Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*

Daron Acemoglu
Department of Economics, MIT

Simon Johnson
Sloan School of Management, MIT

James A. Robinson
Department of Political Science, Berkeley

July 5, 2001

Abstract

Among countries colonized by European powers during the past 500 years those that were relatively rich in 1500 are now relatively poor. We document this reversal using data on urbanization patterns, which, we argue, proxy for economic prosperity. This reversal is inconsistent with a view that links economic development to geographic factors. According to the geography view, societies that were relatively rich in 1500 should also be relatively rich today. In contrast, the reversal is consistent with the role of institutions in economic development. The expansion of European overseas empire starting in the 15th century led to a major change in the institutions of the societies they colonized. In fact, the European intervention appears to have created an "institutional reversal" among these societies, in the sense that Europeans were more likely to impose "extractive" institutions in densely settled and prosperous areas, while introducing institutions encouraging investment in regions that were previously poor. This institutional reversal accounts for the reversal in relative incomes. We provide further support for this view by documenting that the reversal in relative incomes took place during the 19th century, and resulted from societies with extractive institutions failing to take advantage of industrialization opportunities.

In Progress.

*We thank Joshua Angrist, Abhijit Banerjee, Olivier Blanchard, Jan de Vries, and seminar participants at the All-UC History Conference at Berkeley, the conference on "Globalization and Marginalization" in Bergen, Canadian Institute of Advanced Research, Chicago, Columbia, DELTA, MIT, Stanford, and Yale for useful comments.

1 Introduction

The "geography hypothesis" explains the bulk of the differences in economic performance across countries by geographic, climatic or ecological features of the environment. Motivating this view is a fact that Galbraith summarized succinctly in 1951: if "one marks a belt a couple thousand miles in width encircling the earth at the equator one finds within it no developed countries...". The list of scholars who have emphasized the importance of geographic factors is long and distinguished, and includes, inter alia, Niccolò Machiavelli, Charles de Montesquieu, Arnold Toynbee, Alfred Marshall, Ellsworth Huntington, and Gunnar Myrdal. All of these authors viewed climate as a key determinant of work effort, productivity, and ultimately, the success of nations. For example, Marshall (1890) wrote: "vigor depends partly on race qualities: but these, so far as they can be explained at all, seem to be chiefly due to climate" (p. 195). In a recent influential book, Jared Diamond (1997) has argued for the importance of geographic determinants of the Neolithic revolution, and linked modern prosperity to the timing of the emergence of settled agriculture. He forcefully states that "the striking differences between the long-term histories of peoples of different continents have been... [due to]... the differences in their environments" (p. 405). Jeffrey Sachs, on the other hand, has argued for the importance of technology, disease environment and transport costs, which are determined by physical geography and climate. For example, Sachs (2001) writes "By the start of the era of modern economic growth, if not much earlier, temperate-zone technologies were more productive than tropical-zone technologies..." (p. 2).

An alternative view, instead, relates differences in economic performance to the organization of the society. Societies that provide incentives and opportunities for investment will be richer than those who fail to do so. This view dates back at least to John Locke, who argued for the necessity of property rights for productive activities, and to Adam Smith, who stressed the role of "peace, easy taxes, and a tolerable administration of justice" in generating prosperity (quoted in Jones, 1981, p. 235). Recent proponents of this approach have been economists and historians emphasizing the importance of property rights and institutions (e.g., North and Thomas, 1973, North and Weingast, 1989, North, 1990). For this reason, we refer to this approach as the "institutions hypothesis".

In this paper, we attempt to distinguish between these two broad hypotheses. If geography is the key determinant of income differences across countries, economic performance should be *highly persistent*, since geographic factors did not change much during recent history. To the extent that other factors also matter for income, persistence will not be perfect, but we should expect rich countries today to have been, on average, richer 100, 200, 500 or even 1000 years ago (see, e.g., Diamond, 1997). Since institutions and the way that societies are organized are persistent as well, the institutions hypothesis also predicts persistence in income levels. Nevertheless, if there is a major change in institutions, then we should expect a significant change in the distribution of income across countries.

The expansion of the European overseas empire starting at the end of the 15th century provides an appealing "natural experiment" to distinguish between these two contrasting predictions. In the subsequent four and a half centuries, a large part of the globe that was previously ruled by independent polities came under European control, transforming their social organization and institutions in fundamental ways. Despite these radical social changes, the geography view predicts persistence in relative incomes: the same geographic, climatic and ecological factors making countries prosperous before European colonialism should also contribute to prosperity after the arrival of the Europeans. In contrast, if European dominance came with a major change in the organization of these societies, the institutions hypothesis implies that that there should not necessarily be such persistence.

Historical and econometric evidence suggests not only that European colonialism caused a major change in the social organization of these regions, but also that it created an "institutional reversal"—Europeans imposed worse institutions in previously prosperous areas. The main reason for the institutional reversal is that relatively poor areas were sparsely settled, and this enabled or induced Europeans to develop settler colonies with institutions encouraging investment and commerce by a broad cross-section of the society. In contrast, a large population and relative prosperity made it profitable for the Europeans to set up extractive institutions, with political power concentrated in the hands of a small elite. High population density, for example, meant a large supply of labor that the Europeans could force to work in mines or plantations, or tax heavily by taking over existing tribute systems. Furthermore, high population density and the associated diseases often made it less attractive for Europeans to settle. Because Europeans were more likely to set up extractive institutions in places they did not settle, high population density also made the development of extractive institutions more likely through this channel. The institutions hypothesis, together with the institutional reversal caused by European colonialism, suggests the possibility of a reversal among the European colonies: countries that were relatively rich in 1500 should be relatively poor today.

The major finding of this paper is that there is a reversal in relative incomes among the former European colonies. For example, the Mughal Empire in India, the Aztec and the Inca Empires in America were among the richest civilizations in 1500, while the civilizations in North America, New Zealand and Australia were highly primitive. Today the U.S., Canada, New Zealand and Australia are orders of magnitude richer than the countries now occupying the territories of the Mughal, Aztec and Inca Empires, such as India or Ecuador. This reversal is inconsistent with the geography hypothesis, but consistent with the institutions hypothesis. We also show that the timing and the pattern of the reversal is closely linked to the "institutions" imposed by the European powers.

The obvious difficulty in our empirical investigation is lack of data on economic prosperity in 1500. The first contribution of our paper is to propose and use urbanization rates as a proxy for cross-country differences in economic prosperity during preindustrial periods. Bairoch (1988) argues that only areas with high agricultural productivity could support large urban populations, while de Vries (1976, p.164) emphasizes the necessity of improvements in transportation and fuel technology to provide sufficient energy supplies for cities as they grow. Similarly, many economic historians subscribe to the view that increasing urbanization is associated with economic development (see, for example, Kuznets, 1968, de Vries, 1984, Bairoch 1988, Tilly, 1990, De Long and Shleifer, 1993). We also present evidence that both in the time series and the cross section there is a close association between urbanization and income per capita.[1]

As an additional proxy for prosperity we use population density, on which there are relatively more extensive data (e.g. McEvedy and Jones, 1975). Although the theoretical relationship between population density and prosperity is more complex, it seems clear that during the preindustrial periods only relatively prosperous areas could support dense populations.

Figure 1 here: GDP per capita today vs. urbanization in 1500 for ex-colonies

With either measure, there is a negative relationship between economic prosperity in

[1]By economic prosperity or income per capita in 1500, we do not refer to the economic or social conditions of the masses, but to a measure of total production in the economy relative to the number of inhabitants. Although urbanization is likely to have been associated with relatively high output per capita, the majority of urban dwellers lived in poverty and died young because of poor sanitary conditions (see for example Bairoch, 1988, chapter 12).

1500 and today. Figure 1 shows a negative relationship between percent of the population living in towns with more than 5000 inhabitants in 1500 and income per capita today (see below for data details). The big gap in this figure is sub-Saharan Africa for which there are no reliable urbanization data. Figure 2 shows the same negative relationship between population density (number of inhabitants per square km) in 1500 and income per capita today, including data from sub-Saharan African countries. The relationships shown in Figures 1 and 2 are robust. They are unchanged when we control for continent dummies, the identity of the colonial power, geography-related variables, including temperature, humidity, and distance from the equator, and when we exclude the pure settler colonies that Alfred Crosby (1986) refers to as Neo-Europes—the U.S., New Zealand, Canada and Australia.

Figure 2 here: GDP per capita today vs. population density in 1500 for ex-colonies

Some degree of caution is required in interpreting these results since both urbanization and population density estimates for 1500 are likely to be error ridden. Yet the direct effect of measurement error is to create an attenuation bias towards 0 in the correlation between urbanization or population density in 1500 and income today. So measurement error alone is unlikely to account for our results. Naturally, our findings could be explained by errors in the estimates of urbanization or population density that are correlated with current economic development. This possibility seems unlikely, however, since the basic patterns we document are robust to a variety of alternative estimates, and are consistent with qualitative assessments of historians. We therefore conclude that the data point to a reversal among former colonies.

While the reversal in the relative economic position of the colonies weighs strongly against a simple geography view, one might advance a more sophisticated geography theory, which we dub the "Northern drift hypothesis". According to this hypothesis, the center of gravity of economic activity has been gradually shifting to the North, or at least away from the equator. Sachs (2001), for example, implies such a view by pointing out that temperate-zone areas have been growing faster during the past 200 years than tropical areas. In fact, in 1500 the tropical areas were relatively rich, and today they are among the poorest places in the world. It can be argued that areas in the tropics had an early advantage, but later agricultural technologies, such as the heavy plow, crop rotation systems, use of domesticated animals and high-yield crops, have favored countries in the temperate areas. However, the timing of the reversal in relative incomes appears to be inconsistent with this hypothesis. First, the reversal in relative incomes seems to be related to population density before Europeans arrived, *not* to any geographic characteristics of the area. More important, there is no compelling reason why climate and geography should matter more for industry than for agriculture. So according to this hypothesis, the reversal should occur when European agricultural technology spread to the colonies. Yet, in practice the reversal occurred mostly during the 19th century, while the introduction of European agricultural techniques, at least in North America, took place earlier. Moreover, the reversal seems to be closely related to industrialization: ex-colonies that could take advantage of industrial technology surged ahead, while those that failed to do so lagged behind.

Is the reversal related to institutions? The answer appears to be yes. We document that differences in the institutions that the Europeans introduced statistically account for the reversal in relative incomes from 1500 to today. Moreover, the institutions hypothesis suggests that institutional differences should matter more when new technologies requiring investments from a broad cross-section of the society become available. We therefore expect societies with institutions of private property to take advantage of industrialization opportunities, while societies with extractive institutions where political power is concentrated in the hands of the small elite fail to do so. The data support this prediction.

We are unaware of any other work in the economics, political science or history literatures that has noticed or documented this change in the distribution of economic prosperity, or used the experiences of the former colonies to distinguish between the geography and institutions hypotheses. Nevertheless, there are many country experiences consistent with our general picture. For example, many scholars, including McNeill (1967), Jones (1981), Kennedy (1987), Chaudhuri (1990), Wong (1997) and Pomeranz (2000), emphasize that in 1500 the Mughal, Ottoman and Chinese Empires were highly properous,[2] but grew slowly during the next 500 years. Coatsworth (1993) and Engerman and Sokoloff (1997, 2000) document that North America was no more developed than South America in the early 18th century and the data presented by Eltis (1995) and Engerman and Sokoloff (2000) suggest that Carribean islands such as Haiti and Barbados were considerably richer than the United States during early colonial times. Our work therefore also establishes

[2]In addition to data on urbanization and population density that we emphasize, these authors also present extensive documentation that patterns of health, life expectancy and stature differed little in South and East Asia from those prevailing in Europe at the same time. Reid (1988, p. 50) concludes his survey of the evidence by noting "the relatively good health of Southeast Asians...should not surprise us if we compare their diet, medicine, and hygiene with those of contemporary Europeans."

that the pattern pointed out by Coatsworth, Eltis, and Engerman and Sokoloff applies for comparisons of economic prosperity in 1500, i.e., before colonial times, and for the whole sample of former European colonies. More importantly, our larger sample enables us to investigate the robustness of this result and to investigate the relative merits of the geography and institutions hypotheses.

The view that the colonial experience of many countries has retarded their economic development is put forward by many Marxist historians and dependency theorists, for example, Williams (1944), Rodney (1972), Wallerstein (1974-1980) and Frank (1978). The work by Engerman and Sokoloff (1997, 2000) is more closely related because, like us, they emphasize the long-run adverse consequences of the "plantation complex" and the associated institutions in Latin America (see also work by historians Genovese, 1965, Beckford, 1972, Dunn, 1972, Curtin, 1990 and Coatsworth, 1999). Our approach differs from these contributions in two important dimensions: first, we view European colonialism as a "natural experiment" potentially distinguishing between the geography and institutions hypotheses. Therefore, we emphasize *not* the negative effects of European colonialism relative to what would have happened without colonialism, but the differential effect of European colonialism in some countries vs. others: societies where colonialism led to the establishment of good institutions prospered relative to those where colonialism imposed extractive institutions. Second, in our theory the negative effects of colonialism do not result from the plunder of the colonies by the Europeans or dependency as emphasized by Williams, Rodney or Frank, but because the extractive colonial institutions stacked the cards against industrialization.[3]

Our overall interpretation of comparative development in the former colonies is related to our previous paper, Acemoglu, Johnson and Robinson (2001), where we proposed the disease environment at the time Europeans arrived as an instrument for European settlements and subsequent institutional development of the former colonies, and used this to estimate the causal effect of institutional differences on economic performance. The current paper has a different focus and a number of innovations relative to our earlier work. First, our focus here is on the persistence of economic performance, which we argue can distinguish between the geography and institutions hypotheses. Second, we point out and document the major reversal in the distribution of economic prosperity among the

[3]Put differently, according to Williams, Rodney or Frank, countries in Central America, the Caribbean and Africa are poor because of "too much capitalism," whereas in our thesis, they are poor because of "the wrong type of capitalism".

former colonies. Third, our thesis here emphasizes the direct effect of population density, and more generally prosperity at the time Europeans arrived, on the policies pursued by the Europeans, for example, by encouraging labor-oppressive production methods, and the takeover of existing tax and tribute systems. Finally, we suggest and document that the interaction between the set of institutions brought by Europeans and the opportunity to industrialize during the 19th century played a central role in the long-run development of the former colonies.

Our results are also relevant to the literature on the relationship between population and growth. The recent consensus is that population density encourages the discovery and exchange of ideas, and contributes to growth. This view goes back to Boserup (1965), Jacobs (1967) and Kuznets (1968), and was elaborated by Simon (1977). The recent endogenous growth literature also emphasizes the beneficial effects of high population through scale effects (e.g., Romer, 1986, 1990, Grossman and Helpman, 1991, Aghion and Howitt, 1992, Kremer, 1993, Jones, 1999). Our evidence points to a major historical episode of 500 years where high population density was detrimental to economic development, and therefore sheds doubt on the general applicability of this recent consensus.

The rest of the paper is organized as follows. The next section discusses the construction of urbanization and population density data, and provides evidence that these are good proxies for economic prosperity before the modern era. In Section 3, we outline the geography and institutions hypotheses, explain why we should expect an institutional reversal resulting from European colonialism, and present historical and econometric evidence supporting this notion. Section 4 documents the "Reversal of Fortune"—the negative relationship between economic prosperity in 1500 and income per capita today among the former colonies. Section 5 discusses the Northern drift hypothesis, and shows that the reversal took place during the 19th century, and was linked to industrialization, which we interpret as evidence against this hypothesis. Section 6 documents that the institutional reversal caused by European colonialism statistically accounts for the reversal in relative incomes between 1500 and today, and that institutions started playing a much more important role in the age of industry. Section 7 concludes, while the Appendix describes how urbanization and population density data are constructed.

2 Urbanization and Population Density

A measure of economic prosperity in 1500 is crucial for our investigation. In this section, we argue that urbanization and population density provide good proxies for income per capita and/or productivity, and explain how these data are constructed. More detailed discussion of data construction and alternative series are provided in Appendix A.

2.1 Data on Urbanization

Bairoch (1988) provides the best single collection and assessment of urbanization research. Our base data for 1500 consist of Bairoch's (1988) urbanization estimates augmented by the work of Eggimann (1999). We also construct a longer time-series data for a subset of countries to study the timing of the reversal and whether there was a similar reversal before 1500.

Merging the Eggimann and Bairoch series requires us to convert Eggimann's estimates, which are based on a minimum population threshold of 20,000, into Bairoch-equivalent urbanization estimates, which use a minimum population threshold of 5,000. We use a number of different methods to convert between the two sets of estimates, all with similar results. For our base estimates, we run a regression of Bairoch estimates on Eggimann estimates for all countries where they overlap in 1900 (the year for which we have most Bairoch estimates for non-European countries). This regression yields a constant of 6.6 and a coefficient of 0.67. We use these estimates below to generate Bairoch-equivalent urbanization estimates. Alternatively, we converted the Eggimann estimates using country-specific conversion ratios and conversion ratios at the regional level (e.g., separately for North Africa and the Andean region) using information suggested by Bairoch's analysis, and also we checked the robustness of our results using a uniform conversion rate of 2 as suggested by Davis' and Zipf's Laws (see Bairoch, 1988, chapter 9). We also used Bairoch's overall assessment of urbanization for broad regions, e.g., Asia, without the more detailed information from Eggimann, in all cases with similar results. Details and robustness results are reported in the Appendix.

Finally, we also constructed three alternative series without combining estimates from different sources. One of these is based mainly on Chandler's data, the other on Bairoch, and the final one on Eggimann. Results using these three different measures are reported in Table 4.

While the data on sub-Saharan Africa are worse than for any other region, it is clear that urbanization before 1500 was at a higher level than in North America or Australia. Bairoch claims that by 1500 urbanization was "well-established" in sub-Saharan Africa.[4] Yet by 1900, sub-Saharan Africa certainly was less urbanized than the European settler colonies. Because there are no detailed urbanization data for Africa, we leave this region out of the regression analysis when we use urbanization data, though it is included in our regressions using population density.

Table 1 gives descriptive statistics for the key variables of interest, separately for the whole world, for the sample of ex-colonies for which we have population density data in 1500, and for the sample of ex-colonies for which we have urbanization data in 1500.

Table 1 here

2.2 Urbanization and Income

There are good reasons to presume that urbanization and income are positively related. Kuznets (1968, p. 1) opens his book by stating:

> "we identify the economic growth of nations as a sustained increase in per-capita or per-worker product, most often accompanied by an increase in population and usually by sweeping structural changes....in the distribution of population between the countryside and the cities, the process of urbanization."

Bairoch (1988) gives detailed evidence that only relatively rich countries could afford a high fraction of the population living in cities. For example, he points out that during preindustrial periods, a large fraction of the agricultural surplus was likely to be spent on transportation, so both relatively high agricultural surplus and a developed transport system were necessary for large urban populations (see Bairoch 1988, chapter 1, de Vries 1976, p. 164). Bairoch writes "the existence of true urban centers presupposes not only a surplus of agricultural produce, but also the possibility of using this surplus in trade" (p. 11). Moreover, he emphasizes that an increase in agricultural productivity almost always tended to cause increased urbanization:

[4] Sahelian trading cities such as Timbuktu, Gao and Djenne (all in modern Mali) were very large in the middle ages with populations as high as 80,000. Kano (in modern Nigeria) had a population of 30,000 in the early 19th century, and Yorubaland (also in Nigeria) was highly urbanized with a dozen towns with populations of over 20,000 and its capital Ibadan possibly had 70,000 people. For these numbers and more, see Hopkins (1973, Ch. 2).

> "For while it is true that urbanization could not get underway without the concentration of population and the surplus of food resulting from agriculture, it is equally true that the emergence of agriculture set in motion forces that sooner or later led to the growth of cities." (p. 94).

This view, that urbanization and income (productivity) are closely related, is shared by many other historians (see, for example, Tilly, 1990, De Long and Shleifer, 1993, and Tilly and Brockmans, 1994).[5]

A large history literature also documents how urbanization accelerated in Europe during periods of economic expansion (e.g., Pirenne 1956, Postan and Rich 1966, Duby 1974). For example, the period between the beginning of the 11th century and mid-14th century is an era of rapid increase in agricultural productivity and industrial output. The same period also witnessed a proliferation of cities. Bairoch (1988), for example, estimates that the number of cities with more than 20,000 inhabitants increased from around 35-44 to 100-110.

We supplement this argument by empirically investigating the link between urbanization and income. Figures A1 and A2 in the Appendix show the time-series relationship between urbanization and income per capita for a number of countries. In all cases, changes in urbanization and income are highly correlated. In Table 2, we provide regression evidence also pointing in the same direction. Columns 1-6 in Panel A present cross-sectional regressions. Column 1 is for the earliest date for which we have data on urbanization and income per capita for a large number of countries, circa 1900. The regression coefficient, 0.038, is highly significant, with a standard error of 0.006. It implies that a country with 10 percentage points higher urbanization has, on average, 46 percent greater income per capita (38 log points). Column 2 reports a similar result using data for 1950. Column 3 uses current data and shows that even today there is a strong relationship between income per capita and urbanization for a large sample of countries. The coefficient is similar, 0.036, and very precisely estimated, with a standard error of 0.002. This relationship is also shown diagramatically in Figure 3.

Figure 3: cross-sectional graph of income per capita today vs. urbanization today

[5] De Long and Shleifer (1993) write "The larger preindustrial cities were nodes of information, industry, and exchange in areas where the growth of agricultural productivity and economic specialization had advanced far enough to support them. They could not exist without a productive countryside and a flourishing trade network. The population of Europe's preindustrial cities is a rough indicator of economic prosperity" (p. 675).

