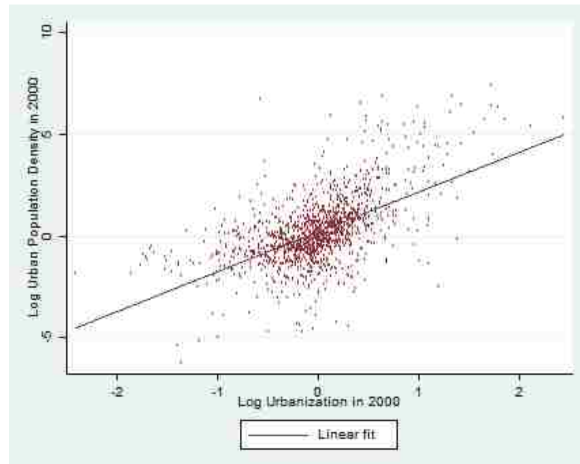
(b) Ln(Urban Population Density) & Ln(Urbanization), in 1850
Correlation: 0.4477, P-value: 0.0000

Figure 2.19: Log Urban Population Density, Urbanization Rate, and Log Population Density, 251 Sovereign Economies

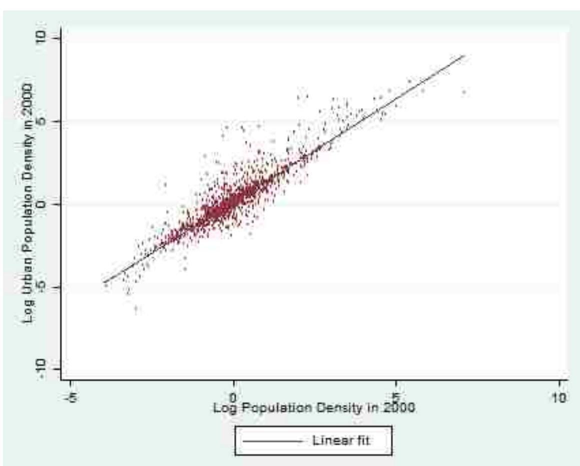
Note: A country's urban population in 1850 is the total population living in cities of the country. City in 1850 is defined with a minimum population threshold of 5,000. Data for total population in 1850 at the country level are collected from of world population history (1978). Unit of population density and urban population density is 100 persons per square kilometer.



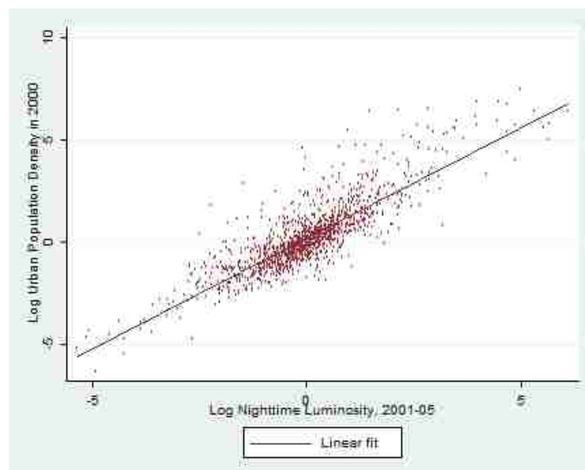
(a) Ln(Urb. Pop. Den.) & Ln(GDP per Capita)
Correlation: 0.3539, P-value: 0.0000



(b) Ln(Urb. Pop. Den.) & Ln(Urbanization)
Correlation: 0.5771, P-value: 0.0000



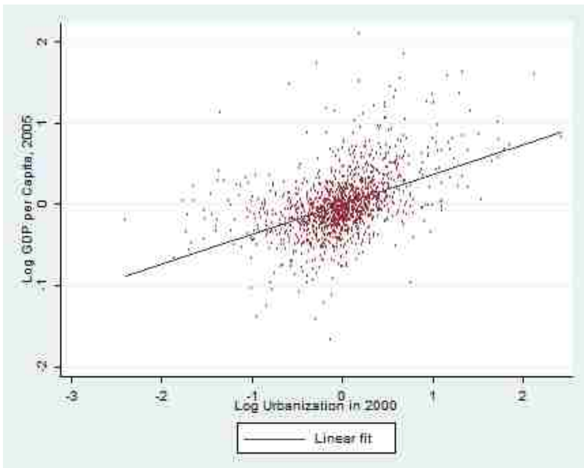
(c) Ln(Urb. Pop. Den.) & Ln(Pop. Den.)
Correlation: 0.9026, P-value: 0.0000



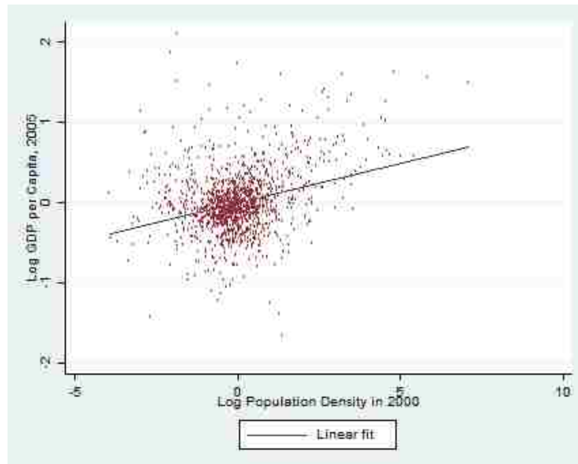
(d) Ln(Urb. Pop. Den.) & Ln(Nighttime Lights)
Correlation: 0.8482, P-value: 0.0000

Figure 2.20: Scatterplots of Country-demeaned Log Urban Population Density against Country-demeaned Log GDP per Capita, Log Urbanization, Log Population Density, and Log Nighttime Luminosity, 2000 - 2005, 1,380 regions

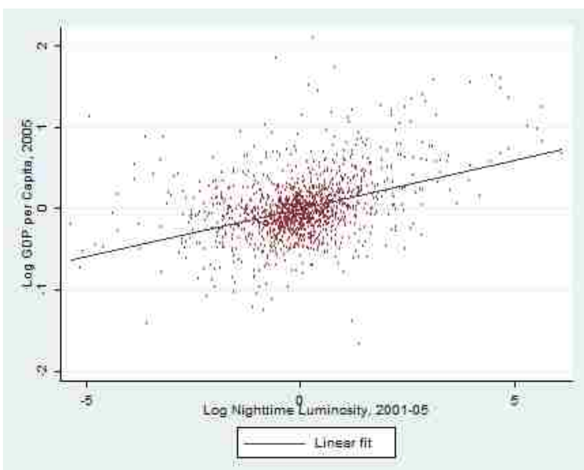
Note: Unit of population density and urban population density is 100 persons per square kilometer.



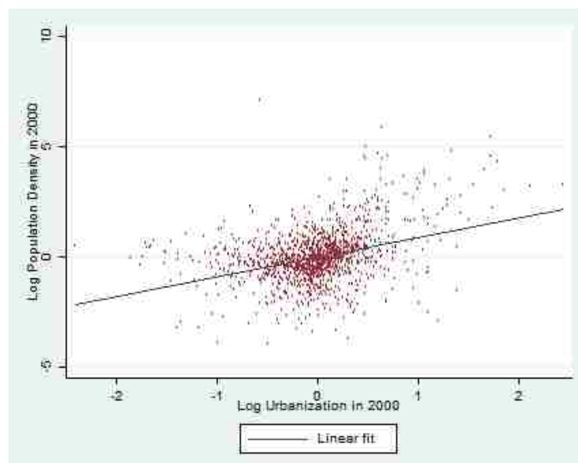
(a) Ln(GDP per Capita) & Ln(Urbanization) Correlation: 0.4597, P-value: 0.0000



(b) Ln(GDP per Capita) & Ln(Pop. Den.) Correlation: 0.3032, P-value: 0.0000



(c) Ln(GDP per Capita) & Ln(Nighttime Lights) Correlation: 0.3981, P-value: 0.0000



(d) Ln(Pop. Den.) & Ln(Urbanization) Correlation: 0.3660, P-value: 0.0000

Figure 2.21: Scatterplots of Country-demeaned Log GDP per Capita against Country-demeaned Log Urbanization, Log Population Density, and Log Nighttime Luminosity, 2000 - 2005, 1,380 regions

Note: Unit of population density and urban population density is 100 persons per square kilometer.

Chapter 3. Investigating Persistence, 1850-2000

3.1 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, if not millennia. At the sub-national level, 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. Despite this, it is often observed that the 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. San Francisco experienced a devastating earthquake in 1906 in which about 200,000 inhabitants were left homeless, but this had little effect on long run population growth (Vigdor, 2008). Similarly, 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] On the other hand, there are examples of regions like Louisiana in the US and the state of West Bengal in India, which, while having some of the highest levels of economic development in the past, have experienced relative declines within the past century.[2] Given the variety of experiences, in this paper I empirically explore the extent to which regional inequalities persist globally; whether they

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

[2] Easterlin (1960, p. 97) 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 has by all accounts experienced deindustrialization since India's independence in 1947.

are driven by geographical differences, whether they vary by continent, and various other such groupings. While there are numerous studies on persistence at the regional level, to the best of my knowledge I am the first to explore the same employing a global sample while simultaneously encompassing a period of one hundred and fifty years.

More specifically, I examine the relationship between contemporary and 1850 measures of regional economic development using a sample that covers 2,058 sub-national regions from 135 countries. For the year 1850, I construct a measure of urban population density - urban population relative to total land area - based on various sources of estimates of historical settlements such as Chandler (1987), Bairoch (1998), and Eggimann (1994). I supplement this measure with indicators to capture the existence of urban areas within a region, as well as its neighboring regions, urban population densities in neighboring regions, as well as quadratic versions of the density variables to capture non-linearities. My results overwhelmingly support worldwide "persistence of fortunes" at the sub-national level during the past 150 years. The existence of larger urban populations 150 years ago is significantly associated with higher regional income per capita in 2005 as well as other proxies of contemporary economic development such as urbanization rates and night-time light density. I control for country fixed effects and a large range of geographic factors commonly used in the literature. The results are also robust across different samples of countries grouped by continent, by their colonization history, and also semi-contemporary controls. I also briefly look for mechanisms through which urbanization 150 years ago affects current economic performance at the sub-national level. While not conclusive, I find that both human capital and physical capital, as measured by infrastructure, are more strongly associated with historical urban density than are cultural or institutional factors. I also find that regions in the US and Canada are exceptions to such persistence.

My 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 three reasons - 1) the number of cities covered by any source or even a combination of sources is likely to get more and more unreliable, 2) even if a city is recorded, population estimates are likely to be increasingly inaccurate as I go further back in time. Indeed, if I go back to 1750 or even 1800, the historical compilations are missing 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, and Tehran in Iran, to name a few.) It is also likely that more developed regions kept longer and more complete historical statistics records. In that case my 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, I believe 1850 remains instructive as a starting point since most countries had only just begun industrializing and integrating into the rest of the world. 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.

Theories that explain regional disparities in economic development emphasize the role of physical geography and the economics of agglomeration, both of which have implications for 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, it is considered that 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, Ottaviano, and Temple, 2013).

The economics of agglomeration postulates that there are advantages to agglomerations 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, Ottaviano, and Temple, 2013). These externalities 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) 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.

My research is inspired by recent advances in the regional economics literature. 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 in Latin America cross-municipality labor income differences within a country is twice as large as cross-country differences. With use of access to paved roads as a proxy for local institutions' efficiency in providing public goods, they show that such huge between-municipality differences are potentially attributed to varying quality of municipal institutions. Tabellini (2010), on the other hand, suggests that variation in institutions may be important in explaining cross-country inequality 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 sizable effect of education on regional GDP per capita (25 - 35 percent) 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. Based on a similar coverage of regions, Mitton (2013) finds no evidence of a positive effect of institutions on development. Acemoglu, Gallego, and Robinson (2014), on the other hand, argue that findings of Gennaioli et al. (2013) on the effects of education and institutions on regional economy are not reasonable and largely result from "bad control" documented by Angrist and Pischke (2008). By instrumenting for the current average years of schooling with the share of Protestant missionaries per 10,000 people in the early 20th century, they claim that the effect of human capital on income per capita returns to the reasonable range of 6 - 10 percent in regions within former colonial countries.

The remainder of the paper is organized as follows. In Section 3.2 I describe my regional measure of development in 1850, measures of contemporary development around 2000-2005, and the empirical framework. In Section 3.3 I present my results. In Section 3.4 I look at potential mechanisms for persistence. In Section 3.5, I briefly investigate persistence over 500 years. Section 3.6 concludes.

3.2 Subnational Data and Empirical Strategy

3.2.1 Measuring Development at the Regional level in 1850

Following Gennaioli et al. (2013), I define sub-national regions as first-level administrative divisions. The geographic boundaries are procured from the Database of Global Administrative Areas Map version 2 (GADMv2). I then construct my measure of development in 1850, urban population density, by using population estimates of urban settlements from various sources. I have provided a detailed explanation of the construction of urban population density in Chapter 2 Section 2.2.

In order to create a measure of urban population density, I need to first define what constitutes an urban location. 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 this work, I include any location that has a recorded population of 5,000 or more in 1850 from my sources.[3] I follow Acemoglu, Gallego, and Robinson (2014), Acemoglu et al. (2011), and Cantoni (2014) in this regard. With all the data sources taken together, I identify 2,803 settlements with populations of 5,000 or more in 1850 spanning 141 contemporary countries. However, there is nothing sacrosanct about the threshold value of 5000. Indeed, there are historical studies that use other thresholds. For example, when studying cities for the period 800-1800, Bosker, Buringh, and van Zanden (2013) only consider those that had at least 10,000 inhabitants. Nunn and Qian (2011) constructs national urbanization numbers for the period 1000-1900 using a much higher threshold of 40,000. In a later section of the paper, I examine the robustness of my results when using higher thresholds.[4] Mapping these settlements into my regional GIS map yields 766 sub-national regions with non-zero urban populations in 1850. For these 766 regions, the average urban population density is 33.6 persons per square kilometer with a standard error of 178.5 persons per square kilometer. In Figure 2.8, I 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. I report summary statistics of urban population density in Table 3.1.[5]

3.2.2 Measuring Outcomes

I use the logarithm of GDP per capita in 2005 as my main measure of contemporary regional prosperity. The data, which comes from Gennaioli et al, overlaps with 92 countries for which

[3] I are grateful to Omer Ozak and David Weil for sharing their data compiled from Bairoch (1998) and Eggimann (1994).

[4] One might wonder why I 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 settlements is far greater thus making my measure even more noisy.

[5] The distribution of cities across countries and continents according to different minimum population thresholds is summarized in Appendix Table A.1.The listing of all settlements with estimated population are displayed in an online appendix of this paper which can be downloaded at http://www.dachaoruan.com/#!research/clvf

I also have urban population data from 1850. This in turn yields 1,395 regions which forms my baseline sample. Of these, 668 had cities with populations greater than 5,000 in 1850. In Figure 3.1, I display boundaries of all subdivisions across the world. The areas shaded in dark red are the ones for which I have GDP per capita data. The areas with stripes are ones for which I have no information on settlements. If an entire present day country had no settlement recorded in any of my sources, it was completely dropped. I also dropped city-states which comprise of only one region.[6] Thus, my baseline analysis is based on the regions which are marked in red and and not striped.

Relying on GDP per capita alone means that I have fewer regions with contemporary income than I have with 1850 urban population data. Moreover, it is known that GDP is not accurately measured, especially in developing countries. One would expect this problem to be more severe at the sub-national level. Within a country GDP in richer regions may be more accurately reported than in poorer regions. To ensure that my conclusions are not driven by the drawbacks of regional income measurement, I use three additional measures of development. These are the log average nighttime light intensity using satellite data, the fraction of population living in cities (i.e. urbanization), and log population density. I have already discussed the merits of the last two. Nighttime luminosity using satellite data has become increasingly popular as a way to circumvent some of the problems related to measurement error in GDP. Henderson, Storeygard, and Weil (2012) and Hodler and Raschky (2014) have documented a positive correlation between GDP and nighttime luminosity at the country level and regional level, respectively. An increasing number of studies focusing on research questions at the sub-national level also rely on satellite data.[7] Using these outcomes allows us to expand the coverage back to 135 countries.

[6]However, when examining geographical spillovers, I, of course, include the information from these two groups.

[7]For example, Storeygard (2013) and Alesina, Papaioannou, and Michalopoulos (2012) use nighttime luminosity to study urbanization and ethnic divisions in Sub-Saharan Africa.

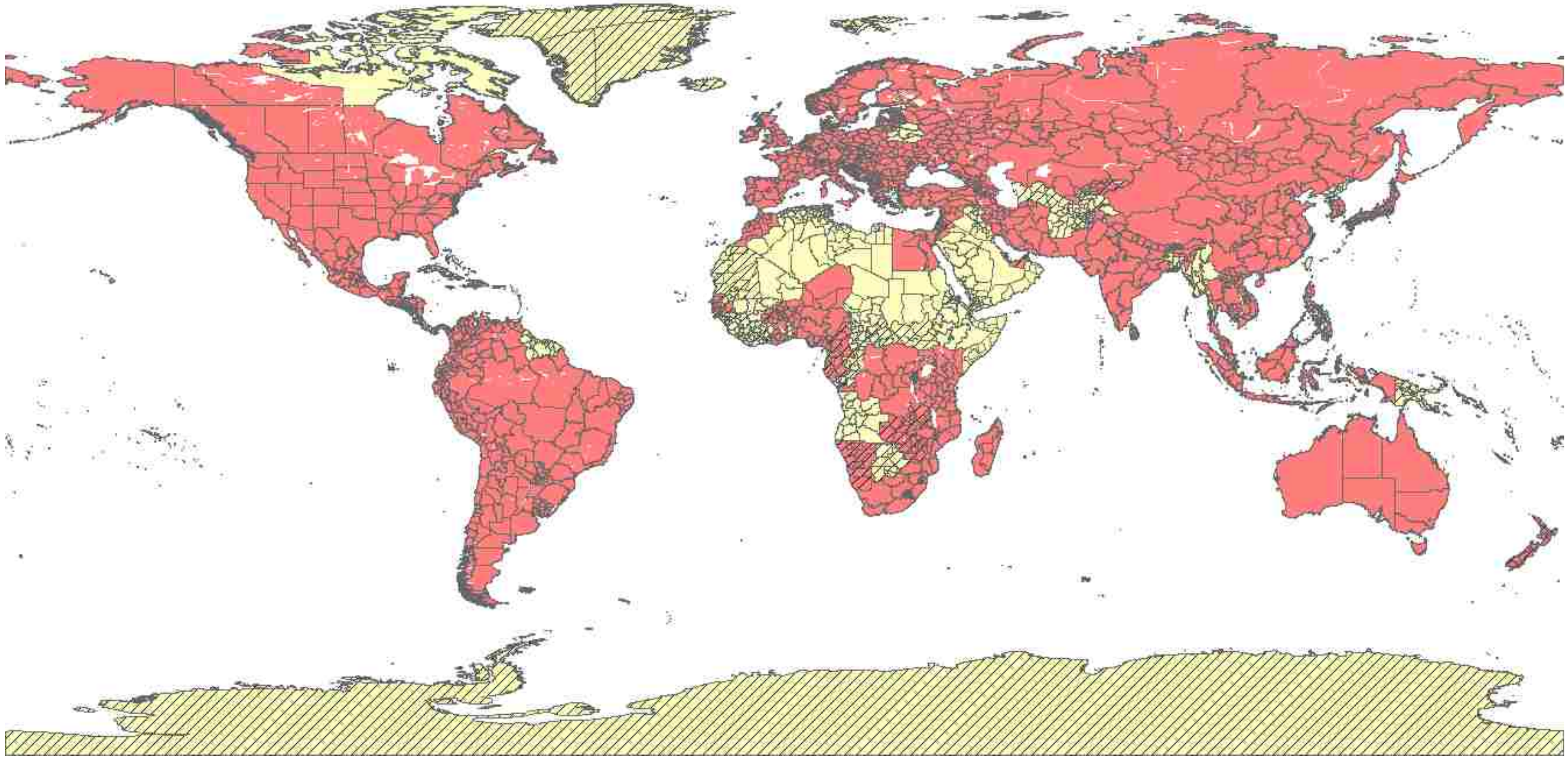


Figure 3.1: Subdivisions across The World

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

Table 3.1 lists summary statistics for these four outcomes. Among them, for the 1,395 regions for which I have GDP per capita data, the mean in 2005 (PPP) is 12,652 US dollars

Table 3.1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
1850 Urbanization Measures:					
Existence of a City (1850) (A)	0.372	0.484	0.000	1.000	2058
- Regions with Regional Income, 2005	0.479	0.499	0.000	1.000	1395
Urban Population Density in 1850 (A)	0.336	1.785	0.000	34.027	766
- Regions with Regional Income, 2005	0.232	1.126	0.000	15.587	668
Existence of a City in Neighboring Regions (1850) (A)	0.467	0.499	0.000	1.000	2058
Urban Pop. Den. in Neighboring Regions, 1850 (A)	1.503	9.548	0.000	198.034	962
Dependent Variables:					
Regional GDP pc in 1,000 USD, Constant 2005 PPP (B)	12.652	13.387	0.070	143.483	1395
Ln(Avg. Nighttime Light Density), 2001-05 (C)	0.257	2.494	-10.776	4.143	2044
Urbanization Rate, 2000 (D)	0.432	0.288	0.000	1.000	2050
Population Density, 2000 (D)	2.855	10.259	0.000	219.105	2058
Urban Population Density, 2000 (D)	2.818	19.720	0.000	674.283	2058
Regional Controls:					
Land Suitability (E)	0.359	0.318	0.000	0.998	2058
Temperature in Celsius (F)	16.719	8.419	-15.421	29.588	2058
Altitude in 100 Meters (F)	5.480	6.366	-0.138	48.786	2058
Rainfall in Meter (F)	1.094	0.746	0.001	5.405	2058
Ruggedness in 100 Meters (G)	1.363	1.354	0.000	9.990	2058
Absolute Value of Latitude (H)	28.902	16.842	0.000	71.000	2058
Proximity to the Coast (C)	0.838	0.162	0.327	1.000	2058
Proximity to Rivers (C)	0.832	0.163	0.210	1.000	2058
Largest National City in 1850 (A)	0.069	0.254	0.000	1.000	2058
Additional Regional Controls:					
Presence of National Capital (I)	0.063	0.243	0.000	1.000	2058
Proximity to Capital City (I)	0.761	0.196	0.076	1.000	2058
Proximity to Borders (C)	0.823	0.161	0.163	1.000	2058
Presence of Diamond Mines (J)	0.067	0.250	0.000	1.000	2058
Ln(Oil Production per Capita) (B)	0.108	0.409	0.000	4.161	1395
Years of Education, 2005 (B)	7.302	3.073	0.252	13.210	1358

Note: Unit of population density is 100 persons per square kilometers. Existence of a City (1850) is a dummy variable taking the value one if there is at least one urban settlement with a population of 5,000 in 1850. Sources of data are listed as below, and a detailed explanation on these variables is in Appendix Table A.2. A: Chandler (1987), Bairoch (1988), and Eggiman (1994). B: Gennaioli et al. (2013). C: the National Geophysical Data Center (NOAA). D: the Center for International Earth Science Information Network (CIESIN). E: Atlas of the Biosphere. F: Global Climate Data (WorldClim). G: Nunn and Puga (2012). H: Global Administrative Areas (GADM). I: World Urbanization Prospects. J: Peace Research Institute Oslo (PRIO).

with a standard deviation of 13,387 dollars. The mean value of luminosity is 0.257 and standard deviation of it is 2.494. Urbanization in 2000 has a mean of 0.432 and a stand deviation of 0.288. Population density in 2000 has a mean of 286 persons per square kilometer and a standard deviation of 1,026 persons per square kilometer. I provide detailed descriptions and the sources of those variables in Appendix Table A.2.

3.2.3 Empirical Strategy

My 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. At the same time more than half the regions in my sample have zero urban population density. To ensure that my results are not driven by this demarcation, I also use a dummy variable taking a value of one if urban density is greater than zero. I will refer to this loosely as the 1850 city dummy. Nevertheless, one should be cautious in the interpretation of the coefficient of this variable. It does not mean that these areas had zero urban populations with zero variation. Secondly, my definition of a region is based on current maps and not those of 1850. Indeed, many of my regions did not exist in their current boundaries one and a half centuries ago, that is if they did at all. Hence, my 1850 density may not be the relevant measure. Even if they did exist, spatial spillovers between adjoining regions is a well documented phenomenon. To deal with these issues of mis-measurement and spatial correlation, I add two more variables. First, I add a dummy variable identifying whether one or more cities existed within 25 miles geodesic distance from the border of the current regions, hereafter year 1850 neighboring city dummy.[8] I use 25 miles as a range of neighboring areas since there are no theoretical reasons to favor other distances. In the appendix, I also report results using 50 miles, 75 miles, and 100 miles as alternative ranges of neighboring areas for robustness checks. Based on those surrounding cities, I generate population density that equals the ratio of aggregated

[8]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.

population in neighboring cities to land area of regions, hereafter year 1850 neighboring urban population density. Positive (negative) spillovers from neighboring cities suggests positive (negative) signs for both the year 1850 neighboring city dummy and neighboring urban population density in 1850.