Below, we draw a distinction between countries colonized by European powers and those never colonized. Column 4 and 5 report the same regression separately for these two samples. The estimates are very similar: 0.037 for the former colonies sample, and 0.033 for the rest of the countries. Finally, in column 6, we add continent dummies to the same regression. This leads to only a slightly smaller coefficient of 0.030, with a standard error of 0.002. This result demonstrates that the correlation between urbanization and income per capita is not driven by differences across continents. We therefore conclude that there is a robust cross-sectional relationship between urbanization and income per capita.

The second panel of the table is more speculative. Here we use estimates from Gregory King and Paul Bairoch to construct a small panel data set of urbanization and income per capita spanning over 200 years with sporadic observations. Columns 7-10 report regressions from this dataset. Remarkably, with or without country and period dummies, the estimates are very similar to those shown in Panel A. They indicate that a 10 percentage point increase in urbanization is associated with an approximately 40 percent increase in income per capita. Overall, we conclude that urbanization is a good proxy for income.

Table 2 here.

2.3 Data on Population Density

The most comprehensive data on population since 1AD comes from McEvedy and Jones (1975). They provide estimates based on censuses and published secondary sources. While some individual country numbers have since been revised and others remain contentious (particularly for pre-Colombian Meso-America), their estimates are still widely respected (see, e.g., the assessment by the Bureau of the Census, www.census.gov/ipc/www/worldhis.html). We use McEvedy and Jones (1975) for our baseline estimates, and indicate where we have tested the effect of using alternative assumptions (e.g., lower or higher population estimates for Mexico and its neighbors before the arrival of Cortes).

We calculate population density by dividing total population by arable land (also estimated by McEvedy and Jones). This excludes primarily desert, inland water, and tundra (we also test for the robustness of including this area). As much as possible, we use the land area of a country at the date which we are considering (e.g., in 1500). We also checked the robustness of our results to changing this assumption and none of our findings are affected by this.

2.4 Population Density and Income

The theoretical relationship between population density and income is more nuanced than that between urbanization and income. With similar reasoning, it seems natural to think that only relatively rich areas could afford dense populations (see Bairoch, 1988). This is also in line with Malthus's classic work. Malthus argued that high productivity increases population by raising birth rates and lowering death rates. However, the main thrust of Malthus's work was on how a higher than equilibrium level of population increases death rates and reduces birth rates to correct itself. Therefore, a high population could also be reflecting an "excess" of population, causing low income per capita.

To clarify the main issues, it is useful to express these relationships mathematically. Let us denote population in country (area) j at time t by $P_j(t)$, land area by L_j, and the level of multifactor productivity by A_j (which is assumed to be time invariant for simplicity). Suppose

$$Y_j(t) = A_j \cdot P_j(t)^{1-\theta} \cdot L_j^{\theta}, \tag{1}$$

where $\theta \in (0,1)$. So greater population increases output, but because of decreasing returns to land, it does so less than proportionately. Dividing both sides of the equation by $P_j(t)$, we obtain

$$y_j(t) = A_j \cdot p_j(t)^{-\theta}, \tag{2}$$

where y is per capita output and $p = P/L$ is population density. Next, we need an equation linking growth of population to income per capita. Suppose that this takes the form

$$p_j(t+1) = \rho \cdot p_j(t) + \lambda \cdot (y_j(t) - \bar{y}) + \varepsilon_j(t) \tag{3}$$

where $\bar{y}$ is subsistence income and ε is a random disturbance term. Equation (3) implies that the rate of growth of population is related to the gap between actual incomes and the subsistence level of income, and when $\rho < 1$, there is also mean reversion in population. As long as $\rho < 1$, there will be a steady-state level of income per capita, y_j^*, and steady-state population density, p_j^*, both strictly increasing in productivity A_j. Therefore, cross-country variation induced by differences in productivity or technology will lead to a positive association between population density and income.

In contrast, consider two areas with the same productivity, A_j, but one of them has higher population density because of differences in other factors captured by ε_j in equation (3). Then, equation (2) implies that the country with higher population will have lower income per capita. So there is an identification problem: it is unclear whether an area with higher population density is in fact richer or not.

When high population density corresponds to lower income because of transitory differences, we should observe a subsequent decline in population density in the more densely settled areas—population there is above its long-run equilibrium level. We show below that differences in population density before 1500 were highly persistent—the more densely settled areas in 1500 were also more densely settled in 1000 or 1A.D.. This still leaves differences in other factors, leading to permanent differences in population density, so caution is required in interpreting population density as a proxy for income per capita.[6]

The empirical evidence regarding the relationship between population density and income is also less clear-cut than the relationship between urbanization and income. Figures A3 and A4 in the appendix show that population density and income increased concurrently in many instances. Nevertheless, there is no similar cross-sectional relationship in recent data, most likely because of the demographic transition—it is no longer true that high population density is associated with high income per capita because the relationship between income and the number of children has changed (e.g., Notestein, 1945, or Cipolla, 1974).

Despite these reservations regarding population density, we present results using population density, as well as urbanization, as a proxy for income per capita. This is motivated by three considerations. First, population density data are often more extensive, so the use of population density data is a useful check on our results using urbanization data. Second, as argued by Bairoch, population density is closely related to urbanization, and in fact, our measures are highly correlated. Third, variation in population density will play an important role not only in documenting the reversal, but also in explaining it. In any case, the relationship between population density in the past and income per capita today parallels the relationship between urbanization in 1500 and income today.

[6]This empirical framework also enables us to discuss another issue of interpretation. A common current interpretation of Malthus is that population dynamics should take all countries down to the subsistence level. This corresponds to the case in which $\rho = 1$ in our framework. In this case, more productive areas (high A_j) still have higher population densities, but all areas have the same long-run equilibrium income per capita, $y_j = \bar{y}$. Although we could still use population density as a proxy for land productivity, it would not be a proxy for income per capita. However, we believe that the case $\rho < 1$, which implies that there can be long-run differences in per capita income across countries, is more plausible, and consistent with the evidence regarding differences in cross-country living standards before the demographic transition.

3 Hypotheses

In this section, we outline the two major hypotheses investigated in this paper.

3.1 The geography hypothesis

The geography hypothesis claims that geographic, climatic and ecological factors have a first-order effect on prosperity and economic development. There are many different versions of this hypothesis.

Perhaps the most common is the view that climate has a direct effect on income through its influence on work effort. This idea dates back to Machiavelli (1519) and Montesquieu (1748). During the early 20th century, the geographer Ellsworth Huntington (e.g. 1915, 1945) pursued this idea further, and even attempted experiments to show the effect of climate on work effort. Both Toynbee (1934, volume 1) and Marshall (1890) similarly emphasized the importance of climate, both on work effort, and more generally, on productivity. One of the pioneers of development economics, Gunnar Myrdal (1968), also placed considerable emphasis on the effect of geography on agricultural productivity. He wrote, for example, that: "climate exerts everywhere a powerful influence on all forms of life," and that "serious study of the problems of underdevelopment... should take into account the climate and its impacts on soil, vegetation, animals, humans and physical assets— in short, on living conditions in economic development."

Jared Diamond has espoused a different theory of development, where the timing of the Neolithic revolution has had a long lasting effect by determining which societies were the first ones to develop strong armies, technology and even political and social organization. For example, he states that: "...proximate factors behind Europe's conquest of the Americas were the differences in all aspects of technology. These differences stemmed ultimately from Eurasia's much longer history of densely populated..." societies dependent on food production (1997, p. 358). According to Diamond, differences in the nature and history of food production, in turn, were due to the types of crops, domesticated animals, and the axis of agricultural technology diffusion in different continents, all of which are geographically determined characteristics.

More recently, Jeffrey Sachs has argued for the importance of geography through its effect on the disease environment, transport costs, and technology. He points out that "Economies in tropical ecozones are nearly everywhere poor, while those in temperate ecozones are generally rich" (2001, p. 1), and explains this by arguing that "Certain parts of the world are geographically favored. Geographical advantages might include access to key natural resources, access to the coastline and sea— navigable rivers, proximity to other successful economies, advantageous conditions for agriculture, advantageous conditions for human health." (2000, p. 30). He further suggests that "Tropical agriculture faces several problems that lead to reduced productivity of perennial crops in general and of staple food crops in particular" (2000, p. 32), and that "The burden of infectious disease is similarly higher in the tropics than in the temperate zones" (2000, p. 32). Finally, Sachs argues that the greater population in the temperate areas over the past centuries led to more rapid advances in technologies appropriate for these areas relative to technologies necessary for development in the tropics (2001, p. 3 and 2000, pp. 33-34).

Despite important differences between these versions of the geography hypothesis, for our purposes they all share the same persistence prediction: to the extent that geographic factors do not change over periods of 500 years or more, countries that were relatively rich 500 years ago should be relatively rich today. This prediction holds also when we consider the period of European colonialism. If climate, transport costs, the timing of the Neolithic revolution and diseases matter for income today, they should have mattered at least as much before Europeans came into contact with the inhabitants of the colonies.

3.2 The institutions hypothesis

According to the institutions hypothesis, societies with a social organization that provides encouragement for investment will prosper. There are also many different versions of this view. John Locke ([1690], 1980) was perhaps the first to clearly articulate the importance of property rights for production. Of land and productive assets, Locke wrote "...there must of necessity be *a means to appropriate them* some way or other, before they can be of any use, or at all beneficial to any particular man" (emphasis in the original, p. 10). He further argued that the main purpose of government was "the preservation of the property of ... members of the society" (p. 47). Similarly, Adam Smith and Frederick von Hayek, among many others, emphasized the importance of property rights for the success of nations.

The argument by some Marxist historians, for example Brenner (1976), on whether the capitalist class had the social and economic power to make the transition to capitalist agriculture is also related. Here the organization of the society, through its effect on the organization of productive activity, determines how productive agriculture is and whether new technologies are adopted.

More recently, economists and historians have emphasized the importance of institutions that guarantee property rights. For example, Douglass North starts his 1990 book by stating (p. 3): "That institutions affect the performance of economies is hardly controversial," and identifies property rights as a key aspect of good organization of society and institutions (see also North and Thomas, 1973).

For the purposes here, we take a "good" organization of society to correspond to a cluster of institutions ensuring that a broad cross-section of the society has effective property rights. We refer to this cluster as *institutions of private property*, and contrast them with *extractive institutions,* where the majority of the population faces a high risk of expropriation by the government, the ruling elite or other agents. Two requirements are implicit in this definition of institutions of private property. First, institutions should provide secure property rights, so that, as emphasized by Locke, those with productive opportunities expect to receive returns from their investments, and are encouraged to undertake such investments. The second requirement, which we believe is equally important, is embedded in the emphasis on "a broad cross-section of the society" having the opportunity to invest. A society in which a very small fraction of the population, for example, a class of landowners, holds all the wealth and political power may not be the ideal environment for investment, even if the property rights of this elite are secure. In such a society, many of the agents with investment opportunities and the entrepreneurial human capital may be those without the effective property rights protection. In particular, the concentration of political and social power in the hands of a small elite implies that the majority of the population does not have secure property rights, and probably risks being *held up* by the powerful elite. This is also consistent with North and Weingast's (1989, p. 805-806) emphasis that what matters is: "... whether the state produces rules and regulations that benefit a small elite and so provide little prospect for long-run growth, or whether it produces rules that foster long-term growth". Whether political power is broad-based or concentrated in the hands of a small elite is crucial in evaluating the role of institutions in the experiences of the Caribbean or India during colonial times, where the property rights of a small elite were well enforced, but the majority of the population had no civil rights or any effective property rights.

The organization of society and institutions also persist (see, for example, the evidence presented in Acemoglu, Johnson and Robinson, 2000). Therefore, the institutions hypothesis also suggests that, most of the time, societies that are prosperous today should continue to be so in the future. However, if a major shock disrupts the organization of a group of societies, we should expect much less persistence. Historical evidence suggests that European colonialism not only disrupted the social organization of a large number of countries, but in fact led to the establishment of extractive institutions in previously prosperous areas and to the development of institutions of private property in previously poor areas. Therefore, we suggest that there was an *institutional reversal* that resulted from European colonialism.[7] The institutions hypothesis, combined with the institutional reversal, predicts a reversal in the relative rankings among countries affected by European colonialism.

3.3 The institutional reversal

In this subsection, we argue that European colonialism had a fundamental effect on the social organization of the colonies, causing an institutional reversal: previously prosperous areas were more likely to end up with extractive institutions, while there often developed institutions of private property in previously poor areas.

Historical evidence is relatively unambiguous that Europeans used very different strategies in different colonies and that European colonialism had a fundamental effect on the organization of the colonial societies. While in a number of colonies such as the U.S., Canada, Australia, New Zealand, Hong Kong and Singapore, Europeans established institutions of private property, in many others they set up highly extractive institutions. Examples of extraction by Europeans include the transfer of gold and silver from Latin America in the 17th and 18th centuries and of natural resources from Africa in the 19th and 20th centuries, the Atlantic slave trade, the plantation agriculture in the Caribbean, Brazil and French Indochina, the rule of the British East India Company in India, and the rule of the Dutch East India Company in Indonesia from the early 17th century on. The distinguishing feature of the institutions that Europeans established in these countries was a high concentration of political power in the hands of a few who used their power to extract resources from the rest of the population. For example, the main objective of the Spanish and the Portuguese colonization was to obtain silver, gold and other valuables from America, and throughout they monopolized military power to enable the extraction

[7] By institutional reversal, we simply mean that countries that were relatively prosperous before the arrival of Europeans were more likely to end up with extractive institutions under European rule.

of these resources. The mining network set up for this reason was based on forced labor and the oppression of the native population. Similarly, the British West Indies in the 17th and 18th centuries was controlled by a small group of planters (e.g., Williams, 1970, Dunn, 1972). Political power was important to the planters in the West Indies, and to other elites in the colonies dominated by plantation agriculture, because it enabled them to force large masses of natives or African slaves to work for low wages.

Europeans running the Atlantic slave trade, despite their small numbers, also had disproportionate power in Africa. The consensus view amongst historians appears to be that the slave trade fundamentally altered the organization of the society in Africa, leading to state centralization and warfare as African states competed and expanded to supply the European slavers. Manning (1990, p. 147) describes this situation as follows: "with the allure of imported goods and the brutality of capture, slave traders broke down barriers isolating Africans in their communities. Merchants and warlords spread the tentacles of their influence into almost every corner of the continent. By the 19th century, much of the continent was militarized; great kingdoms and powerful warlords rose and fell, their fate linked to fluctuations in the slave trade."[8]

What determines whether Europeans pursued an extractive strategy or introduced institutions of private property? And why was extraction more likely in relatively prosperous areas? Two factors appear crucial:

1. *The economic profitability of alternative policies*: When extractive institutions were more profitable, Europeans were more likely to opt for them. High population density, by providing a supply of labor that could be forced to work in agriculture or mining, made the development of extractive institutions, with poltical power concentrated in the hands of a small elite, more profitable. For example, the presence of abundant Amerindian labor in the Meso-America region was conducive to the establishment of large plantation complexes, while the high population of West Africa created a profit opportunity for slave traders, supplying additional labor to the plantations in the Americas. The experience of the Spanish conquest around the La Plata river (current day Argentina) during the early 16th century gives an interesting example of how population density affected European colonization (see Lockhart and Schwartz, 1983, pp. 259-60, for details). Early in 1536, a very large

[8]There are many detailed studies of different parts of Africa supporting this claim. See for example, Wilks (1975) for Ghana, Low (1977) for Nigeria, Harms (1981) for the Congo/Zaire, Miller (1988) on Angola.

Spanish expedition arrived in the area, and founded the city of Buenos Aires in the mouth of the river Plata. The area was sparsely inhabited by non-sedentary Indians. The Spaniards could not enslave a sufficient number of Indians for food production. Starvation forced them to abandon Buenos Aires and retreat up the river to a post at Asuncion (current day Paraguay). This area was more densely settled by semi-sedentary Indians. These Indians were enslaved by the Spaniards for food production, and the colony of Paraguay, with relatively extractive institutions, was founded. In contrast, Argentina was finally colonized later, with greater European settlements, no forced labor, and generally better institutions.

Other types of extractive institutions were also more profitable in densely-settled and prosperous areas, where there was more to be directly plundered by European colonists. More important, in these densely-settled areas there was an existing system of tax administration or tribute, and the large population made it profitable for the Europeans to take control of the existing tax and tribute systems and to increase taxes on the native population (see, e.g., Wiegersma, 1988, p. 69, on French policies in Vietnam, or Marshall, 1998, pp. 492-497, on British policies in India).

2. *Whether Europeans could settle or not*: Europeans were more likely to develop institutions of private property when they settled in large numbers, for the natural reason that they themselves were affected by these institutions.[9] Moreover, when a large number of European settled, the lower strata of the settlers demanded rights and protection similar to, or even better than, those in the home country. This made the development of effective property rights for a broad cross-section section of the society more likely. European settlements, in turn, were affected by population density both directly and indirectly. Population density had a direct effect on settlements, since Europeans could easily settle in large numbers in sparsely inhabited areas. The indirect effect, on the other hand, worked through the disease environment. In many of the densely settled areas, there were diseases—in particular, malaria and yellow fever—for which Europeans did not have an immunity.[10]

[9]Extraction and European settlement patterns were mutually self-reinforcing. In areas areas where extractive policies were pursued, the authorities also actively discouraged settlements by Europeans, presumably because this would intervene with the effective extraction of resources from the locals (e.g. Coatsworth, 1982).

[10]Most diseases require a dense human population to act as carriers. In Acemoglu, Johnson and

> European settlements shaped both the type of institutions that developed and the structure of production. For example, while in Potosí (Bolivia) mining employed forced labor (Cole 1985), and in Brazil and the Caribbean sugar was produced by African slaves, in the U.S. and Australia mining companies employed free migrant labor, and sugar was grown by smallholders in Queensland, Australia (Denoon, 1983). Consequently, in Bolivia, Brazil and the Caribbean, political institutions were designed to ensure the control of the laborers and slaves, while in the U.S. and Australia, the workers and smallholders had greater political rights (Cole 1985, Hughes 1988, chapter 10).

The historical evidence is therefore consistent with our notion that European colonialism, by introducing worse institutions in previously densely-settled and prosperous areas and developing institutions of private property in previously poor regions, caused an institutional reversal. We further substantiate this view empirically in the next subsection.

3.4 Econometric evidence on institutional reversal

Table 3 shows the relationship between urbanization or population density in 1500 and institutions using four different measures of institutions. The first two measures refer to current institutions: protection against expropriation risk between 1985 and 1995 from Political and Risk Services, which approximates how secure property rights are, and constraints on the executive in 1990 from Gurr's Polity IV data set, which can be thought of as a proxy for how concentrated power is in the hands of ruling groups. Columns 1-6 of Table 3 show a negative relationship between our measures of prosperity in 1500 and current institutions.

It is perhaps more relevant to know whether there is such an institutional reversal during the colonial times or shortly after independence. Since the Gurr data set does not contain information for non-independent countries, we can only look at this after independence. Columns 7-12 show the relationship between our measures of prosperity in 1500 and measures of early institutions. The first measure we use is constraints on the

Robinson (2000), we documented that Europeans were less likely to settle in areas infected by malaria and yellow fever, and both of these diseases were absent in sparsely settled areas such as Australia or New Zealand. The diseases in the New World did not initially cause high European mortality, but malaria and yellow fever developed soon after African slaves were brought to the continent. In any case, the correlation between population density and the European settler mortality variable we constructed in Acemoglu, Johnson and Robinson (2000) is 0.33.

executive in 1900. Since colonial rule often concentrated political power in the hands of a small elite, it seems reasonable to assign the lowest score to countries still under colonial rule in 1900. As an alternative variable, we use constraints of the executive in the first year of independence from the same data set, while also controlling for time since independence as an additional covariate. Notice that when both urbanization and log population density in 1500 are included, it is the population density variable that is significant. This supports the interpretation that it was the differences between densely and sparsely settled areas that was crucial in determining what type of institutions Europeans introduced (though it may also reflect the fact that the population density variable is measured with less measurement error). Finally, the second panel of the table includes (the absolute value of) latitude as an additional control, showing that the institutional reversal does not reflect some geographic pattern of institutional change.

Figure 4 shows the negative relationship between constraints on the executive in the first year of independence and our measures of prosperity in 1500 diagrammatically (the effect of time since independence is controlled for by orthogonalizing both the left-hand side and the right-hand side variables with respect to time since independence). Many of the colonies such as Canada, the United States, New Zealand and Australia that were relatively poor before Europeans arrived became independent with relatively good institutions, which contrasts with the experience of many Latin American, African and Asian countries, which were relatively prosperous in 1500, but inherited extractive institutions from colonial powers.

Table 3: Institutions vs. urbanization in 1500 and population density in 1500

Figure 4

Overall, we interpret both the historical and the econometric evidence as supporting the notion that European colonialism caused an institutional reversal. Therefore, the institutions hypothesis suggests the possibility of a reversal in relative prosperity among European colonies between 1500 and today, while the geography hypothesis predicts a high degree of persistence during the same period.

4 The Reversal of Fortune

4.1 Results with urbanization

This section presents our main results. Figure 1 in the introduction depicts the relationship between urbanization 1500 and income per capita today. Table 4 reports regressions documenting the same relationship. Column 1 is our most parsimonious specification, regressing log income per capita in 1995 (PPP basis) on urbanization rates in 1500 for our sample of former colonies.[11] The coefficient is -0.08 with a standard error of 0.03. This coefficient implies that a 10 percent lower urbanization in 1500 is associated with approximately twice as high GDP per capita today (70 log points≈100 percent).[12] It is important to note that this is not simply mean reversion—i.e., richer than average countries reverting back to the mean. It is a *reversal*. To illustrate this, let us compare Chile and India. The native population in Chile basically had no urbanization, while India had an urbanization rate of at least 8.5 percent (though there are also estimates as high as 15 percent for urbanization in India, see Bairoch, 1988). The estimate in column 1 of Table 2, 0.038, for the relationship between income and urbanization implies that India at the time was at least 33 percent richer than Chile (if we take higher estimates of urbanization in India, this gap would be correspondingly greater). Today according to our estimate in column 1 of Table 4, we expect Chile to be approximately 70 percent richer than India (in practice, Chile is about 6 times as rich as India).