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

$$\begin{aligned}\mathrm{Y}_{i,2005} = \alpha + \beta_1 \text{City Dummy}_{i,1850} + \beta_2 \text{UrbPopDensity}_{i,1850} + \beta_3 \text{UrbPopDensity}^2_{i,1850} \\ + \beta_4 \text{NeibCityDummy}_{i,1850} + \beta_5 \text{NeibUrbPopDensity}_{i,1850} \\ + \beta_6 \text{NeibUrbPopDensity}^2_{i,1850} + X_i \delta + \mu_c + \varepsilon_i\end{aligned} \tag{3.1}$$

where $\mathrm{Y}_{i,2005}$ mainly represents log GDP per capita for region i in year 2005. I also use log average nighttime light intensity over 2001-2005, the degree of urbanization in 2000, and the log population density in 2000 as alternative outcome variables. $\text{CityDummy}_{i,1850}$ is the year 1850 urban dummy of the i^{th} region. $\text{UrbPopDensity}_{i,1850}$ is the year 1850 urban population density of the i^{th} region. $\text{NeibCityDummy}_{i,1850}$ is the year 1850 neighboring city dummy of the i^{th} region. $\text{NeibUrbPopDensity}_{i,1850}$ is the year 1850 neighboring urban population density of the i^{th} region. The vector X_i represents a comprehensive set of regional geographic factors commonly used in the literature including land suitability, temperatures in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in meters, absolute value of latitude, inverse distance to the coast, and inverse distance to a river. The term μ_c represents country fixed effects. To account for nonlinearity in the relationship between urban population density and income per capita, quadratic terms for both year 1850 urban population density and year 1850 neighboring urban population density, are all included in the equation. In a supplementary specification, I also include a dummy variable that equals one if a nation's most populous city in 1850 was in that region. Regions having such cities might be economically and politically important and have a relatively large urban population density compared to other regions within countries in 1850. Doing this enables us to see the extent to which my results are driven by this small group of regions.[9] Nevertheless, I report some of these regressions to show there are generally consistent elasticities of income today with respect to the year 1850 urbanization density variables in Table A.7.

I check the robustness of results through several strategies. First, I replace per capita GDP in 2005 with other contemporary measures of regional development during 2000 to 2005 such as log average nighttime light intensity, urbanization in 2000, and population density in 2000. Second, I include additional controls such as the inverse distance to capital city, an indicator that the capital city exists in a region, the inverse distance to borders, an indicator that the largest city in 1850 within a contemporary country existed in a region, and an indicator that diamond mines exist in a region. Third, I stratify regions with urban population in 1850 into 5 groups based on urban population density in 1850, and replace urban population density in 1850 with these group dummies. Fourth, I reconstruct my urbanization variables based on different minimum population thresholds and consider neighboring urbanization variables in various distances from the border. In addition, I investigate the existence of persistence in various samples of nations according to continent groups and ex-colonial status. I also look at how the effect of urban population density is differentiated by the largest city size in regions and size of urban population. In the last check, I run quantile regressions.

Following the investigation of early urbanization effects on contemporary economic development, I look for potential channels of persistence. I regress contemporary variables of education, culture, institutions, and infrastructure on year 1850 urbanization variables based on similar model specifications to Equation 3.1.

[9]One might wonder why, in the initial scatterplots, I used log urban population density and yet in the regressions I do not use logarithmic values. The simple reason is the presence of zero values. Nevertheless, as I show later, my results hold when I restrict the sample to only those regions with positive urban population density and used logs instead. Another alternative is to I substitute urban population density in 1850, neighboring urban population density in 1850, and their quadratic terms into log (urban population density in 1850 + 0.00001) and log (neighboring urban population density in 1850 + 0.00001). Year 1850 city dummy and year 1850 neighboring city dummy remain but are not reported in tables.

Finally, I attempt to go back further to year 1750 and year 1500, and look for a link between the past and today spanning a longer time horizon.

3.3 Results

3.3.1 Baseline Results

Table 3.2 presents four model specifications regressing log income per capita in 2005 on the year 1850 urbanization variables for regions whose income data in 2005 are available and whose countries had settlements in 1850 according to my city data. All estimates include country fixed effects and robust standard errors clustered at the country level are shown in parentheses. I present both within-country and between-country R^2 in regressions. Column (1) is the most parsimonious model in which I capture the early urbanization effect on log income per capita in 2005 through both the year 1850 urban dummy and year 1850 urban population density. The coefficient of the dummy is 0.087 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 results suggest that regions that had cities in 1850 were likely to record 9 percent greater GDP per capita in 2005. Furthermore, among the regions that did have cities, every additional 100 urban residents per square kilometer was associated with another 10 percent higher GDP per capita. The two urbanization variables together explain 4 percent of within-country variation of income per capita in 2005.

In column (2) I consider the contribution of urbanization of surrounding cities in 1850 to income per capita today, and therefore I add a year 1850 neighboring urban dummy and year 1850 neighboring urban population density. Coefficients of both variables are small in magnitude and insignificant. Coefficients of year 1850 urban dummy and year 1850 urban population density remain close to their values in column (1). Both within-country and between-country R^2 show little change compared to column (1). Columns (3) and (4) assume quadratic effects for both year 1850 urban population density and year 1850 neighboring urban population density. Negative signs of squared density variables indicate that the effects of year 1850 urban population density and year 1850 neighboring urban population density on per capita GDP in 2005 are concave. Substantial increases in within-country R^2 compared to columns (1) and (2) also support models with the quadratic forms.

Table 3.2: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	Dependent Variable: Log of Regional GDP per Capita (PPP), 2005			
	(1)	(2)	(3)	(4)
Existence of a City (1850)	0.086***	0.097***	0.058*	0.070***
	(0.029)	(0.025)	(0.029)	(0.027)
Urban Population Density 1850	0.095***	0.089***	0.353***	0.335***
	(0.024)	(0.024)	(0.064)	(0.065)
Square Urban Pop. Den. 1850			-0.023***	-0.023***
			(0.006)	(0.006)
City in Neighboring Regions (1850)		-0.049		-0.045
		(0.049)		(0.046)
Urb. Pop. Den. in Neib. 1850		0.009		0.065**
		(0.008)		(0.031)
Squ. Urb. Pop. Den. in Neib. 1850				-0.002**
				(0.001)
Countries	92	92	92	92
Observations	1395	1395	1395	1395
within R^2	0.04	0.05	0.08	0.09
between R^2	0.24	0.17	0.25	0.19

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. No controls are included. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I present results of my favored model specification in column (4). Estimated coefficients suggest that regions had cities in 1850 were associated with higher GDP per capita in 2005, which is significant at 5 percent. Among regions with urban population in 1850, every additional 100 urban residents per square kilometer (about one standard deviation) was correlated with another 36.6 percent (33.5 minus 2.3 log points) higher GDP per capita, with a significance of 1 percent. Spatial correlation that refers to spillovers of urban development in neighboring areas (using 25 miles away from regions' boundaries) is captured by year 1850 neighboring urban population density and year 1850 neighboring city dummy. My estimates suggest that a unit change in neighboring urban population density in 1850 was associated with 6.5 percent (6.5 minus 0.2 log points) higher GDP per capita, with a significance of 5 percent. But whether there existed a city in neighborhood areas whose population in 1850 is slightly higher than 5,000 had no prediction about regional income differences in 2005. Positive spillovers are generally supported by results.

I further restrict the sample to regions with a city in 1850 in Appendix Tables A.3. The results remain consistent. In addition, I use 50 miles, 75 miles, and 100 miles as alternative ranges of neighboring areas and reconstruct neighboring urban population density and neighboring city dummy for estimation. I report results in Appendix Table A.6, and the magnitude of coefficients on urban population density and its significance change moderately. Moreover, the coefficients of neighboring city and neighboring urban population density are attenuated with a longer distance used for defining neighboring area. For example, using 100 miles from regions' boundaries to construct neighboring area, the coefficient of neighboring urban population density in 1850 decreases from 0.065 to about 0.006 and become insignificant.

3.3.2 Urbanization in 1850, Physical Geography, and Development

One might be concerned that the association between contemporary income and my early urbanization variables may simply represent the influence of environmental characteristics on contemporary income. The importance of physical geography in explaining economic activity has been extensively studied in the literature. Physical geography shapes contemporary income inequalities through a number of channels. First, some geographic and climatic characteristics have a direct impact on economic activities over centuries (Davis and Weinstein, 2002). Second, some of them have played an important role in shaping culture and social norms which persists over hundreds of years (Breinlich, Ottaviano, and Temple, 2013). Third, they may have triggered path dependence in agglomerations hundreds of years ago though their economic advantages may have long faded away (Bleakley and Lin, 2012). While I do not have a way to completely disentangle these channels, nevertheless I can control for a host of variables connected to physical geography, and see if the persistence result continues to survive.

Physical geography can be captured in many ways, among which temperatures, land suitability, ruggedness of terrain, latitude, and proximity to the coast are highlighted in recent studies. For example, Dell, Jones, and Olken (2012) find cross-country evidence that higher temperatures have negative effects on agricultural output, industrial output, and political stability. In addition, temperatures and annual precipitation are negatively associated with growth rates. Proximity to the coast measures ease of ocean access, and a shorter distance to coast is often regarded as an advantage to external trade (for example, Frankel and Romer (1999)). Ruggedness is expected to adversely affect productivity. For example, high elevation and ruggedness means higher costs of economic activities such as construction and transportation. Nunn and Puga (2012) find evidence showing a negative impact of ruggedness on economic development is generally true across countries in the world. Absolute value of latitude measures the general distance away from the equator. A longer distance to the equator relates less severe disease environment, less tropical area, and lower temperatures which are beneficial to development (see e.g. Acemoglu, Johnson, and Robinson (2002)). While the number of ways to capture geographic diversity has proliferated, my core set of variables are drawn from Michalopoulos (2012) who uses an exhaustive palette.

I investigate the concern about physical geography in Table 3.3. In column (1), I report impacts of geographic and climatic characteristics on log income per capital in 2005 without including measures of development in 1850. Temperatures and rainfall both have negative impacts on regional income as expected, though the effect of temperatures is insignificant and rainfall is only significant at 10 percent. Land suitability has a negative and significant impact on income today and is consistent with recent regional studies (Mitton, 2013; Maloney and Valencia Caicedo, 2015). Elevation and terrain ruggedness both have expected effects on income. Nunn and Puga (2012) and Mitton (2013) both find significantly negative impact of ruggedness on regional income. An expected positive correlation between proximity to ocean and regional income is also supported in my findings. The coefficient of inverse distance to river is positive but insignificant. All together, the 8 geographic and climatic variables explain 15 percent of within-country variation and 52 percent of between-country variation.

In column (2), I include a dummy variable indicating regions in which nations' most populous cities in 1850 existed. This small group of regions might be political and economically crucial to their countries and have a relatively high urban population density to other regions within countries in 1850. The dummy therefore enables us to observe the extent to which my results are driven by these regions. My estimates in column (2) indicate that including the dummy affects geography coefficients - both rainfall and ruggedness become insignificant.

I include my measures of development in 1850 in columns (3) and (4). Persistence remain significant but has slightly lower magnitude in column (3) than in column (4) of Table 3.2. The effect of the year 1850 city dummy is basically unchanged (the coefficient decreases from 0.07 to 0.068) and the impact of urban population density in 1850 declines from 0.335 to around 0.265. However, some geography factors - ruggedness, absolute value of latitude, and rainfall - turn insignificant, suggesting these factors are likely to function as a trigger of early development which persists over hundreds of years according to path dependence theory. Models with geography controls have within-country R^2 of around 20 percent and between-country R^2 of around 50 percent. Take two regions in China, Jiangsu and Sichuan, as an example. Jiangsu had an urban population density of 7.9 persons per square kilometer in 1850, while Sichuan had 0.5 persons per square kilometer.

Table 3.3: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850 and Geographic Controls

	Dependent Variable: Log of Regional GDP per Capita (PPP), 2005			
	(1)	(2)	(3)	(4)
Existence of a City (1850)			0.068**	0.018
			(0.028)	(0.029)
Urban Population Density 1850			0.267***	0.189***
			(0.056)	(0.057)
Square Urban Pop. Den. 1850			-0.018***	-0.014***
			(0.005)	(0.005)
City in Neighboring Regions (1850)			-0.035	-0.035
			(0.032)	(0.031)
Urb. Pop. Den. in Neib. 1850			0.047	0.059**
			(0.030)	(0.029)
Squ. Urb. Pop. Den. in Neib. 1850			-0.001	-0.002**
			(0.001)	(0.001)
Temperature	-0.018	-0.018	-0.023	-0.021
	(0.015)	(0.015)	(0.017)	(0.016)
Land suitability	-0.202***	-0.195***	-0.154***	-0.153***
	(0.054)	(0.054)	(0.055)	(0.055)
Elevation (100 meters)	-0.010	-0.011	-0.013	-0.012
	(0.008)	(0.008)	(0.009)	(0.008)
Ruggedness	-0.049**	-0.038	-0.029	-0.027
	(0.023)	(0.023)	(0.022)	(0.022)
Rainfall in meter	-0.080*	-0.067	-0.066	-0.061
	(0.048)	(0.047)	(0.048)	(0.047)
Abs. (latitude)	0.013**	0.012**	0.010	0.011*
	(0.006)	(0.006)	(0.007)	(0.006)
Proximity to the coast	1.012***	0.913***	0.889***	0.869***
	(0.189)	(0.186)	(0.181)	(0.180)
Proximity to a river	0.267	0.205	0.195	0.177
	(0.189)	(0.199)	(0.194)	(0.197)
Largest National City in 1850		0.362***		0.278***
		(0.046)		(0.058)
Countries	92	92	92	92
Observations	1395	1395	1395	1395
within R^2	0.15	0.21	0.20	0.23
between R^2	0.52	0.52	0.56	0.54

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

In column (4), the coefficient of dummy that nation's largest city in 1850 existed in a region is 0.278 and is significant at 1 percent. This is evidence of persistence for those regions. Including the dummy lowers coefficients of year 1850 city dummy and urban population density in 1850. For example the impact of urban population density in 1850 declines to 0.189 but it is still significant at 1 percent. Therefore, persistence is evident in many other regions than just a small group of prominent regions.

As an additional robustness check, I consider five additional contemporary controls - inverse distance to capital, inverse distance to borderlines, an indicator equal to one if national capital city exists in a region, an indicator that diamond mines exist, and log regional population in 2000. I perform a similar exercise as in Table 3.3 and report estimates in Table 3.4. I first display their effects on regional income excluding development in 1850 in columns (1) and (2), and show results based on these variables and development in 1850 in columns (3) and (4). My results show that the indicator for existence of the national capital city is the only one that has a statistically significant impact on income today. The coefficient of year 1850 city dummy is close to 0 and becomes insignificant, while the coefficient of urban population density in 1850 remain significant at 5 percent though its magnitude falls substantially. These two coefficients are likely to be downward biased as the additional 5 contemporary controls are included as most of them are potentially endogenous and positive correlated early urbanization.

The coefficient on the dummy for the nation's most populous city in 1850 declines substantially and becomes insignificant in columns (2) and (4). Its effect is likely to be taken by the dummy of capital city, as among the 92 capital cities 57 were the largest cities within countries in 1850. Nevertheless, including the 5 contemporary controls that are potentially endogenous does not alter the existence of persistence.

Table 3.4: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850 and Other Contemporary Controls

	Dependent Variable: Log of Regional GDP per Capita (PPP), 2005			
	(1)	(2)	(3)	(4)
Existence of a City (1850)			0.003	-0.002
			(0.024)	(0.024)
Urban Population Density 1850			0.118**	0.112**
			(0.050)	(0.051)
Square Urban Pop. Den. 1850			-0.009**	-0.009**
			(0.004)	(0.004)
City in Neighboring Regions (1850)			-0.037	-0.038
			(0.036)	(0.036)
Urb. Pop. Den. in Neib. 1850			0.067**	0.068**
			(0.028)	(0.028)
Squ. Urb. Pop. Den. in Neib. 1850			-0.002**	-0.002**
			(0.001)	(0.001)
Proximity to Capital City	0.008	0.011	-0.010	-0.005
	(0.125)	(0.125)	(0.118)	(0.119)
Proximity to Borders	-0.129	-0.131	-0.137	-0.138
	(0.278)	(0.280)	(0.283)	(0.284)
Presence of National Capital	0.580***	0.542***	0.535***	0.517***
	(0.047)	(0.061)	(0.054)	(0.059)
Presence of Diamond Mines	0.027	0.024	0.028	0.026
	(0.062)	(0.062)	(0.063)	(0.063)
Ln(Population)	-0.001	-0.003	0.004	0.003
	(0.023)	(0.024)	(0.023)	(0.024)
Largest National City in 1850		0.070		0.040
		(0.052)		(0.053)
Countries	92	92	92	92
Observations	1395	1395	1395	1395
within R^2	0.29	0.29	0.30	0.30

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.3.3 Alternative Measures of Economic Development

To address drawbacks of using log GDP per capita in 2005 as a outcome - limitations on sample size and varying degrees of measurement error across regions, I use three alternative measures of contemporary development commonly used in regional economics, e.g. log average nighttime light intensity (Hodler and Raschky, 2014), urbanization rate, and log population density (Rappaport and Sachs, 2003). In order for persistence to be supported, positive relationships between urbanization variables in 1850 and the level of contemporary development using alternative measures are expected.

Estimates using alternative outcomes are displayed in Table 3.6. In any case, regions that had cities in 1850 are associated with higher level of contemporary development and among regions with urban population in 1850, every additional 100 urban residents per square kilometer was correlated with a higher level of development with a quadratic effect. The coefficients are significant at 1 percent. Positive spillovers are supported. Overall, using alternative measures of economic development leads to the same conclusion as using log GDP per capita does. I also reduce my sample to 1,395 regions in which regional income is available and the conclusions remain the same.

3.3.4 Is The Relationship Monotonic ?

Various results thus far have revealed a positive and concave relationship between urban population density in 1850 and the level of development around 2000 to 2005. However, the evidence of persistence is not widely guaranteed for all regions in the distribution of urban population density. For example, what if the positive correlation is driven by extremely high and low levels of urban population density? If the relationship is generally continuous, one would see a pattern in a pair of numerical coordinates that a region's contemporary

Table 3.5: Regressions of Other Development Outcomes on Urbanization in 1850

	Panel A: Ln(Average Nighttime Light Density, Averaged over 2001-2005)			Panel B: Fraction of Population Living in Cities in 2000		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City	0.979***	0.812***	0.569***	0.123***	0.114***	0.073***
1850	(0.107)	(0.098)	(0.089)	(0.019)	(0.019)	(0.017)
Urban Population	0.395***	0.321***	0.171**	0.082***	0.069***	0.043**
Density 1850	(0.091)	(0.082)	(0.075)	(0.018)	(0.017)	(0.017)
Square Urban Pop.	-0.013***	-0.010***	-0.006**	-0.003***	-0.003***	-0.002***
Den. 1850	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
City in Neighboring	0.618***	0.455***	0.486***	-0.021	-0.023*	-0.018
Regions 1850	(0.096)	(0.085)	(0.083)	(0.014)	(0.014)	(0.013)
Urb. Pop. Den. in	0.031*	0.026*	0.028*	-0.001	-0.002	-0.001
Neib. 1850	(0.017)	(0.016)	(0.016)	(0.006)	(0.006)	(0.006)
Squ. Urb. Pop. Den.	-0.000**	-0.000*	-0.000**	-0.000	-0.000	-0.000
in Neib. 1850	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Largest National			1.020***			0.173***
City in 1850			(0.170)			(0.028)
Baseline Controls	No	Yes	Yes	No	Yes	Yes
Countries	135	135	135	135	135	135
Observations	2044	2044	2044	2050	2050	2050
within R^2	0.14	0.25	0.27	0.10	0.13	0.16

	Panel C: Ln(Population Density in 2000)		
	(1)	(2)	(3)
Existence of a City	0.882***	0.751***	0.518***
1850	(0.099)	(0.090)	(0.091)
Urban Population	0.660***	0.622***	0.478***
Density 1850	(0.119)	(0.118)	(0.112)
Square Urban Pop.	-0.021***	-0.019***	-0.016***
Den. 1850	(0.004)	(0.004)	(0.004)
City in Neighboring	0.621***	0.462***	0.492***
Regions 1850	(0.097)	(0.078)	(0.076)
Urb. Pop. Den. in	0.044*	0.042*	0.044*
	(0.025)	(0.024)	(0.025)
Urb. Pop. Den.	-0.000*	-0.000*	-0.000*
in Neib. 1850	(0.000)	(0.000)	(0.000)
Largest National			0.981***
City in 1850			(0.159)
Baseline Controls	No	Yes	Yes
Countries	135	135	135
Observations	2058	2058	2058
within R^2	0.22	0.32	0.35

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

development around 2000 to 2005 increases with the region's urban population density in 1850. To verify the existence of the pattern, I apply the following strategy. I stratify regions into 6 groups according to urban population density in 1850, indexed starting from 0 for regions with 0 values of urban population density to 5 for regions with highest values of urban population density. I regress log GDP per capita in 2005 and the other three alternative outcomes on the 6 groups controlling for regional spillovers and the 8 geography factors. A higher coefficient for a larger group number is therefore evidence supporting a positive relationship between outcomes and urban population density in 1850.

In Panel A of Table 3.7, regions with positive urban population density in 1850 were divided into 5 equal groups, and cutoffs between groups are therefore arbitrary. In Panel B, regions with positive urban population density in 1850 were divided into 5 groups with cutoffs at one sixth of the mean of urban population density - 0.063, one third of the mean - 0.125, one third of the mean plus one standard deviation - 1.226, and one third of the mean plus two standard deviations - 2.326. The base group consists of regions in which urban population density is 0. In almost all cases, coefficients of dummies are positive and are ascending with density groups, suggesting that the effect of urban population density in 1850 on development today is continuous. The evidence of persistence is therefore generally applicable to all regions.

Furthermore, I investigate the concern that my evidence for the relationship between urban population density and income per capita might be driven by regions with the super cities or regions with huge urban populations. I interact urban population density in 1850 with region groups according to the size of the largest city within regions in Panel A of Appendix Table A.4, and the size of regional urban population in Panel B, respectively. Overall, estimates show that a positive and concave relationship between urban population density and contemporary development is mostly supported in all groups. In addition, I find no evidence that the magnitude of the association is monotonic to either regional population size or population size of regions' largest city.

Table 3.6: Regressions of Other Development Outcomes on Urbanization in 1850

	Panel A: Ln(Average Nighttime Light Density, Averaged over 2001-2005)			Panel B: Fraction of Population Living in Cities in 2000		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City 1850	0.979***	0.812***	0.569***	0.123***	0.114***	0.073***
	(0.107)	(0.098)	(0.089)	(0.019)	(0.019)	(0.017)
Urban Population Density 1850	0.395***	0.321***	0.171**	0.082***	0.069***	0.043**
	(0.091)	(0.082)	(0.075)	(0.018)	(0.017)	(0.017)
Square Urban Pop. Den. 1850	-0.013***	-0.010***	-0.006**	-0.003***	-0.003***	-0.002***
	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
City in Neighboring Regions 1850	0.618***	0.455***	0.486***	-0.021	-0.023*	-0.018
	(0.096)	(0.085)	(0.083)	(0.014)	(0.014)	(0.013)
Urb. Pop. Den. in Neib. 1850	0.031*	0.026*	0.028*	-0.001	-0.002	-0.001
	(0.017)	(0.016)	(0.016)	(0.006)	(0.006)	(0.006)
Squ. Urb. Pop. Den. in Neib. 1850	-0.000**	-0.000*	-0.000**	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Largest National City in 1850			1.020***			0.173***
			(0.170)			(0.028)
Baseline Controls	No	Yes	Yes	No	Yes	Yes
Countries	135	135	135	135	135	135
Observations	2044	2044	2044	2050	2050	2050
within R^2	0.14	0.25	0.27	0.10	0.13	0.16

	Panel C: Ln(Population Density in 2000)		
	(1)	(2)	(3)
Existence of a City 1850	0.882***	0.751***	0.518***
	(0.099)	(0.090)	(0.091)
Urban Population Density 1850	0.660***	0.622***	0.478***
	(0.119)	(0.118)	(0.112)
Square Urban Pop. Den. 1850	-0.021***	-0.019***	-0.016***
	(0.004)	(0.004)	(0.004)
City in Neighboring Regions 1850	0.621***	0.462***	0.492***
	(0.097)	(0.078)	(0.076)
Urb. Pop. Den. in	0.044*	0.042*	0.044*
	(0.025)	(0.024)	(0.025)
Urb. Pop. Den. in Neib. 1850	-0.000*	-0.000*	-0.000*
	(0.000)	(0.000)	(0.000)
Largest National City in 1850			0.981***
			(0.159)
Baseline Controls	No	Yes	Yes
Countries	135	135	135
Observations	2058	2058	2058
within R^2	0.22	0.32	0.35

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.7: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	Ln(GDP per capita, 2005)	Ln(nighttime lights), 2001-05	Urban., 2000 Rate	Ln(Pop. Den.), 2000
	Panel A: Quintiles of Regions for Urb. Pop. Den., 1850			
	(1)	(2)	(3)	(4)
Quintile with Smallest Non-zero	-0.033	0.160	0.020	0.062
Urb. Pop. Den. in 1850	(0.038)	(0.139)	(0.015)	(0.123)
The 2nd Smallest Quintile	0.009	0.734***	0.071***	0.601***
	(0.036)	(0.099)	(0.019)	(0.099)
The 3rd Quintile	0.094**	0.800***	0.091***	0.696***
	(0.038)	(0.106)	(0.016)	(0.112)
The 4th Quintile	0.208***	1.063***	0.181***	1.069***
	(0.043)	(0.177)	(0.028)	(0.170)
Quintile with Largest	0.461***	2.027***	0.339***	2.540***
Urb. Pop. Den. in 1850	(0.055)	(0.217)	(0.037)	(0.227)
City in Neighboring Regions (1850)	-0.055*	0.455***	-0.024*	0.443***
	(0.031)	(0.083)	(0.013)	(0.073)
Urb. Pop. Den. in Neib. 1850	0.064**	0.037**	0.000	0.059**
	(0.027)	(0.017)	(0.006)	(0.027)
Squ. Urb. Pop. Den. in Neib. 1850	-0.002*	-0.000**	-0.000	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)
Countries	92	135	135	135
Observations	1395	2044	2050	2058
within R^2	0.23	0.28	0.17	0.38
	Panel B: Groups of Regions by Urb. Pop. Den., 1850			
	(1)	(2)	(3)	(4)
between 0 to	0.035	0.604***	0.073***	0.481***
0.063 (or $\frac{1}{6}mean$)	(0.029)	(0.084)	(0.017)	(0.080)
between 0.063 to	0.228***	0.970***	0.159***	0.966***
0.125(or $\frac{1}{3}mean$)	(0.052)	(0.181)	(0.035)	(0.169)
between 0.125 to	0.387***	1.814***	0.302***	2.064***
1.226(or $\frac{1}{3}mean + std.dev.$)	(0.067)	(0.219)	(0.041)	(0.209)
between 1.226 to	0.546***	2.810***	0.516***	3.807***
2.326(or $\frac{1}{3}mean + 2 \times std.dev.$)	(0.132)	(0.416)	(0.051)	(0.338)
greater than 2.326	0.667***	2.241***	0.361***	3.538***
	(0.099)	(0.293)	(0.066)	(0.392)
City in Neighboring Regions (1850)	-0.042	0.476***	-0.020	0.493***
	(0.030)	(0.083)	(0.013)	(0.074)
Urb. Pop. Den. in Neib. 1850	0.053*	0.034**	-0.000	0.054**
	(0.028)	(0.016)	(0.006)	(0.026)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001*	-0.000**	-0.000	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)
Countries	92	135	135	135
Observations	1395	2044	2050	2058
within R^2	0.23	0.28	0.17	0.39

Note: The unit of observation is a subnational region. RRobust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Appendix Table A.5, I report quantile regressions of log income per capita in 2005 on urbanization in 1850 for quantiles 0.1 in Panel A, 0.25 in Panel B, 0.5 in Panel C, 0.75 in Panel D, 0.9 in Panel E based on the whole sample. I observe a pattern of persistence in each quantile, although magnitudes vary. The median regression estimates (in quantile of 0.5, Panel C) are close to the OLS regression estimates. Overall, my quantile regressions suggest that my conclusions based on OLS estimation are less likely to be driven by regions with unusually low/high income per captia in 1850.

3.3.5 Alternative Minimum Population Threshold

In order to construct measures of development in 1850, I define cities in 1850 using a minimum population of 5,000 as the threshold. One might be worried that the number is so small that many settlements in 1850 with population slightly higher than 5,000 may not be available in any record leading to a measurement error of urban population density in 1850. I do find that for some continents or countries only settlements whose estimated population reaches a much higher number than 5,000 are available in my city data. For example, most settlements in 1850 in Africa and Asia in my data have a population size higher than 15,000.[10] However, if settlements within each country are completely recorded based on a consistent population threshold, then country fixed effects will mitigate the impact of losing of small cities on estimation.

To investigate the potential effect of using various minimum population thresholds on my estimation, I reconstruct variables measuring development in 1850 by using minimum population thresholds of 20,000, 50,000, and 100,000 respectively. I start with a threshold of 20,000 for the reason that Chandler (1987), one of the most influential source of historical

[10] When I raise threshold from 5,000 to 20,000, I find that the number of regions with urban population and the total urban population remain steady. The number of regions with cities in 1850 declines from 205 to 178 for Asia and from 49 to 32 for Africa, and aggregate urban population decreases from 28,878 to 26,846 for Asia and from 3,149 to 2,799 for Africa. However, both numbers drop substantially for Europe and the Americas.

cities and the benchmark of many others' work, is based on the same threshold. I report evidence in Table 3.8. I find that results based on a minimum population threshold of 20,000 are very similar to the threshold of 5,000. However, the coefficient of urban population density in 1850 diminishes quite a bit when threshold increases from 20,000 to 50,000, and to 100,000. This may suggest that the coefficient may vary according to various regional characteristics such as continent, size of the largest city within regions, and so on. For example, the number of regions whose urban population density is positive drops more quickly in the Americas and Africa than in Asia and Europe when a higher threshold is used. There are 164 out of 196 regions with urban population in 1850 from Asia or Europe when 50,000 is used as a threshold to define city, and 76 out of 90 regions when 100,000 is chosen. I will discuss these in the rest of this section. In sum, though the magnitude of the coefficients of the urbanization variables vary across different thresholds, the pattern of persistence of economic activities across regions remains robust.

3.3.6 Evidence in Subsamples

I also check whether the evidence for persistence is driven by regions in a small group of countries characterized by similar characteristics such as countries by various continents or countries by different income groups. I divide the sample into various groups according to different criteria. Table 3.9 reports regressions for regions in different continent groups. Results on African countries using log GDP per capita and log average nighttime luminosity as dependent variables are reported in Panels A and B, respectively. I have discussed that unavailability and low accuracy are two drawbacks of income data at the regional level. The drawbacks are magnified in Africa. For example, only 123 regions from 13 countries, about one third of regions in Africa, are included in estimation when I use log GDP per capita as dependent variable. Nighttime luminosity is used as a popular substitute of GDP in recent studies focusing on Africa (Henderson, Storeygard, and Weil, 2012; Alesina, Papaioannou, and Michalopoulos, 2012; Storeygard, 2013). The evidence for persistence in Africa is supported with use of log average luminosity as a measure of development.

I report West European countries in Panel C of Table 3.9, other European countries in D, the Americas in E, American countries excluding the US and Canada in F, and Asia in G. The effects of early development on log GDP per capita across all country groups follow the same pattern that is found in the whole sample. Magnitudes of the effects vary greatly across groups; the coefficient of urban population density in 1850, for example, is lowest for regions from Western European countries, is moderate for regions in Asia, and highest for regions in the Americas and Non-West Europe. This may partially result from the concavity of the relationship between urban population density in 1850 and log income per capita in 2005. As displayed in Figure 2.8, regions with densest population in 1850 are mainly from West Europe and least dense population from the Americas and East Europe excluding Africa and Oceania. Partly due to the concavity, the coefficient of urban population density is lower if I mainly focus on regions with a higher density. To further support it, I substitute year 1850 urban population density, year 1850 neighboring urban population density and their quadratic forms with logs of both variables. Results of various groups are shown in Panels A - F of Appendix Table A.7. The coefficient of the log urban population density in 1850 is in a narrow range of 0.08 - 0.11 across different groups.

Because results using log regional income in 2005 as the dependent variable do not support persistence in Africa, excluding African regions from the whole sample should not dramatically change my conclusions based on the whole sample. In Panel A of Table 3.10, I use log GDP per capita in 2005 as dependent variable and exclude African countries. Estimates are close to those in the whole sample.

Due to the remarkable movement of goods and services and production factors in the US and Canada, regions and cities in these two countries have experienced lots of ups and downs during the 150 years. For example, the US states such as California and Texas that were

Table 3.8: Alternative Minimum Population Thresholds in Creating Urbanization Variables, Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	Panel A: Based on localities with a minimum population of 20,000			Panel B: Based on localities with a minimum population of 50,000		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.080***	0.064**	0.003	0.138***	0.106**	0.030
	(0.030)	(0.032)	(0.031)	(0.049)	(0.050)	(0.046)
Urban Population Density 1850	0.345***	0.280***	0.211***	0.306***	0.239***	0.182**
	(0.068)	(0.061)	(0.060)	(0.079)	(0.072)	(0.073)
Square Urban Pop. Den. 1850	-0.023***	-0.019***	-0.015***	-0.020***	-0.015***	-0.013**
	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)
City in Neighboring Regions (1850)	0.006	0.003	0.004	0.013	-0.021	-0.008
	(0.026)	(0.026)	(0.025)	(0.051)	(0.037)	(0.034)
Urb. Pop. Den. in Neib. 1850	0.118*	0.083	0.091	0.032	0.023	0.023
	(0.060)	(0.066)	(0.063)	(0.048)	(0.056)	(0.054)
Squ. Urb. Pop. Den. in Neib. 1850	-0.010*	-0.007	-0.007	-0.002	-0.002	-0.002
	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)
Largest National City in 1850			0.281***			0.293***
			(0.058)			(0.055)
Observations	1395	1395	1395	1395	1395	1395
within R^2	0.08	0.20	0.23	0.06	0.19	0.22

	Panel C: Based on localities with a minimum population of 100,000		
	(1)	(2)	(3)
Existence of a City (1850)	0.128**	0.097	0.020
	(0.063)	(0.074)	(0.077)
Urban Population Density 1850	0.246***	0.189**	0.120
	(0.080)	(0.079)	(0.082)
Square Urban Pop. Den. 1850	-0.016***	-0.013**	-0.009
	(0.006)	(0.006)	(0.006)
City in Neighboring Regions (1850)	0.006	-0.034	-0.015
	(0.067)	(0.063)	(0.060)
Urb. Pop. Den. in Neib. 1850	0.041	0.022	0.021
	(0.079)	(0.088)	(0.082)
Squ. Urb. Pop. Den. in Neib. 1850	-0.003	-0.002	-0.002
	(0.006)	(0.007)	(0.006)
Largest National City in 1850			0.326***
			(0.053)
Observations	1395	1395	1395
within R^2	0.04	0.17	0.21

Note: All regressions are based on a sample of 1,395 regions from 92 countries. Robust standard errors clustered at the country level are shown in parentheses. Columns (2) and (3) include land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.9: Robustness to Country Groups Based on Continent, Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	Panel A: Africa			Panel B: Africa - luminosity		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.022	0.029	-0.243	0.987**	0.555*	-0.253
	(0.166)	(0.168)	(0.226)	(0.425)	(0.300)	(0.399)
Urban Population Density 1850	3.882	-1.760	1.003	1.338**	0.813**	0.501
	(5.537)	(5.047)	(4.725)	(0.489)	(0.367)	(0.346)
Square Urban Pop. Den. 1850	-5.446	7.212	-5.466	-0.095**	-0.059*	-0.039
	(17.525)	(15.953)	(15.127)	(0.039)	(0.030)	(0.028)
City in Neighboring Regions (1850)	-0.119	-0.085	-0.147	0.856**	0.047	0.030
	(0.098)	(0.116)	(0.128)	(0.342)	(0.235)	(0.238)
Urb. Pop. Den. in Neib. 1850	0.542	-2.240	-1.623	0.617	0.322	0.304
	(2.023)	(2.542)	(2.405)	(0.610)	(0.661)	(0.650)
Squ. Urb. Pop. Den. in Neib. 1850	-0.059	0.500	0.376	0.052	0.035	0.043
	(0.413)	(0.519)	(0.491)	(0.138)	(0.123)	(0.121)
Largest National City in 1850			0.412			1.554***
			(0.256)			(0.546)
Baseline Controls Included	No	Yes	Yes	No	Yes	Yes
Countries	13	13	13	28	28	28
Observations	123	123	123	357	357	357
within R^2	0.06	0.35	0.39	0.11	0.44	0.47
	Panel C: West Europe			Panel D: Rest Europe		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.084*	0.096*	0.091**	0.147***	0.174**	0.133*
	(0.047)	(0.049)	(0.040)	(0.043)	(0.063)	(0.066)
Urban Population Density 1850	0.113***	0.131***	0.071**	0.844*	0.790	0.311
	(0.035)	(0.035)	(0.028)	(0.449)	(0.491)	(0.582)
Square Urban Pop. Den. 1850	-0.006**	-0.007**	-0.004**	-0.169	-0.154	-0.020
	(0.003)	(0.003)	(0.002)	(0.138)	(0.156)	(0.179)
City in Neighboring Regions (1850)	-0.007	-0.033	-0.055	-0.285*	-0.154**	-0.151**
	(0.025)	(0.034)	(0.032)	(0.150)	(0.072)	(0.064)
Urb. Pop. Den. in Neib. 1850	0.046***	0.055***	0.073***	0.921	0.288	0.448
	(0.015)	(0.017)	(0.018)	(0.626)	(0.440)	(0.390)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001***	-0.002***	-0.002***	-1.107	-0.302	-0.455
	(0.000)	(0.000)	(0.001)	(0.741)	(0.482)	(0.455)
Largest National City in 1850			0.226**			0.349**
			(0.103)			(0.147)
Baseline Controls Included	No	Yes	Yes	No	Yes	Yes
Countries	16	16	16	18	18	18
Observations	214	214	214	290	290	290
within R^2	0.26	0.32	0.37	0.18	0.49	0.52

Continued on next page...

Table 3.9 – Continued

	Panel E: The Americas			Panel F: The Americas no US & Canada		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.037	0.048	0.021	0.046	0.045	0.009
	(0.046)	(0.047)	(0.052)	(0.059)	(0.065)	(0.070)
Urban Population Density 1850	0.781***	0.710***	0.528**	0.850***	0.756**	0.440
	(0.182)	(0.239)	(0.217)	(0.257)	(0.323)	(0.308)
Square Urban Pop. Den. 1850	-0.103**	-0.090	-0.057	-0.116*	-0.097	-0.041
	(0.042)	(0.054)	(0.049)	(0.055)	(0.068)	(0.064)
City in Neighboring Regions (1850)	-0.081	-0.047	-0.051	-0.087	-0.058	-0.062
	(0.058)	(0.055)	(0.053)	(0.069)	(0.066)	(0.063)
Urb. Pop. Den. in Neib. 1850	0.565*	0.399	0.407*	0.446	0.324	0.378
	(0.296)	(0.234)	(0.231)	(0.449)	(0.369)	(0.391)
Squ. Urb. Pop. Den. in Neib. 1850	-0.094	-0.059	-0.060	-0.067	-0.041	-0.052
	(0.067)	(0.053)	(0.053)	(0.102)	(0.084)	(0.089)
Largest National City in 1850			0.227**			0.280**
			(0.094)			(0.119)
Baseline Controls Included	No	Yes	Yes	No	Yes	Yes
Countries	20	20	20	18	18	18
Observations	387	387	387	324	324	324
within R^2	0.12	0.31	0.32	0.10	0.30	0.32

	Panel G: Asia		
	(1)	(2)	(3)
Existence of a City (1850)	0.035	0.043	-0.004
	(0.047)	(0.049)	(0.042)
Urban Population Density 1850	0.481***	0.420***	0.332**
	(0.142)	(0.119)	(0.148)
Square Urban Pop. Den. 1850	-0.037***	-0.033***	-0.027**
	(0.012)	(0.010)	(0.012)
City in Neighboring Regions (1850)	0.080	0.071	0.077
	(0.054)	(0.055)	(0.054)
Urb. Pop. Den. in Neib. 1850	0.159	0.124	0.124
	(0.179)	(0.208)	(0.201)
Squ. Urb. Pop. Den. in Neib. 1850	-0.012	-0.010	-0.010
	(0.014)	(0.016)	(0.016)
Largest National City in 1850			0.242*
			(0.120)
Baseline Controls Included	No	Yes	Yes
Countries	24	24	24
Observations	373	373	373
within R^2	0.12	0.22	0.24

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

underdeveloped 150 years ago have been growing rapidly. On the other hand, Louisiana, a state that were prosperous before, is recently one of the poorest states in the US. One would expect the US and Canada would be exceptions to the persistent regional disparities. I investigate this two countries in Panel B of Table 3.10. Considering that persistence in the two countries may exist in a different model specification, I regress log GDP per capita on year 1850 city dummy and urban population density in 1850 with including or excluding spillover effects from neighboring cities or/and quadratic terms of urban population density in 1850. None of the results support persistent regional inequalities for regions in the US and Canada during the 150 years. As expected, the US and Canada are exceptions to persistence.

Table 3.10: Evidence in Subsample: Non African Countries, and US & Canada

	Panel A: Excluding Africa			Panel B: US and Canada excl. DC			
	(1)	(2)	(3)	(1)	(2)	(3)	(4)
Existence of a City 1850	0.069**	0.072**	0.030	0.076*	0.053	0.070	0.051
	(0.028)	(0.030)	(0.029)	(0.008)	(0.019)	(0.019)	(0.027)
Urban Population Density 1850	0.337***	0.268***	0.187***	0.199	-0.075	0.906	-0.126
	(0.065)	(0.056)	(0.057)	(0.350)	(0.132)	(1.889)	(0.703)
Square Urban Pop. Den. 1850	-0.023***	-0.018***	-0.014***			-2.073	-0.284
	(0.006)	(0.005)	(0.005)			(4.590)	(3.138)
City in Neighboring Regions 1850	-0.043	-0.026	-0.025		0.053		0.047
	(0.048)	(0.034)	(0.033)		(0.053)		(0.029)
Urb. Pop. Den. in Neib. 1850	0.049*	0.035	0.047*		0.474**		0.832
	(0.028)	(0.029)	(0.028)		(0.022)		(1.250)
Squ. Urb. Pop. Den. in Neib. 1850	-0.002*	-0.001	-0.001*				-0.421
	(0.001)	(0.001)	(0.001)				(1.452)
Largest National City in 1850			0.280***				
			(0.061)				
Baseline Controls	No	Yes	Yes	Yes	Yes	Yes	Yes
Countries	79	79	79	2	2	2	2
Observations	1272	1272	1272	62	62	62	62
within R^2	0.09	0.21	0.24	0.38	0.46	0.38	0.46

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. For the US and Canada in Panel B, District of Colombia is excluded. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.4 Potential Mechanisms

All results so far report persistence in the long run development at the regional level over the past 150 years or longer, and such results are robust to controlling for a comprehensive set of geographic factors, using alternative measures of contemporary economic development, and alternative samples. The interesting question is through what channels is early development linked to income today at the regional level. Many cross-country studies have emphasized the importance of geography, institutions, and culture in determining income differences. However, institutions and culture are less likely to vary much within a country. Conditional on country fixed effects, institutions and culture are unlikely to be the main driving forces behind the link at the regional level.

I use a similar exercise to Putterman and Weil (2010) to look for the potential channels in Table 3.11. I look at the relationship between urbanization in 1850 and years of education in 2005 without taking any geographic controls in column (1) of Panel A. The coefficient of year 1850 urban dummy is 0.273 with a standard error of 0.08. The coefficient of year 1850 urban population density is 0.599 significant at 1 percent, and its quadratic form is -0.042 significant at 1 percent. Intuitively, residents of regions in which cities existed in 1850 are expected to have more years education today, and an additional 100 inhabitants per square kilometer living urban area in 1850 predicts 0.56 more average years of education in the region. The coefficient of year 1850 neighboring urban dummy is small in magnitude with a negative sign and it is insignificant. Coefficients of year 1850 neighboring urban population density and its quadratic form have the magnitudes about half of those within the region. Both are significant at 5 percent level. The early urbanization variables together explain 10 percent of within-country variation of years of education and 21 percent of between-country variation.

Table 3.11: Impact of Urbanization in 1850 on Contemporary Education, Culture, Institution, and Infrastructure

	Dependent Variable:						
	Years of Educ. in 2000	Trust in others	Informal Payments	Access to Financing	Ln Days of no electricity	Ln power line density	Ln travel time
	Panel A: Without baseline controls						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Existence of a City 1850	0.273***	0.006	0.105	0.044*	0.270**	0.243***	-0.367***
	(0.080)	(0.010)	(0.141)	(0.024)	(0.105)	(0.045)	(0.069)
Urban Population Density 1850	0.600***	-0.018	0.132	-0.014	-0.028	0.193	-0.819***
	(0.141)	(0.012)	(0.101)	(0.017)	(0.088)	(0.135)	(0.131)
Square Urban Pop. Den. 1850	-0.042***	0.001*	-0.003	0.001	-0.002	-0.025*	0.055***
	(0.013)	(0.001)	(0.008)	(0.001)	(0.007)	(0.014)	(0.014)
City in Neighboring Regions 1850	-0.041	-0.017**	0.116	-0.023	-0.055	0.294***	-0.373***
	(0.057)	(0.008)	(0.110)	(0.023)	(0.117)	(0.053)	(0.064)
Urb. Pop. Den. in Neib. 1850	0.267**	-0.001	-1.117**	0.254**	-0.008	0.088	-0.248**
	(0.122)	(0.008)	(0.452)	(0.113)	(0.413)	(0.106)	(0.094)
Squ. Urb. Pop. Den. in Neib. 1850	-0.009**	-0.000	0.239**	-0.051**	-0.010	-0.001	0.007**
	(0.004)	(0.000)	(0.100)	(0.024)	(0.091)	(0.003)	(0.003)
Countries	90	61	65	68	64	92	92
Observations	1358	665	331	372	203	1395	1395
within R^2	0.10	0.01	0.02	0.04	0.04	0.10	0.29
	Panel B: With baseline controls and regional population						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Existence of a City 1850	0.122*	0.009	0.124	0.037	0.309**	0.159***	-0.234***
	(0.070)	(0.011)	(0.153)	(0.028)	(0.131)	(0.040)	(0.056)
Urban Population Density 1850	0.284**	-0.020	0.077	-0.015	-0.064	0.088	-0.666***
	(0.129)	(0.014)	(0.125)	(0.020)	(0.123)	(0.139)	(0.141)
Square Urban Pop. Den. 1850	-0.024**	0.001*	0.002	0.001	-0.000	-0.018	0.045***
	(0.010)	(0.001)	(0.010)	(0.002)	(0.010)	(0.013)	(0.013)
City in Neighboring Regions 1850	-0.031	-0.016*	0.083	-0.030	-0.139	0.232***	-0.257***
	(0.052)	(0.008)	(0.107)	(0.024)	(0.144)	(0.053)	(0.046)
Urb. Pop. Den. in Neib. 1850	0.275**	-0.003	-1.214**	0.259**	-0.155	0.083	-0.215**
	(0.109)	(0.009)	(0.501)	(0.098)	(0.385)	(0.103)	(0.097)
Squ. Urb. Pop. Den. in Neib. 1850	-0.010***	-0.000	0.265**	-0.051**	0.028	-0.000	0.006**
	(0.003)	(0.000)	(0.110)	(0.021)	(0.086)	(0.003)	(0.003)
Largest National City in 1850	0.817***	-0.006	0.000	0.017	0.010	0.171*	-0.119
	(0.120)	(0.012)	(0.093)	(0.020)	(0.097)	(0.096)	(0.088)
Countries	90	61	65	68	64	92	92
Observations	1358	665	331	372	203	1395	1395
within R^2	0.21	0.02	0.05	0.06	0.14	0.15	0.44

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Column (2) of Panel A considers an indicator of culture, trust in others. The urbanization coefficients are close to zero and none of them are significant. Predictive power is also close to zero according to within and between R^2.

In columns (3) - (5) of Panel A, I regress three outcomes of regional institutions - informal payments, access to financing, and log days without electricity - on the year 1850 urbanization variables. The correlations are mostly insignificant and difficult to explain. Access to financing reported in column (4) is positively correlated with year 1850 urban dummy, significantly at 10 percent level, reflecting a weak positive impact of early urbanization on contemporary institutions. However, log of days without electricity is positively associated with year 1850 urban dummy, suggesting a negative effect of early urbanization on institutions today.

The remaining two columns of Panel A report the effect of urbanization in 1850 on infrastructure measured by the log power line density in column (6) and log travel time in column (7). Power line density is more likely to reflect the scale of infrastructure while the travel time captures the quality of infrastructure. Both columns show that regions with a higher level of early urbanization in 1850 tend to have larger and more efficient infrastructure. Urbanization in 1850 explains 10 percent of within and 40 percent of between R^2 for log power line density, and 29 percent and 59 percent for log travel time.