Table 4 here.

The second column of Table 4 excludes North African countries for which data quality may be lower. The result is unchanged, with a coefficient of -0.11 and standard error of 0.03. Column 3 drops the Americas, which increases both the coefficient and the standard error, but the estimate remains highly significant. Column 4 adds continent dummies, to check whether the relationship is being driven by differences across continents. Al-

[11] Here we limit ourselves to looking at log GDP per capita in 1995 as the dependent variable. For completeness, the top panel of Table 8 shows the relationship between urbanization in 1995 and urbanization in 1500.

[12] Because China was never a formal colony, we do not include it in our sample of ex-colonies. Adding China strengthens the results further. For example, with China, the baseline estimates changes from -0.0815 (s.e.=0.0258) to -0.0822 (s.e.=0.0255). Furthermore, our sample excludes counties that were colonized by European powers briefly during the 20th century, such as Iran, Saudi Arabia and Syria. If we include these observations, the results are unchanged. For example, the baseline estimate changes to -0.074 (s.e.=0.024).

though continent dummies are jointly significant, the coefficient on urbanization in 1500 is unaffected: it is -0.09 with a standard error of 0.03.

One might also be concerned that the relationship is being driven mainly by the Neo-Europes; USA, Canada, New Zealand and Australia. These countries are settler colonies built on lands that were inhabited by relatively undeveloped civilizations. Although the contrast between the development experiences of these areas and the relatively advanced civilizations of India or Central America is of independent interest, one would like to know whether there is anything more than this contrast in the results of Table 4. In column 5, we drop these observations. The relationship is now weaker, but still negative and statistically significant at the 6 percent level.

Is the reversal in relative incomes related to geography? To investigate this issue, in columns 6 and 7, we add a variety of geography-related variables, including distance from the equator, and differences in temperature, humidity, soil quality, natural resource endowments. These variables themselves are insignificant, and have almost no effect on the pattern of the reversal. This indicates that the reversal is not apparently related to geography.

Finally, in columns 8 and 9, we add the identity of the colonial power (or legal origin), which is emphasized by Hayek (1960) and La Porta et al. (1998) as important determinants of development, and religion, stressed by Max Weber. These also have little effect on our estimate.

The urbanization variable used in Table 4 relies on the work by Bairoch and Eggimann, and as explained in the previous section, we had to make a number of assumptions to combine estimates from these researchers. In Table 5, we use separately data from Bairoch and Eggimann, as well as data from Chandler, who provided the starting point for Bairoch's data. We repeat the regressions of Table 4 using these three alternative series. The results are very similar to the baseline estimates reported in Table 5: in all cases, there is a negative relationship between urbanization in 1500 and income per capita today, and in almost all cases, this relationship is statistically significant at the 5 percent level.

Table 5 here.

4.2 Results with population density

A big gap in our urbanization data is for sub-Saharan Africa. Bairoch (1988) refers to our knowledge regarding Africa as: "Black Africa: An Urban History That Remains to Be Written" (p. 52). Nevertheless, we have population density data for these countries. In Panel A of Table 6, we regress income per capita today on population density in 1500. The results are quite similar to those in Table 4. In all specifications, we find that countries with higher population density in 1500 are substantially poorer today. The coefficient of -0.38 in column 1 implies that a 10 percent higher population density in 1500 is associated with a 4 percent lower income per capita today. For example, the area now corresponding to Bolivia was seven times more densely settled than the area corresponding to Argentina, so on the basis of this regression, we expect Argentina to be three times as rich as Bolivia, which is more or less the current gap in income between these countries.

Table 6 here.

The remaining columns perform the same robustness checks as in Table 4 and 5, and show that including geography-related variables, the identity of the colonial power, religion variables, or dropping the Americas, the Neo-Europes, or North Africa has very little effect on the results.

The estimates in the top panel of Table 6 use variation in population density. This reflects two components; differences in population and differences in arable land area. In Panel B, we separate the effects of these two components, and find that they come in with equal and opposite signs. This suggests that our specification with population density alone was appropriate.

Finally, in Panel C we estimate the same relationships as in Panel A, but using population density in 1000 as an instrument for population density in 1500. This is useful since, as discussed in Section 2.4, it is differences in long-run population density that are likely to be better proxies for income per capita. Instrumenting for population density in 1500 with population density in 1000 isolates the long-run component of population density differences across countries (i.e., the component of population density 1500 that is correlated with population density in 1000). The Two-Stage Least Squares (2SLS) results in Panel C using this instrumental variables strategy are very similar to the OLS results in Panel A.

4.3 Further results, robustness checks and discussion

Caution is required in interpreting the results presented in Tables 4-6. Estimates of urbanization and population more than 500 years ago are likely to be ridden with error. Nevertheless, the first effect of measurement error would be to create an attenuation bias towards 0. Therefore, one might think that the negative coefficients in Tables 4-6 are, if anything, underestimates. A more serious problem would be if errors in the urbanization and population density estimates were not random, but correlated with current income in some systematic way. So it is important to ensure that our results are robust to plausible variations in data. The results with alternative urbanization estimates in Table 5 suggest that this is not major problem. We investigate this issue further in Table 7. We use a variety of different estimates on urbanization and population density, such as those assigning lower urbanization and population density numbers to the Americas, North Africa and India. The results are robust to these changes. Panel B of this table also reports results weighted by population in 1500, with very similar results.

Much of the variation in urbanization or population density is not at the country level, but at the level of "civilizations". For example, in 1500 there were many fewer separate civilizations in the Americas, and perhaps even in Asia. For this reason, we repeat our key regressions using variation only at the civilization level. First, in column 7 we use only data corresponding to the areas occupied by the 7 civilizations that are identified by Arnold Toynbee in *A Study of History* and are in our ex-colonies sample. In column 8, we then use an extended classification with 14 civilizations.[13] The results confirm our basic findings. With only the 7 civilizations identified by Toynbee, there is a negative, but statistically insignificant, relationship (which is natural given the number of observations). When we use all 14 civilizations, this relationship becomes significant at the 5 percent level.

Table 7 here.

Finally, in the bottom panel of Table 7 we include urbanization and population density simultaneously in these regressions. In all cases, population density is negative and highly significant, while urbanization is insignificant. This is again consistent with the notion

[13]The 7 civilizations identified by Toynbee which are relevant to our study, are Andean, Yucatec, Mexican, Arabic, Hindu, Eskimos and Polynesian. We follow McNeill (1967) and add to this list, North American Indian, Non-Andean South American Indian, Australian Aboriginal, Malaysian-Indonesian, Filipino, Vietnam-Cambodia-Laos, Caribbean to arrive to 14 civilizations.

that differences in population density played a key role in the institutional reversal, and therefore in causing the reversal in relative incomes among the colonies (though it may also reflect measurement error in the urbanization estimates).

Is the reversal shown in Figures 1 and 2 and Tables 4, 5 and 6 consistent with other evidence? There seems to be little doubt from the literature on the history of civilizations, e.g. Toynbee (1934-61), and also from recent qualitative assessments, e.g. McNeill (1967), that 500 years ago, many parts of Asia were highly prosperous (perhaps as prosperous as Western Europe), and civilizations in Meso-America and North Africa were relatively developed. In contrast, there was little agriculture in most of North America, Australia and New Zealand, at most consistent with population density of 0.1 people per square kilometer (see the map in McEvedy and Jones p. 273, reproduced below in the Appendix). McEvedy and Jones (1975, p. 322) describe the state of Australia at this time as "an unchanging palaeolithic backwater". In fact, because of the relative backwardness of these areas, European powers did not view them as valuable colonies. Voltaire is often quoted as referring to Canada as "few acres of snow", and the European powers at the time paid little attention to Canada relative to the colonies in the West Indies. In a few parts of North America, along the East Coast and in the South-West, there was settled agriculture, supporting a population density of approximately 0.4 people per square kilometer, but this was certainly much less than that in the Aztec and Inca Empires, which had a fully developed agriculture with a population density of between 1 and 3 people (or even higher) per kilometer squared, and also much less than the corresponding numbers in Asia and Africa.

Overall we conclude that the evidence points to a reversal in relative incomes among former European colonies. This reversal is inconsistent with geographic factors being the major cause of income differences across countries. But it is consistent with the institutions hypothesis. As argued in the previous section, historical and econometric evidence suggests that European colonialism led to an institutional reversal, creating worse institutions in previously prosperous areas. Therefore, if institutions are a crucial determinant of economic performance, we should expect a reversal in relative economic prosperity. In this light, we interpret the evidence in this section as providing support to the institutions hypothesis against the simple geography hypothesis.

4.4 Persistence before 1500 and Persistence among the non-colonies

Like the geography hypothesis, the institutions view predicts persistence in economic prosperity as long as there is no major shock to the social organization of the societies in question. It is therefore instructive to briefly look at patterns of persistence among non-colonized countries and among our sample of ex-colonies during the period before 1500. Table 8 reports a range of relevant results.

Table 8 here.

Two important patterns emerge from this table. First, among European countries, or all countries that were not colonized by Western Europe (including European countries), there is persistence between 1500 and today (columns 6-9). Second, both in the sample of non-colonized countries and in our sample of former colonies, there is persistence between 1000 and 1500, or even between 1AD and 1000 when we use log population density (columns 1-5). These results suggest that the reversal we document in this section indeed reflects an unusual event, and gives further support to the idea that the reversal is related to European colonialism.

5 The Northern Drift, The Timing of the Reversal and Industrialization

The evidence presented so far weighs against the view that geography is a major determinant of economic outcomes. Nevertheless, the reversal does not necessarily reject a more sophisticated geography view, which emphasizes the Northern (or away from the equator) shift in the center of economic gravity over time. According to this view, which we call "the Northern drift" hypothesis, geography becomes important when it interacts with the presence of certain technologies. For example, one can argue that tropical areas provided the best environment for early civilizations—after all humans evolved in the tropics and the required calorie intake is lower in warmer areas. But with the arrival of "appropriate" technologies, temperate areas became more productive (e.g., White, 1962). The technologies that were crucial for the rise of civilizations in temperate areas include the heavy plow, systems of crop rotation, domesticated animals such as cattle and sheep, and some of the high productivity European crops, such as wheat and barley. Despite the importance of these technologies for development in the temperate areas, they have had much less of an effect on tropical zones. Although we are not aware of any social scientist who has explicitly formulated this view, Jeffrey Sachs implies it in his recent paper when he writes "Since technologies in the critical areas of agriculture, health, and related areas could diffuse *within* ecological zones, but not across ecological zones, economic development spread through the temperate zones but not through the tropical regions" (2001, p. 12, italics in the original).

The reversal in relative incomes does not rule out an explanation for the reversal based on the Northern drift hypothesis. Many of the colonies that subsequently became rich, such as North America or Australia, were situated in temperate climates. The populations in these areas did not have access to the agricultural technology in use in Europe at the time, and the spread of this technology, which came with European colonialism, may have enriched these areas, causing the reversal.

The Northern drift view is plausible in the context of agricultural technology, since different climates favor different types of crops and harvesting methods. Yet there seems to be no compelling case that climate and geography should matter more for industry than for agriculture. Therefore, the Northern drift hypothesis suggests that the reversal should be associated with the spread of European agricultural technologies.[14]

The timing and the nature of the reversal in relative incomes do not support the Northern drift hypothesis, however. First, the results presented so far give no indication that the reversal is related to any geographic characteristics (e.g., see Table 4). Perhaps more important, European agricultural technology spread to the colonies between the 16th and 18th centuries (see for example McCusker and Menard, 1985, Chapter 3 for North America). So the Northern drift hypothesis would suggest that the reversal should take place during this period. In contrast, the evidence indicates that the reversal in relative incomes is largely a 19th-century phenomenon. This is documented in Figures 5 and 6.

Figure 5 shows that among the ex-colonies there appears to be no reversal between 1500 and 1700, while the reversal is quite apparent between 1700 and today. Figure 6 looks in detail at the evolution of urbanization in a number of countries to date the timing of the reversal. In Figure 6a, we classify ex-colonies according to their level of

[14]There is another line of argument which might suggest that geography matters more for industry than for agriculture. Imagine there is more room for specialization in industry, but this requires trade, and countries differ according to their transport costs. This may create an advantage for countries with low transport costs once industrialization becomes possible. This argument is not very convincing, however, since many of the previously prosperous colonies that failed to industrialize include islands such as the Caribbean, or countries with natural ports such as those in Central America, India or Indonesia.

urbanization in 1500.[15] The figure shows the higher level of urbanization among the previously high urbanization colonies until 1850, and then a big surge in urbanization among the previously low urbanization colonies starting in 1850. This is largely driven by the U.S., but also by Canada, and Argentina. Figure 6b shows a comparison of Mexico, India and the U.S. as three important and large colonies, reiterating the pattern in Figure 6a. Figure 6c shows the comparison of Mexico, the U.S., Brazil and Peru, this time using data from Mitchell (1983) and (1995).[16]

Finally, Figure 6d shows per capita industrial production for the U.S., Canada, New Zealand, Australia, Brazil, Mexico, and India using data from Bairoch (1982). This figure shows the takeoff in industrial production in the U.S., Australia, Canada and New Zealand relative to Brazil, Mexico and India. Although it is difficult to see it in the figure, in fact in 1750, per capita industrial production was higher in India, 7, relative to the U.S., 4 (these numbers are an index with UK industrial production per capita in 1900 normalized to 100). Bairoch (1982) also reports that in 1750, China had industrial production per capita twice the level of the U.S., 8. Yet, as Figure 6d shows, over the next 200 years, there was a much larger increase in industrial production in the U.S. than in India (and also than in China).

This general interpretation that the reversal in relative incomes took place during the 19th century, and was linked to industrialization, is also consistent with the fragmentary evidence we have on other measures of income per capita and industrialization. For example, Coatsworth (1993), Eltis (1995) and Engerman and Sokoloff (1997) provide evidence that much of Spanish America and the Caribbean was more prosperous (higher per capita income) than British North America until 1700. The future United States rose in per capita income during the 1700s relative to the Caribbean and South America, but only really pulled ahead in the 100 years after 1800 (Coatsworth, 1993). McCusker and Menard (1985) and Galenson (1996) both emphasize that productivity and income growth

[15]Former colonies with high urbanization in 1500 are Ecuador, Mexico, Guatemala, Honduras, Nicaragua, Panama, El Salvador, Belize, Algeria, Tunisia, Egypt, Morocco, and India. Former colonies with low urbanization in 1500 are Argentina, Brazil, Canada, Chile, Paraguay, Uruguay, and the U.S. The data used in this graph are from Chandler through 1861 and Bairoch for the twentieth century. Australia and New Zealand are not included in this graph, since they were settled only in the 19th century. Including these countries, assuming constant urbanization before European settlements, does not affect the pattern shown in the figure.

[16]Mitchell provides a continuous series at approximately 10 year intervals for urbanization from the late 18th to 20th century. In contrast most of Chandler's data end in 1850, Bairoch reports estimates only for every 50 years, and Eggimann does not provide numbers for some key countries, such as the U.S.. Where they overlap, Mitchell's estimates are consistent with those from our other sources.

in North America before the 18th century was limited.[17] Using data from Maddison (1995), Pritchett (1997) documents the large divergence in per capita income from 1870 onwards.

The available data also suggest that during the critical period of growth in the U.S., between 1840 and 1900, there was modest growth in agricultural output per capita, and very rapid growth in industrial output per capita. Gallman (2000) reports that in 1840, agriculture was 41 percent of GDP, while manufacturing was 17 percent of GDP. By 1900, these numbers had changed to 19 and 31 respectively. He also reports that GNP per capita in 1860 dollars was approximately 96.5 around 1840, and 269.1 around 1900. These numbers imply that between 1840 and 1900, agricultural product per capita increased roughly from 39.5 to 51.1, which is very small relative to the increase in manufacturing output per capita, which went up from 16.4 in 1840 to 83.4 in 1900.

Overall this evidence suggests that previously poor colonies became relatively rich during the 19th century, and this was associated with rapid industrialization in these areas. We interpret this pattern as evidence against the Northern drift hypothesis.[18] We next present econometric evidence that is consistent with the role of institutions in the reversal.[19]

Figures 5 and 6.

[17]McCusker and Menard (1985, p. 270) argue "although there was some gains in agriculture, manufacturing and shipping, improvments in technology had only a minor impact on the colonial economy [1607-1785]....improvements in economic organization, specifically shifts away from self-sufficiency and toward production for exchange, and the development of more efficient markets, were the principal sources of growth in the colonies of British America."

[18]The timing and nature of the reversal also weigh against two other potential hypotheses. The first, inspired by Williams (1944), Rodney (1972), Frank (1978) and Wallerstein (1974-1980), would be that the Europeans extracted resources from, or more generally plundered, the formerly rich colonies, and this plunder may have caused the reversal. Since the extraction of resources was most intense shortly after initial colonization, this hypothesis would also suggest an early reversal. An alternative possible thesis is that the reversal reflects the direct effect of the diseases Europeans brought to the New World (e.g., McNeill, 1976, Crosby, 1986, Diamond, 1997). Since the remaining inhabitants of the Americas had developed immunities to these diseases by the 17th century, this view also suggests we should observe a relatively early timing for the reversal.

[19]Another hypothesis with a flavor similar to the geography views would be that different paths of development reflect the direct influence of Europeans. Places where there are more Europeans have become richer, either because Europeans brought certain values conducive to development (e.g. Landes, 1998), or because having more Europeans confers certain benefits (e.g., through trade with Europe or because Europeans are more productive). Since Europeans settled in greater numbers in previously sparsely settled areas, the reversal may reflect this effect. In Acemoglu, Johnson and Robinson (2001) we show that once the effect of institutions are taken into account appropriately, the current racial composition of a country has no effect on income. So we do not pursue this explanation further here.

6 Institutions and The Making of the Modern World Income Distribution

We have so far documented the reversal in economic prosperity among former European colonies between 1500 and today, and argued that this is inconsistent both with the simple and sophisticated geography views. We have also linked the reversal in economic performance to the institutional reversal among these countries caused by European colonialism. The next step is to argue that this institutional reversal and the institutional differences between these countries account for their different economic trajectories, and to investigate the mechanism through which institutions have affected economic development. In this section, we first show that institutional differences statistically account for the reversal between 1500 and today. We next document that institutions started playing a much more important role in the age of industry.

6.1 Institutions and the reversal

If the institutional reversal is the reason why there was a reversal in income levels among the former colonies, then once we account for the role of institutions appropriately, the reversal should disappear. That is, according to this view, the reversal documented in Figures 1 and 2 and Tables 4-7 reflects the correlation between economic prosperity in 1500 and income today working through the intervening variable, institutions.

How do we establish that an intervening variable X is responsible for the correlation between variable Z and Y? To answer this question, suppose that the true relationship between Y, and X and Z is

$$Y = \alpha \cdot X + \beta \cdot Z + \varepsilon \tag{4}$$

where α and β are coefficients and ε is a disturbance term. In our case, we can think of Y as income per capita today, X as measures of institutions, and Z as population density (or urbanization) in 1500. The variable Z is included in equation (4) either because it has a direct effect on Y or because it has an effect through some other variables not included in the analysis. The hypothesis we are interested in is that $\beta = 0$—that is, the only channel through which population density or urbanization has an effect on income today is through institutions.

This hypothesis obviously requires that there is a statistical relationship between X and Z. So we postulate that

$$X = \lambda \cdot Z + \nu.$$

To start with, suppose that ε is independent of X and Z and that v is independent of Z. Now imagine a regression of Y on Z only (in our context, of income today on prosperity in 1500), similar to those we reported in Tables 4-7:

$$Y = b \cdot Z + u_1.$$

As is well-known, the probability limit of the OLS estimate from this regression, $\widehat{b}$, is

$$\text{plim}\widehat{b} = \beta + \alpha \cdot \lambda.$$

So the results in the regressions of Tables 4-7 are consistent with $\beta = 0$ as long as $\alpha \neq 0$ and $\lambda \neq 0$. In this case, we could be capturing the effect of Z (population density or urbanization) on income working solely through institutions. This is the hypothesis that we are interested in testing. Under the assumptions regarding the independence of Z from ν and ε, and of X from ε, there is a simple way of testing this hypothesis, which is to run an OLS regression of Y on Z and X:

$$Y = a \cdot X + b \cdot Z + u_2, \tag{5}$$

to obtain the estimates $\widehat{a}$ and $\widehat{b}$. The fact that ε in (4) is independent of both X and Z rules out omitted variable bias, so $\text{plim}\widehat{a} = \alpha$ and $\text{plim}\widehat{b} = \beta$. Hence, a simple test of whether $\widehat{b} = 0$ is all that is required to test our hypothesis that the effect of Z is through X alone.

In practice, there are likely to be many omitted variables, endogeneity bias because Y has an effect on X, and attenuation bias because X is measured with error or corresponds poorly to the real concept that is relevant to development.[20] So the above procedure is not possible. However, the exact logic applies as long as we have a valid instrument, M, for X, such that

$$X = \gamma \cdot M + \zeta,$$

and M is independent of ε in (4). We can then simply estimate (5) using Two-Stage Least Squares (2SLS) with the first-stage

$$X = c \cdot M + d \cdot Z + u_3.$$

[20]For example, plausibly what matters for economic development is a broad range of institutions, whereas we only have an index for a particular type of institutions, e.g., enforcement of private property rights.

Testing our hypothesis that Z has an effect on Y only through its effect on X, then amounts to testing that the 2SLS estimates of b, $\widetilde{b}$ is equal to 0. Intuitively, the 2SLS procedure ensures that we obtain a consistent estimate of α, thus enabling us to appropriately test whether Z has a direct effect.

The key to the success of this strategy is obviously a good instrument for X. Here we refer to our previous work, Acemoglu, Johnson and Robinson (2001), where we suggested and documented that mortality rates faced by settlers are a good instrument for settlements of Europeans in the colonies and the subsequent institutional development of these countries. Table 9 reports results from this type of 2SLS tests using the log of potential settler mortality rates as an instrument for institutional development.[21]

Table 9 here

Panel A reports results from regressions that enter urbanization and log population density in 1500 as separate (exogenous) regressors in the first and the second stages, while Panel B reports the corresponding first stages. Different columns correspond to different institutions variables, or different specifications. For comparison, Panel C reports the 2SLS coefficient with exactly the same sample as the corresponding column, but without urbanization or population density. We look at the same four institutions variables used in Table 3; protection against expropriation risk from Political and Risk Services, and constraints on the executive in 1990, in 1900 and in the first year of independence from Gurr's Polity IV data set. The results are consistent with our hypothesis. In all columns, we never reject the hypothesis that urbanization 1500 or population density in 1500 have *no* direct effect once we control for the effect of institutions on income per capita, and the addition of these variables has little effect on the 2SLS estimate of the effect of institutions on income per capita. This supports our notion that the reversal in economic prosperity reflects the effect of early prosperity and population density working through the institutions and policies introduced by European colonialists.

6.2 Institutions and industrialization

The historical evidence suggests that the institutional reversal—the introduction of extractive institutions in previously prosperous areas and institutions of private property in

[21]See Acemoglu, Johnson and Robinson (2001) for details of how the settler mortality data are constructed and supporting evidence that it is a valid instrument for differences in institutions.

previously poor areas—was well underway by the 19th century. So why did the reversal take place during the 19th century?

To answer this question, let us first reiterate that institutions of private property require a broad cross-section of the society to have effective property rights. Imagine a society similar to the colonies in the Caribbean during the 17th and 18th centuries where a small elite controls all the political power. The revolution in Haiti notwithstanding, the property rights of this elite are relatively well protected, but the rest of the population has no effective property rights. According to our definition, this is not a society with institutions of private property. Nevertheless, when the major investment opportunities are in agriculture, this lack of effective private property enforcement for the majority of the population may not matter too much, since the elite can invest in the land and employ the rest of the population, and so will have relatively good incentives to increase output. Now imagine the arrival of a new technology, for example, industrial opportunities. If the elite could undertake industrial investments without losing its political power, we may expect them to take advantage of these opportunities. However, in practice there will be two major problems. First, the people with the entrepreneurial skills and ideas may not be members of the elite. These potential entrepreneurs will not undertake the necessary investments, because they do not have secure property rights. They correctly anticipate that they will be held up once they undertake these investments because political power still rests in the hands of the elite. Second, the elite may want to *block* investments in new industrial activities, fearing the political turbulence and the threat to their political power that new technologies will bring (see Acemoglu and Robinson, 2001).

This reasoning suggests that whether a society has institutions of private property or extractive institutions may matter much more when new technologies require a broad-based economic participation. Industrialization is precisely such a process, requiring investments from a large number of agents who were not previously part of the ruling elite. Therefore, there are natural reasons to expect that institutional differences will matter much more during the age of industry. This is consistent with the evidence discussed in the last subsection. We now investigate this issue in more detail.

If this hypothesis is correct, we should expect societies with good institutions to take better advantage of industrialization opportunities. We can test this using data on institutions, industrialization and GDP from the 19th and early 20th centuries. Bairoch (1982) presents detailed evidence on industrial output for a number of countries at a variety of dates, and Maddison (1995) has estimates of GDP for a larger group of countries. We take Bairoch's estimates of UK industrial output as a proxy for industrialization opportunities for the whole world, since during this period the UK was the industrial leader. We then run a panel data regression of the following form:

$$y_{it} = \delta_t + \delta_i + \pi \cdot X_{it} + \phi \cdot X_{it} \cdot UKIND_t + \varepsilon_{it}, \tag{6}$$

where y_{it} is the outcome variable of interest in country i at date t. We consider industrial output per capita and income per capita as two different measures of economic success during the 19th century. In addition, δ_t's are a set of time effects and δ_i's denote a set of country effects, X_{it} denotes the measure of institutions in country i at date t, and $UKIND_t$ is industrial output per capita in the UK at date t. The coefficient of interest is ϕ, which reflects whether there is an interaction between good institutions and industrialization opportunities. A positive and significant ϕ is interpreted as evidence in favor of the view that countries with institutions of private property took better advantage of the opportunities to industrialize. The parameter π measures the direct effect of institutions on industrialization, and is evaluated at the mean value of $UKIND_t$.

The top panel of Table 10 reports regressions of equation (6) with industrial output per capita as the left-hand side variable. Here we have data for the U.S., Canada, New Zealand, Australia, South Africa, Brazil, Mexico, and India. The bottom panel considers log GDP per capita as the dependent variable, and in addition, includes data also from Argentina, Bangladesh, Burma/Myanmar, Chile, Colombia, Egypt, Ethiopia, Ghana, India, Indonesia, Ivory Coast, Kenya, Morocco, Nigeria, Pakistan, Peru, Tanzania, Venezuela, and Zaire.[22] We also have data for the following dates: 1830, 1860, 1880, 1900, 1913, 1928, 1953, and 1980 (though not for all countries for all dates).

Column 1 reports a regression of (6) using only pre-1950 data. The interaction term, ϕ, is estimated to be 0.13, and is highly significant with a standard error of 0.03. To calculate the implication of this estimate, note that Bairoch's estimate of total UK industrialization rose from 16 to 115 between 1800 and 1913 (UK industrial production per capita in 1900 normalized to 100). In the meantime, the U.S. per capita production grew from 9 to 126, whereas India's per capita industrial production *fell* from 6 to 2. Since the average difference between the constraints on the executive in the U.S. and India over this period is approximately 6, the estimate implies that the U.S. industrial output per capita should

[22]In the top panel we did not use logs, since industrial output per capita was 0 for some of the observations.

have increased by 77 more than India's, which is over half the actual difference. If instead we use the instrumental variables (IV) estimate from column 7, 0.25, the predicted increase turns out to be somewhat larger than the actual increase.

In column 2, we drop the country dummies, with little effect on the results. In column 3, we use extended data through 1980, again with no effect on the results. In columns 4 and 5, we use average institutions for each country, $\bar{X}_i$, rather than institutions at date t, so the equation becomes $y_{it} = \delta_t + \delta_i + \phi \cdot \bar{X}_i \cdot UKIND_t + \varepsilon_{it}$. This specification may give more sensible results if either variations in institutions from year to year are endogenous with respect to changes in industrialization or income, or are subject to measurement error. Interestingly, ϕ is now estimated to be larger, suggesting that measurement error is a more important problem than the endogeneity of the changes in institutions across time periods.

The real advantage of the specifications in columns 4 and 5 is that they allow us to instrument for the regressor of interest $\bar{X}_i \cdot UKIND_t$, using the interaction between log settler mortality (which we used in Table 9) and UK industrialization per capita, $M_i \cdot UKIND_t$, to deal with the issue of endogeneity of institutions.[23] Once again, institutions might differ across countries because more productive or otherwise different countries have different institutions, and in this case, our interaction of industrialization and institutions could be capturing the effects of these characteristics on economic performance. To the extent that mortality rates faced by the early settlers is a good instrument for institutions (see Acemoglu, Johnson and Robinson, 2001), the interaction between log settler mortality and UK industrialization will be a good instrument for the interaction between institutions and UK industrialization. In this case, the IV procedure will deal both with the endogeneity and omitted variables biases. The IV estimates reported in column 6 and 7 are very close to the OLS estimates in columns 4 and 5, again suggesting that the endogeneity of $\bar{X}_i$ is not important.

In columns 8 and 9, we add the interaction between latitude and industrialization. This is useful because, if the reason why the U.S. surged ahead relative to Mexico and India during the 19th century were its geographic advantage, our measures of institutions might be proxying for this, incorrectly assigning the role of geography to institutions. The results give no support to this view: the estimates of ϕ are affected little and remain significant, while the interaction between industrialization and latitude is insignificant.[24]

[23]We could not instrument for $X_{it} \cdot UKIND_t$, since our potential instrument for X_{it} is not time varying.

[24]It is important to note that we do not have data from Singapore, Hong Kong and Malaysia, three

For completeness, column 10 reports results using European countries. This again leads to a positive and significant estimate of ϕ, which is somewhat smaller than the estimate from the former colonies sample. We conjecture that this is because there is less variation in institutions in the European sample. The second panel of Table 10 repeats these regressions using log GDP per capita as the left-hand side variable, and the results are similar.[25]

Overall, we interpret these results as providing support for the view that institutions played an important role in the surge in industrialization and income of the formerly poor areas.

7 Conclusion

Among countries colonized by European powers during the past 500 years those that were relatively rich in 1500 are now relatively poor. We document that both today and in the past there is a close association between urbanization rates and economic prosperity. Using both population density and urbanization as proxies for economic prosperity, we find a robust reversal in relative incomes among the former colonies between 1500 and today.

We argue that this reversal is inconsistent with the geography hypothesis, which explains the bulk of the income differences we observe across countries by geographic differences, and consequently predicts a high degree of persistence in economic outcomes. In contrast, it is quite consistent with the institutions hypothesis, combined with the fact that European colonialism caused an "institutional reversal", introducing worse institutions in previously prosperous areas. We also show that the timing and the nature of the reversal is not consistent with a more sophisticated geography hypothesis, the Northern drift, which argues that the center of gravity of economic activity has been shifting to the North (or away from the equator), and that contact with the Europeans may have led to such a process among the colonies.

Instead, the facts are consistent with the importance of institutions. First, the differences in institutions (and the institutional reversal caused by European colonialism)

former colonies countries that have prospered, most probably thanks to relatively secure property rights, and that are closer to the equator than the industrialized countries in our sample. We conjecture that with data from these countries, institutions would be even stronger relative to latitude.

[25] In the bottom panel, the interaction term is parameterized as $M_i \cdot \ln(UKIND_t)$ since the left-hand side variable is log of GDP per capita.

statistically account for reversal in relative incomes over the past 500 years. Second, we documented that societies with institutions of private property took better advantage of the opportunities to industrialize.

Given the quality of data on urbanization and population density more than 500 years ago, some degree of caution is required. Nevertheless, the broad patterns in the data seem uncontroversial: civilizations in Meso-America and India were undoubtedly much more urbanized and richer than those located in North America, Australia, New Zealand or the southern cone of Latin America. The intervention of Europe reversed this pattern. This is a first-order fact for understanding the process of economic and political development over the past 500 years. Our empirical work and reading of historical evidence suggest that the interaction between institutions brought by Europeans and industrialization was essential. Much more work is still required to understand this reversal and the associated development experiences of the former colonies fully.

8 References

Acemoglu, Daron, Simon Johnson and James A. Robinson (2001) "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, forthcoming.

Acemoglu, Daron and James A. Robinson (2001) "Economic Backwardness in Political Perspective," Unpublished.

Aghion, Philippe and Peter W. Howitt (1992) "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323-331.

Bairoch, Paul (1982) "International Industrialization Levels from 1750 to 1980," *Journal of Economic History*, 11, 269-333.

Bairoch, Paul (1988) *Cities and Economic Development: From the Dawn of History to the Present*, Chiacgo; University of Chicago Press.

Bairoch, Paul, Jean Batou and Pierre Chevre (1988) *La Population des villes Europeenees de 800 a 1850: Banque de Données et Analyse Sommaire des Résultats*, Centre d'histoire économique Internationale de l'Uni. de Genève, Libraire Droz.

Beckford, George L. (1972) *Persistent poverty; Underdevelopment in Plantation Economies of the Third World*, New York; Oxford University Press.

Boserup, Ester (1965) *The Conditions of Agricultural Growth; the Economics of Agrarian Change under Population Pressure*, Chicago; Aldine Pub. Co.

Brenner, Robert (1976) "Agrarian Class Structure and Economic Development in Preindustrial Europe" *Past and Present*, No. 70.

Cain, Philip J. and Anthony G. Hopkins (1993) *British Imperialism: Innovation and Expansion 1688-1914*, Longman, New York.

Chandler, Tertius (1987) *Four Thousand Years of Urban Growth: An Historical Census*, Lewiston, N.Y.; St. David's University Press.

Chandler, Tertius and Gerald Fox (1974) *Three Thousand Years of Urban Growth*, New York; Academic Press.

Chaudhuri, Kirti N. (1990) *Asia Before Europe: Economy and Civilization of the Indian Ocean from the Rise of Islam to 1750*, New York, Cambridge University Press.

Cipolla, Carlo M. (1974) *The Economic History of World Population*, 6th edition, Baltimore; Penguin.

Coatsworth, John H. (1982) "The Limits of Colonial Absolutism: Mexico in the 18th Century," in Karen Spalding ed. *Essays in the Political, Economic and Social History* *of Latin America*, University of Delaware Press, Newark DL.

Coatsworth, John H. (1993) "Notes on the Comparative Economic History of Latin America and the United States," in Walter L. Bernecker and Hans Werner Tobler eds. *Development and Underdevelopment in America: Contrasts in Economic Growth in North and Latin America in Historical Perpsective*, New York; Walter de Gruyter.

Coatsworth, John H. (1999) "Economic and Institutional Trajectories in 19th Century Latin America," in John H. Coatsworth and Alan M. Taylor eds. *Latin America and the World Economy since 1800*, Harvard University Press, Cambridge MA.

Cole, Jeffrey A. (1985) *The Potosi Mita, 1573-1700: Compulsory Indian Labor in the Andes*, Stanford University Press, Palo Alto.

Crosby, Alfred (1986) *Ecological Imperialism: The Biological Expansion of Europe 900-1900*, Cambridge University Press, New York NY.

Curtin, Philip D. (1990) *The Rise and Fall of the Plantation Complex: Essays in Atlantic History*, Cambridge University Press, New York NY.

De Long, Bradford J. and Andrei Shleifer (1993) "Princes and Merchants: European City Growth before the Industrial Revolution, *Journal of Law and Economics*, 36, October, 671-702.

de Vries, Jan (1976) *The Economy of Europe in an Age of Crisis, 1600-1750*, Cambridge University Press.

de Vries, Jan (1984) *European Urbanization, 1500-1800*, Boston; Harvard University Press.

Denoon, Donald (1983) *Settler Capitalism: The Dynamics of Dependent Development in the Southern Hemisphere*, Clarendon Press, Oxford UK.

Diamond, Jared M. (1997) *Guns, Germs and Steel: The Fate of Human Societies*, W.W. Norton & Co., New York NY.

Domar, Evsey (1970) "The Causes of Slavery and Serfdom: A Hypothesis," *Journal of Economic History*, 30, 18-32.

Duby, Georges (1974) *The Early Growth of the European Economy; Warriors and Peasants from the Seventh to the Twelfth Century*, Ithaca; Cornell University Press

Dunn, Richard S. (1972) *Sugar and Slaves: The Rise of the Planter Class in the English West Indies 1624-1713*, Chapel Hill; University of North Carolina Press.

Eggimann, Gilbert (1999) *La Population des villes des Tiers-Mondes, 1500-1950*, Centre d'histoire économique Internationale de l'Uni. de Genève, Libraire Droz.

Eltis, David (1995) "The Total Product of Barbados, 1664-1701," *Journal of Economic History*, 55, 321-336.

Engerman, Stanley L. and Kenneth L. Sokoloff (1997) "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies," in S.H. Haber ed. *How Latin America Fell Behind*, Stanford University Press, Stanford CA.

Engerman, Stanley L. and Kenneth L. Sokoloff (2000) "Institutions, Factor Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*, 3, 217-232.

Frank, Andre Gunder (1978) *Dependent Accumulation and Underdevelopment*, London; Macmillan.

Galbraith, James K. (1951) "Conditions for Economic Change in Underdeveloped Countries," *American Journal of Farm Economuics*, 33.

Galenson, David W. (1996) "The Settlement and Growth of the Colonies: Population, Labor and Economic Development," in Stanley L. Engerman and Robert E. Gallman eds. *The Cambridge Economic History of the United States,* Volume I, New York; Cambridge University Press

Gallman Robert E. (2000) "Economic Growth and Structural Change in the Long Nineteenth Century," in Stanley L. Engerman and Robert E. Gallman eds. *The Cambridge Economic History of the United States,* Volume II, New York; Cambridge University Press.

Genovese, Eugene D. (1965) *The Political Economy of Slavery: Studies in the Economy and Society of the Slave South*, New York; Pantheon Books.

Grossman, Gene M. and Elhanan Helpman (1991) *Innovation and Growth in the Global Economy*, Cambridge; MIT Press.

Halperin Donghi, Tulio (1993) *The Contemporary History of Latin America*, Durham; Duke University Press.

Harms, Robert C. (1981) *River of Wealth, River of Sorrow: The Central Zaire Basin in the era of the Slave and Ivory Trade, 1500-1891*, New Haven; Yale University Press.

Hayek, Freidrich von (1960) *The Constitution of Liberty*, Chicago; University of Chicago Press.

Hohenberg, Paul M. and Lynn Hollen Lees (1985) *The Making of Urban Europe, 1000-1950*, Cambridge MA; Harvard University Press.

Hopkins, Anthony G. (1973) *An Economic History of West Africa*, New York;

AddisonWesley Longman.

Hughes, Robert (1988) *The Fatal Shore*, New York, Vintage Books.

Huntington, Ellsworth (1915) *Civilization and Climate*, New Haven, Yale University Press.

Huntington, Ellsworth (1945) *Mainsprings of Civilization*, New York, J. Wiley and Sons.

Jacobs, Jane (1969) *The Economy of Cities*, New York; Random House.

Jones, Charles I. (1998) "Population and Ideas: A Theory of Endogenous Growth," http://www.stanford.edu/~chadj/pop400.pdf.

Jones, Eric L. (1981) *The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia*, New York; Cambridge University Press.

Kennedy, Paul M. (1987) *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*, New York; Random House.

Kremer, Michael (1993) "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics*, 1087, 681-716.

Kuznets, Simon (1968) *Modern Economic Growth: Rate Structure and Spread*, New Haven; Yale University Press.

Landes, David S. (1998) *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*, W.W. Norton & Co., New York.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer and Robert W. Vishny (1998). "Law and Finance," *Journal of Political Economy*, 106, 1113-55.

Law, Robin C.C. (1977) *The Oyo Empire 1600-1836: A West Africa Imperialism in the Era of the Atlantic Slave Trade*, New York; Oxford University Press.

Locke, John [1690] **(1980)** *Two Treatises of Government*, Indianapolis; Hackett.

Lockhart, James and Stuart B. Schwartz (1983) *Early Latin America*, Cambridge University Press, New York NY.

Machiavelli, Niccolò [1519] **(1987)** *Discourses on Livy*, New York; Oxford University Press.

Maddison, Angus (1995) *Monitoring the World Economy*, 1820-1992, Development Centre of the Organization for Economic Cooperation and Development.

Manning, Patrick (1988) *Francophone Sub-Saharan Africa, 1880-1995*, Cambridge University Press, New York NY.

Manning, Patrick (1990) *Slavery and African Life: Occidental, Oriental and African*

Slave Trades, New York; Cambridge University Press.

Marshall, Alfred [1890] (1949) *Principles of Economics*, London; Macmillan.

Marshall, P.J. (1998) "The British in Asia: Trade to Dominion, 1700-1765," in P.J. Marshall ed. *The Oxford History of the British Empire, Volume II The 18th Century*, Oxford University Press, New York.

McCusker, John J. and Russell R. Menard (1985) *The Economy of British America, 1607-1785*, Chapel Hill; University of North Carolina Press.

McEvedy, Colin (1967), *The Penguin Atlas of Ancient History*, Penguin Books, London.

McEvedy, Colin (1992), *The New Penguin Atlas of Medieval History*, Penguin Books, London.

McEvedy, Colin and Richard Jones (1975), *Atlas of World Population History*, Facts on File, New York.

McNeill, William H. (1976) *Plagues and Peoples*, Anchor Press, Garden City NJ.

Miller, Joseph C. (1988) *Way of Death: Merchant Capitalism and the Angolan Slave Trade, 1730-1830*, Madison; University of Wisconsin Press.

Mitchell, B.R. (1993) *International Historical Statistics, The Americas 1750-1988*, Second Edition, Stockton Press, New York.

Mitchell, B.R. (1995) *International Historical Statistics, Africa, Asia & Oceania 1750-1988*, Second Edition, Stockton Press, New York.

Montesquieu, Charles de Secondat [1748] (1989) *The Spirit of the Laws*, Cambridge University Press, New York NY.

Myrdal, Gunnar (1968) *Asian Drama; An Inquiry into the Poverty of Nations*, 3 Volumes, New York; Twentieth Century Fund.

North, Douglass C. (1990) *Institutions, Institutional change, and Economic Performance*, New York; Cambridge University Press.

North, Douglass C. and Robert P. Thomas (1973) *The Rise of the Western World: A New Economic History*, Cambridge University Press, Cambridge UK.

North, Douglass C. and Barry R. Weingast (1989) "Constitutions and Commitment: Evolution of Institutions Governing Public Choic in Seventeenth Century England, *Journal of Economic History*, 49, 803-832.

Notestein, Frank W. (1945) "Population: The Long View," in Theodore W. Schultze ed. *Food for the World*, Chicago; University of Chicago Press.