In Panel B I show regressions of the same regional outcomes on urbanization in 1850 while including my baseline geographic controls and the log of regional population. Geographic factors are controlled for so as to rule out the possibility that early urbanization captures advantageous geographic or natural environments that favor economic development. For example, an ideal geographic condition in the plain helped to build city hundreds of years ago also means a relatively low cost to construct modern infrastructure such as schools, roads, etc. Controlling for current regional population rules out the concern that early urbanization is purely picking up scale effect of population size which may plausibly persist over the past 150 years. My results in Panel B show that 1) including these controls lowers the effect of urbanization in 1850 by different magnitudes for different outcomes, and 2) urbanization in 1850 is still strongly correlated with contemporary years of education and infrastructure.

My evidence suggests that both path dependence theory and physical geography are important to understand persistence of economic disparities across regions over the past 150 years. More exactly, dependence theory in my context is closely related to accumulations of human capital and infrastructure over time. For example, in Gennaioli et al. (2013), the role of human capital in regional disparities suggests a long run accumulation of human capital. In their structural model, individuals decide where to live - productive region or unproductive region - subject to a moving cost, and whether to be entrepreneur or workers. A higher human capital stock is expected to be found in a more productive region and contributes regional economy through differing roles of individuals - as workers or entrepreneur - and human capital externalizes.

However, the results of the exercise are only suggestive as one can make the reverse inference that early urban development influences current income level that favors the quality of infrastructure and level of education.

3.5 Going Back Further

One would expect that contemporary regional disparities might originate even earlier than 1850. I therefore extend the time horizon of this study to 500 years ago. However, Acemoglu, Johnson, and Robinson (2002) find a reversal of fortune among ex-colonies at the country level. They argue that it is a result of different settlement strategies adopted by European settlers according to population density in colonies in 1500; extractive institutions were more likely to be introduced in places where population were more dense 500 hundred years ago. Moreover, the reversal was almost complete prior by the middle of the 19th century. Therefore, for ex-colonies, reversal of prosperity of regions may not be observed until post 19th-century if there exists.

I verify these statements in this section. Instead of looking at all countries in the sample, I separate countries based on their ex-colony status. Table 3.12 regresses per capita GDP in 2005 on urbanization in 1850 in both colonial countries and other countries separately. My results show that regions with a higher urbanization 150 years ago tend to be richer today in either ex-colonial countries or other countries. I therefore find evidence that there exists persistence during post-industrialization period. For the time before industrialization, I regress the year 1850 urban population density on the year 1500 urbanization variables in Table 3.13. The persistence still exists in non-colonial countries. However, there is no evidence indicating regions that were more urbanized in 1500 were still richer than others in 1850 within an ex-colonial country. Overall, these results suggest that economic prosperity can persist for much longer time than 150 years unless the there is institutional reversal.

3.6 Conclusion

The debate regarding the 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. Previous research also has suggested early development favors long term economic growth through developing growth-promoting elements, such as human capital, culture, and institutions. For inequality of economic development at the sub-national level, however, most of studies are restricted to a single country or several ones, and only a few of them have looked at the roles of historical factors, mainly due to limited availability of sub-national data, especially historical data.

In this paper, I construct urban population density in 1850 to study regional disparities over the past 150 years. This study complements the literature on the long run within-country differences by introducing a new proxy for regional development in 1850, and by

Table 3.12: Persistence with Colonized Countries, Regressions of Log GDP per Capita in 2005 on Urbanization in 1850

	Colonized countries					
	Panel A: Fixed-effects			Panel B: Control for 1500 population density		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City 1850	0.032	0.021	-0.030	0.040	0.028	-0.019
	(0.037)	(0.033)	(0.038)	(0.037)	(0.034)	(0.038)
Urban Population Density 1850	0.805***	0.625***	0.512**	0.810***	0.634***	0.528***
	(0.172)	(0.178)	(0.195)	(0.171)	(0.177)	(0.192)
Square Urban Pop. Den. 1850	-0.116***	-0.088**	-0.070	-0.118***	-0.091**	-0.074*
	(0.040)	(0.042)	(0.043)	(0.040)	(0.042)	(0.043)
City in Neighboring Regions 1850	-0.048	-0.036	-0.040	-0.041	-0.028	-0.031
	(0.044)	(0.042)	(0.042)	(0.044)	(0.041)	(0.041)
Urb. Pop. Den. in Neib. 1850	0.533***	0.383**	0.374**	0.543***	0.392**	0.383**
	(0.126)	(0.169)	(0.156)	(0.125)	(0.173)	(0.161)
Squ. Urb. Pop. Den. in Neib. 1850	-0.070**	-0.042	-0.040	-0.072**	-0.045	-0.042
	(0.030)	(0.037)	(0.035)	(0.030)	(0.038)	(0.035)
Largest National City in 1850			0.244***			0.228***
			(0.078)			(0.076)
Log Population Density 1500				-0.368***	-0.296***	-0.299***
				(0.090)	(0.072)	(0.072)
Countries	43	43	43	43	43	43
Observations	658	658	658	658	658	658
within R^2	0.11	0.20	0.22	0.11	0.20	0.22

	Panel C: Non colonized countries		
	(1)	(2)	(3)
Existence of a City 1850	0.103**	0.127***	0.081**
	(0.039)	(0.043)	(0.039)
Urban Population Density 1850	0.258***	0.205***	0.117**
	(0.063)	(0.055)	(0.058)
Square Urban Pop. Den. 1850	-0.017***	-0.013***	-0.008*
	(0.005)	(0.005)	(0.004)
City in Neighboring Regions 1850	-0.077	-0.070*	-0.062
	(0.081)	(0.041)	(0.041)
Urb. Pop. Den. in Neib. 1850	0.025**	0.020	0.036**
	(0.012)	(0.013)	(0.017)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001**	-0.001*	-0.001**
	(0.000)	(0.000)	(0.000)
Largest National City in 1850			0.303***
			(0.083)
Countries	49	49	49
Observations	737	737	737
within R^2	0.11	0.32	0.35

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Columns (2) and (3) include land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. Fixed-effects estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.13: Persistence with Colonized Countries, Regressions of Urban Population Density in 1850 on Urbanization in 1500

	Dependent Variable: Urban Population Density in 1850					
	Panel A: Colonized countries				Panel B: Non colonized countries	
	Fixed-effects		Random-effects			
	(1)	(2)	(3)	(4)	(1)	(2)
Existence of a City 1500	96.539 (64.204)	84.061 (55.888)	105.564* (62.918)	94.282* (55.238)	84.771*** (24.686)	80.317*** (24.956)
Urban Population Density 1500	-73.672 (305.957)	-77.246 (327.503)	-55.838 (315.405)	-70.910 (323.880)	858.119** (329.808)	830.115** (325.682)
Square Urban Pop. Den. 1500	33.108 (113.126)	36.271 (123.198)	23.622 (116.931)	28.736 (121.757)	-349.755** (138.496)	-339.341** (135.946)
City in Neighboring Regions 1500	72.467 (62.095)	66.114 (50.217)	65.137 (53.979)	59.038 (46.693)	22.177 (20.466)	15.977 (17.132)
Urb. Pop. Den. in Neib. 1500	-366.713 (312.695)	-348.027 (255.514)	-414.339 (330.433)	-405.785 (291.503)	1.439 (1.793)	1.200 (1.759)
Squ. Urb. Pop. Den. in Neib. 1500	285.567 (235.102)	271.904 (198.234)	323.374 (253.964)	315.101 (226.030)	-0.016 (0.021)	-0.013 (0.020)
Log Population Density 1500			-2.225 (7.923)	2.877 (9.935)		
Baseline Controls	No	Yes	No	Yes	No	Yes
Countries	30	30	30	30	46	46
Observations	598	598	598	598	710	710
within R^2	0.10	0.16	0.10	0.14	0.16	0.18

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. Fixed-effects estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

covering regions from most countries in the world. I find widespread evidence in the world that regions had cities in 1850 are associated with higher development today and among regions with urban population in 1850, regions with more dense population are correlated with a higher level of development. I also document that there exist small positive spillovers of urban development in 1850 from neighboring areas.

I briefly look for potential paths of the persistence, such as human capital, culture, institutions, and infrastructure. While not conclusive, urbanization in 150 years ago affects cross-region variations of current human capital and infrastructure. In the last section of the study, I extend the time horizon to 500 years ago. My results suggest regional economic disparities may persist for 500 years or longer but not for ex-colonial countries.

Chapter 4. Regional Convergence over the Past 150 Years

4.1 Introduction

There is worldwide evidence that the difference in development across regions can persist for decades or even hundreds of years (Michalopoulos and Papaioannou, 2013; Davis and Weinstein, 2002; Maloney and Valencia Caicedo, 2015). On the other hand, a large number of empirical studies have documented the existence of regional convergence, as predicted by the neoclassical model of growth, using regional income data covering the past decades. For example, Gennaioli et al. (2014) estimate the speed of convergence covering a large sample of regions worldwide over the past 50 years. They claim that the convergence rate is about 2 % per year towards the steady state. Given evidence of both persistence and convergence in the growth literature, one would like to see if it is possible to reconcile the two. After all, persistence does not by nature parallel convergence. In this paper, I empirically explore the existence of regional convergence over a much longer period than any other in the literature. Together with the pattern of persistence in regional development over the same period documented by Chapter 3, I describe a complete picture of the dynamic of long run regional growth.

In this chapter, I explore convergence of sub-national development by using a global sample over the years 1850, 1900, 1950, and 2000 such that there are three 50-year intervals that are coincident with the three critical phases of world development in modern times. The first 50-year interval, 1850 - 1900, is an integral part of the "industrial revolution" starting from around 1750 to 1914. Although the Industrial Revolution already ended around 1840, the widespread use of railroads, machinery, steam, and oil, and the introduction of inventions such as electricity, telegraph, and telephone dated back to this 50 years. Furthermore, the standard of living began to grow substantially and consistently for the first time in history in the late 19th century (Robert E. Lucas, 2002; Feinstein, 1998; Szreter and Mooney, 1998). In the second 50-year interval, 1900 - 1950, additionally to the faster-growing economy and technologies, the world also experienced worldwide turmoils such as two world wars and the Great Depression. The average growth over 1900 - 1950 was slightly lower than the twenty years prior to the World War I. During the last 50-year interval, 1950 - 2000, a multilateral world system has evolved with establishment of international organizations such as the United Nations, International Monetary Fund, World Bank, and World Trade Organization. The world system together with new transportation and communication technologies has made the world more integrated than ever. The global economy in this period has achieved growth of the level of the standard of living that was higher all previous eras combined. Figure 4.1 illustrates the world economy by looking at GDP per capita between 1700 and 2000. The world experienced little improvement in GDP per capita before the late 19th century (though there was a slight increase in growth starting around 1820), a consistent and significantly higher growth during the end of the 19th century to around 1950, and an astonishing increase in growth after 1950. Given the vast differences in the world economy over the three 50-year intervals, in this paper, I investigate how regional convergence differs from one subperiod to another.

In contrast to previous studies that look at regional convergence by using GDP per capita, evidence of convergence in this study is based on urban population density as a measure of development. For the years 1850, 1900, and 1950, I construct urban population density - urban population divided by land area, based on various sources of estimates of historical settlements such as Eggimann (1994), Modelski (2003), Bairoch (1998), and Chandler (1987). For the empirical strategy, I mainly reference the framework of Gennaioli et al. (2014) that estimates convergence of regional GDP per capita by considering interregional flows of capital brought in by openness. The average annual growth rate depends on the initial level of

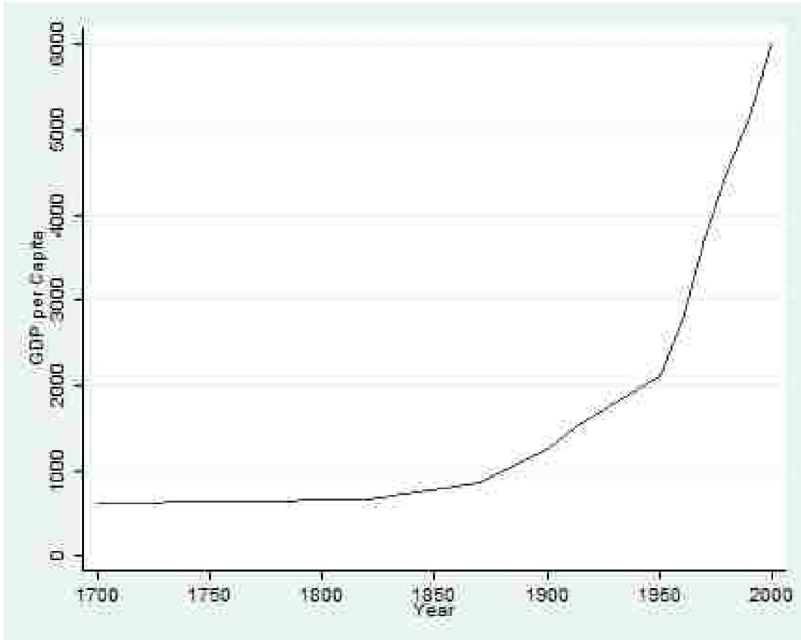


Figure 4.1: World GDP per Capital, 1700 - 2000 (1990 International Dollars)
Note: The graph is produced by using GDP per capita data from New Maddison Project Database by Bolt and van Zanden (2014).

capital, the initial level of national GDP per capita, and regional characteristics. In addition, I also include spatial interaction effects from neighboring areas and country fixed effects.

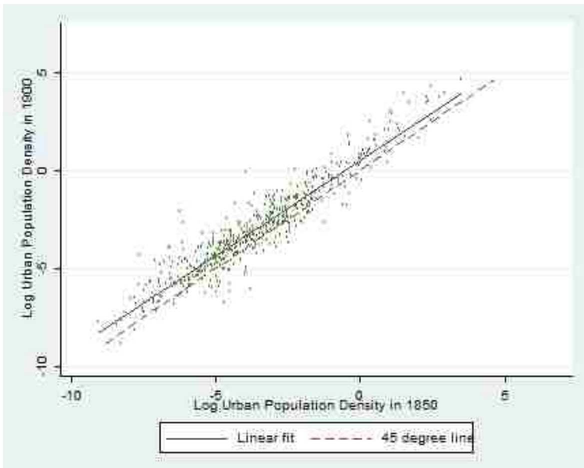
I find remarkably robust evidence of convergence across regions worldwide over the past century. The level of urban population density at the beginning of the period is significantly associated with growth of urban population density. I explore the difference in regional convergence across three 50-year subperiods during 1850 and 2000. My results show that regional convergence started around the beginning of 20th century and has been accelerating over time, which is generally consistent with the date of the growth of the standard of living. These results are also robust across different samples of countries grouped according to continent, the within-region largest city size, and regional urban population quintile, using different definitions of cities, and using alternative empirical strategies. I find that regional convergence varies across different countries but without a clear pattern. The cross-county differences in regional convergence rate are inconsistent over various subperiods. For example, the finding of Gennaioli et al. (2014), suggesting that countries with better market infrastructure have faster convergence of GDP per capital during 1960 and 2010, applies only during 1950 - 2000 but not the other two subperiods.

My findings directly contribute to the literature on regional convergence. Many studies found evidence of convergence across regions within a single or several countries such as the United States (Barro and Sala i Martin, 1991, 1992, 2004; Garofalo and Yamarik, 2002), Japan (Shioji, 1996), Canada (Coulombe and Tremblay, 2001), Australia (Cashdm, 1995), and European countries (Durlauf and Quah, 1999). As an increasing availability of sub-national income data, beyond industrialized countries, Gennaioli et al. (2014) study regional convergence by using a large sample of regions from over 80 countries since the 1950s. Moreover, the estimated speed of β-convergence is often around 2 % per year in the literature of regional convergence (Magrini, 2004).

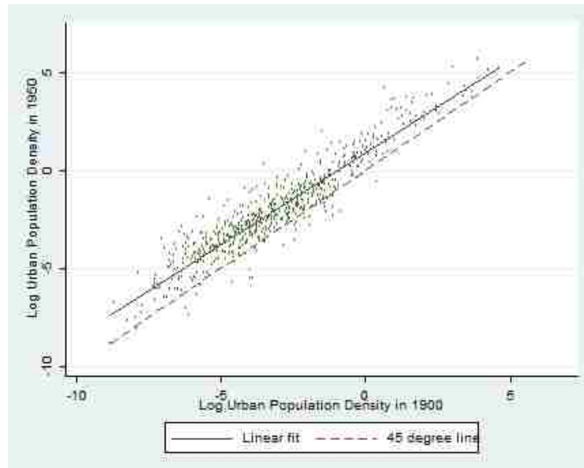
The estimated speed of convergence is around 0.34 - 0.42 % per year which is much lower than the 2 % convergence rate of GDP per capita documented in the literature. I argue that three major factors mainly drive this difference. First, the documented convergence rates are primarily based on income data during the most recent decades. As I show that regional convergence has been growing over time, the difference in the estimated rates is partially explained by the much longer period used in this study. Second, the estimated convergence rate in this study is mainly based on regions with non-zero urban population and, therefore, is likely to downward bias the convergence rate across all regions. That is because regions with missing urban population in a period are more likely to be the least developed regions within a country and experience relatively higher growth rates in the subsequent years, and excluding these regions may potentially lower the overall convergence rates. As a compromise, I include these dropped regions in analysis by imputing their development levels using several conservative methods, and including these regions gives rise to substantially higher convergence rates. In addition, in a smaller sample of regions from Gennaioli et al. (2014) where both urban population and GDP per capita are available around 1950 and 2000, I find that the estimated convergence rate using GDP per capita is around 50 % higher than using urban population density. The last source of the difference in the estimates of regional convergence may be explained by measurement error in regional income since GDP is understated in poor regions. It is likely that this measurement error was more severe in the past, therefore, the speed of catch-up of poorer regions to richer regions would partially reflect the improvement in data accuracy in GDP per capita.

In Chapter 3, I find that both geographic advantages and path dependence are major explanations for persistence. Moreover, human capital and physical capital are the channels. I use this urban population density as a proxy for development and discover strong and robust persistence in regional disparities across the world over the past 150 years. I complement their findings with Figure 4.2 in which I construct urban population density using the same method as Chapter 3 and plot log urban population density in 1900 against 1850, 1950 against 1900, and 2000 against 1950, respectively, for regions I have urban population data in two consecutive points in time. The positive fitted (green) line in each plot reveals that regions that had relative higher urban population density tend to have higher urban population density 50 years later. More interestingly, the vertical difference between the fitted (green) line and 45-degree (red) line is exactly the average 50-year growth of urban population density in any given level of urban population density at the beginning of the period. Thus, along x-axis, if regions that are close to the original point tend to have larger vertical differences between the fitted line and 45-degree line, then regions with lower urban population density experience higher growth, which is exactly evidence of convergence. In Figure 4.2 (a), there is no evidence of convergence because the growth during 1850 - 1900 was generally small and almost constant across all the urban population density levels in 1850. Figures 4.2 (a) and (b) illustrate strong convergence of urban population density. Furthermore, convergence is stronger during 1950 - 2000 than 1900 - 1950.

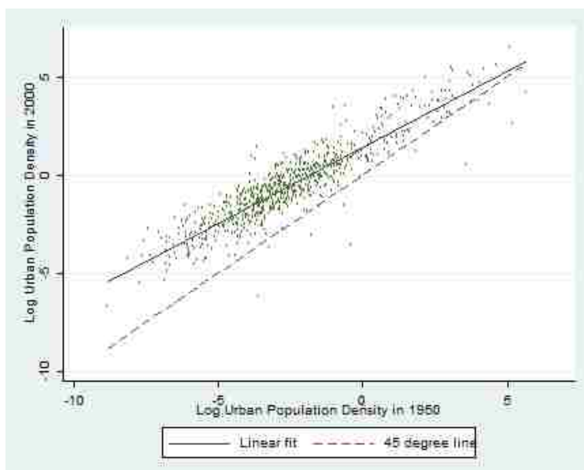
Combining evidence of persistence in Chapter 3 and regional convergence everywhere in the world over the past 150 years in this paper, I am now able to depict a complete



(a) Log Urban Population Density in 1850 & 1900



(b) Log Urban Population Density in 1900 & 1950



(c) Log Urban Population Density in 1950 & 2000

Figure 4.2: Scatterplots of Log Urban Population Density across 1850, 1900, 1950, and 2000

Note: A region's urban population is the total population living in cities. Population estimates of settlements are mainly from Eggimann (1994), Chandler (1987), Bairoch (1998), and Modelski (2003) in 1850, 1900, and 1950, and the Socioeconomic Data Applications Center (SEDAC) (http://sedac.ciesin.columbia.edu/) in 2000. City in 1850, 1900, and 1950 is defined with a minimum population threshold of 5,000. Unit of urban population density is 100 persons per square kilometer. For any comparison, only regions whose urban population are recorded in both years are included.

pattern of regional disparity over the past century and predict where it is heading. There has been persistence in regional development for over 150 years or even longer. However, both persistence and convergence have coexisted since around the late 19th or early 20th centuries. In other words, though regions that were relatively more developed in 1850 tend to be relatively richer today, the gaps in prosperity across regions has been declining at an increasing rate over the past 150 years. I expect that the coexistence of persistence and convergence of regional development will be continuously observed in the future. Persistence in regional prosperity is being weakened by the accelerating speed of convergence rate. As markets within a country get more and more integrated and mobility of capital becomes easier with technological improvements in information systems, transportation, and communication, divergence in regional development is unlikely to happen. In fact, developed economies, on average, have much lower regional disparities.

The remainder of the paper is organized as follows. In Section 4.2 I discuss the empirical framework and describe regional measure of urban population density in 1850, 1900, 1950, and 2000, and control variables. Section 4.3 I present my results. In Section 4.4 I discuss how the estimated convergence rates are comparable to previous studies. Section 4.5 concludes.

4.2 The Model and Data

To study the speed of convergence, one needs a reliable measure of development. GDP per capita, as suggested by the traditional neoclassical model of growth, is the most commonly used. However, as the concept of GDP was not developed until 1934 and became popular worldwide at a much later time, GDP data for most countries is only available after 1940th.[1] The data at the sub-national level is even scarcer, and, in fact, it was only recently that Gennaioli et al. (2013, 2014) compiled regional GDP per capita for the late 20th century and

[1] Simon Kuznets developed the concept of GDP in 1934.

early 21st century. Because of this data availability problem, it is impossible to explore the evolution of regional inequality for an extended period such as a century using GDP data.

In this study, I consider urban population density, the population living in urban areas divided by land area, as an alternative measure of development and study regional convergence by looking at the effect of the initial level of development on growth. Theoretically, urban areas are where goods and services are intensively traded, a variety of jobs are created, and schools and factories are constructed. Trade allows urban residents to specialize in professions where they have comparative advantages so that the process of production is highly efficient. Therefore, the sign of the urban population, which is expected to be positively correlated with the frequency of trade, intensity of specialization, and efficiency in production, should reflect the level of development. In the urban economics literature, urban populations are routinely used to compare relative prosperity. I divide the urban population by land area at the sub-national level to deflate the effect of regional size on urban population (as large administrative areas are more likely to include more cities and urban population). In Chapter 3, I construct this variable and use it to study persistence in regional development over the past 150 years and find evidence that this urban population density is positively associated with GDP per capita, the fraction of people living in cities, population density, and nighttime light intensity. In addition, I find that it is significantly correlated with human capital as measured by average years of schooling and physical capital as measured by infrastructure.

4.2.1 The Model

In the main model specification, the annual growth of urban population density is determined with the following equation:

$$\begin{aligned}\ln UrbPopDensity_{i,t+1} - \ln UrbPopDensity_{i,t} = \alpha - \beta \ln UrbPopDensity_{i,t} \\ + A_i\delta + \lambda \ln y_t + \gamma \ln NeibUrbPopDensity_{i,t} + \mu_c + \varepsilon_{i,t+1}\end{aligned} \tag{4.1}$$

where the coefficient of the initial level of regional composite capital, β, is the estimated speed of convergence. The vector A_i represents a comprehensive set of regional characteristics that determine total factor productivity (TFP) in region i. I include eight geographic and climatic factors: temperature in Celsius, rainfall in meters, altitude in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, the proximity to the coast, and proximity to a river, and a dummy indicating national capital exists in a region, the proximity to the national capital, and the proximity to national borders. To account for country-level time-invariant characteristics such as culture, social norms, etc. that may affect the estimated convergence rate, I include country fixed effects μ_c.

I include log the initial level of national GDP per capita in order to control interregional spillovers brought in the growth framework suggested by Gennaioli et al. (2014). However, it is unlikely that spatial correlation effects received by each region within a country are the same, while regions are likely to be more intensively interacted with their neighboring regions. To avoid misidentification of spatial correlation, I include log urban population density in neighboring areas. However, including national GDP per capita and urban population density in neighboring areas substantially lowers the sample size, as a compromise, I use log national urban population density to substitute log national GDP per capita and a dummy variable indicating whether cities existed in neighboring areas to replace for log urban population density in neighboring areas for most regressions.