Patterson, Orlando (1982) *Slavery and Social Death: A Comparative Study*, Cambridge; Harvard University Press.

Pirenne, Henri (1956) [1925] *Medieval Cities; Their Origins and the Revival of Trade*, New York; Doubleday.

Pomeranz, Kenneth (2000) *The Great Divergence: Europe, China, and the Making of the Modern World Economy*, Princeton; Princeton University Press.

Postan, M.M. and E.E. Rich (1966) *The Cambridge Economic History of Europe: Volume 2, Trade and Industry in the Middle Ages*, Cambridge; Cambridge University Press.

Pritchett, Lant (1997) "Divergence, Big Time" *Journal of Economic Perspectives,* 11, 3-18.

Reid, Anthony (1988) *Southeast Asia in the Age of Commerce: Volume 1, The Lands Below the Winds*, New Haven, Yale University Press.

Reid, Anthony (1993) *Southeast Asia in the Age of Commerce: Volume 2, The Age of Crisis*, New Haven, Yale University Press.

Rodney, Walter (1972) *How Europe Underdeveloped Africa*, London; Bogle-L'Ouverture Publications

Romer, Paul M. (1986) "Increasing Returns and Long-Run Growth," *Journal of Political Economy*, 94, 1002-1037.

Romer, Paul M. (1990) "Endogenous Technical Change," *Journal of Political Economy*, 98, S71-S102.

Sachs, Jeffrey D. (2000) "Notes on a New Sociology of Economic Development" in Lawrence E. Harrison and Samuel P. Huntington eds. *Culture Matters: How Values Shape Human Progress*, New York; Basic Books.

Sachs, Jeffrey D. (2001) "Tropical Underdevelopment," NBER Working Paper #8119.

Showers, Victor (1979) *World Facts and Figures*, New York; Wiley.

Simon, Julian L. (1977) *The Economics of Population Growth*, Princeton; Princeton University Press.

Tilly, Charles (1990) *Coercion, Capital, and European States, AD 990-1990*, Cambridge MA; Basil Blackwell.

Tilly, Charles and Wim P. Blockmans eds. (1994) *Cities and the Rise of States in Europe, A.D. 1000 to 1800*, Boulder, Westview Press.

Toynbee, Arnold J. (1934-1961) *A Study of History*, 12 Volumes, Oxford; Oxford University Press.

Wallerstein Immanuel M. (1974-1980) *The Modern World-System*, 3 Volumes, New York, Academic Press.

Wiegersma, Nancy (1988) *Vietnam: Peasant Land, Peasant Revolution*, New York; St. Martin's Press.

White, Lynn, Jr. (1962) *Medieval Technology and Social Change*, London: Oxford University Press.

Wilks, Ivor (1975) *Asante in the 19th Century: The Structure and Evolution of a Political Order*, New York, Cambridge University Press.

Williams, Eric E. (1944) *Capitalism and Slavery*, Chapel Hill; University of North Carolina Press.

Williams, Eric E. (1970) *From Columbus to Castro: A History of the Caribbean*, London; Andre Deutsch.

Wong, R. Bin (1997) *China Transformed: Historical Change and the Limits of European Experience*, Ithaca; Cornell University Press.

Young, Crawford (1994) *The African Colonial State in Comparative Perspective*, Yale University Press, New Haven CT.

Table 1
Descriptive Statistics

	Whole World	Base Sample for population density	Base Sample for urbanization
Log GDP per capita (PPP) in 1995	8.3	7.8	8.5
	(1.1)	(1.0)	(0.9)
Log Population Density in 1500	1	0.59	0.29
	(1.6)	(1.5)	(1.9)
Urbanization in 1500	7.4	6.7	6.3
	(5.0)	(4.9)	(4.9)
Log Population Density in 1000	0.6	0.11	0.11
	(1.5)	(1.6)	(2.0)
Urbanization in 1000	4.7	4.1	4
	(4.4)	(4.0)	(4.0)
Urbanization in 1995	53	43.5	57.5
	(23.8)	(21.3)	(22.4)
European Settlements in 1900	29.6	12.3	23.2
	(41.7)	(22.9)	(28.7)
Constraint on the Executive in 1990	3.6	3.7	4.9
	(2.3)	(2.3)	(2.1)
Constraint on the Executive in 1900	1.9	2	2.7
	(1.8)	(1.9)	(2.3)
Constraint on the Executive in	3.6	3.4	3.3
First Year of Independence	(2.4)	(2.3)	(2.5)
Number of Observations	162	80	41

Standard deviations are in parentheses. For detailed sources and descriptions see Appendix Table 1.

Table 2
Urbanization and Per Capita Income

	Cross-sectional regression in 1900, all countries	Cross-sectional regression in 1950, all countries	Cross-sectional regression in 1995, all countries	Cross-sectional regression in 1995, only for excolonies	Cross-sectional regression in 1995, never colonized countries only	Cross-sectional regression in 1995, all countries, with continent dummies
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Dependent Variable is Log GDP per capita						
Urbanization	0.038	0.026	0.036	0.037	0.033	0.030
	(0.006)	(0.002)	(0.002)	(0.003)	(0.007)	(0.002)
R-Squared	0.69	0.57	0.63	0.69	0.34	0.68
Number of Observations	22	128	162	93	51	162

	Using panel data set through 1900	Using panel data set through 1900	Using panel data set through 1900	Using panel data set through 1900
	(7)	(8)	(9)	(10)
Panel B: Dependent Variable is Log GDP per capita				
Urbanization	0.034	0.030	0.041	0.026
	(0.004)	(0.004)	(0.004)	(0.004)
Country Dummies	No	No	Yes	Yes
Period Dummies	No	Yes	No	Yes
R-Squared	0.52	0.71	0.87	0.96
Number of Observations	62	62	62	62

GDP per capita through 1900 is from Bairoch (1978), with the addition of Gregory King for England before 1700. Urbanization is percent of population living in towns with at least 5,000 people, from Bairoch (1988) through 1900 (see appendix for details). GDP per capita in 1950 is from Maddison (1995); this regression uses urbanization in 1960 from the World Bank's World Development Indicators 1999. GDP per capita and Urbanization data for 1995 are from the World Bank's World Development Indicators 1999. Population density is total population divided by land area usable for agriculture, both from McEvedy and Jones (1978). For detailed sources and descriptions see Appendix Table 1.

Table 3
Urbanization, Population Density and Institutions

	Dependent Variable is: Average Protection Against Expropriation Risk, 1985-95			Constraint on Executive in 1990			Constraint on Executive in 1900			Constraint on Executive in First Year of Independence		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Without Additional Controls												
Urbanization in 1500	-0.11		-0.01	-0.16		-0.04	-0.12		0.03	-0.14		0.01
	(0.04)		(0.06)	(0.07)		(0.10)	(0.07)		(0.10)	(0.07)		(0.11)
Log Population Density in 1500		-0.37	-0.36		-0.49	-0.40		-0.35	-0.52		-0.33	-0.53
		(0.10)	(0.15)		(0.15)	(0.25)		(0.12)	(0.26)		(0.15)	(0.28)
R-Squared	0.14	0.16	0.25	0.12	0.12	0.18	0.06	0.09	0.15	0.31	0.16	0.37
Number of Observations	42	75	42	41	84	41	43	88	43	42	85	42
Panel B: Controlling for Latitude												
Urbanization in 1500	-0.10		-0.003	-0.16		-0.04	-0.10		0.04	-0.13		0.02
	(0.04)		(0.06)	(0.07)		(0.10)	(0.07)		(0.10)	(0.07)		(0.11)
Log Population Density in 1500		-0.31	-0.34		-0.45	-0.42		-0.27	-0.48		-0.30	-0.53
		(0.10)	(0.15)		(0.16)	(0.25)		(0.12)	(0.25)		(0.16)	(0.29)
Latitude	2.85	3.53	2.57	-1.51	2.63	-1.87	5.82	4.83	5.45	1.49	2.68	1.47
	(1.48)	(1.25)	(1.41)	(2.39)	(2.01)	(2.34)	(2.37)	(1.54)	(2.30)	(2.53)	(1.17)	(2.46)
R-Squared	0.22	0.24	0.31	0.13	0.13	0.19	0.19	0.19	0.26	0.32	0.17	0.38
Number of Observations	42	75	42	41	84	41	43	87	43	42	84	42

Regressions use data on all former colonies for which information on urbanization and population density in 1500 is available, as explained in the text. Urbanization in 1500 is percent of the population living in towns with 5,000 or more people, from Bairoch (1988), with supplementary sources as described in the appendix. Population density in 1500 is total population divided by land area usable for agriculture, from McEvedy and Jones (1978). Regressions with constraint on executive in first year of independence also include time since independence as a regressor. For detailed sources and descriptions see Appendix Table 1.

Table 4

Urbanization in 1500 and GDP per capita in 1995 for former European colonies

	Dependent Variable is log GDP per capita (PPP basis) in 1995								
	base sample	without North Africa	without the Americas	with continent dummies	without neo-Europes	controlling for latitude	controlling for geography	controlling for colonial origin	controlling for religion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Urbanization in 1500	-0.08	-0.11	-0.12	-0.09	-0.05	-0.08	-0.09	-0.07	-0.07
	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
Asia Dummy				-1.30					
				(0.60)					
Africa Dummy				-0.50					
				(0.77)					
America Dummy				-0.94					
				(0.57)					
Former French Colony								-0.58	
								(0.39)	
Former Spanish Colony								0.07	
								(0.29)	
P-Value for Temperature							[0.24]		
P-Value for Humidity							[0.64]		
P-Value for Soil Quality							[0.96]		
P-Value for Natural Resources							[0.93]		
Latitude						1.4			
						(0.91)			
P-value for Religion									[0.49]
R-Squared	0.20	0.24	0.26	0.34	0.09	0.25	0.62	0.27	0.26
Number of Observations	41	37	17	41	37	41	41	41	41

Dependent Variable: Log GDP per capita (PPP basis) in 1995, current prices (from the World Bank's World Development Indicators 1999). Base sample is all former colonies for which we have data and excludes countries in Africa south of the Sahara. Urbanization in 1500 is percent of the population living in towns with 5,000 or more people, from Bairoch (1988), with supplementary sources as described in the appendix. The neo-Europes are the USA, Canada, Australia and New Zealand. The regression that controls for geography also includes a dummy for whether or not the country is landlocked; this is not significant. The regression that controls for colonial origin (column 8) includes dummies for former French colony, Spanish colony, Portuguese colony, Belgian colony, Italian colony, German colony, and Dutch colony. British colonies are the base case. The religion variables are percent of the population who are Muslim, Catholic, and "other"; percent Protestant is omitted. For detailed sources and descriptions see Appendix Table 1.

Table 5

Alternative Measures of Urbanization

Dependent Variable is log GDP per capita (PPP basis) in 1995

	base sample	without North Africa	without the Americas	with continent dummies	without neo-Europes	controlling for latitude	controlling for geography	controlling for colonial origin	controlling for religion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Using the measure in Table 3									
Urbanization in 1500	-0.08	-0.11	-0.12	-0.09	-0.05	-0.08	-0.09	-0.07	-0.07
	(0.03)	(0.03)	(0.05)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)
R-Squared	0.20	0.24	0.26	0.34	0.09	0.25	0.62	0.27	0.26
Number of Observations	41	37	17	41	37	41	41	41	41
Panel B: Using only Bairoch's estimates									
Urbanization in 1500	-0.14	-0.14	-0.25	-0.10	-0.10	-0.14	-0.10	-0.13	-0.08
	(0.03)	(0.03)	(0.07)	(0.03)	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)
R-Squared	0.46	0.46	0.59	0.59	0.30	0.46	0.89	0.54	0.67
Number of Observations	30	30	10	30	27	30	30	30	30
Panel C: Using only Eggimann's estimates									
Urbanization in 1500	-0.04	-0.04	-0.10	-0.04	-0.02	-0.04	-0.05	-0.04	-0.03
	(0.02)	(0.02)	(0.05)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
R-Squared	0.11	0.09	0.24	0.29	0.05	0.17	0.62	0.24	0.25
Number of Observations	40	37	16	40	36	40	40	40	40
Panel D: Using only Chandler's estimates									
Urbanization in 1500	-0.06	-0.06	-0.06	-0.07	-0.04	-0.05	-0.04	-0.05	-0.06
	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)
R-Squared	0.25	0.23	0.17	0.42	0.17	0.32	0.84	0.44	0.57
Number of Observations	26	22	12	26	23	26	26	26	26

Dependent Variable: Log GDP per capita (PPP basis) in 1995, current prices (from the World Bank's World Development Indicators 1999). Base sample is all former colonies for which we have data and excludes countries in Africa south of the Sahara. Urbanization in 1500 is from alternative sources in the four panels, as described in the text and the appendix. The neo-Europes are the USA, Canada, Australia and New Zealand. The regression that controls for geography also includes a dummy for whether or not the country is landlocked; this is not significant. The religion variables are percent of the population who are Muslim, Catholic, and "other"; percent Protestant is omitted. The regression that controls for colonial origin (column 8) includes dummies for former French colony, Spanish colony, Portuguese colony, Belgian colony, Italian colony, German colony, and Dutch colony; British colonies are the base case. For detailed sources and descriptions see Appendix Table 1.

Table 6
Population Density and GDP per capita in former European colonies (including sub-Saharan Africa)

	Dependent Variable is log GDP per capita (PPP basis) in 1995								
	base sample	without Africa	without the Americas	with continent dummies	without neo-Europes	controlling for latitude	controlling for geography	controlling for colonial origin	controlling for religion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Log population density in 1500 as independent variable									
Log Population Density in 1500	-0.38	-0.40	-0.32	-0.26	-0.31	-0.34	-0.32	-0.32	-0.37
	(0.06)	(0.05)	(0.07)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)
Asia Dummy				-0.91					
				(0.55)					
Africa Dummy				-1.7					
				(0.52)					
America Dummy				-0.7					
				(0.52)					
Former French Colony								-0.48	
								(0.21)	
Former Spanish Colony								0.26	
								(0.23)	
P-Value for Temperature							[0.32]		
P-Value for Humidity							[0.05]		
P-Value for Soil Quality							[0.33]		
P-Value for Natural Resources							[0.76]		
Latitude						2.1			
						(0.75)			
P-Value for Religion									[0.74]
R-Squared	0.33	0.55	0.27	0.55	0.22	0.39	0.61	0.47	0.35
Number of Observations	90	46	58	90	86	90	84	90	84
Panel B: Log population and log land in 1500 as separate independent variables									
Log Population in 1500	-0.40	-0.41	-0.32	-0.28	-0.31	-0.35	-0.33	-0.34	-0.39
	(0.06)	(0.05)	(0.07)	(0.05)	(0.06)	(0.05)	(0.07)	(0.06)	(0.07)
Log Arable Land in 1500	0.31	0.37	0.21	0.23	0.20	0.26	0.24	0.27	0.31
	(0.06)	(0.06)	(0.09)	(0.06)	(0.07)	(0.06)	(0.09)	(0.06)	(0.08)
R-Squared	0.37	0.57	0.31	0.57	0.31	0.44	0.62	0.50	0.38
Number of Observations	90	46	58	90	86	90	84	90	84
Panel C: Using population density in 1000 AD as an instrument for population density in 1500 AD									
Log Population Density in 1500	-0.31	-0.4	-0.15	-0.18	-0.22	-0.27	-0.24	-0.26	-0.25
	(0.06)	(0.06)	(0.08)	(0.07)	(0.08)	(0.06)	(0.09)	(0.06)	(0.08)
Number of Observations	83	43	51	83	80	83	78	83	77

Dependent Variable: Log GDP per capita (PPP basis) in 1995, current prices (from the World Bank's World Development Indicators 1999). Population density in 1500 is total population divided by land area usable for agriculture, from McEvedy and Jones (1978). The neo-Europes are the USA, Canada, Australia and New Zealand. The regression that controls for geography also includes a dummy for whether or not the country is landlocked; this is not significant. The regression that controls for colonial origin (column 8) includes dummies for French colony, Spanish colony, Portuguese colony, Belgian colony, Italian colony, German colony, and Dutch colony; British colonies are the base case. The religion variables are percent of the population who are Muslim, Catholic, and "other"; percent Protestant is omitted. For detailed sources and descriptions see Appendix Table 1.

Table 7
Robustness Checks for Urbanization and Log Population Density

	Dependent Variable is log GDP per capita (PPP basis) in 1995								
	Base Sample	Assuming lower urbanization in the Americas (9% in the Andes and Central America)	Assuming lower urbanization in North Africa (10%)	Assuming lower urbanization in Indian subcontinent (6%)	Using least favorable combination of assumptions	Using basic Toynbee definition of civilization (excolonies only)	Using augmented Toynbee definition of civilization (excolonies only)	Using land area in 1995	Alternative assumptions for log population density (halve estimates for Africa)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel A: Unweighted Regressions								
Urbanization in 1500	-0.08	-0.09	-0.10	-0.07	-0.11	-0.10	-0.12		
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.06)	(0.05)		
Log Population Density in 1500								-0.42	-0.34
								(0.06)	(0.07)
R-Squared	0.20	0.22	0.24	0.16	0.21	0.45	0.30	0.38	0.24
Number of Observations	41	41	41	41	41	7	14	82	82
	Panel B: Regressions weighted using log population in 1500								
Urbanization in 1500	-0.07	-0.08	-0.10	-0.06	-0.10	-0.09	-0.12		
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.07)	(0.05)		
Log Population Density in 1500								-0.39	-0.30
								(0.06)	(0.07)
R-Squared	0.19	0.21	0.23	0.13	0.19	0.27	0.29	0.34	0.20
Number of Observations	40	40	40	40	40	7	14	82	82
	Panel C: Including both urbanization and log population density as independent variables								
Urbanization in 1500	0.04	0.04	0.02	0.04	0.02	0.07	0.07	0.01	0.01
	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.05)	(0.05)	(0.02)	(0.02)
Log Population Density in 1500	-0.41	-0.41	-0.37	-0.40	-0.37	-0.43	-0.48	-0.43	-0.41
	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.11)	(0.10)	(0.07)	(0.07)
R-Squared	0.55	0.55	0.54	0.56	0.54	0.85	0.79	0.61	0.60
Number of Observations	40	40	40	40	40	7	14	40	40

Dependent Variable: Log GDP per capita (PPP basis) in 1995, current prices (from the World Bank's World Development Indicators 1999). In our base sample, urbanization in 1500 is percent of the population living in towns with 5,000 or more people, from Bairoch (1988), with supplementary sources as described in the appendix. Alternative assumptions are as described at the top of each column. Population density in 1500 is total population divided by land area usable for agriculture. For detailed sources and descriptions see Appendix Table 1.

Table 8
Persistence of Urbanization and Population Density

	Base Sample	Base Sample (without Africa)	Base Sample (without America)	Base Sample with continent dummies	Using augmented Toynbee definition of civilization (excolonies only)	Just Europe	All countries never colonized by Western Europe	All countries never colonized by Western Europe, without the Russian Empire	All countries never colonized by Western Europe, with continent dummies
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				Panel A: Dependent Variable is Urbanization in 1995					
Urbanization in 1500	-1.80	-2.70	-1.80	-2.21	-2.70	1.30	1.10	1.30	1.20
	(0.70)	(0.87)	(1.40)	(0.76)	(2.20)	(0.40)	(0.47)	(0.54)	(0.47)
R-Squared	0.14	0.20	0.09	0.39	0.13	0.23	0.11	0.14	0.14
Number of Observations	44	39	20	44	12	34	46	36	46
				Panel B: Dependent Variable is Urbanization in 1500					
Urbanization in 1000	0.94	0.86	1.50	0.87	0.82	0.46	0.42	0.46	0.45
	(0.16)	(0.14)	(0.41)	(0.13)	(0.16)	(0.21)	(0.17)	(0.18)	(0.19)
R-Squared	0.50	0.58	0.53	0.74	0.73	0.14	0.13	0.17	0.14
Number of Observations	35	31	14	35	12	31	42	35	42
				Panel C: Dependent Variable is Log Population Density in 1500					
Log Population Density in 1000	0.91	0.97	0.86	0.83	0.82	1.20	1.10	1.13	1.03
	(0.04)	(0.05)	(0.04)	(0.04)	(0.16)	(0.10)	(0.06)	(0.09)	(0.06)
R-Squared	0.89	0.92	0.90	0.92	0.70	0.83	0.85	0.79	0.85
Number of Observations	80	37	55	80	12	33	60	44	59
				Panel D: Dependent Variable is Log Population Density in 1000					
Log Population Density in 1 AD	0.94	1.10	0.86	0.89	0.98	0.82	0.82	0.65	0.83
	(0.03)	(0.02)	(0.04)	(0.03)	(0.23)	(0.04)	(0.06)	(0.10)	(0.06)
R-Squared	0.94	0.99	0.92	0.95	0.67	0.92	0.79	0.54	0.80
Number of Observations	70	36	46	70	11	32	55	39	54

Urbanization is the percent of the population living in cities with more than 5,000 inhabitants (sources are described in the appendix.) Population density in 1AD, 1000 and 1500 is total population divided by land area usable for agriculture. Column 8 includes a dummy for Asia; Europe is the base case. For detailed sources and descriptions see Appendix Table 1.

Table 9
GDP per capita and Institutions

	Dependent Variable is log GDP per capita (PPP basis) in 1995							
	Average Protection Against Expropriation Risk, 1985-95		Constraint on Executive in 1990		Constraint on Executive in 1900		Constraint on Executive in First Year of Independence	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Second Stage Regressions							
Institutions	0.52	0.88	0.83	0.50	0.45	0.51	0.37	0.46
	(0.10)	(0.21)	(0.47)	(0.11)	(0.16)	(0.14)	(0.12)	(0.16)
Urbanization in 1500	-0.03		0.03		-0.02		-0.02	
	(0.02)		(0.08)		(0.04)		(0.03)	
Log Population Density in 1500		-0.08		-0.10		-0.16		-0.13
		(0.10)		(0.10)		(0.10)		(0.10)
	Panel B: First Stage Regressions							
Log Settler Mortality	-1.20	-0.47	-0.75	-0.88	-1.30	-0.81	-1.80	-0.78
	(0.23)	(0.14)	(0.44)	(0.20)	(0.46)	(0.19)	(0.40)	(0.25)
Urbanization in 1500	-0.05		-0.09		-0.06		-0.05	
	(0.04)		(0.07)		(0.07)		(0.06)	
Log Population Density in 1500		-0.21		-0.35		-0.20		-0.24
		(0.11)		(0.15)		(0.14)		(0.17)
R-Squared	0.53	0.29	0.12	0.37	0.25	0.31	0.56	0.23
Number of Observations	38	64	37	67	39	70	38	67
	Panel C: Coefficient on Institutions without Urbanization or Population Density in 1500							
Institutions	0.56	0.96	0.77	0.54	0.48	0.61	0.39	0.52
	(0.09)	(0.17)	(0.33)	(0.09)	(0.14)	(0.13)	(0.11)	(0.15)

Dependent Variable: Log GDP per capita (PPP basis) in 1995, current prices, (from the World Bank's World Development Indicators 1999). The measure of institutions used in each regression is indicated at the head of each column. Regressions with constraint on the executive in the first year of independence also include time since independence as a regressor. Urbanization in 1500 is percent of the population living in towns with 5,000 or more people, from Bairoch (1988), with supplementary sources as described in the appendix. Population density is calculated as total population divided by land area usable for agriculture (both from McEvedy and Jones 1978). Constraint on the executive in 1990, 1900 and the first year of independence are all from the Polity III dataset. For detailed sources and descriptions see Appendix Table 1.

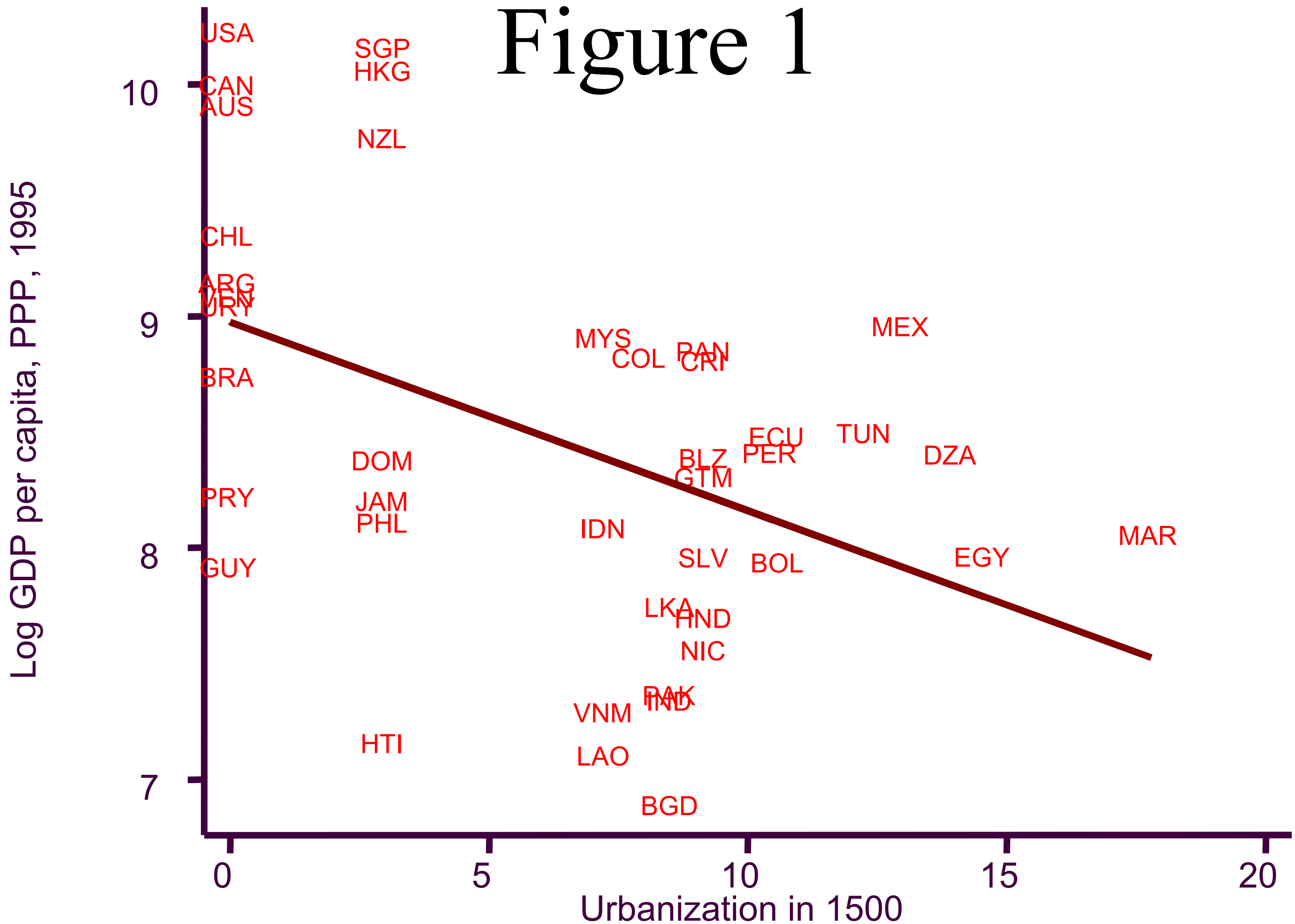
Table 10

The Interaction of UK Industrialization and Local Institutions

	Former Colonies, using only pre-1950 data	Former Colonies, using only pre-1950 data	Former Colonies, using data through 1980 (all data)	Former Colonies, using average institutions for each country, using only pre-1950 data	Former Colonies, using average institutions for each country, using all data	Former Colonies, using average institutions for each country, instrumenting using settler mortality, only pre-1950 data	Former Colonies, using average institutions for each country, instrumenting using settler mortality, all data	Former Colonies, using average institutions for each country, instrumenting using settler mortality, only pre-1950 data	Former Colonies, using average institutions for each country, instrumenting using settler mortality, all data	All European countries, using data through 1950
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Dependent Variable is Industrial Production Per Capita										
UK Industrialization*Institutions	0.13	0.08	0.13	0.21	0.28	0.16	0.25	0.13	0.20	0.06
	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.06)	(0.05)	(0.01)
Institutions	8.97	9.71	-3.36							4.34
	(2.30)	(2.58)	(4.46)							(0.96)
UK Industrialization*Latitude								0.30	0.52	
								(0.49)	(0.36)	
R-Squared	0.75	0.44	0.74	0.87	0.94	0.85	0.94	0.83	0.93	0.78
Number of Observations	59	59	75	59	75	59	75	59	75	138
Panel B: Dependent Variable is Log GDP Per Capita										
Log UK Industrialization*Institutions	0.08	0.10	0.06	0.13	0.12	0.15	0.10	0.09	0.07	0.02
	(0.02)	(0.05)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.05)	(0.05)	(0.01)
Institutions	-0.03	0.16	-0.08							-0.01
	(0.03)	(0.04)	(0.03)							(0.01)
Log UK Industrialization*Latitude								0.59	0.27	
								(0.40)	(0.38)	
R-Squared	0.95	0.59	0.92	0.96	0.93	0.96	0.93	0.96	0.93	0.95
Number of Observations	79	79	131	79	131	79	131	79	131	103
Country Dummies	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel regressions; country and period dummies are included as indicated at the foot of each column. Dependent variable in Panel A is industrial output per capita, from Bairoch (1982), with up to 10 observations per country, 1750-1980. Dependent variable in Panel B is log GDP per capita, from Maddison (1995). Independent variable is the UK industrial output per capita (from Bairoch 1982) interacted with constraint on the executive from the Polity IV dataset. The main effect of institutions is evaluated at the mean value of UK idustrialization. Polity IV provides information only for independent countries; if a country was a colony at a particular date, we assign the lowest value of constraint on the executive. For detailed sources and descriptions see Appendix Table 1.

Figure 1



Log GDP per capita, PPP, 1995
Urbanization in 1500
10
9
8
7
0
5
10
15
20
USA
SGP
HKG
CAN
AUS
NZL
CHL
ARG
VEN
URY
BRA
PRY
GUY
DOM
JAM
PHL
HTI
MYS
COL
PAN
CRI
MEX
ECU
TUN
BLZ
PER
DZA
GTM
IDN
SLV
BOL
EGY
MAR
LKA
HND
NIC
PAK
IND
VNM
LAO
BGD

Figure 2

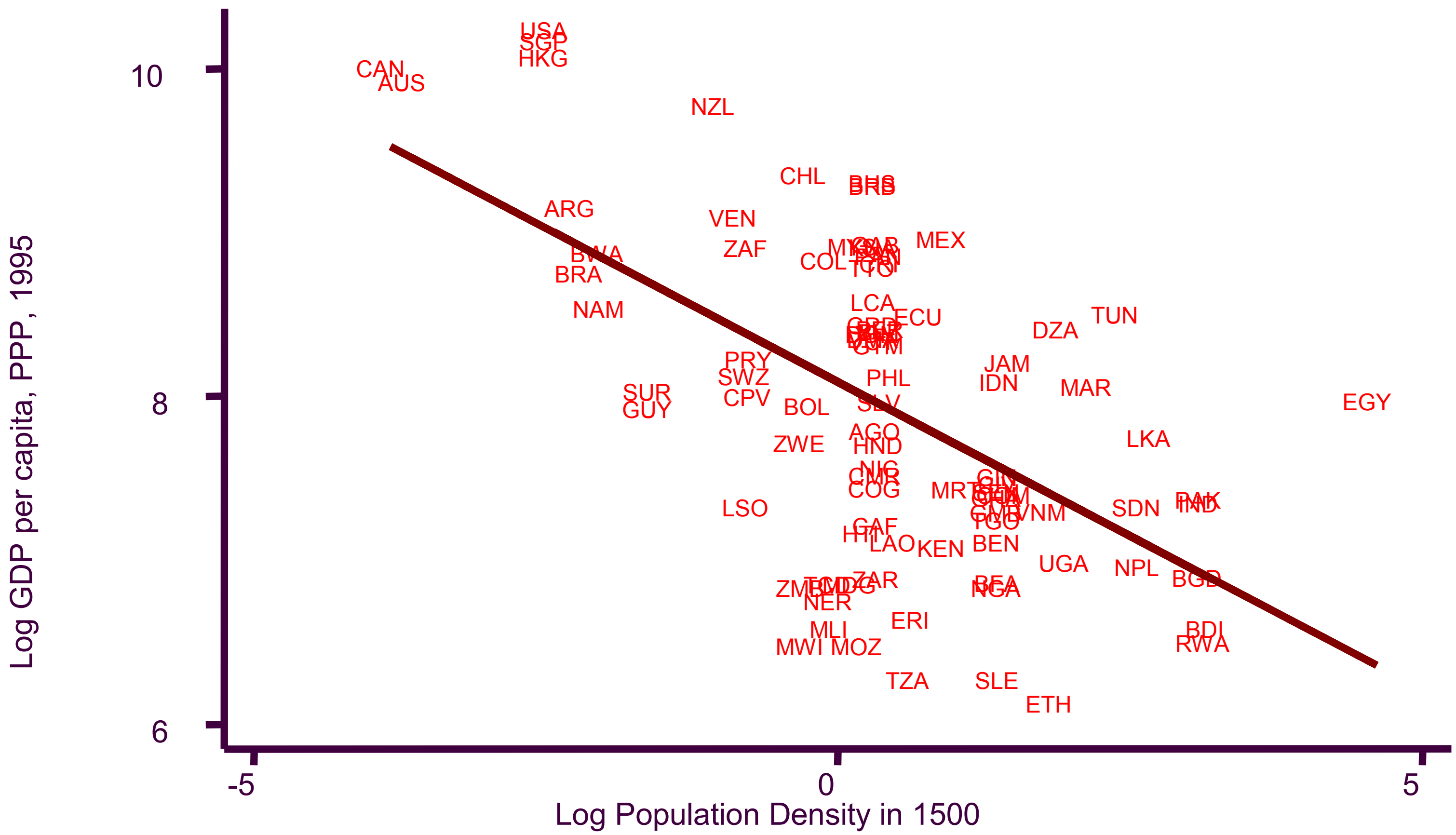

Log GDP per capita, PPP, 1995
10
8
6
-5
0
5
Log Population Density in 1500
USA
HKG
CAN
AUS
NZL
CHL
ARG
VEN
ZAF
MEX
COL
BWA
BRA
NAM
LCA
ECU
DZA
TUN
PRY
SWZ
CPV
SUR
GUY
BOL
PHL
JAM
IDN
MAR
EGY
ZWE
AGO
HND
LKA
COG
LSO
SDN
CAF
LAO
KEN
BEN
UGA
NPL
NER
ERI
MLI
MWI
MOZ
RWA
TZA
SLE
ETH

Figure 3

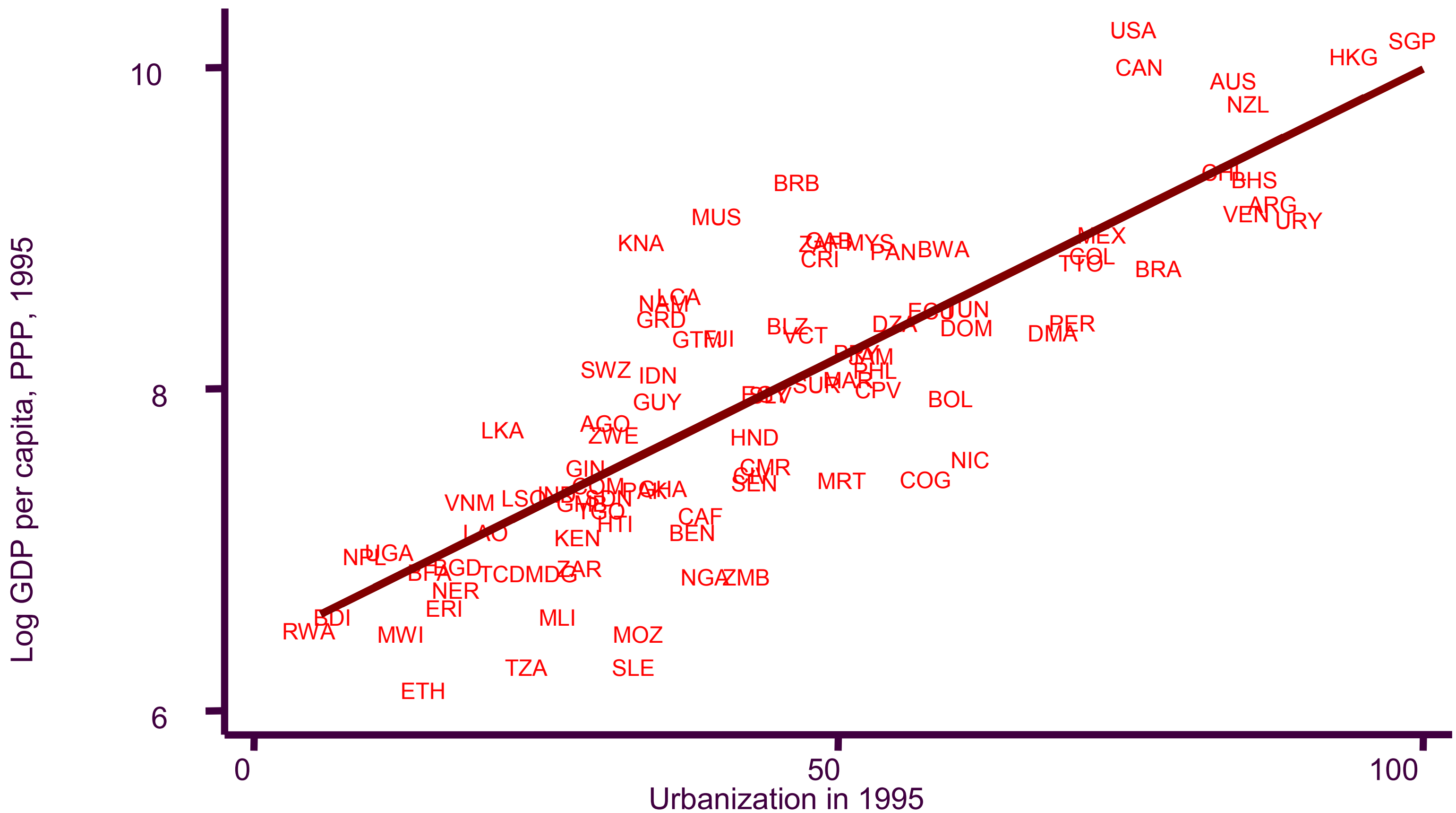

Log GDP per capita, PPP, 1995
Urbanization in 1995
10
8
6
0
50
100

Figure 4A


Residuals from Regression of Constraint on Executive in First Year
of Independence on Years Since Independence
5
0
-5
-10
0
10
Residuals from Regression of Urbanization in 1500 on Years Since Independence
USA
NZL
CAN
AUS
GUY
PNG
MYS
BGD
LAO
COL
BOL
ECU
MEX
DOM
PAN
PHL
LBY
HTI
PER
EGY
MAR
DZA
PAK
TUN