I include a large range of robustness checks. For example, I investigate the existence of convergence in sub-samples of countries or regions based on the continent, the largest city size, and regional population size. I use higher minimum population thresholds to reconstruct urban population density variables in 1850, 1900, and 1950. I report results using quantile regressions. Finally, I create group dummies, for regions with lowest urban population density to the highest and check whether regions in relatively higher density group experience relatively lower growth rates in the next 50 years, as suggested by convergence.

4.2.2 Data

The unit of observation at each point in time is a sub-national region. I use the same regions as those in Chapter 3, which are mainly first-level administrative divisions. The data for regions' geographic boundaries are derived from the Database of Global Administrative Areas Map version 2 (GADMv2).[2]

To construct urban population density, I first aggregate population estimates of urban settlements in each time point. Population estimates of settlements are mainly from Eggimann (1994), Chandler (1987), Bairoch (1998), and Modelski (2003) for 1850, 1900, and 1950. Population estimates for 2000 are from the Socioeconomic Data Applications Center (SEDAC).[3]

I define an urban location according to the size of the population inhabiting an area - whether a location has a recorded population of 5,000 or more in 1850, 1900, 1950. I identify 4,223 settlements with a population of 5,000 or more spanning 160 contemporary countries. Mapping these urban locations into the data for sub-national boundaries yields 538 sub-national regions with positive urban populations in 1850, 596 regions in 1900, and 754 regions in 1950 and 2000 for the main analysis. However, there is no theory that suggests a threshold value of 5,000. Some historical studies have used higher limits. Bosker, Buringh, and van Zanden (2013) study cities for the period 800 - 1800 including settlements in which inhabitants are greater than 10,000. Nunn and Qian (2011) uses settlements with a population greater than 40,000 to calculate national urbanization for the period 1000 - 1900. In a later section, I analyze how my results change in response to using higher thresholds.

For the 538 regions in which at least a city existed in 1850, the average urban population density is 42.9 persons per square kilometer with a standard deviation of 203.3 persons per square kilometer; for the 596 regions in 1900, the average is 110.2 persons per square kilometer with a standard deviation of 647.9; for the 754 regions in 1950, the average is

[2]Source: www.gadm.org.
[3]http://sedac.ciesin.columbia.edu/

239.4 and standard deviation is 1,403.5; and for the 754 regions in 2000, the average is 649.7 persons per square kilometer and standard deviation is 3,185.2.

Based on data for urban population density in 1850, 1900, 1950, and 2000, I calculate the 50-year average annual growth rate of urban population density as the difference in log urban population density for every two consecutive points in time divided by 50. The average 50-year average annual growth rate for the 538 regions between 1850 and 1900 covered in this study is 1.2 percent with a standard deviation of 1.6 percent; the average growth rate for the 596 regions between 1900 and 1950 is 2.1 percent with a standard deviation of 1.7 percent; and the average growth rate for the 754 regions between 1950 and 2000 is 3.9 percent with a standard deviation of 2.0 percent.

In addition, to control for spatial correlation, I construct urban population density in neighboring areas. I use 25 miles geodesic distance from regions' boundaries as a range of neighboring areas, based on which I aggregate population within the neighboring areas and divide the total population by the land area of the region. 61.5 percent of or 331 regions in 1850 had at least a city in the neighboring areas; 51.7 percent of or 308 regions in 1900 in 1900; 53.8 percent of or 406 regions in 1950.

I report summary statistics of urban population density, its 50-year average annual growth as well as other regional characteristics in Table 4.1.[4]

4.3 Results

I now report the estimated speed of regional conditional convergence of urban population density. A consistent estimator requires that factors that contribute to the steady-state level of development must not be excluded. To avoid omitted variable problems, I first estimate Equation 4.1 including all eight geographic and climatic variables. I then include country fixed-effects and variables indicating distance to the national capital within each country,

[4]The listing of all regions and years are displayed in an online appendix of this paper which can be downloaded at www.dachaoruan.com

and control for spatial correlation by considering urban population density in neighboring areas.

Table 4.1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
50-Year Average Annual Growth Rate (A):					
1850 - 1900	0.012	0.016	-0.046	0.083	538
1900 - 1950	0.021	0.017	-0.033	0.078	596
1950 - 2000	0.039	0.020	-0.064	0.101	754
Urban Population Density at the beginning of the period (A):					
in 1850	0.429	2.033	0.000	34.027	538
in 1900	1.102	6.479	0.000	102.081	596
in 1950	2.394	14.035	0.000	283.496	754
in 2000	6.497	31.852	0.001	674.283	754
Cities exist in Neighboring Areas (A):					
in 1850	0.615	0.487	0.000	1.000	538
in 1900	0.517	0.500	0.000	1.000	596
in 1950	0.538	0.499	0.000	1.000	754
Urban Population Density in Neighboring Areas (A):					
in 1850	0.296	1.523	0.000	21.64593	331
in 1900	0.533	2.482	0.000	31.79526	308
in 1950	1.488	7.804	0.000	108.8663	406
Regional Controls:					
Presence of National Capital (I)	0.154	0.361	0	1	827
Proximity to Capital City (I)	0.733	0.202	0.094	1	827
Proximity to Borders (B)	0.789	0.164	0.172	0.999	827
Temperature (F)	16.912	7.748	-10.228	29.284	827
Rainfall in Meter (F)	1.074	0.73	0.006	3.993	827
Elevation (100 meters) (F)	5.643	6.834	-0.138	48.786	827
Ruggedness (G)	1.247	1.136	0.012	5.766	827
Land Suitability (E)	0.432	0.308	0	0.998	827
Absolute Value of Latitude (H)	28.265	16.19	0	67	827
Proximity to the Coast (C)	0.834	0.16	0.327	1	827
Proximity to Rivers (C)	0.86	0.144	0.21	1	827

Note: A region's urban population is the total population living in cities. Unit of population density is 100 persons per square kilometers. Sources of data are listed as below, and a detailed explanation on these variables is in Appendix Table A.2.
A: Chandler (1987), Bairoch (1998), Eggimann (1994), and Modelski (2003). B: Gennaioli et al. (2013). C: the National Geophysical Data Center (NOAA). E: Atlas of the Biosphere. F: Global Climate Data (WorldClim). G: Nunn and Puga (2012). H: Global Administrative Areas (GADM). I: World Urbanization Prospects.

I investigate differences in the estimated speed of convergence across various subperiods of the past 150 years - 1850-1900, 1900-1950, 1950-2000, and combinations of the three, and across subsamples of countries according to continents. For robustness checks, I reconstruct urbanization variables using alternative definitions of cities, report results based on alternative model specifications, and look at convergence in different subsamples based on largest city size and urban population size. In the final analysis, I impute missing urban population density. Though imputation nevertheless overstates the level of development for those regions with missing values for urban population density, regressions including those imputed regions provide a lower bound of the estimated speed of convergence rates.

4.3.1 Basic Results

Table 4.2 presents OLS estimates in the form of Equation 4.1 for 1,036 regions from 153 countries across four time points - 1850, 1900, 1950, and 2000. The dependent variable is 50-year average annual growth rates of urban population density, and the coefficient on log urban population density displayed in the first row is the estimated speed of convergence over the past 150 years.

As discussed in Chapter 3 Section 3.3, it has been extensively documented in the literature that physical geography plays an essential role in shaping long run economic disparities. Consequently, in estimating Equation 4.1, I include geographic and climatic controls.

In column (1), I report impacts of log urban population density at the beginning of the period as well as geographic and climatic characteristics on 50-year average annual growth rates of urban population density during 1850 - 2000. Robust standard errors clustered at the country level are shown in parentheses. The estimated convergence rate of 0.09 % is significant at the 1 percent level. Also, temperature, rainfall, elevation, land suitability, latitude, and proximity to the coast are negatively correlated with the growth though the effect of rainfall is not significant. The ruggedness of terrain and proximity to rivers are positively but insignificantly associated with the growth. However, effects of ruggedness, latitude, and proximity to the coast on the growth are the opposite of their effects on the level of development.

To account for within-country spillovers suggested by Gennaioli et al. (2014), I control for country level urban population density at the beginning of the period and report results in column (2). In the spirit of the results in Gennaioli et al (2014), I find the coefficient of the level of development measured by urban population density is positive, and adding country level development raises the estimated speed of convergence. As country-level GDP per capita is available in 1850, 1900, or 1950 for some countries in my sample, I check whether using national GDP per capita generates the similar results in column (4).[5] In column (3), I use model specification in column (2) and the restricted sample of regions and years in which national GDP per capita is available. The number of observations drops substantially from 2201 to 1399. The estimated convergence rate is, however, 0.15 %, which is essentially the same as in column (2). Column (4) presents a regression using country GDP per capita rather than urban population density as a proxy for the level of national development. Coefficient of country GDP is positive and significant, and the use of country GDP per capita does not alter the convergence rate, estimated at 0.14 %.

Since urban population density is derived from historical data, there may exist measurement error in the variable. The estimated convergence rate, subject to measurement error in urban population density, is therefore biased toward zero (Barro and Sala i Martin, 2004). I follow Barro and Sala i Martin (2004) and address measurement error using lagged urban population density as instruments in regression. I first display OLS results using a sample of restricted regions and years where lagged regional urban population density is available in column (5). The estimated rate of convergence rises to 0.21 %. In column (6), I show results based on IV regression and obtain a slightly higher estimated convergence than column (5), while the both are greater than other columns.

[5] The historical data for national GDP per capita is from Bolt and van Zanden (2014).

It is noteworthy that in columns (5) and (6) observations from the year of 1850 are dropped. This increases the estimated convergence rates in these two columns may suggest the speed of convergence may vary over different time periods. Later, I investigate regional convergence rates in subperiods in Section 4.3.4.

Table 4.2: Regional Convergence of Urban Population Density 1850-2000, OLS Estimation

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0009***	-0.0016***	-0.0015***	-0.0014***	-0.0021***	-0.0025***
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
Ln Urb. Pop. Den. (Country)		0.0020***	0.0006		0.0014***	0.0016***
		(0.0003)	(0.0004)		(0.0005)	(0.0005)
Ln GDP pc (Country)				0.0032***		
				(0.0007)		
Temperature	-0.0005***	-0.0007***	0.0002	0.0003	-0.0003	-0.0003
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Rainfall in Meter	-0.0014	-0.0026***	-0.0038***	-0.0032***	-0.0032**	-0.0032**
	(0.0009)	(0.0009)	(0.0011)	(0.0010)	(0.0013)	(0.0013)
Elevation (100 meters)	-0.0004***	-0.0006***	-0.0002	-0.0002	-0.0004***	-0.0004***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Ruggedness	0.0007*	0.0006	0.0013**	0.0016***	0.0012**	0.0012*
	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0006)	(0.0006)
Land Suitability	-0.0030**	-0.0049***	-0.0046***	-0.0040**	-0.0014	-0.0021
	(0.0014)	(0.0014)	(0.0017)	(0.0016)	(0.0020)	(0.0020)
Absolute Value of Latitude	-0.0005***	-0.0007***	-0.0005***	-0.0005***	-0.0006***	-0.0006***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Proximity to the Coast	-0.0077**	-0.0098***	-0.0059	-0.0082*	-0.0072	-0.0054
	(0.0034)	(0.0034)	(0.0045)	(0.0044)	(0.0045)	(0.0045)
Proximity to Rivers	0.0026	0.0035	0.0015	0.0009	0.0042	0.0047
	(0.0034)	(0.0033)	(0.0039)	(0.0038)	(0.0044)	(0.0044)
Instrument						Lagged Urb Pop. Den.
Countries	144	144	118	118	129	129
Observations	1888	1888	1255	1255	1061	1061
R^2	0.08	0.10	0.21	0.22	0.15	0.15

Note: The unit of observation is a subnational region. All regressions are OLS estimations with robust standard errors clustered at the country level. Columns (1) and (2) are regressions based on the whole sample. Columns (3) and (4) are based on the restricted sample of regions in which Log national GDP per capita in the beginning of the period is available. Columns (5) and (6) are based on the restricted sample of regions in which lagged urban population density is available. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.3.2 Additional Controls

Table 4.3 presents results controlling for country fixed effects. By including country fixed effects, I am able to control for time-invariant country characteristics such as culture, social norms, etc. that potentially correlated with both the initial level of development and growth of regions. For example, the same level of urban population density in two regions from different countries, say the US and India, does not necessarily mean these two regions have the same level of development. The measure does not reflect the fact that a large portion of urban residents in South Asia area are slum dwellers. The country fixed effects addresses this problem to some extent.

Column (1) presents country fixed effects results including the same independent variables as column (2) of Table 4.2. The estimated regional convergence rate goes up from 0.16 % to 0.22 % per year. The coefficient of country level urban population density is also much higher than OLS estimate. In addition, there are changes in the effects of the eight geographic and climatic controls. Only elevation, land suitability, proximity to the coast, and proximity to rivers have significant effects; elevation and land suitability are negatively correlated while proximity to the coast and proximity to rivers are positively correlated with average annual growth of urban population density.

Michalopoulos and Papaioannou (2014) find cross-ethnicity evidence in Africa that the correlation between national institutions and sub-national development depends on the distance from the national capital; the effect of national institutions on development of a region is weaker as the region is spatially more distant from its national capital. Therefore, nationwide characteristics such as national institutions and the law of origin may have different influences on development and economic growth across regions according to the distance from the national political center. For this reason, I include an indicator that the national capital exists in a region in column (2) and proximity to the capital city

Table 4.3: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation with Additional Controls

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0022***	-0.0032***	-0.0027***	-0.0034***	-0.0022***	-0.0034***
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
Ln Urb. Pop. Den. (Country)	0.0097***	0.0106***	0.0103***	0.0108***	0.0097***	0.0108***
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
Presence of National Capital		0.0099***		0.0088***		0.0088***
		(0.0013)		(0.0014)		(0.0014)
Proximity to Capital City			0.0140***	0.0076**		0.0079**
			(0.0037)	(0.0037)		(0.0038)
Proximity to Borders					-0.0004	0.0034
					(0.0031)	(0.0036)
Temperature	-0.0002	-0.0001	-0.0002	-0.0001	-0.0002	-0.0001
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Rainfall in Meter	0.0001	0.0005	0.0000	0.0004	0.0001	0.0004
	(0.0009)	(0.0008)	(0.0009)	(0.0008)	(0.0009)	(0.0008)
Elevation (100 meters)	-0.0003*	-0.0003*	-0.0003*	-0.0003*	-0.0003*	-0.0003*
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Ruggedness	0.0002	0.0004	0.0003	0.0004	0.0002	0.0004
	(0.0004)	(0.0005)	(0.0004)	(0.0005)	(0.0004)	(0.0005)
Land Suitability	-0.0020*	-0.0016	-0.0032**	-0.0023*	-0.0021*	-0.0022*
	(0.0012)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0013)
Absolute Value of Latitude	-0.0001	0.0000	-0.0002	-0.0001	-0.0001	-0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Proximity to the Coast	0.0083**	0.0130***	0.0084**	0.0126***	0.0083**	0.0128***
	(0.0035)	(0.0040)	(0.0042)	(0.0042)	(0.0036)	(0.0043)
Proximity to Rivers	0.0109***	0.0103***	0.0075***	0.0085***	0.0110***	0.0075**
	(0.0032)	(0.0032)	(0.0029)	(0.0029)	(0.0033)	(0.0030)
Countries	144	144	144	144	144	144
Observations	1888	1888	1888	1888	1888	1888
within R^2	0.17	0.19	0.18	0.19	0.17	0.19

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in column (3). Both are positively correlated with regional growth and include each of them substantially increases the estimated speed of convergence, estimated at 0.32 % in column (2) and 0.27 % in column (3). I include, both, the presence of national capital and the proximity to capital city and report results in column (4). Both variables remain positive and significant, indicating some nationwide factors have declining positive effects on regional growth for regions further from the national political center.

To make sure that it is the distance to national capital that matters rather than the distance to national borders, I include the proximity to national borders as an additional explanatory variable. Based on the results in columns (5) and (6), I find no evidence that proximity to national borders has any effect on the average annual growth of urban population density.

4.3.3 Spatial Interaction Effects

Due to technological spillovers and mobility of human capital and physical capital, economic prosperity in one region may be closely related to the state of development and the characteristics of neighboring regions . Failure to model spatial dependence between regions leads to omitted variable problems. In this study, I address potential spatial correlation by assuming that spatial correlation arises from the spillover effect of cities in neighboring areas on regions. I add a dummy which equals to one if at least a city existed within 25 miles of a region. Based on those surrounding cities, I create population density in the neighboring areas that equals the ratio of aggregated population in neighboring cities to land area of regions, from now on urban population density in neighboring areas. A positive spillover from neighboring cities suggests positive signs for both urban population density and a dummy for cities in neighboring areas.

In column (1) of Table 4.4, I present a regression that includes eight geographic and climatic characteristics, three variables identifying the distance from the national political center, as well as country fixed effects. The estimated speed of convergence is 0.34 % per year. In column (2), I use a smaller sample for those regions had at least a city in neighboring areas at the beginning of the period, and the estimated rate of convergence

Table 4.4: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0034***	-0.0039***	-0.0047***	-0.0038***	-0.0034***	-0.0039***
	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0002)
Ln Urb. Pop. Den. (Country)	0.0108***	0.0121***	0.0114***	0.0122***	0.0107***	-0.0070***
	(0.0014)	(0.0020)	(0.0020)	(0.0020)	(0.0014)	(0.0016)
Ln Urb. Pop. Den. in Neighboring Areas			0.0018***			
			(0.0005)			
Urb. Pop. Den. in Neighboring Areas				-0.0001		
				(0.0002)		
City in Neighboring Regions					0.0027***	
					(0.0010)	
Presence of National Capital	0.0088***	0.0097***	0.0118***	0.0095***	0.0095***	0.0074***
	(0.0014)	(0.0022)	(0.0023)	(0.0022)	(0.0015)	(0.0015)
Proximity to Capital City	0.0079**	0.0078	0.0016	0.0079	0.0058	0.0106**
	(0.0038)	(0.0049)	(0.0048)	(0.0050)	(0.0039)	(0.0041)
Proximity to Borders	0.0034	0.0062	0.0068*	0.0064	0.0035	0.0012
	(0.0036)	(0.0044)	(0.0041)	(0.0044)	(0.0036)	(0.0036)
Temperature	-0.0001	0.0002	0.0002	0.0002	-0.0001	-0.0000
	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Rainfall in Meter	0.0004	0.0013	0.0012	0.0012	0.0004	0.0005
	(0.0008)	(0.0014)	(0.0014)	(0.0014)	(0.0008)	(0.0008)
Elevation (100 meters)	-0.0003*	-0.0002	-0.0002	-0.0002	-0.0003*	-0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Ruggedness	0.0004	0.0000	0.0002	-0.0000	0.0003	0.0004
	(0.0005)	(0.0006)	(0.0005)	(0.0006)	(0.0004)	(0.0005)
Land Suitability	-0.0022*	-0.0028	-0.0029	-0.0030	-0.0029**	-0.0028*
	(0.0013)	(0.0020)	(0.0020)	(0.0021)	(0.0013)	(0.0015)
Absolute Latitude	-0.0001	0.0000	0.0001	-0.0000	-0.0001	-0.0000
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Proximity to the Coast	0.0128***	0.0079*	0.0052	0.0078*	0.0122***	0.0133***
	(0.0043)	(0.0043)	(0.0041)	(0.0043)	(0.0042)	(0.0045)
Proximity to Rivers	0.0075**	0.0055	0.0040	0.0055	0.0073**	0.0076**
	(0.0030)	(0.0064)	(0.0058)	(0.0064)	(0.0030)	(0.0035)
Year Fixed-Effects						Yes
Country Fixed-Effects	Yes	Yes	Yes	Yes	Yes	Yes
Countries	144	94	94	94	144	144
Observations	1888	1045	1045	1045	1888	1888
within R^2	0.19	0.21	0.22	0.21	0.20	0.40

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

increases to 0.39 %. Column (3) presents regressions controlling for log urban population density in the neighboring area. Every additional 1 % increase in urban population density in neighboring areas is, on average, associated with 0.18 %, higher the annual growth rate of urban population density. The correlation is significant at the 1 % level and including it raises the estimated convergence rate by 0.08 % indicating the existence of spatial dependence. However, when I include urban population density in the neighboring area instead of its log transformation in column (4), the spatial spillover effects disappear.

As an alternative way to model spatial correlation, I use a dummy that equals one if one or more cities existed within neighboring areas and report results based on the whole sample in column (5). Having at least a city in neighboring areas at the beginning of the period is, on average, correlated with 0.27 % higher the annual growth of urban population density, though adding the dummy variable does not change the estimate of the convergence rate (if comparing to the convergence rate in column (1) that is based on the whole sample). I use regression in column (5) as my baseline for a larger sample in the rest of this paper, although the inappropriate way to model spatial correlation may contribute to a downward biased estimate of convergence rate.

In the last column, I include year fixed effects based on the regression in column (1), the estimated convergence rate raises to 0.39 % per year indicating the convergence rate may be changing over the past 150 years. Therefore, I investigate the difference in convergence rate across various subperiods in Section 4.3.4.

4.3.4 Evidence in Subperiods of the Past 150 Years

With the four time points - 1850, 1900, 1950, and 2000, available in my data, the past 150 years can be divided into three 50-year subperiods, namely, 1850 - 1900, 1900 - 1950, and 1950 - 2000. The first subperiod started at the end of the first industrial revolution that began in Britain and spread to West European countries, the United States, and Japan. The standard of living started to improve during the period (Robert E. Lucas, 2002; Feinstein, 1998; Szreter and Mooney, 1998). In the next 50 years, from 1900 to 1950, a world system had emerged and linked most peoples in the world. For example, events such as the Great Depression and the two world wars had worldwide effects. Both total population and urban population grew rapidly. Between 1913 to 1950, world population increased from 1.8 billion to 2.5 billion while the percentage of people living cities rose from 18 % to 30 % (McNeill, 2001; Maddison, 2006). The last subperiod, 1950 to 2000, overlaps the post-war era. The world economy has been better integrated with establishment of the United Nations, the International Monetary Fund, the World Bank, etc. Given the variety of experiences across the three subperiods, it is reasonable to believe that regional convergence might behave differently from one subperiod to another. Also, previous regressions have also suggested that the speed of regional convergence varies in different subperiods of the past 150 years. I investigate this hypothesis in this section.

In Table 4.5, I regress the 50-year average annual growth rate of urban population density on log urban population density, log national urban population density, a dummy that equals one if cities existed in neighboring areas of a region at the beginning of the period and zero otherwise, three variables capturing the distance from the national political center, eight geographic and climatic controls, and country fixed effects, for various subperiods. Results based on all years, 1850, 1900, 1950, and 2000, are shown in column (1) where the estimated convergence rate is 0.34 %. In column (2), I exclude the year of 1850 in the regression, the estimated speed of convergence raises by 0.08 percent points, suggesting that including 1850 slows down the convergence rate. In column (3), instead, I exclude the year of 2000 in the regression, the estimated speed of convergence substantially drops to the half the level using all years, indicating an accelerating speed of convergence in the past 50 years.

In addition, I report results based on every two time points from 1850 - 1900 in column (4), 1900 - 1950 in column (5), and 1950 - 2000 in column (6). The estimated speed of regional convergence is 0.08 %, but insignificantly different from zero, during the first 50

Table 4.5: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation across Subperiods

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	All Years	1900-1950 -2000	1850-1900 -1950	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0034***	-0.0042***	-0.0018***	-0.0008	-0.0022***	-0.0049***
	(0.0003)	(0.0003)	(0.0004)	(0.0006)	(0.0005)	(0.0005)
Ln Urb. Pop. Den. (Country)	0.0107***	0.0105***	0.0019			
	(0.0014)	(0.0018)	(0.0023)			
City in Neighboring Regions	0.0027***	0.0035***	0.0017	0.0020*	0.0023*	0.0029***
	(0.0010)	(0.0010)	(0.0011)	(0.0011)	(0.0013)	(0.0011)
Presence of National Capital	0.0095***	0.0105***	0.0123***	0.0099***	0.0146***	0.0067**
	(0.0015)	(0.0020)	(0.0021)	(0.0027)	(0.0027)	(0.0026)
Proximity to Capital City	0.0058	0.0083*	-0.0067	-0.0086	-0.0081	0.0129***
	(0.0039)	(0.0045)	(0.0064)	(0.0098)	(0.0063)	(0.0048)
Proximity to Borders	0.0035	0.0040	0.0046	-0.0044	0.0107	-0.0063
	(0.0036)	(0.0039)	(0.0052)	(0.0069)	(0.0066)	(0.0039)
Temperature	-0.0001	0.0003	-0.0009*	-0.0006*	-0.0006	0.0007*
	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0006)	(0.0004)
Rainfall in Meter	0.0004	-0.0005	0.0005	0.0030	-0.0017	0.0006
	(0.0008)	(0.0008)	(0.0012)	(0.0018)	(0.0015)	(0.0012)
Elevation (100 meters)	-0.0003*	-0.0001	-0.0006***	-0.0004**	-0.0006**	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0002)
Ruggedness	0.0003	0.0003	0.0004	0.0002	0.0005	0.0006
	(0.0004)	(0.0005)	(0.0007)	(0.0007)	(0.0008)	(0.0009)
Land Suitability	-0.0029**	-0.0028**	-0.0048**	-0.0011	-0.0046**	-0.0022
	(0.0013)	(0.0014)	(0.0021)	(0.0037)	(0.0020)	(0.0019)
Absolute Latitude	-0.0001	-0.0000	-0.0001	0.0001	-0.0000	-0.0000
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0003)	(0.0002)
Proximity to the Coast	0.0122***	0.0137***	0.0160**	0.0015	0.0241***	0.0019
	(0.0042)	(0.0043)	(0.0068)	(0.0116)	(0.0049)	(0.0063)
Proximity to Rivers	0.0073**	0.0052	0.0132***	0.0243***	0.0065	0.0092**
	(0.0030)	(0.0041)	(0.0046)	(0.0087)	(0.0086)	(0.0046)
Countries	144	142	132	108	130	140
Observations	1888	1350	1134	538	596	754
within R^2	0.20	0.18	0.07	0.08	0.14	0.30

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

years of 1850 - 2000, 0.22 % during the subsequent 50 years, and 0.49 % during the last 50 years. The results reveal a pattern that regions started to converge mainly during the period between 1900 and 1950, and the speed of convergence has been increasing since then. This pattern is consistent with the change in living standard that began in the late 19th century. As the change in living standard varies in various countries, I investigate whether I can observe the similar pattern in different sub-samples of countries in the next section.