Figure 4B

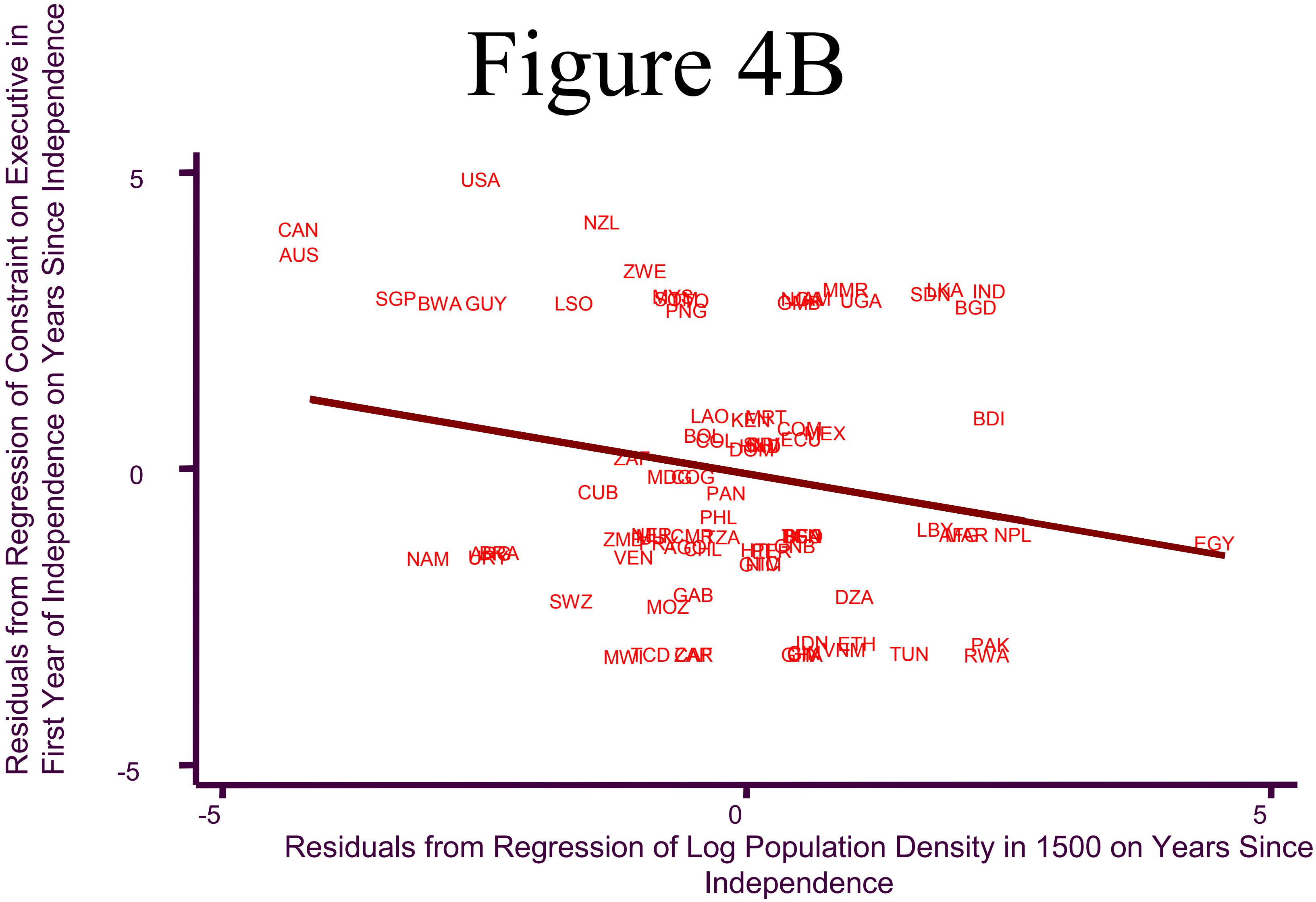
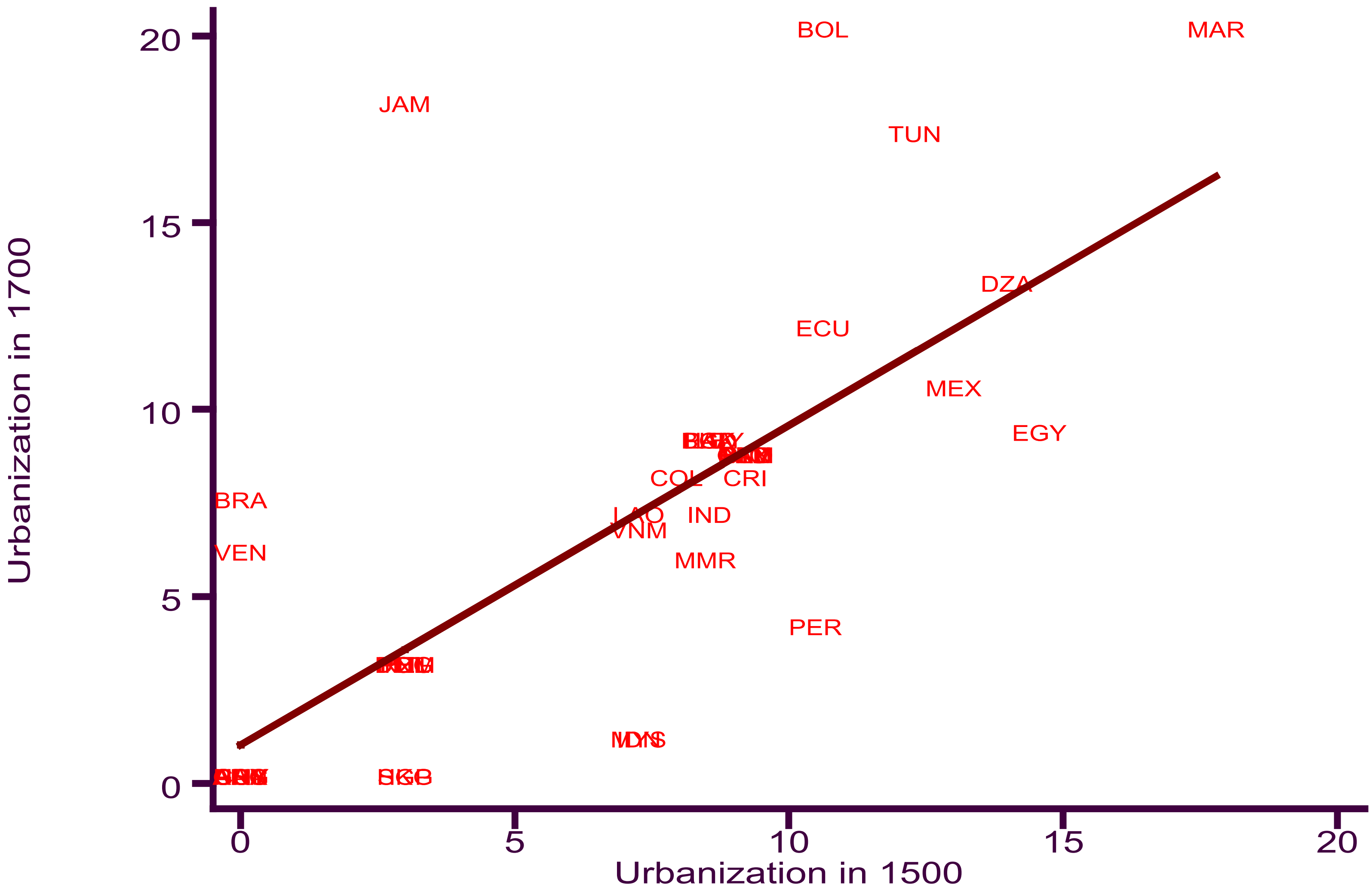
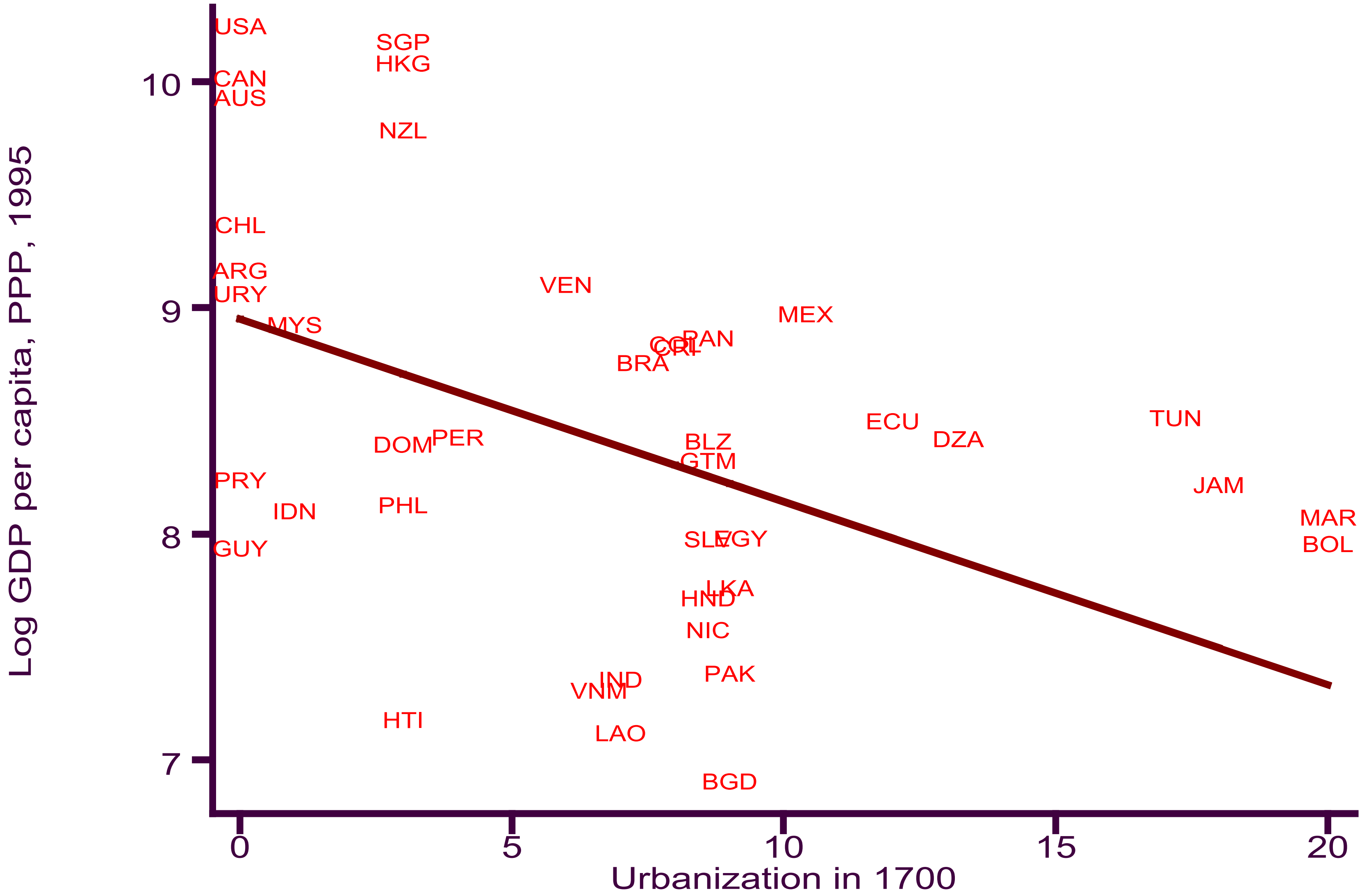

Residuals from Regression of Constraint on Executive in First Year of Independence on Years Since Independence
Residuals from Regression of Log Population Density in 1500 on Years Since Independence
5
0
-5
-5
0
5

Figure 5A



Urbanization in 1700
Urbanization in 1500
0
5
10
15
20
0
5
10
15
20
BOL
MAR
JAM
TUN
DZA
ECU
MEX
EGY
COL
CRI
BRA
LAO
IND
VNM
VEN
MMR
PER

Figure 5B



Log GDP per capita, PPP, 1995
Urbanization in 1700
10
9
8
7
0
5
10
15
20
USA
SGP
HKG
CAN
AUS
NZL
CHL
ARG
URY
VEN
MYS
MEX
COL
CRI
PAN
BRA
ECU
TUN
DZA
DOM
PER
BLZ
GTM
PRY
JAM
IDN
PHL
MAR
BOL
GUY
SLE
EGY
LKA
HND
NIC
PAK
IND
VNM
HTI
LAO
BGD

Figure 6A

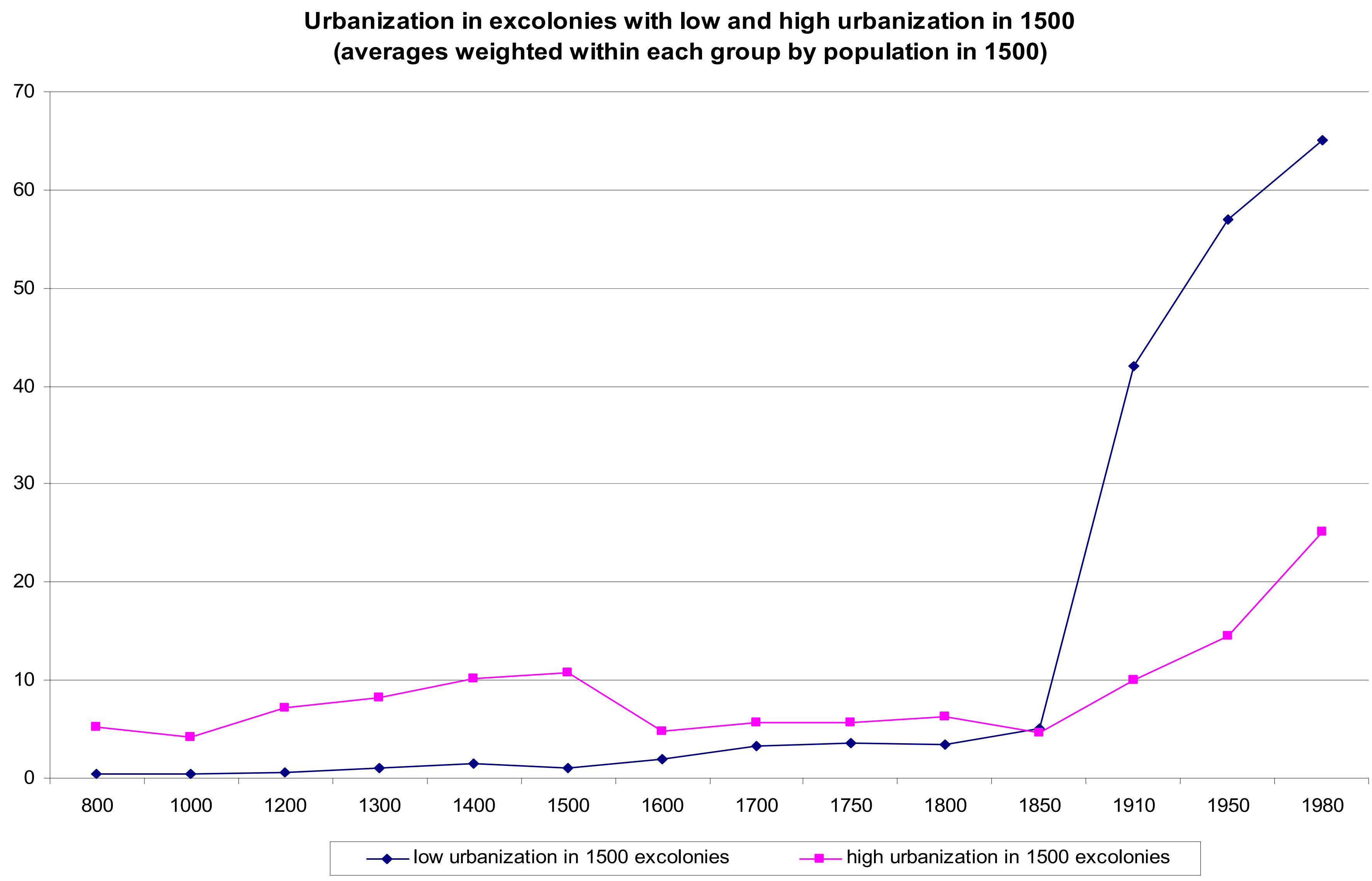
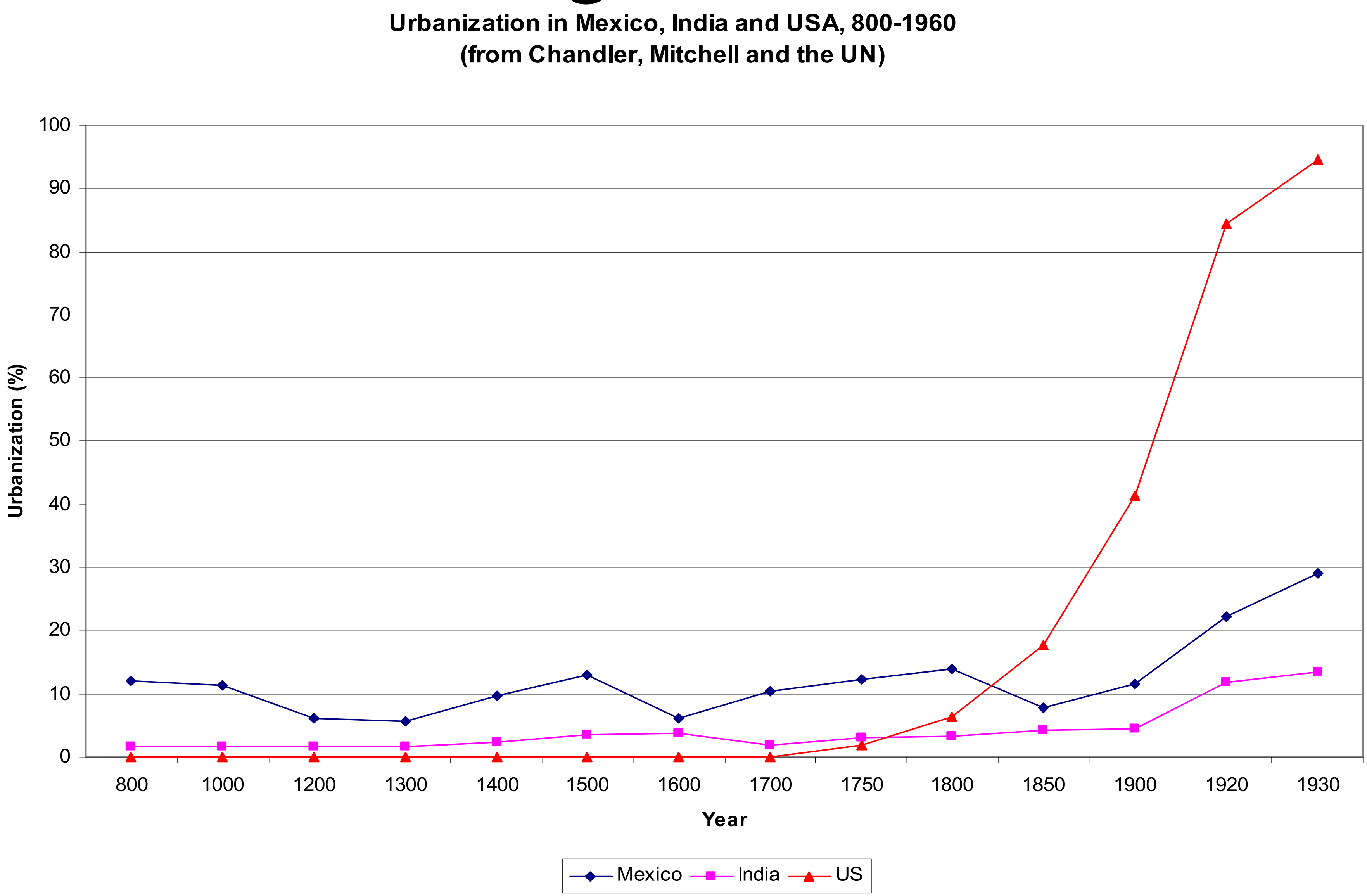

Urbanization in excolonies with low and high urbanization in 1500
(averages weighted within each group by population in 1500)
70
60
50
40
30
20
10
0
800
1000
1200
1300
1400
1500
1600
1700
1750
1800
1850
1910
1950
1980
low urbanization in 1500 excolonies
high urbanization in 1500 excolonies

Figure 6B



Urbanization in Mexico, India and USA, 800-1960
(from Chandler, Mitchell and the UN)
Urbanization (%)
100
90
80
70
60
50
40
30
20
10
0
800
1000
1200
1300
1400
1500
1600
1700
1750
1800
1850
1900
1920
1930
Year
Mexico
India
US

Figure 6C

Just the Americas (1750-1930)
(Mitchell Series)

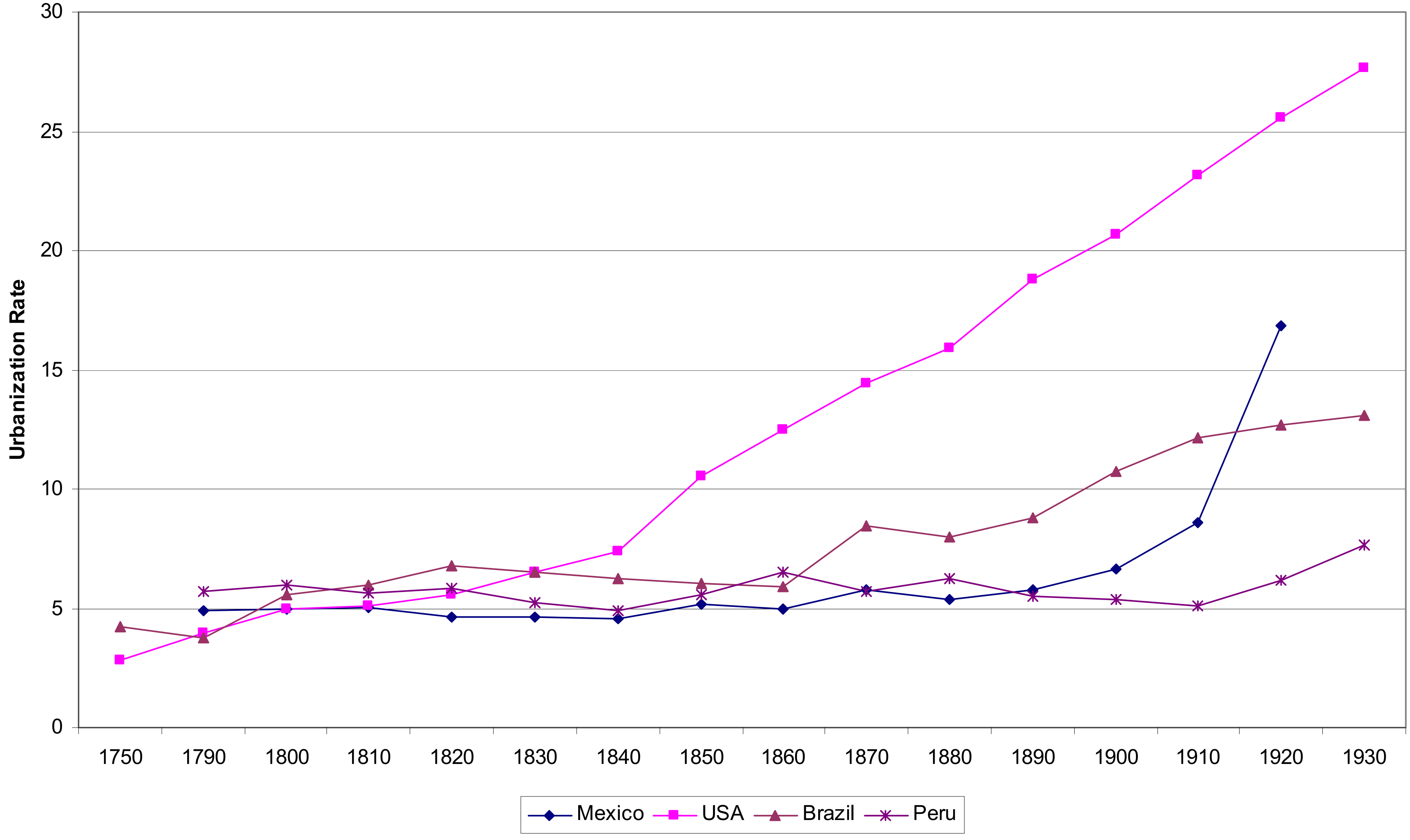
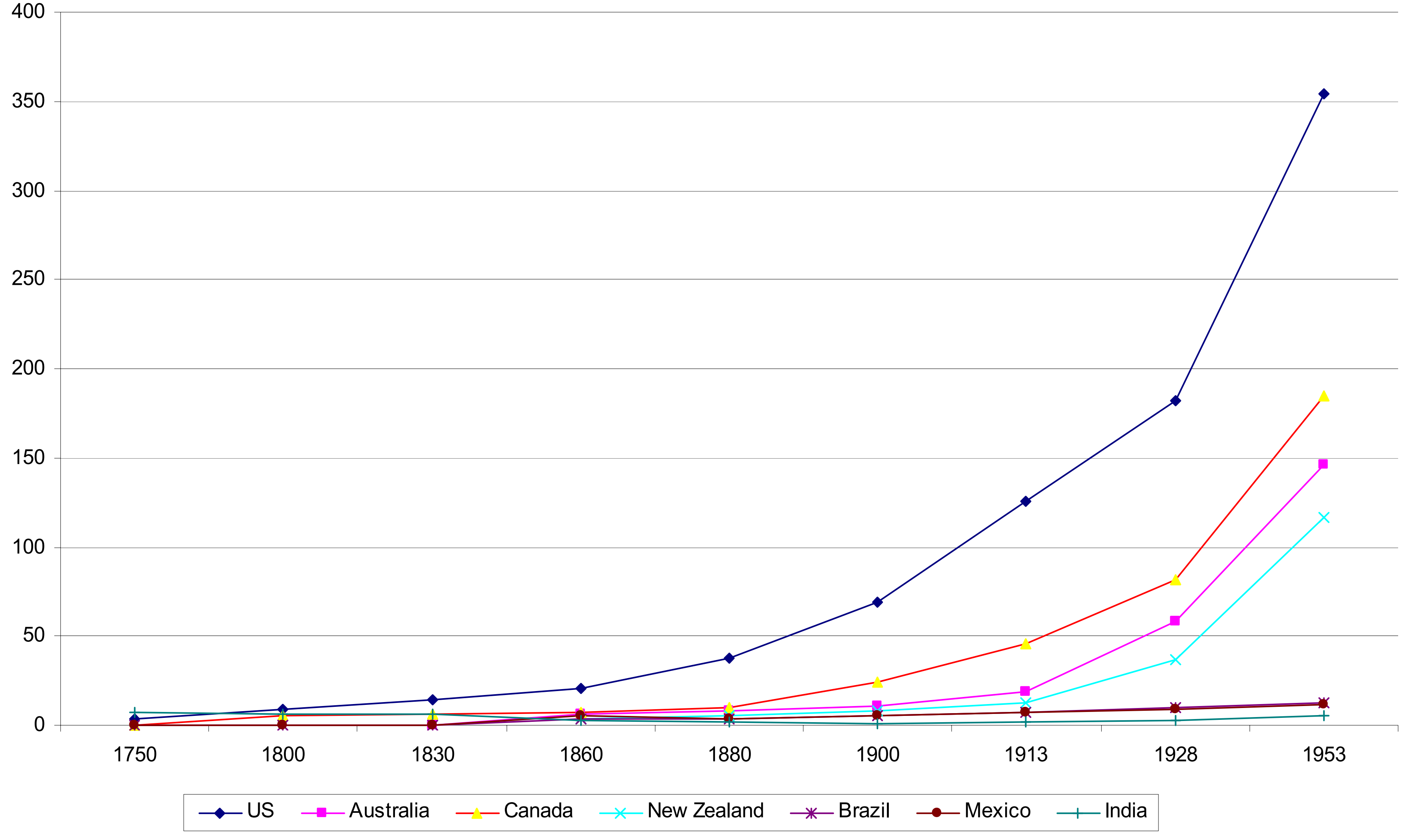