4.3.5 Evidence in Subsamples

Gennaioli et al. (2014) argue that the regional convergence rate of a country is correlated with the national market infrastructure, such as financial market regulation and international trade. They find cross-region evidence during 1950 - 2010 that countries with better economic and financial development or fewer barriers to international trade tend to have higher regional convergence rates in terms of GDP per capita. In this section, I also look for differences in the speed of regional convergence of urban population density across various continents in the world over the past 150 years.

Table 4.6 A displays regressions for regions from different continents for all years, For regressions covering all time points - 1850, 1900, 1950, and 2000, the estimated regional convergence rate for the United States excluding Washington, D.C. is 0.53 % in column (1), Latin America is 0.39 % in column (2), West European countries is 0.30 % in column (3), Other European countries is 0.33 % in column (4), Asia is 0.39 % in column (5), and Africa is 0.31 % in column (6). In general, the estimated regional convergence rates remain within a narrow range around 0.35 % across different areas in the world, which is comparable to the estimate based on the whole sample displayed in column (1) of Table 4.5.

Table 4.6 B investigates the regional convergence rates in each sub-sample across the three 50-year subperiods during the past 150 years, ordered from the first 50-year subperiod to the last in rows. According to the results, regions from non-West European and Asian

Table 4.6: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation across Country Groups Based on Continent

	US excl. DC	Latin America	West EU	Rest EU	Asia	Africa
	Panel A: Estimated Convergence Rates including All Years					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0063***	-0.0039***	-0.0017*	-0.0040***	-0.0040***	-0.0031**
	(0.0021)	(0.0007)	(0.0009)	(0.0007)	(0.0005)	(0.0014)
Ln Urb. Pop. Den.	0.0027	0.0131***	0.0042*	0.0023	0.0134***	0.0113***
(Country)	(0.0024)	(0.0025)	(0.0020)	(0.0021)	(0.0028)	(0.0027)
City in Neighboring	0.0071	0.0026	-0.0061	-0.0078**	0.0052***	0.0018
Regions	(0.0069)	(0.0019)	(0.0041)	(0.0036)	(0.0016)	(0.0036)
Presence of National		0.0105***	0.0122*	0.0200***	0.0076***	0.0036
Capital		(0.0021)	(0.0058)	(0.0062)	(0.0022)	(0.0058)
Proximity to Capital	-0.0071	0.0189***	-0.0170	-0.0358***	0.0176***	0.0257
City	(0.0169)	(0.0058)	(0.0223)	(0.0050)	(0.0051)	(0.0161)
Proximity to Borders	-0.0021	0.0123*	-0.0034	-0.0157	-0.0011	0.0296
	(0.0231)	(0.0070)	(0.0157)	(0.0091)	(0.0076)	(0.0177)
Countries	1	29	16	14	38	42
Observations	86	519	245	153	597	259
R^2	0.37	0.30	0.10	0.18	0.27	0.25
	Panel B: Estimated Convergence Rates in Subperiods					
	(1)	(2)	(3)	(4)	(5)	(6)
1850-1900	-0.0004	-0.0011	0.0030***	-0.0039**	-0.0022***	0.0000
	(0.0051)	(0.0010)	(0.0009)	(0.0016)	(0.0007)	(0.0014)
Countries	1	27	16	14	31	17
Observations	24	148	100	63	164	35
R^2	0.73	0.15	0.43	0.34	0.14	0.79
1900-1950	-0.0034	-0.0017***	-0.0002	-0.0025***	-0.0029***	-0.0023
	(0.0023)	(0.0006)	(0.0011)	(0.0007)	(0.0009)	(0.0017)
Countries	1	29	16	12	33	35
Observations	29	178	72	36	185	85
R^2	0.74	0.31	0.18	0.71	0.17	0.42
1950-2000	-0.0122***	-0.0062***	-0.0067***	-0.0040***	-0.0051***	-0.0032**
	(0.0018)	(0.0007)	(0.0004)	(0.0005)	(0.0008)	(0.0012)
Countries	1	27	16	12	38	42
Observations	33	193	73	54	248	139
R^2	0.85	0.43	0.73	0.62	0.31	0.30

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results in Panel B are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions except those in columns (1) include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

countries have evidence of convergence during 1850 to 1900. The estimated convergence rate is 0.39 % in non-West European regions and 0.22 % for regions in Asia. Regional convergence is either small in magnitude or insignificant for regions in the US, Latin America, and Africa. However, West European regions experienced surprisingly high divergence during the same period. In the second 50-year subperiod, regional convergence in West Europe was small in magnitude and insignificant. Other than that, regional convergence ranged from 0.17 % (in Latin America) to 0.34 % (in the United States), though estimated rates in the United States and Africa are not statistically different from zero. In the last 50-year subperiod, all sub-samples experienced convergence rates that are significantly different from zero, of which the highest is the United States, estimated at 1.22 %, and the lowest is Africa, estimated at 0.32 %. The speed of regional convergence for each continent during this subperiod is higher than ever before in the past 150 years.

The heterogeneity in regional convergence rate across various sub-samples indicates that convergence may be jointly driven by unobserved and hard-to-measure factors related to national characteristics. For example, Gennaioli et al. (2014) investigate regional convergence of GDP per capita during 1960 to 2010, and find that a national market infrastructure such as financial development, international trade, and government transfers raises the speed of convergence. This explanation can also be applied to the convergence rates of urban population density during 1950 to 2000, the period comparable to Gennaioli et al. (2014). The US, which has the best market infrastructure, has the highest regional convergence while African countries that have the poorest have the lowest regional convergence. However, the same explanation is not applicable to the results in other two subperiods of my sample.

However, I find that the changes in regional convergence rates across different subperiods are consistent with increases in the standard of living. For example, convergence rates in the United States and West European countries increased by over 0.90 % from the first 50-year subperiod to the third, while convergence rates in other countries increased by around the half of the level.

4.3.6 Robustness Checks

In order to ensure that measurement errors do not drive all the above conclusions, I try the following strategies: 1) I use alternative minimum population thresholds to reconstruct urban population density and the growth rate; 2) I report results based on quantile regressions; 3) I check whether regions in higher urban population density group are associated with lower growth rates, and 4) I look for evidence of convergence in sub-samples of regions according to the largest city size and regional population size.

Alternative Minimum Population Threshold

Urban population density in 1850, 1900, and 1950 is derived from historical data on population estimates of cities. To construct urban population density in these three time points, I first define cities using a minimum population of 5,000 as the threshold. However, for many countries (especially in Africa and Asia), only settlements whose estimated population reaches a much higher level, say 15,000 or 20,000, can be found in my city data. Therefore, the inconsistent definition of cities across different countries may contribute to a measurement error of urban population density although the use of country fixed effects theoretically addresses the problem. To investigate the problem associated with using various minimum population thresholds in my estimation, I reconstruct urban population density and the growth rate in these three years using higher minimum population thresholds of 20,000, 50,000, and 100,000 respectively. I report evidence in Table 4.7.

According to the results, the use of various minimum population thresholds has a minor impact on convergence. For regressions based on all years in panel A, convergence rates based on higher minimum population thresholds range from 0.33 % to 0.40 % while the original estimate was 0.34 %. For regressions based on the three subperiods in panel B, the estimated convergence rates with the use of higher minimum population thresholds are close to the original estimates.

Table 4.7: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation Using Alternative Minimum Population Thresholds in Constructing Urbanization Variables

	Dependent Variable: Average Annual Growth Rate of Urban Population Density			
	Based on localities with a minimum population of			
	5,000	20,000	50,000	100,000
	Panel A: Estimated Convergence Rates including All Years			
	(1)	(2)	(3)	(4)
Ln Urban Pop. Den.	-0.0034***	-0.0033***	-0.0035***	-0.0040***
	(0.0003)	(0.0003)	(0.0004)	(0.0006)
Countries	144	140	123	103
Observations	1888	1569	1011	698
within R^2	0.20	0.21	0.19	0.24
	Panel B: Estimated Convergence Rates in Subperiods			
	(1)	(2)	(3)	(4)
1850-1900	-0.0008	-0.0003	-0.0005	-0.0012
	(0.0006)	(0.0005)	(0.0006)	(0.0011)
Countries	108	91	54	32
Observations	538	391	180	88
within R^2	0.08	0.12	0.13	0.19
1900-1950	-0.0022***	-0.0020***	-0.0025***	-0.0028***
	(0.0005)	(0.0006)	(0.0007)	(0.0007)
Countries	130	108	85	59
Observations	596	452	306	207
within R^2	0.14	0.14	0.20	0.24
1950-2000	-0.0049***	-0.0048***	-0.0047***	-0.0054***
	(0.0005)	(0.0005)	(0.0006)	(0.0007)
Countries	140	137	121	103
Observations	754	726	525	403
within R^2	0.30	0.30	0.34	0.45

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Quantile Regressions

Quantile regression estimates parameters by minimizing the sum of absolute weighted deviations of the observed responses from the regression mean. The main purpose of quantile regression here is that quantile regression estimates are more robust against observations with unusually low or high annual growth rates of urban population density. Also, the use of quantile regression enables me to investigate the existence of regional convergence across various quantiles of the annual growth rate.

In Table 4.8, I report quantile regressions of the 50-year average annual growth rate of urban population density on log urban population density, log country level urban population, the presence of cities in neighboring areas at the beginning of the period, eight geographic and climatic controls, three variables measuring the proximity to the national capital center, and country fixed effects. The speed of convergence estimated using quantile regressions for quantiles 0.1 is displayed in column (1), 0.25 in column (2), 0.5 in column (3), 0.75 in column (4), and 0.9 in column (5). For results based on the whole sample displayed in Panel A, the estimates remain within a narrow range between 0.30 % to 0.34 %, which includes the rate estimated in the baseline model.

Table 4.8: Regional Convergence of Urban Population Density 1850-2000, Quantile Regression Estimation

	Quantile Regression at				
	Panel A: Estimated Convergence Rates including All Years				
	0.1	0.25	0.5	0.75	0.9
	(1)	(2)	(3)	(4)	(5)
Ln Urban Pop. Den.	-0.0030***	-0.0031***	-0.0032***	-0.0033***	-0.0034***
	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
Countries	144	144	144	144	144
Observations	1888	1888	1888	1888	1888
	Panel B: Estimated Convergence Rates in Subperiods				
	(1)	(2)	(3)	(4)	(5)
1850-1900	-0.0002	-0.0000	-0.0002	-0.0009**	-0.0020***
	(0.0006)	(0.0005)	(0.0005)	(0.0004)	(0.0006)
Observations	538	538	538	538	538
1900-1950	-0.0009*	-0.0013***	-0.0017***	-0.0031***	-0.0033***
	(0.0006)	(0.0005)	(0.0005)	(0.0004)	(0.0005)
Observations	596	596	596	596	596
1950-2000	-0.0047***	-0.0043***	-0.0056***	-0.0056***	-0.0058***
	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Observations	754	754	754	754	754

Note: The unit of observation is a subnational region. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Panel B, I report quantile regressions in the three subperiods - from 1850 to 1900, 1900 to 1950, and 1950 to 2000. For each subperiod, regions in upper quantiles have a relatively higher speed of convergence; however, the difference between quantiles declines over time. In addition, for regressions based on the same quantile, I show an increasing rate of convergence over time, which is consistent with the baseline results.

Overall, my quantile regressions suggest that my previous conclusions are unlikely to be driven by regions with unusually low/high growth rate of urban population density.

Monotonicity

The findings of regional convergence are derived from the estimated parameter of urban population density at the beginning of the period in the baseline model specification; if the coefficient is negative and significantly different from zero, there is convergence of regional development. The negative correlation means that poorer regions should grow at faster rates than richer regions. Therefore, a more straightforward way to verify the existence of convergence is by investigating whether regions with lower urban population density at the beginning of the period tend to have higher average annual growth rate of urban population density. I create six group dummy variables for regions with lowest urban population density to the highest in 1850, 1900, and 1950, and regress the average annual growth of urban population density on the group dummies controlling for log country level urban population, the presence of cities in neighboring areas at the beginning of the period, eight geographic and climatic controls. Results based on this strategy is reported in Table 4.9.

Regions in each subperiod are divided into six equal groups with arbitrary thresholds in Panel A and divided into six unequal groups according to selected thresholds calculated based on the mean and standard deviation of urban population density. In either scenario, the coefficients of dummies are negative and ascending in magnitude with density groups for periods 1900 - 1950 and 1950 - 2000, supporting a systematic convergence across all regions during the two subperiods. For the subperiod 1850 - 1900, however, convergence is only observed in regions with the lowest urban population density group. In general, findings using this strategy are consistent with the conclusions based on the baseline model.

Additional Robustness

In the last robustness check, I investigate whether the evidence of convergence is driven by regional characteristics associated with the size of the biggest city or total urban population, and therefore interact urban population density at the beginning of the period with these two features, respectively, and explore whether regional convergence still universally exists. Results are shown in Table 4.10.

In Panel A, for each subperiod, regions are separated into four groups based on the size of the biggest city, such as whether population in the regional largest city is greater than 100,000, between 50,000 and 100,000, between 20,000 and 50,000, or between 5,000 and 20,000. In Panel B, regions are equally divided into five groups according to the total urban population, from highest to the lowest. Coefficients of interactions between urban population density and group dummies are reported. Again, I find that convergence universally exists for periods 1900 - 1950 and 1950 - 2000, regardless of the size of the city or urban population.

4.4 Convergence Rates

Because of the strategy that uses the ratio of the total population living in cities to land area as the measure of regional development, the level of development for regions that had no cities is unknown and excluded in the analysis. All evidence so far is based on regions that had a settlement in 1850, 1900, or 1950. However, it would be interesting to know what the speed of regional convergence could be and how the convergence patterns would change in a larger sample that includes all regions.

Table 4.9: Correlation between Average Annual Growth Rate of Urban Population Density and Quintiles of Regions for Urban Population Density, Country Fixed Effects Estimation across Subperiods

	Dependent Variable: Annual Growth of Urban Population Density		
	Panel A: Quintiles of Regions for Urban Population Density		
	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)
Quintile with Smallest Urb. Pop. Den.	-0.0073**	-0.0056**	-0.0060**
	(0.0029)	(0.0023)	(0.0024)
The 2nd Smallest Quintile	-0.0072***	-0.0094***	-0.0113***
	(0.0026)	(0.0026)	(0.0022)
The 3rd Quintile	-0.0103***	-0.0121***	-0.0146***
	(0.0025)	(0.0034)	(0.0026)
The 4th Quintile	-0.0083***	-0.0134***	-0.0211***
	(0.0027)	(0.0039)	(0.0028)
Quintile with Largest Urb. Pop. Den.	-0.0060*	-0.0129***	-0.0310***
	(0.0031)	(0.0042)	(0.0037)
Countries	108	130	140
Observations	538	596	754
within R^2	0.11	0.16	0.26
	Panel B: Alternative Groups of Regions by Urban Population Density		
	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)
$\frac{1}{100}mean$ - $\frac{1}{48}mean$	-0.0064**	-0.0049***	-0.0074***
	(0.0030)	(0.0016)	(0.0016)
$\frac{1}{48}mean$ - $\frac{1}{24}mean$	-0.0089***	-0.0088***	-0.0099***
	(0.0029)	(0.0021)	(0.0020)
$\frac{1}{24}mean$ - $\frac{1}{8}mean$	-0.0089***	-0.0104***	-0.0155***
	(0.0025)	(0.0030)	(0.0022)
$\frac{1}{8}mean$ - $\frac{1}{8}mean + \frac{1}{6}S.D.$	-0.0083***	-0.0123***	-0.0206***
	(0.0027)	(0.0033)	(0.0027)
greater than $\frac{1}{8}mean + \frac{1}{6}S.D.$	-0.0061*	-0.0119***	-0.0319***
	(0.0036)	(0.0033)	(0.0046)
Countries	108	130	140
Observations	538	596	754
within R^2	0.11	0.15	0.25

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.10: Regional Convergence of Urban Population Density by City Size and Regional Urbanization Level, Country Fixed Effects Estimation across Subperiods

	Panel A: Interactions b/w Urb. Pop. Den. & Biggest City Sizes		
	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)
Regions within which population in biggest city > 1000			
Ln Urban Pop. Den.	-0.0004	-0.0014**	-0.0043***
	(0.0012)	(0.0005)	(0.0007)
Regions within which population in biggest city was between 50,000 - 100,000			
Ln Urban Pop. Den.	0.0007	-0.0018***	-0.0038***
	(0.0007)	(0.0006)	(0.0007)
Regions within which population in biggest city was between 20,000 - 50,000			
Ln Urban Pop. Den.	-0.0002	-0.0013**	-0.0050***
	(0.0006)	(0.0006)	(0.0005)
Regions within which population in biggest city was between 5,000 - 20,000			
Ln Urban Pop. Den.	-0.0013**	-0.0026***	-0.0068***
	(0.0006)	(0.0006)	(0.0011)
Countries	108	130	140
Observations	538	596	754
within R^2	0.11	0.17	0.32
	Panel B: Interactions b/w Urb. Pop. Den. and Quintiles of Regions for Urban Population		
	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)
Quintile with Largest Urban Population			
Ln Urban Pop. Den.	0.0011	-0.0009	-0.0053***
	(0.0010)	(0.0006)	(0.0009)
The 2nd Quintile			
Ln Urban Pop. Den.	0.0011	-0.0013**	-0.0040***
	(0.0008)	(0.0005)	(0.0007)
The 3rd Quintile			
Ln Urban Pop. Den.	0.0002	-0.0023***	-0.0041***
	(0.0005)	(0.0007)	(0.0006)
The 4th Quintile			
Ln Urban Pop. Den.	-0.0009*	-0.0012**	-0.0038***
	(0.0005)	(0.0005)	(0.0006)
Quintile with Smallest Urban Population			
Ln Urban Pop. Den.	-0.0015**	-0.0026***	-0.0057***
	(0.0007)	(0.0006)	(0.0006)
Countries	108	130	140
Observations	538	596	754
within R^2	0.14	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. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To answer these questions, I need to give a value of urban population density for regions with unknown historical urban population density. A small value will, however, inflate the estimated speed of convergence.[6] I use the national lowest urban population of a year to impute missing urban population and further replace missing urban population density by dividing the imputed urban population by regional land area. This value provides a conservative estimate of the speed of convergence for the whole sample, for the reason that the level of development for regions without urban population density is overstated and the overstatement lowers the estimated convergence rate.

I redo Tables 4.5 and 4.6 using a much larger sample including regions with the imputed values of urban population density and report results in Tables 4.11 and 4.12. The number of observations increases from 1,888 from 144 countries to 5,729 from 150 countries. The estimated convergence rate across all regions over the past 150 years, as displayed in column 1 of Table 4.11, is 0.61 % per year, almost the twice as large as in the smaller sample. Higher estimated convergence rates are observed across different areas in the world, as shown in Table 4.12. The rates in Latin America and Asia remain close to those based on the smaller sample, estimated at 0.43 % per year, although the number of observations in either area doubles. The rates in West European countries and Africa increase to over 0.70 % and other European countries is over 1 %. However, all previous patterns derived from the smaller sample still apply to here.

Furthermore, I also apply less aggressive ways to enlarge the baseline sample. In Tables A.9 and A.10, I assign the national lowest urban population of a year to regions with missing urban population only when they had urban population in the next available time point. For example, if I can observe urban population of a region in the year of 1950 but 1850 and 1900, I will only impute the region's urban population in 1900 using the national lowest urban population in 1900. In Tables A.11 and A.12, I do the similar thing; however, I

[6] If I uniformly assign a low value of the level of development of regions whose historical urban population density is unknown, the average annual growth rate of urban population density would be very high for those regions even if the level of contemporary development is not high. The estimated speed of convergence therefore increases when those regions are included.

directly use the national lowest urban population density of a year to impute the missing urban population density. The highest estimated speed of convergence is close to 0.90 % for all years and 1.04 % for the period between 1950 and 2000.

Table 4.11: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation across Subperiods

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	All Years	1900-2000	1850-1950	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0061***	-0.0062***	-0.0029***	-0.0010***	-0.0011***	-0.0070***
	(0.0006)	(0.0006)	(0.0007)	(0.0004)	(0.0004)	(0.0009)
Countries	150	145	142	120	135	142
Observations	5729	3958	3717	1771	1946	2012
within R^2	0.17	0.16	0.06	0.04	0.05	0.24

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. I use the lowest urban population within countries of a year to impute missing urban population and impute missing urban population density based on $\frac{ImputedUrbanPopulation}{LandArea}$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To further compare regional convergence rates estimated based on urban population density and GDP per capita, I work with a smaller sample of regions from Gennaioli et al. (2014) where both urban population and GDP per capita are available around 1950 and 2000. For regions without GDP per capita in 1950, I use GDP per capita in the closest year when the data is available. I report results in Table 4.13, and display the estimated speed of regional convergence of urban population density in Panel A and the one of GDP per capita in Panel B. For the comparison based on all available regions in columns (1) and (2) - 667 regions from 138 countries for urban population density and 537 regions from 34 countries for GDP per capita, the two types of convergence rates are not in the same magnitude - the rates based on regional income is much higher than that based on urban population density. However, if I estimate regional convergence rates based on 279 regions where both regional income and urban population density are available between 1950 and 2000, convergence rates estimated by income is around 50 % higher than that by urban population density. The rest of differences may be contributed by measurement error in regional GDP per capita. It is well-known that GDP is understated for poor regions but overstated for rich regions, and it is very likely that this problem get worse as one go further back in time, which gives rise to an overestimated convergence rate using GDP per capita.

Table 4.12: Regional Convergence of Urban Population Density 1850-2000, Country Fixed Effects Estimation across Country Groups Based on Continent

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	US excl. DC	Americas excl. US & Canada	West EU	Rest EU	Asia	Africa
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0083***	-0.0041***	-0.0079***	-0.0103***	-0.0045***	-0.0075***
	(0.0020)	(0.0009)	(0.0011)	(0.0010)	(0.0006)	(0.0018)
Countries	1	29	16	20	38	42
Observations	150	1408	633	793	1495	1118
R^2	0.21	0.08	0.39	0.47	0.29	0.26
	Estimated Convergence Rates in Subperiods					
	(1)	(2)	(3)	(4)	(5)	(6)
1850-1900	0.0018	-0.0005	-0.0009	-0.0022**	-0.0018***	-0.0006
	(0.0026)	(0.0006)	(0.0015)	(0.0008)	(0.0005)	(0.0007)
Countries	1	27	16	20	34	20
Observations	50	454	211	294	482	258
R^2	0.32	0.04	0.27	0.14	0.09	0.09
1900-1950	0.0048	-0.0012**	-0.0015	-0.0054**	-0.0015**	0.0007
	(0.0031)	(0.0005)	(0.0012)	(0.0024)	(0.0006)	(0.0008)
Countries	1	29	16	15	35	35
Observations	50	477	211	261	490	402
R^2	0.38	0.16	0.18	0.20	0.10	0.20
1950-2000	-0.0130***	-0.0059***	-0.0091***	-0.0135***	-0.0060***	-0.0040***
	(0.0012)	(0.0012)	(0.0011)	(0.0003)	(0.0009)	(0.0013)
Countries	1	29	16	12	38	42
Observations	50	477	211	238	523	458
R^2	0.88	0.16	0.47	0.74	0.22	0.22

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions except those in columns (1) include country fixed effects. I use the lowest urban population within countries of a year to impute missing urban population and impute missing urban population density based on $\frac{ImputedUrbanPopulation}{LandArea}$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.5 Conclusions

The debate regarding whether there exists convergence of development has shifted its attention from between-country disparity to cross-region inequality within countries. The most recent work on regional convergence, by Gennaioli et al. (2014), using newly constructed data for regional income for regions from 83 countries over the past 60 years, documents that regions have been converging to each other with a rate of around 2 % per year. No other work has covered regions in so many countries and spanned such a long time. In this paper, I complement the literature by investigating regional convergence worldwide over a much longer period.

Table 4.13: Regional Convergence of GDP per Capita and Urban Population Density 1950 - 2000, for Selected Regions

	Dependent Variable: 50-Year Average Annual Growth Rate of			
	All available regions		Shared regions	
	OLE	FE	OLS	FE
	Panel A: Urban Population Density			
	(1)	(2)	(3)	(4)
Ln Urban Pop. Den.	-0.0023***	-0.0023***	-0.0059***	-0.0066***
	(0.0004)	(0.0005)	(0.0011)	(0.0010)
Countries	138	138	29	29
Observations	667	667	279	279
R^2	0.27	0.09	0.34	0.24
	Panel B: GDP per capita			
	(1)	(2)	(3)	(4)
Ln GDP per capita	-0.0113***	-0.0132***	-0.0093***	-0.0106***
	(0.0013)	(0.0022)	(0.0015)	(0.0014)
Countries	34	34	29	29
Observations	537	537	279	279
R^2	0.48	0.41	0.52	0.29

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions except those in columns (1) and (3) include country fixed effects.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I use urban population density as a measure of development and investigate the existence of regional convergence. With the use of historical population estimates of cities, I construct the measure of development in 1850, 1900, and 1950, and estimate the speed of convergence of regional development covering regions from 145 countries in the world during 1850 to 2000. I find that worldwide convergence of regional development has mainly occurred in the last 100 years and has been increasing over time. I find evidence that changes in the speed of regional convergence are correlated with increases in overall productivity over the past 150 years. My results are comparable to Gennaioli et al. (2014). For the subperiod between 1950 and 2000, regions from country groups with better financial regulation and fewer barriers to international trade tend to have relatively higher convergence rates.

Combing the results in this paper with Chapter 3 which finds persistence in regional development in the past 150 years. I find that both convergence of regional development and persistence in regional prosperity exist over the past 100 years. However, the accelerating convergence rates everywhere in the world suggests that regional disparity is getting smaller and smaller, and the regional disparity is diminishing at an increasing rate.

The estimated convergence rate based on the main sample is around 0.35 %, much less than 2 % documented in the literature. There are three explanations for the difference. First, the measure of development, urban population density, is different from the traditional measure of development - GDP per capita used in convergence studies. Since many regions did not have a city until decades ago, regions that might have small and unknown values of development can not be measured with urban population density and are therefore excluded in this study. Excluding observations with the small level of development is expected to lower the estimated convergence rates. I use a conservative method to impute the degree of development of those regions and include them in regressions, and receive a much higher regional convergence rate. Second, I study regional convergence spanning a much longer period. As I have previously showed that convergence has been increasing over the past 150 years, the convergence rate based on recent years should be higher. The convergence rate based on my primary sample between 1950 and 2000 in this study is around 0.5 %, higher than estimated using all years. The last source of the difference may be explained by measurement error in GDP per capita. My evidence indicates that convergence rates estimated using both variables are comparable.

Chapter 5. Conclusion

The use of Geographic Information Systems (GIS) has vastly expanded the frontier of economics research for the past few years. This dissertation serves as an additional example of the use of GIS in research to uncover historical patterns and fills the void of lack of appropriate data in research on long-run regional growth. In addition, GIS enables me to study long-run regional growth within the framework of spatial model and utilize abundant geographic and climatic spatial data. I primarily investigate two major questions exist in regional economics over the past 150 years using a global sample: 1) the extent to which regional disparity persists, and 2) the existence of regional convergence.

For the first question, I find global evidence that regions that were relatively more developed in 1850 are, on average, richer today. I reveal two drivers of the persistence pattern: 1) locational advantages that favor economic development, and 2) a consequence of path dependence.

For the question on convergence, I document that worldwide convergence of regional development had not occurred until the beginning of the 20th century, coincident with the time when living standards began to consistently grow for the first time in the history. Moreover, the convergence rate has been increasing for most countries during the past 100 years.

Combining the two patterns, one can have a more clear picture on the regional disparity over the past 100 years and foresee where it heads. It is clear that the dramatic development of transportation and communication technologies has reshaped this world into a more integrated one, in which, boundaries of regions are fading and costs of mobility are no long as important.

References

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

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

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

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 107 (4):1231–1294.

Alesina, Alberto, Elias Papaioannou, and Stelios Michalopoulos. 2012. "Ethnic Inequality." NBER Working Papers 18512, National Bureau of Economic Research.

Angrist, Joshua D. and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

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.

Bairoch, Paul. 1998. *Population des villes européennes de 800 à 1850 : banque de données et analyse sommaire des résultats (la).* Librairie Droz.

Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries." 106 (2):407–443.

———. 2012. "Convergence and Modernization Revisited." Working Paper 18295, National Bureau of Economic Research.

Barro, Robert J. and Xavier Sala i Martin. 1991. "Convergence Across States and Regions." *Brookings Papers on Economic Activity* 1991 (1):pp. 107–182.

———. 1992. "Convergence." *Journal of Political Economy* 100 (2):223–251.

———. 2004. *Economic Growth.* McGraw-Hill.

Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127 (2):587–644.

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.

Bolt, J. and J. L. van Zanden. 2014. “The Maddison Project: collaborative research on historical national accounts.” *The Economic History Review* 67 (3):627Ű651.

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.

Breinlich, Holger, Gianmarco I. P. Ottaviano, and Jonathan R. W. Temple. 2013. “Regional Growth and Regional Decline.” Cep discussion papers, Centre for Economic Performance, LSE.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. “The Political Economy of Deforestation in the Tropics*.” 127 (4):1707–1754.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel. 2013. “The Value of Democracy: Evidence from Road Building in Kenya.” Working Paper 19398, National Bureau of Economic Research.

Cantoni, Davide. 2014. “The Economic Effects of The Protestant Reformation: Testing The Weber Hypothesis in The German Lands.” *Journal of the European Economic Association* 00 (00):1–38.

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort. 1996. “Reopening the Convergence Debate: A New Look at Cross-Country Grwoth Empirics.” *Journal of Economic Growth* 1 (3):363–389.

Cashdm, Paul. 1995. “Economic Growth and Convergence Across the Seven Colonies of Australasia: 1861Ű1991.” *Economic Record* 71 (2):132–144.

Chanda, Areendam, C. Justin Cook, and Louis Putterman. 2014. “Persistence of Fortune: Accounting for Population Movements, There Was No Post-Columbian Reversal.” *American Economic Journal: Macroeconomics* 6 (3):1–28.

Chandler, Tertius. 1987. *Four Thousand Years of Urban Growth: An Historical Census*. St. David’s University Press.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2010. *Estimating Agglomeration Economies with History, Geology, and Worker Effects*. University of Chicago Press, 15–66.

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.

Coulombe, Serge and Jean-François Tremblay. 2001. “Human capital and regional convergence in Canada.” *Journal of Economic Studies* 28 (3):154–180.

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.

Dell, Melissa. 2009. *GIS Analysis for Applied Economists*. Manuscript.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. "Temperature Shocks and Economic Growth: Evidence from the Last Half Century." *American Economic Journal: Macroeconomics* 4 (3):66–95.

Donaldson, Dave. 2010. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." Tech. rep.

Douglas A. Hibbs, 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 (10):3715–3720.

Durlauf, Steven N. and Danny T. Quah. 1999. "The new empirics of economic growth." In *Handbook of Macroeconomics*, edited by J. B. Taylor and M. Woodford. Elsevier.

Easterlin, Richard. 1960. "Interregional Differences in Per Capita Income, Population, and Total Income, 1840-1950." In *Trends in the American Economy in the Nineteenth Century*. National Bureau of Economic Research, Inc.

Eggimann, Gilbert. 1994. *La population des villes des tiers-mondes, 1500-1950: banque de données*. Librairie Droz.

Feinstein, Charles H. 1998. "Pessimism Perpetuated: Real Wages and the Standard of Living in Britain during and after the Industrial Revolution." *The Journal of Economic History* 58 (3):625–658.

Feyrer, James and Bruce Sacerdote. 2009. "Colonialism and Modern Income: Islands as Natural Experiments." *The Review of Economics and Statistics* 91 (2):245–262.

Frankel, Jeffrey A. and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89 (3):379–399.

Galor, Oded and Omer Moav. 2007. "The Neolithic Revolution and Contemporary Variations in Life Expectancy." Tech. rep.

Garofalo, Gasper A. and Steven Yamarik. 2002. "Regional Convergence: Evidence From A New State-By-State Capital Stock Series." *The Review of Economics and Statistics* 84 (2):316–323.

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.

———. 2014. "Growth in Regions." *Journal of Economic growth* 19 (3):259–309.

Hall, Robert E. and Charles I. Jones. 1999. "Why Do Some Countries Produce so much more Output per Worker than Others?" *Quarterly Journal of Economics* 114 (1):83–116.

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

Hodler, Roland and Paul A. Raschky. 2014. "Regional Favoritism." *The Quarterly Journal of Economics* 129 (2):995–1033.

Linden, Leigh and Jonah E. Rockoff. 2008. "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *American Economic Review* 98 (3):1103–1127.

Lowe, Matt. 2014. *Night Lights and ArcGIS: A Brief Guide*. Manuscript.

Maddison, Angus. 2006. *The World Economy*. OECD Publishing.

Magrini, Stefano. 2004. "Regional (di)convergence." In *Handbook of Regional and Urban Economics*, *Handbook of Regional and Urban Economics*, vol. 4, edited by J. V. Henderson and J. F. Thisse, chap. 62. Elsevier, 2741–2796.

Maloney, William F. and Felipe Valencia Caicedo. 2015. "The Persistence of (Subnational) Fortune." *The Economic Journal* :n/a–n/a.

Mankiw, N. Gregory, David Romer, and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107 (2):407–437.

McNeill, J. R. 2001. *Something New under the Sun: An Environmental History of the Twentieth-Century World*. W. W. Norton & Company.

Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review* 102 (4):1508–1539.

Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81 (1):113–152.

———. 2014. "National Institutions and Subnational Development in Africa." 129 (1):151–213.

Mitton, Todd. 2013. "The Wealth of Subnations: Geography, Institutions and Within-country Development."

Modelski, George. 2003. *World Cities: -3000 to 2000*. Washington, DC: FAROS2000.

Nunn, Nathan. 2008. "The Long-term Effects of Africa's Slave Trades." 123 (1):139–176.

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

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

of world population history, Atlas. 1978. *Colin McEvedy and Richard Jones*. Penguin.

Olsson, Ola and Douglas A. Hibbs. 2005. "Biogeography and Long-Run Economic Development." *European Economic Review* 49:909–938.

Putterman, Louis and 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.

Rafael, Edward Glaeser, Rafael La Porta, Florencio Lopez de silanes, and Andrei Shleifer. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9:271–303.

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

Rappaport, Jordan and Jeffrey D. Sachs. 2003. "The United States as a coastal nation." *Journal of Economic Growth* 8 (1):5–46.

Robert E. Lucas, Jr. 2002. *Lectures on Economic Growth*. Harvard University Press.

Rosenthal, Stuart S. and William C. Strange. 2008. "The attenuation of human capital spillovers." *Journal of Urban Economics* 64 (2):373–389.

Shioji, Etsuro. 1996. "Regional Growth in Japan." Cepr discussion papers, C.E.P.R. Discussion Papers.

Storeygard, Adam. 2013. "Farther on down the road : transport costs, trade and urban growth in Sub-Saharan Africa." Policy research working paper series, The World Bank.

Szreter, Simon and Graham Mooney. 1998. "Urbanization, Mortality, and the Standard of Living Debate: New Estimates of the Expectation of Life at Birth in Nineteenth-century British Cities." *The Economic History Review* 51 (1):84–112.

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

Vigdor, Jacob. 2008. "The Economic Aftermath of Hurricane Katrina." *Journal of Economic Perspectives* 22 (4):135–154.

Appendix: Relevant Extra Material

A Definition of Region

This section describes regions used in this paper. I match the Gennaioli et al. (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 the Gennaioli et al. (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. I 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. I 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, I exclude regions in countries that do not have a single region with settlement data.

B Urban Population

In this study, I include any location that has a recorded population of 5,000 or more in 1850 from my sources. In an effort to enlarge my sample, I also include locations with records from 1825 and 1875 but none in 1850. Melbourne is therefore considered a city in 1850 even though data for its estimated population, 222,000 according to Rozenblat's estimates, is only available in 1875. When all of my data sources are taken together, I 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 I am able to confirm in which region the city locates. 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, 6 are city states - Gibraltar, Guernsey, Hong Kong, Macau, Malta, and Singapore, I drop them in the study. I end up with 766 regions from 128 countries in my whole sample that had urban areas in 1850.

C Why Not A Log-Log Model ?

This section explains why the use of a log-log specification can invalidate my estimates. Consider the simplest case where the only two variables of interest are the dummy variable for the existence of a city and the region's urban population density. My urban population density measure is positive and continuous for some observations (regions with cities) and 0 for others. Thus, depending on the observation(i.e. region), the implied estimation takes one of two forms,

$$\text{LnY}_i = \begin{cases} \beta_0 + \delta_i + \beta_1 \text{Ln} x_i + \mu_i & \text{if region } i \text{ has a positive } x \\ \beta_0 + \mu_i & \text{otherwise,} \end{cases} \tag{C.1}$$

where for the i^{th} observation, β_1 is the elasticity of Y_i with respect to x, δ captures the difference in LnY_i between regions with a positive value of x and regions with a value of 0; I am interested in both variables. The coefficient β_0 is constant which are the same for both types of regions, and μ_i is a white noise.

The unit of variable x is arbitrarily chosen. The unit in 100 persons per square kilometer is not theoretically more correct than 1 person per square kilometer. However, the scaling of x will eventually contaminate the estimated δ. I show this in the following two equations in which I scale up x by 100 times.

$$\mathrm{LnY}_i = \begin{cases} \beta_0 + \delta_i + \beta_1 \mathrm{Ln}(x_i * 100) + \mu_i & \text{if region } i \text{ has a positive } x \\ \beta_0 + \mu_i & \text{otherwise,} \end{cases} \tag{C.2}$$

$$\mathrm{LnY}_i = \begin{cases} \beta_0 + [\delta_i + \beta_1 \mathrm{Ln}(100)] + \beta_1 \mathrm{Ln} x_i + \mu_i & \text{if region } i \text{ has a positive } x \\ \beta_0 + \mu_i & \text{otherwise,} \end{cases} \tag{C.3}$$

In Equation B.2, I scale up x by 100 times. Because $\mathrm{Ln}(x_i * 100)$ is equal to $\mathrm{Ln}x_i$ plus Ln(100), then I have Equation B.3. The estimated β_1 is the same as it is estimated in Equation B.1. However, δ and a constant, Ln(100), resulted from scaling up of x are estimated as a whole. Because the unique 'real' unit of x that does not exist, I therefore are not able to depart the constant from δ.

D Appendix Tables

Table A.1: 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

Continued on next page...

Table A.1 – Continued

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
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

Continued on next page...

Table A.1 – Continued

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
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

Table A.2: Variables, Descriptions, and Sources

Variable	Description	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, I load coordinates of the localities in 1850 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I code 1 for regions contain at least one of these coordinates; 0, otherwise.	Chandler (1987), Bairoch (1998), and Eggimann (1994). GADM 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, I load coordinates of the localities in 1850 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I 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.	Chandler (1987), Bairoch (1998), and Eggimann (1994). GADM database of Global Administrative Areas.
Urban Population Density in 1850	Regional population density in the urban areas in 1850 in 100 urban inhabitants per square kilometer. To generate this variable, I load localities in 1850 with population greater than 5,000 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take ratio of the total population in cities within regions to the land area of the regions.	Chandler (1987), Bairoch (1998), and Eggimann (1994). GADM database of Global Administrative Areas.
Urban Pop. Den. in Neighboring Regions, 1850	100 surrounding urban inhabitants per square kilometer of the region. To generate this variable, I load localities in 1850 with population greater than 5,000 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take ratio of the total population in cities within 25 miles away from regions' boundaries to the land area of the regions.	Chandler (1987), Bairoch (1998), and Eggimann (1994). GADM database of Global Administrative Areas.
	Dependent Variables:	
GDP per capita	Regional income per capita in PPP constant 2005 international dollars in 2005.	Gennaioli et al. (2013).
Ln(Avg. Nighttime Light Density), 2001-05	The logarithm of average nighttime light intensity yearly averaged through 2001 to 2005. To produce the regional numbers, I load the night lights data in 5 years from 2001 to 2005 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take the ratio of total light intensity in each region to the land area of the region.	National Geophysical Data Center (NOAA). GADM database of Global Administrative Areas.

Continued on next page...

Urban Population Density in 2000	Regional population density in the urban areas in 2000 in 100 persons per square kilometer. To produce the numbers, I load global settlement points grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take the ratio of total population living in the urban within a region to the land area of the region.	Center for International Earth Science Information Network (CIESIN). GADM database of Global Administrative Areas.
Trust in Others	Percent of respondents who think most people can be trusted.	Gennaioli et al. (2013).
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.	Gennaioli et al. (2013).
Access to Financing	Percent of respondents think that access to financing is at least a moderate obstacle to business.	Gennaioli et al. (2013).
Ln(Days without Electricity)	The logarithm of 1 plus the regional average of days with no electricity in the past year reported by respondents.	Gennaioli et al. (2013).
Ln(Power Line Density)	The logarithm of 1 plus the length in kilometers of power lines per 10 square kilometers in 2007.	Gennaioli et al. (2013).
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.	Gennaioli et al. (2013).
	Baseline regional controls:	
Temperature	Average temperature during 1950 - 2000 in Celsius. To produce the regional numbers, I load the global temperature grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take average of the temperature within regions.	Global Climate Data (WorldClim). GADM database of Global Administrative Areas.
Land Suitability	An index of the suitability for agriculture based on temperature and soil quality measurements. To produce the regional numbers, I load the world suitability for agriculture grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take average of the index within regions.	Global Climate Data (WorldClim). GADM database of Global Administrative Areas.

Continued on next page...

Table A.2 – Continued

Altitude	Average altitude in regions in 100 meters. To produce the regional numbers, I load the global altitude grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take average of the value within regions.	Global Climate Data (WorldClim). GADM database of Global Administrative Areas.
Ruggedness	Average terrain ruggedness in regions in 100 meters. To produce the regional numbers, I load the global terrain ruggedness index grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take average of the value within regions.	Nunn and Puga (2012). GADM database of Global Administrative Areas.
Rainfall	Average precipitation in regions during 1950 - 2000 in meter. To produce the regional numbers, I load the global precipitation grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I take average of the value within regions.	Global Climate Data (WorldClim). GADM database of Global Administrative Areas.
Absolute Value of Latitude	Absolute value of latitude of regional centroid. To produce the regional numbers, I load the worldwide regions' digital map derived from the Database of Global Administrative Areas. I generate regions' median centroid and keep coordinates of them.	GADM database of Global Administrative Areas.
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, I load the world coastline grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I generate regions' median centroid and keep coordinates of them. I calculate the distance of the centroid to nearest coastlines.	National Geophysical Data Center (NOAA). GADM database of Global Administrative Areas.
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, I load the world river grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I generate regions' median centroid and keep coordinates of them. I calculate the distance of the centroid to nearest rivers.	National Geophysical Data Center (NOAA). GADM database of Global Administrative Areas.
	Other regional controls:	
Proximity to Capital City	The reciprocal of 1 plus the distance of regions' centroid to their own national capitals in 1,000 kilometers. To produce the numbers, I input national capitals' coordinates and make the world capitals grid and load the worldwide regions' digital map derived from the Database of Global Administrative Areas. I generate regions' median centroid and keep coordinates of them. I calculate the distance of the centroid to national capitals.	GADM database of Global Administrative Areas.

Continued on next page...

Proximity to Borders	The reciprocal of 1 plus the distance of regions' centroid to the nearest national borderlines in 1,000 kilometers. To produce the numbers, I load the world national borderlines grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I generate regions' median centroid and keep coordinates of them. I calculate the distance of the centroid to nearest national borderlines.	National Geophysical Data Center (NOAA). GADM database of Global Administrative Areas.
Presence of National Capital	A dummy indicating regions in which national capitals exist. To produce the numbers, I input national capitals' coordinates and make the world capitals grid and load the worldwide regions' digital map derived from the Database of Global Administrative Areas. I code 1 for regions contain national capitals; 0, otherwise.	GADM database of Global Administrative Areas.
Largest National City in 1850	A dummy indicating regions in which the biggest year 1850 locality (localities if there were several with the same population size) within contemporary national boundaries existed. To generate this variable, I load coordinates of all localities in 1850 and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I look for the localities with the largest population size within each contemporary national boundary and code 1 for regions contain any of these localities; 0, otherwise.	Chandler (1987), Bairoch (1998), and Eggimann (1994). GADM database of Global Administrative Areas.
Presence of Diamond Mines	A dummy equals to one if a diamond mine exists in a region. To produce the numbers, I load diamond mine grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I code 1 for regions that intersect with at least a diamond mine.	Peace Research Institute Oslo (PRIO). GADM database of Global Administrative Areas.
Log Population	The number of inhabitants in the region in 2000. To produce the numbers, I load global population grid and the worldwide regions' digital map derived from the Database of Global Administrative Areas. I sum population within regions.	Center for International Earth Science Information Network (CIESIN). GADM database of Global Administrative Areas.
Years of Education	Average years of schooling beyond primary school for those who are 15 years old and older.	Gennaioli et al. (2013).
Ln(Oil Production/Capita)	Logarithm of 1 plus the estimated per capita volume of cumulative oil production and reserves in millions of barrels of oil	Gennaioli et al. (2013).

Table A.3: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850, For Regions with Positive Urban Population

	Dependent Variable: Log of Regional GDP per Capita (PPP), 2005							
	Panel A: Regions with Urbanization				Panel B: Regions without Urbanization			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Existence of a City 1850								
Urban Population Density 1850	0.100***	0.104***	0.358***	0.367***				
	(0.026)	(0.036)	(0.065)	(0.068)				
Square Urban Pop. Den. 1850			-0.023***	-0.024***				
			(0.006)	(0.006)				
City in Neighboring Regions 1850		0.029		0.049		-0.082		-0.098
		(0.048)		(0.042)		(0.068)		(0.068)
Urb. Pop. Den. in Neib. 1850		-0.006		0.025		0.017		0.113*
		(0.020)		(0.024)		(0.014)		(0.062)
Squ. Urb. Pop. Den. in Neib. 1850				-0.002				-0.003*
				(0.001)				(0.002)
Countries	88	88	88	88	77	77	77	77
Observations	668	668	668	668	727	727	727	727
within R^2	0.09	0.09	0.17	0.18	0.00	0.01	0.00	0.02

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. No controls are included. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	Dependent Variable:			
	Ln(GDP per capita, 2005)	Ln(Ave. nighttime luminosity) 2001-2005	Urbanization Rate in 2000	Ln(Pop. Density in 2000)
	Panel A: Quintiles of Regions for Urban Population Density in 1850			
	(1)	(2)	(3)	(4)
Existence of a City (1850)	0.061**	0.729***	0.100***	0.665***
	(0.030)	(0.096)	(0.018)	(0.090)
Regions within which population in biggest city in 1850 > 100,000, 90 regions				
Urban Population Density 1850	0.215***	0.760***	0.140***	1.312***
	(0.059)	(0.144)	(0.019)	(0.173)
Square Urban Pop. Den. 1850	-0.013***	-0.048***	-0.008***	-0.078***
	(0.004)	(0.011)	(0.002)	(0.014)
Regions within which population in biggest city in 1850 was between 50,000 - 100,000, 105 regions				
Urban Population Density 1850	0.782***	2.351***	0.544***	2.672***
	(0.257)	(0.377)	(0.061)	(0.573)
Square Urban Pop. Den. 1850	-0.117**	-0.362***	-0.142***	-0.323***
	(0.056)	(0.082)	(0.014)	(0.123)
Regions within which population in biggest city in 1850 was between 20,000 - 50,000, 296 regions				
Urban Population Density 1850	0.422***	1.078***	0.217***	1.390***
	(0.068)	(0.327)	(0.058)	(0.427)
Square Urban Pop. Den. 1850	-0.035***	-0.091***	-0.016***	-0.103***
	(0.006)	(0.028)	(0.005)	(0.037)
Regions within which population in biggest city in 1850 was between 5,000 - 20,000, 275 regions				
Urban Population Density 1850	0.212	0.022	-0.029	0.058
	(0.470)	(0.097)	(0.030)	(0.141)
Square Urban Pop. Den. 1850	0.049	-0.000	0.000	-0.001
	(0.319)	(0.003)	(0.001)	(0.004)
City in Neighboring Regions (1850)	-0.041	0.472***	-0.020	0.483***
	(0.031)	(0.081)	(0.013)	(0.074)
Urb. Pop. Den. in Neib. 1850	0.057*	0.031**	-0.001	0.048**
	(0.029)	(0.015)	(0.006)	(0.023)
Squ. Urb. Pop. Den. in Neib. 1850	-0.002*	-0.000**	-0.000	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)
Countries	92	135	135	135
Observations	1395	2044	2050	2058
within R^2	0.21	0.27	0.17	0.36

Continued on next page...

Table A.4 – Continued

	Panel B: Alternative Groups of Regions by Urban Population Density in 1850			
	(1)	(2)	(3)	(4)
Existence of a City (1850)	0.046	0.724***	0.098***	0.635***
	(0.030)	(0.096)	(0.018)	(0.089)
Quintile with Largest Urban Population in 1850				
Urban Population Density 1850	0.237***	0.818***	0.145***	1.415***
	(0.065)	(0.165)	(0.024)	(0.191)
Square Urban Pop. Den. 1850	-0.015***	-0.053***	-0.009***	-0.085***
	(0.005)	(0.013)	(0.002)	(0.016)
The 2nd Largest Quintile				
Urban Population Density 1850	0.718***	1.151***	0.138	1.746***
	(0.250)	(0.203)	(0.086)	(0.268)
Square Urban Pop. Den. 1850	-0.109*	-0.092***	-0.010	-0.131***
	(0.059)	(0.017)	(0.007)	(0.022)
The 3rd Quintile				
Urban Population Density 1850	0.446***	1.524***	0.326***	2.032***
	(0.114)	(0.532)	(0.075)	(0.529)
Square Urban Pop. Den. 1850	-0.040***	-0.152***	-0.028***	-0.188***
	(0.011)	(0.056)	(0.008)	(0.054)
The 4th Quintile				
Urban Population Density 1850	2.225***	0.278*	0.054***	0.600***
	(0.732)	(0.157)	(0.013)	(0.155)
Square Urban Pop. Den. 1850	-0.185***	-0.008*	-0.002***	-0.017***
	(0.062)	(0.005)	(0.000)	(0.005)
Quintile with Smallest Urban Population in 1850				
Urban Population Density 1850	0.203	0.664	0.233***	1.192**
	(0.719)	(0.595)	(0.084)	(0.480)
Square Urban Pop. Den. 1850	0.066	-0.050	-0.019***	-0.088**
	(0.488)	(0.042)	(0.006)	(0.034)
City in Neighboring Regions (1850)	-0.033	0.491***	-0.016	0.511***
	(0.033)	(0.083)	(0.013)	(0.075)
Urb. Pop. Den. in Neib. 1850	0.056	0.031*	-0.002	0.047*
	(0.035)	(0.017)	(0.006)	(0.026)
Squ. Urb. Pop. Den. in Neib. 1850	-0.002	-0.000**	-0.000	-0.000**
	(0.001)	(0.000)	(0.000)	(0.000)
Countries	92	135	135	135
Observations	1395	2044	2050	2058
within R^2	0.22	0.27	0.16	0.37

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Quantile Regressions of Log GDP per Capita on Urbanization in 1850

	Panel A: Quantile = 0.1			Panel B: Quantile = 0.25		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.129***	0.119***	0.085**	0.101***	0.077***	0.053*
	(0.031)	(0.033)	(0.034)	(0.028)	(0.026)	(0.028)
Urban Population Density 1850	0.123***	0.130***	0.125**	0.272***	0.246***	0.168***
	(0.045)	(0.048)	(0.050)	(0.040)	(0.038)	(0.040)
Square Urban Pop. Den. 1850	-0.007*	-0.008**	-0.008**	-0.019***	-0.018***	-0.012***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
City in Neighboring Regions (1850)	-0.034	-0.060*	-0.065**	-0.034	-0.079***	-0.075***
	(0.030)	(0.033)	(0.032)	(0.027)	(0.026)	(0.026)
Urb. Pop. Den. in Neib. 1850	0.028	0.024	0.029	0.025	0.021	0.020
	(0.029)	(0.031)	(0.030)	(0.026)	(0.024)	(0.024)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001	-0.001	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Largest National City in 1850			0.106*			0.177***
			(0.056)			(0.045)
	Panel C: Quantile = 0.5			Panel D: Quantile = 0.75		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.043	0.050**	0.027	0.037	0.046	-0.006
	(0.027)	(0.025)	(0.027)	(0.032)	(0.032)	(0.034)
Urban Population Density 1850	0.327***	0.284***	0.147***	0.442***	0.351***	0.160***
	(0.039)	(0.036)	(0.039)	(0.046)	(0.046)	(0.050)
Square Urban Pop. Den. 1850	-0.024***	-0.020***	-0.010***	-0.026***	-0.021***	-0.010***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
City in Neighboring Regions (1850)	-0.022	-0.034	-0.038	-0.035	-0.004	-0.004
	(0.026)	(0.024)	(0.025)	(0.031)	(0.031)	(0.032)
Urb. Pop. Den. in Neib. 1850	0.038	0.029	0.032	0.083***	0.074**	0.088***
	(0.025)	(0.023)	(0.024)	(0.030)	(0.029)	(0.030)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001	-0.001	-0.001	-0.002**	-0.002**	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Largest National City in 1850			0.282***			0.418***
			(0.044)			(0.056)
	Panel E: Quantile = 0.9					
	(1)	(2)	(3)			
Existence of a City (1850)	0.069	0.099	-0.020			
	(0.065)	(0.063)	(0.057)			
Urban Population Density 1850	0.481***	0.385***	0.351***			
	(0.094)	(0.091)	(0.083)			
Square Urban Pop. Den. 1850	-0.029***	-0.024***	-0.023***			
	(0.008)	(0.008)	(0.007)			
City in Neighboring Regions (1850)	-0.061	-0.028	-0.018			
	(0.063)	(0.062)	(0.053)			
Urb. Pop. Den. in Neib. 1850	0.152**	0.142**	0.136***			
	(0.060)	(0.058)	(0.050)			
Squ. Urb. Pop. Den. in Neib. 1850	-0.005**	-0.004**	-0.004**			
	(0.002)	(0.002)	(0.002)			
Largest National City in 1850			0.448***			
			(0.093)			

Note: Regressions are based on 1395 regions from 92 countries with including country fixed effects. Columns (2) and (3) include 8 regional controls listed in Table 3.1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Robustness to Variations in Distance to Neighboring Cities, Regressions of Log Regional GDP per Capita in 2005 on Urbanization in 1850

	Panel A: Neighboring cities within 50 miles away from region			Panel B: Neighboring cities within 75 miles away from region		
	(1)	(2)	(3)	(1)	(2)	(3)
Existence of a City (1850)	0.078***	0.072***	0.023	0.076***	0.071**	0.021
	(0.026)	(0.027)	(0.028)	(0.028)	(0.028)	(0.028)
Urban Population Density 1850	0.338***	0.271***	0.193***	0.326***	0.261***	0.179***
	(0.064)	(0.056)	(0.058)	(0.067)	(0.058)	(0.062)
Square Urban Pop. Den. 1850	-0.022***	-0.018***	-0.014***	-0.023***	-0.018***	-0.014**
	(0.006)	(0.005)	(0.005)	(0.007)	(0.005)	(0.005)
City in Neighboring Regions (1850)	-0.093	-0.080*	-0.066	-0.097	-0.080	-0.066
	(0.058)	(0.046)	(0.046)	(0.068)	(0.061)	(0.058)
Urb. Pop. Den. in Neib. 1850	0.040**	0.025	0.024	0.012	0.009	0.012
	(0.016)	(0.015)	(0.017)	(0.011)	(0.010)	(0.010)
Squ. Urb. Pop. Den. in Neib. 1850	-0.001**	-0.000*	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Largest National City in 1850			0.264***			0.272***
			(0.059)			(0.058)
Baseline Controls Included	No	Yes	Yes	No	Yes	Yes
Countries	92	92	92	92	92	92
Observations	1395	1395	1395	1395	1395	1395
within R^2	0.09	0.21	0.23	0.09	0.21	0.23

	Panel C: Neighboring cities within 100 miles away from region		
	(1)	(2)	(3)
Existence of a City (1850)	0.076***	0.071**	0.021
	(0.028)	(0.027)	(0.028)
Urban Population Density 1850	0.335***	0.268***	0.191***
	(0.067)	(0.057)	(0.059)
Square Urban Pop. Den. 1850	-0.023***	-0.018***	-0.014***
	(0.006)	(0.005)	(0.005)
City in Neighboring Regions (1850)	-0.106	-0.090	-0.074
	(0.078)	(0.067)	(0.064)
Urb. Pop. Den. in Neib. 1850	0.007	0.005	0.006
	(0.005)	(0.005)	(0.005)
Squ. Urb. Pop. Den. in Neib. 1850	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Largest National City in 1850			0.267***
			(0.057)
Baseline Controls Included	No	Yes	Yes
Countries	92	92	92
Observations	1395	1395	1395
within R^2	0.09	0.21	0.23

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Baseline controls included are land suitability, temperature in Celsius, altitude in 100 meters, ruggedness in 100 meters, rainfall in millimeter, absolute value of latitude (integer), proximity to the coast, and proximity to a river. All estimates include country fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Regressions of Log GDP per capita 2005 on Urban Population Density 1850

	Panel A: The whole world			Panel B: Asia		
	(1)	(2)	(3)	(1)	(2)	(3)
ln(Urban Population Density 1850)	0.092***	0.076***	0.054***	0.112***	0.089***	0.069***
	(0.009)	(0.008)	(0.009)	(0.014)	(0.012)	(0.014)
ln(neib. urban pop. den. 1850)	0.043***	0.027**	0.031**	0.065*	0.045	0.048
	(0.013)	(0.013)	(0.013)	(0.036)	(0.040)	(0.040)
Largest National City in 1850			0.249***			0.254**
			(0.053)			(0.102)
Observations	1395	1395	1395	373	373	373
within R^2	0.11	0.22	0.24	0.16	0.23	0.24
	Panel C: Western Europe			Panel D: Non-Western Europe		
	(1)	(2)	(3)	(1)	(2)	(3)
ln(Urban Population Density 1850)	0.076***	0.086***	0.066***	0.110**	0.120***	0.088***
	(0.013)	(0.012)	(0.014)	(0.043)	(0.033)	(0.022)
ln(neib. urban pop. den. 1850)	0.021*	0.020	0.023*	-0.001	-0.022	-0.008
	(0.012)	(0.013)	(0.013)	(0.014)	(0.013)	(0.013)
Largest National City in 1850			0.149			0.270*
			(0.091)			(0.136)
Observations	214	214	214	290	290	290
within R^2	0.24	0.31	0.33	0.17	0.50	0.52
	Panel E: The Americas			Panel F: The Americas no US & Canada		
	(1)	(2)	(3)	(1)	(2)	(3)
ln(Urban Population Density 1850)	0.090***	0.075***	0.048**	0.091***	0.074***	0.037
	(0.013)	(0.015)	(0.019)	(0.017)	(0.019)	(0.025)
ln(neib. urban pop. den. 1850)	0.056**	0.051**	0.056**	0.077**	0.074***	0.074***
	(0.023)	(0.021)	(0.020)	(0.028)	(0.019)	(0.018)
Largest National City in 1850			0.260**			0.308**
			(0.109)			(0.129)
Observations	387	387	387	324	324	324
within R^2	0.11	0.29	0.30	0.10	0.29	0.31
	Panel I: Ex-colonial countries			Panel J: Non ex-colonial countries		
	(1)	(2)	(3)	(1)	(2)	(3)
ln(Urban Population Density 1850)	0.084***	0.057***	0.037***	0.099***	0.083***	0.056***
	(0.012)	(0.011)	(0.014)	(0.015)	(0.013)	(0.013)
ln(neib. urban pop. den. 1850)	0.080***	0.062***	0.065***	0.022**	0.007	0.013*
	(0.024)	(0.023)	(0.022)	(0.008)	(0.008)	(0.007)
Largest National City in 1850			0.270***			0.244***
			(0.081)			(0.073)
Observations	658	658	658	737	737	737
within R^2	0.09	0.18	0.20	0.15	0.34	0.36

Note: Robust standard errors clustered at the country level are shown in parentheses, and country fixed effects are included. Columns (2) and (3) include 8 regional controls listed in Table 3.1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Number of Regions by Country by Year

Contient	*Country*	*1850-1900*	*1900-1950*	*1950-2000*
Africa	Algeria	7	14	17
Africa	Angola	1	1	3
Africa	Benin	0	1	4
Africa	Burkina Faso	0	2	2
Africa	Burundi	0	0	1
Africa	Cameroon	0	3	4
Africa	Central African Republic	0	1	1
Africa	Chad	0	0	1
Africa	Cote d'Ivoire	0	1	2
Africa	Dem. Rep. Congo	0	1	6
Africa	Djibouti	0	1	1
Africa	Egypt	3	3	4
Africa	Eritrea	0	1	1
Africa	Ethiopia	0	1	4
Africa	Gabon	0	0	1
Africa	Gambia	0	1	1
Africa	Ghana	1	3	5
Africa	Guinea	0	1	1
Africa	Kenya	1	1	3
Africa	Liberia	0	1	1
Africa	Libya	1	2	1
Africa	Madagascar	1	3	6
Africa	Malawi	0	1	1
Africa	Mali	1	2	5
Africa	Mauritius	1	1	1
Africa	Morocco	5	10	11
Africa	Mozambique	0	1	2
Africa	Nigeria	5	6	6
Africa	Republic of Congo	0	1	2
Africa	Reunion	3	3	3
Africa	Senegal	0	2	5
Africa	Sierra Leone	1	1	1
Africa	Somalia	0	1	3
Africa	South Africa	1	3	3
Africa	South Sudan	0	0	1
Africa	Sudan	1	2	6
Africa	Tanzania	1	2	2
Africa	Togo	0	0	1
Africa	Tunisia	1	6	7
Africa	Uganda	0	0	1
Africa	Zambia	0	0	4
Africa	Zimbabwe	0	1	4
Americas	Argentina	13	15	19
Americas	Bahamas	1	1	1
Americas	Barbados	1	1	1
Americas	Belize	0	1	0
Americas	Bolivia	5	5	5
Americas	Brazil	19	20	20
Americas	Canada	2	3	5
Americas	Chile	8	10	11

Continued on next page...

Table A.8 – Continued

Americas	Colombia	13	16	19
Americas	Costa Rica	1	1	1
Americas	Cuba	10	12	13
Americas	Dominican Republic	2	2	2
Americas	Ecuador	4	4	4
Americas	El Salvador	3	5	5
Americas	Guatemala	3	5	5
Americas	Guyana	1	1	1
Americas	Haiti	4	4	4
Americas	Honduras	2	2	2
Americas	Jamaica	1	1	1
Americas	Martinique	1	1	1
Americas	Mexico	26	28	29
Americas	Nicaragua	3	4	4
Americas	Panama	1	1	3
Americas	Paraguay	1	2	2
Americas	Peru	10	11	12
Americas	Puerto Rico	2	3	4
Americas	Suriname	1	1	1
Americas	Trinidad - Tobago	1	1	0
Americas	United States	25	30	34
Americas	Uruguay	0	7	7
Americas	Venezuela	11	13	16
Asia	Afghanistan	3	3	10
Asia	Armenia	0	0	1
Asia	Azerbaijan	0	1	1
Asia	Bahrain	1	1	1
Asia	Bangladesh	3	4	4
Asia	Brunei	1	1	1
Asia	Cambodia	1	1	2
Asia	China	31	31	31
Asia	Georgia	1	1	1
Asia	India	21	21	21
Asia	Indonesia	11	14	27
Asia	Iran	12	13	19
Asia	Iraq	4	5	7
Asia	Israel	0	0	1
Asia	Japan	18	15	17
Asia	Jordan	0	0	2
Asia	Kazakhstan	0	0	2
Asia	Kuwait	1	1	1
Asia	Laos	0	0	1
Asia	Lebanon	2	2	2
Asia	Malaysia	2	4	4
Asia	Mongolia	1	1	1
Asia	Myanmar	6	6	9
Asia	Nepal	1	1	1
Asia	North Korea	1	5	9
Asia	Pakistan	4	5	5
Asia	Philippines	3	8	12
Asia	Saudi Arabia	4	4	5
Asia	South Korea	2	4	7

Continued on next page...

Table A.8 – Continued

Asia	Sri Lanka	3	4	4
Asia	Syria	4	5	5
Asia	Taiwan	2	3	3
Asia	Thailand	2	2	5
Asia	Turkey	11	11	12
Asia	United Arab Emirates	0	1	1
Asia	Uzbekistan	3	1	1
Asia	Vietnam	3	4	7
Asia	Yemen	2	2	5
Europe	Austria	3	2	2
Europe	Belarus	2	1	1
Europe	Belgium	6	5	5
Europe	Bulgaria	1	1	1
Europe	Croatia	1	1	1
Europe	Czech Republic	3	3	3
Europe	Denmark	1	1	1
Europe	Estonia	1	1	1
Europe	Finland	1	1	1
Europe	France	19	9	9
Europe	Germany	15	12	13
Europe	Greece	2	2	2
Europe	Hungary	3	1	1
Europe	Ireland	1	1	1
Europe	Italy	14	11	11
Europe	Latvia	2	1	1
Europe	Moldova	1	0	0
Europe	Netherlands	5	3	3
Europe	Norway	2	1	1
Europe	Poland	12	6	6
Europe	Portugal	3	2	2
Europe	Romania	5	1	1
Europe	Russia	17	14	29
Europe	Serbia	2	1	1
Europe	Slovakia	1	0	0
Europe	Spain	10	7	7
Europe	Sweden	3	2	2
Europe	Switzerland	3	2	2
Europe	Ukraine	12	5	8
Europe	United Kingdom	12	11	11
Oceania	Australia	1	4	5
Oceania	New Zealand	0	2	2
Oceania	Papua New Guinea	0	1	1
Africa:		35	85	139
Americas:	Americas	175	211	232
Asia:		164	185	248
Europe:		163	108	127
Oceania:		1	7	8
World Total:		538	596	754

Note: write notes here.

Table A.9: Regional Convergence of Urban Population Density 1850-2000, Fixed Effects Estimation across Subperiods

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	All Years	1900-2000	1850-1950	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0050***	-0.0057***	-0.0025***	-0.0014**	-0.0029***	-0.0065***
	(0.0008)	(0.0009)	(0.0005)	(0.0006)	(0.0006)	(0.0010)
Countries	145	142	134	112	131	141
Observations	2909	2291	1315	618	697	1594
within R^2	0.23	0.21	0.09	0.07	0.18	0.24

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. I use the lowest urban population within countries of a year to impute missing urban population unless the region has no missing 50 years later and impute missing urban population density based on $\frac{ImputedUrbanPopulation}{LandArea}$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Regional Convergence of Urban Population Density 1850-2000, Fixed Effects Estimation across Country Groups Based on Continent

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	US excl. DC	Americas excl. US & Canada	West EU	Rest EU	Asia	Africa
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0091***	-0.0037***	-0.0015	-0.0118***	-0.0046***	-0.0029***
	(0.0014)	(0.0005)	(0.0009)	(0.0004)	(0.0005)	(0.0010)
Countries		29	16	15	38	42
Observations	112	710	285	315	874	570
R^2	0.50	0.28	0.10	0.69	0.26	0.16
	Estimated Convergence Rates in Subperiods					
	(1)	(2)	(3)	(4)	(5)	(6)
1850-1900	0.0023	-0.0009	0.0030***	-0.0054**	-0.0025***	-0.0021
	(0.0030)	(0.0009)	(0.0009)	(0.0019)	(0.0007)	(0.0018)
Countries	1	27	16	15	32	19
Observations	34	173	100	69	186	50
R^2	0.59	0.08	0.43	0.53	0.15	0.32
1900-1950	-0.0055**	-0.0015***	-0.0002	-0.0094**	-0.0031***	-0.0008
	(0.0023)	(0.0005)	(0.0011)	(0.0032)	(0.0008)	(0.0014)
Countries	1	29	16	12	34	35
Observations	33	192	72	45	224	117
R^2	0.67	0.31	0.18	0.73	0.16	0.35
1950-2000	-0.0117***	-0.0056***	-0.0061***	-0.0132***	-0.0057***	-0.0025*
	(0.0011)	(0.0012)	(0.0009)	(0.0003)	(0.0008)	(0.0013)
Countries	1	28	16	12	38	42
Observations	45	345	113	201	464	403
R^2	0.90	0.19	0.46	0.75	0.23	0.21

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions except those in columns (1) include country fixed effects. I use the lowest urban population within countries of a year to impute missing urban population unless the region has no missing 50 years later and impute missing urban population density based on $\frac{ImputedUrbanPopulation}{LandArea}$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Regional Convergence of Urban Population Density 1850-2000, Fixed Effects Estimation across Subperiods

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	All Years	1900-2000	1850-1950	1850-1900	1900-1950	1950-2000
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0088***	-0.0094***	-0.0055***	-0.0048***	-0.0049***	-0.0104***
	(0.0007)	(0.0007)	(0.0007)	(0.0010)	(0.0007)	(0.0006)
Countries	145	142	134	113	131	140
Observations	2851	2226	1320	625	695	1531
within R^2	0.36	0.37	0.18	0.21	0.19	0.41

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions include country fixed effects. I use the lowest urban population density within countries of a year to impute missing urban population density unless the region has no missing 50 years later. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Regional Convergence of Urban Population Density 1850-2000, Fixed Effects Estimation across Country Groups Based on Continent

	Dependent Variable: Average Annual Growth Rate of Urban Population Density					
	US excl. DC	Americas excl. US & Canada	West EU	Rest EU	Asia	Africa
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Urban Pop. Den.	-0.0111***	-0.0080***	-0.0037**	-0.0128***	-0.0086***	-0.0095***
	(0.0017)	(0.0013)	(0.0013)	(0.0010)	(0.0011)	(0.0019)
Countries	1	29	16	15	38	42
Observations	111	715	290	313	878	503
R^2	0.53	0.35	0.30	0.77	0.38	0.43
	Estimated Convergence Rates in Subperiods					
	(1)	(2)	(3)	(4)	(5)	(6)
1850-1900	-0.0008	-0.0041***	0.0030***	-0.0108***	-0.0046***	-0.0089**
	(0.0030)	(0.0014)	(0.0009)	(0.0021)	(0.0012)	(0.0037)
Countries	1	27	16	15	32	20
Observations	34	174	100	69	187	56
R^2	0.61	0.23	0.43	0.73	0.24	0.74
1900-1950	-0.0062**	-0.0026***	-0.0002	-0.0111**	-0.0056***	-0.0065***
	(0.0024)	(0.0008)	(0.0011)	(0.0036)	(0.0010)	(0.0021)
Countries	1	29	16	12	34	35
Observations	33	193	72	45	229	109
R^2	0.83	0.25	0.18	0.66	0.25	0.41
1950-2000	-0.0129***	-0.0115***	-0.0075***	-0.0139***	-0.0096***	-0.0090***
	(0.0014)	(0.0012)	(0.0013)	(0.0008)	(0.0012)	(0.0016)
Countries	1	27	16	12	38	42
Observations	44	348	118	199	462	338
R^2	0.85	0.37	0.56	0.80	0.41	0.50

Note: The unit of observation is a subnational region. Robust standard errors clustered at the country level are shown in parentheses. Results are based on regressions controlling for log national urban population density, presence of city in neighboring areas at the beginning of the period, presence of national capital, proximity to national capital, proximity to borders, temperature in Celsius, rainfall in meter, elevation in 100 meters, ruggedness in 100 meters, land suitability, absolute value of latitude, proximity to the coast, and proximity to a river. All regressions except those in columns (1) include country fixed effects. I use the lowest urban population density within countries of a year to impute missing urban population density unless the region has no missing 50 years later. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Vita

Dachao Ruan was born in Nanjing, Jiangsu, China in 1985. He attended the University of Macau, Macau, China, where he earned a Bachelor of Business Administration with a major in Economics and International Finance in 2007. In 2010, he completed his Master of Social Sciences in Economics at the University of Macau. After that, he entered graduate school in the Department of Economics at Louisiana State University (LSU) and received his Master of Science in Economics in 2012. During graduate school at LSU, he worked as a research assistant in the Division of Economic Development at LSU working on government-sponsored research projects as well as an instructor in the Department of Economics teaching Principles of Microeconomics and Principles of Macroeconomics. He will join Capital One Financial Corporation as a Principle Associate, Quantitative Analysis at Capital One's headquarters in McLean, Virginia, following his anticipated graduation in the Summer of 2015.