Figure 6D

Industrial Production Per Capita, UK in 1900 = 100
(from Bairoch)



Appendix Table A1
Variable Definitions and Sources

Variable	Description	Source
Log GDP per capita (PPP) in 1995	Logarithm of GDP per capita, on Purchasing Power Parity Basis.	World Bank, World Development Indicators, CD-Rom, 1999. Data on Surinam is from the 2000 version of this same source.
Log Population Density in 1500	Logarithm of population density (total population divided by total arable land) in 1500	McEvedy and Jones
Log Population Density in 1000	Logarithm of population density (total population divided by total arable land) in 1000	McEvedy and Jones
Log Population Density in 1 AD	Logarithm of population density (total population divided by total arable land) in 1 AD	McEvedy and Jones
Urbanization in 1995	Percent of population living in urban areas with a population of at least 5,000 in 1995	World Bank, World Development Indicators, CD-Rom, 1999.
Urbanization in 1960	Percent of population living in urban areas with a population of at least 5,000 in 1960	World Bank, World Development Indicators, CD-Rom, 1999.
Urbanization in 1900	Percent of population living in urban areas with a population of at least 5,000 in 1900	Bairoch and supplemental sources, as described in the appendix.
Urbanization in1500	Percent of population living in urban areas with a population of at least 5,000 in 1500	Bairoch and supplemental sources, as described in the appendix.
Urbanization in 1000	Percent of population living in urban areas with a population of at least 5,000 in 1000	Bairoch and supplemental sources, as described in the appendix.
European settlements in 1900	Percent of population that was European or of European descent in 1900. Ranges from 0 to 0.99 in our base sample.	McEvedy and Jones (1978) and other sources listed in Appendix Table 5 of AJR (2000).
European settlements in 1800	Percent of population that was European or of European descent in 1800. Ranges from 0 to 0.99 in our base sample.	McEvedy and Jones (1978) and other sources listed in Appendix Table 5 of AJR (2000).
Average Protection against Expropriation Risk, 1985-95	Risk of expropriation of private foreign investment by government, from 0 to 10, where a higher score means less risk. We calculated the mean value for the scores in all years from 1985 to 1995.	Dataset obtained directly from Political Risk Services, September 1999. This data was previously used by Knack and Keefer (1995) and was organized in electronic form by the IRIS Center (University of Maryland). The original compilers of this data are
Constraint on Executive in 1990	A seven category scale, from 1 to 7, with a higher score indicating more constraints. Score of 1 indicates unlimited authority; score of 3 indicates slight to moderate limitations; score of 5 indicates substantial limitations; score of 7 indicates executive parity or subordination. Scores of 2, 4, and 6 indicate intermediate values.	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Constraint on Executive in 1970	authority; score of 3 indicates slight to moderate limitations; score of 5 indicates substantial limitations; score of 7 indicates executive parity or subordination. Scores of 2, 4, and 6 indicate intermediate values. Set equal to 1 if	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Constraint on Executive in 1900	From 1 to 7. Coding as for Constraint on Executive in 1990 and 1970. Set equal to 1 if country was not independent at that date.	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Constraint on Executive in first year of independence	From 1 to 7. Coding as for Constraint on Executive in 1970. Date of independence is first year that country appears in Polity III dataset.	Polity III dataset, downloaded from Inter-University Consortium for Political and Social Research. Variable described in Gurr 1997.
Percent of European descent 1975	Percent of population that was European or of European descent in 1975. Ranges from 0 to 1 in our base sample.	McEvedy and Jones (1978).
Religion Variables	95 for countries formed more recently). The four classifications are: Roman Catholic, Protestant, Muslim, and "other".	La Porta et al (1999)
Colonial Dummies	Dummy variable indicating whether country was a British, French, German, Spanish, Italian, Belgian, Dutch or Portuguese colony.	La Porta et al (1999)
Temperature Variables	Temperature variables are average temperature, minimum monthly high, maximum monthly high, minimum monthly low, and maximum monthly low, all in centigrade.	Parker (1997)
Humidity Variables	Humidity variables are morning minimum, morning maximum, afternoon minimum, and afternoon maximum, all in percent.	Parker (1997)
Soil Quality	Measures of soil quality/climate are steppe (low latitude), desert (low latitude), steppe (middle latitude), desert (middle latitude), dry steppe wasteland, desert dry winter, and highland.	Parker (1997)
Natural Resources	of world zinc reserves today, number of minerals present in country, and oil resources (thousands of barrels per	Parker (1997)
Dummy for Landlocked	Dummy variable equal to 1 if country does not adjoin the sea.	Parker (1997)
Malaria in 1994	Malaria in 1994 is percent of people living in area where falciparum malaria is endemic	Gallup and Sachs 1998.
Latitude	Absolute value of the latitude of the country, scaled to take values between 0 and 1, where 0 is the equator.	La Porta et al (1999)
Log Mortality	Log of estimated settler mortality. From 1.7 to 6.2.	AJR (2000)

Appendix Table 2

Country name	Abbreviation used in graphs	Log population density in 1500	Urbanization in 1500 (base sample estimate)
Former European Colonies (in our base sample)			
Argentina	ARG	-2.21	0.0
Australia	AUS	-3.65	0.0
Bangladesh	BGD	3.17	8.5
Belize	BLZ	0.43	9.2
Bolivia	BOL	-0.19	10.6
Brazil	BRA	-2.13	0.0
Canada	CAN	-3.83	0.0
Chile	CHL	-0.22	0.0
Colombia	COL	-0.04	7.9
Costa Rica	CRI	0.43	9.2
Dominican Re	DOM	0.38	3.0
Algeria	DZA	1.95	14.0
Ecuador	ECU	0.77	10.6
Egypt	EGY	4.61	14.6
Guatemala	GTM	0.43	9.2
Guyana	GUY	-1.55	0.0
Hong Kong	HKG	-2.44	3.0
Honduras	HND	0.43	9.2
Haiti	HTI	0.28	3.0
Indonesia	IDN	1.45	7.3
India	IND	3.17	8.5
Jamaica	JAM	1.53	3.0
Laos	LAO	0.55	7.3
Sri Lanka	LKA	2.74	8.5
Morocco	MAR	2.21	17.8
Mexico	MEX	0.96	13.0
Malaysia	MYS	0.20	7.3
Nicaragua	NIC	0.43	9.2
New Zealand	NZL	-0.99	3.0
Pakistan	PAK	3.17	8.5
Panama	PAN	0.43	9.2
Peru	PER	0.45	10.5
Philippines	PHL	0.52	3.0
Paraguay	PRY	-0.69	0.0
Singapore	SGP	-2.44	3.0
El Salvador	SLV	0.43	9.2
Tunisia	TUN	2.46	12.3
Uruguay	URY	-2.21	0.0
USA	USA	-2.44	0.0
Venezuela	VEN	-0.82	0.0
Vietnam	VNM	1.82	7.3

Appendix A
Urbanization Estimates

1. General

Sources

We use four main sets of sources for information on urbanization. The first set of sources are publications by Paul Bairoch and his associates. The second are publications by Tertius Chandler, particularly Chandler (1988). The third are urbanization estimates that are more focussed on particular places and time periods, but which fill important gaps in the first two sources or supplement their data in a useful way. Finally, Mitchell (1993 and 1995) provides good urbanization data, but only for about 1750 to the present.

Bairoch (1988) provides the best single collection and assessment of urbanization research.[1] This volume reviewed evidence on urbanization from around the world since the beginning of recorded time and we rely primarily on his assessment of plausible urbanization estimates throughout our analysis. Even when we are forced to use specific numbers from another source, we check these estimates for consistency with Bairoch's various qualitative assessments (as documented in this appendix.)

Bairoch (1988) argues that the correct threshold for measuring urbanization in general is 5,000 inhabitants. There are some exceptions, but in general this level of urban population indicates that people are engaged in non-agricultural activities. On average over the past 200 hundred years urbanization and industrial employment have moved in step (Bairoch, 1988, Table 29.2).

After the publication of Bairoch (1988), Bairoch and his associates continued to develop more detailed country-level data.[2] Two publications are particularly useful. Bairoch, Batou and Chevre (1988) provide a comprehensive dataset on European urbanization since 800 AD. We use this primarily as a check on the European estimates in Bairoch (1988). We find no major revisions, but there is much more detail and documentation. There is also more detailed data on Eastern Europe and the Balkans.

The second useful publication is data on non-European urban populations collected by Gilbert Eggimann and published by Librarie Droz as "La Population des villes des tiers-mondes, 1500-1950" on the web at http://histsociale.isuisse.com/histsociale/ (see also www.droz.org) since fall 1999. The Scientific Editors of this publication were Paul Bairoch and Jean-Claude Toutain. Eggimann draws heavily on Chandler (1987) for information on early cities, and Showers (1979) for more recent data. He also makes extensive use of country-specific sources that are well-documented in his bibliography. The main drawback is that Eggimann's data only covers cities that crossed the threshold of 20,000 inhabitants at some point between 1500 and 1950.[3] Eggimann also does not cover the former colonies that have become rich (the US, Canada, Australia and New Zealand).

The second main source of urban population estimates is work by Tertius Chandler. Bairoch (1988) reports that he used Chandler and Fox (1974) as the basis of his first estimates.[4] Independently of

[1] The original French publication was in 1985.

[2] Professor Bairoch sadly passed away in 1999. We understand that a more complete global urbanization data was under preparation at the time of his death, but have not been able to locate it. We rely instead on official publications of his Center of International Economic History which moved to the University of Lausanne in September 1999. Many of its publications are available through Librairie Droz (Geneva), under the direction of Jean Batou and Bouda Etemad. For helpful communications on these issues, we thank Professor Etemad.

[3] Eggimann originally intended to look at cities with populations above 10,000, but this proved to be much more work. See the section on his website under "Pourquoi avoir abandonné la limite de 10'000 pour se limiter à 20'000?"

[4] Bairoch (1988) writes, "Indeed, had it not been for the publication of the investigations of Chandler and Fox (1974), I would never have done the research reported here, at least not in its present audacious form" (p.519). Bairoch explains that his task was to enlarge the list of cities and using more recent

Bairoch's later work, Chandler (1987) extended and refined these estimates. Chandler's data is attractive because it is more detailed for non-European countries than what has been published by Bairoch's group (with the exception of Eggimann, who only covers countries that are now part of the "Third World"). In particular, Chandler reports numbers by individual city since 800 A.D. as much as possible. He also has information on cities before 800 A.D., although this data is less comprehensive. Eggimann (1987) cites Chandler as a source for some countries.

However, there are three limitations of the Chandler data that make us prefer Bairoch's estimates (augmented with Eggimann's data). First, Chandler's minimum city size for Europe and the Americas is 20,000 (although he has estimates for some years with a lower threshold). For the reasons outlined above, this is not as satisfactory as Bairoch's 5,000 inhabitants threshold. Second, for Asia, Chandler's minimum city size is 40,000, making it difficult to compare precisely with his own estimates for Europe and the Americas. For example, this leads to a much lower (and rather implausible) estimate of urbanization in India and China compared with what he finds for Europe and the Americas, and compared with the assessments in Bairoch (1988). Third, Chandler's data stops in 1850 (1861 for the Americas). Given the importance for our analysis of the urbanization that came with industrialization, we prefer to use data that is consistent into the twentieth century. Nevertheless, Chandler (1987) provides a very useful check on the Bairoch-school, particularly because it was an important source for that work. Chandler (1987) also provides usefully detailed urbanization estimates across almost all countries.

Our third set of urbanization estimates comes from scholars who have looked a urbanization in particular places or periods. De Vries (1984) provides a detailed set of estimates for urbanization in Europe, with a minimum city size of 10,000. We use this series as a check on the Bairoch-school, and find some differences. However, the differences do not affect our main results (e.g., concerning the persistence of prosperity within Europe and the absence of a reversal of fortune within Europe since 1500). Hohenberg and Lees (1985) also assess urbanization in Europe over the past 1000 years. McEvedy (1967, 1992) reports approximate urbanization estimates for Europe and the Middle East from the neolithic revolution to 1000 AD. He also provides a useful series of maps that show how the development of large urban centers related to trade flows. We use these numbers in our assessment of the persistence of urbanization.

Calculating Urbanization

Bairoch generally provides an estimate of urbanization (i.e., percent of the total population living in urban areas). However, most other scholars of non-European urbanization, including Eggimann, Chandler and Fox, and Chandler, provide only estimates of total urban population (in cities about their minimum population threshold.) To convert these into urbanization estimates, we use the population figures in McEvedy and Jones (175). We also use the land area and population estimates in McEvedy and Jones (1975) to calculate population density.[5]

We report our urbanization estimates in three sections. First, we explain in detail our baseline estimates for 1500, including both our main sources (usually from Bairoch-Eggimann) and supporting

---

sources. "This project was conducted relatively in depth for the following regions and periods: Europe from 800 to 1800 (ten periods), Latin America from 1600 to 1950 (nine periods), Africa from 1900 to 1950 (five periods), and the larger countries of Asia from 1500 to 1950 (nine periods).

[5] The McEvedy and Jones work is nearly 25 years old but as far as we know it has not been superceded as a single comparable set of estimates across all countries for the past 2000 years. Some estimates are still controversial (e.g., on pre-European America, but McEvedy and Jones are rather conservative and higher population estimates for that region would help our results) and estimates for Africa must be treated with particular caution (as McEvedy and Jones acknowledge.) In a recent assessment, the US Bureau of the Census (see the main text for the precise reference) showed that McEvedy and Jones's aggregate estimates compare favorably with more recent estimates. Using the McEvedy and Jones data is standard in the economics literature; see for example Kremer (1993).

evidence (usually from Chandler). Second, we provide evidence on the development of urbanization from 1500. Third, we provide evidence on urbanization before 1500.

2. Base Estimates for 1500

Our base data consists of Bairoch's (1988) urbanization augmented by the work of Eggimann (1999). Merging these two series requires us to convert Eggimann's estimates, based on a minimum population threshold of 20,000, into Bairoch-equivalent urbanization estimates, based on a minimum population threshold of 5,000. It is not appealing to look only at towns above the threshold of 20,000 people, as this would ignore even substantial urbanization levels. For example, England before 1500 had only one city with population over 20,000, but an urbanization rate of 8%; see Chandler 1987, p.19, Bairoch 1988, p.179, and Bairoch, Batou and Chevre 1988, pp.32-35.

To convert between the two sets of estimates, we run a regression of Bairoch estimates on Eggimann estimates for all countries where they overlap in 1900 (the year for which we have most Bairoch estimates for non-European countries).[6] There are thirteen countries for which we have good overlapping data. This regression yields a constant of 6.6 and a coefficient of 0.67. We use these estimates below to generate Bairoch-equivalent urbanization estimates.

We have also checked the robustness of our results using alternative methods of converting Eggimann estimates into Bairoch-equivalent numbers. We have calculated country-specific conversion ratios and conversion ratios at the regional level (e.g., North Africa and the Andean region separately). We have also used Bairoch's overall assessment of urbanization for broad regions, e.g., Asia, without the more detailed information from Eggimann. None of these alternative methods affect any of our main results.

We provide urbanization estimates for all countries that now have a population over 500,000. Most countries with a population under 500,000 either lack reliable data or are not fully independent (e.g., this is true of many small islands that are nominally independent in the Caribbean or the Pacific) or both.

Europe in 1500

The best available urbanization estimates are those for Europe. After more than 40 years of detailed quantitative studies, there has been remarkable convergence in the opinion of scholars about the extent of urbanization in the late Middle Ages.

Bairoch (1988, table 11.2, p.179) reports urbanization in Europe in 1500, using a population of 5,000 as the minimum threshold. These estimates, together with his numbers for 1300 and 1700 are reproduced in Appendix Table 3. Through 1800 this data is in the form of a range of estimates and we use the mid-point of this range in our base data.[7]

Bairoch, Batou and Chevre (1988) provide estimates for the population of every European town that had "at some time between 800 and 1800, 5,000 or more inhabitants" (p.ix). We use these estimates where Bairoch (1988) does not provide data, and as a check on Bairoch (1988). However, because Bairoch, Batou and Chevre (1988) do not provide any non-European data, we prefer to use Bairoch (1988) as the basis for our comparative analysis.

One important question is which urbanization estimate to use when borders have changed. For example, for the Baltic states (Latvia, Lithuania and Estonia) should we use their urbanization or the urbanization of Russia-in-Europe (arguably the relevant geopolitical entity in 1500)? We use the political entity selected by Bairoch and his associates to report urbanization numbers. We also report alternative estimates wherever available. None of these alternative estimates affect our main results.

---

[6] Alternatively, two sets of estimates indicate that a rough and conservative conversion factor is to multiply the Eggimann urbanization estimates by two. We have constructed an alternative set of estimates using this method and none of our main results were affected.

[7] Slightly different estimates are reported in Bairoch, Batou and Chevre (1988), but using these does not affect our results.

We have direct estimates from Bairoch (1988) for France (9-12%), Germany (7-9%), Belgium (30-45%), the Netherlands (20-26%), Portugal (11-13%), Spain (10-16%), and Switzerland (6-8%).[8] Land area has not changed significantly, except that Germany lost some of its relatively less urbanized eastern area (i.e., the area occupied by Germany today almost certainly had a higher urbanization in 1500 than did the Germany of 1500). For Britain we use Bairoch's (1988, p.179) estimate for England (7-9%).[9]

We use the Scandinavia estimate (6.5%) for Sweden, Norway, Denmark and Finland. This is probably high for Finland.[10] We use the Russia estimate (4.5%) for Russia, Ukraine, Belarus, Estonia, Latvia and Lithuania.[11] We use the Austria-Hungary (6.5%) estimate for Austria, Slovakia, Hungary and the Czech Republic.[12]

Bairoch (1988) does not provide an estimate for Romania or Poland. We use the estimates of 3% (Romania) and 6% (Poland) urbanization from Bairoch, Batou and Chevre (p.259).

From Bairoch, Batou and Chevre, we know the urban population of Greece in 1500 was 35,000. The population from McEvedy and Jones (p.113) was 1m. So we use an urbanization rate of 3.5% in our base sample.

From Bairoch, Batou and Chevre (p.13), the urban population of Bulgaria in 1500 was 97,000. The population from McEvedy and Jones (p.113) was 800,000. We use an urbanization rate of 12.1% in our base sample. Note that this was by far the most urbanized country in the Balkans (with the exception of Turkey-in-Europe.)

From Bairoch, Batou and Chevre (p.3), the urban population of Albania in 1500 was 5,000. The population from McEvedy and Jones (p.113) was 200,000. We therefore use an urbanization rate of 2.5% in our base sample.

From Bairoch, Batou and Chevre (p.69), the urban population of the countries that later constituted Yugoslavia was 112,000. The population from McEvedy and Jones (p.113) was 2.25m. We use an urbanization rate of 5% for Macedonia, Croatia, and Slovenia. Note that Bosnia-Hercegovina and Serbia are not in our dataset as we do not have data for GDP per capita in 1995.[13]

---

[8] The estimates from Bairoch, Batou and Chevre (1988, p.259) are quite similar: France, 8.8%; Germany, 8.2%; Belgium, 28%; the Netherlands, 29.5%; Portugal, 15.0%; Spain, 18.4%; and Switzerland, 6.8%.

[9] Urbanization in the UK as a whole was almost certainly lower; Bairoch, Batou and Chevre give an estimate of 4.6%. We do not have reliable urbanization estimates for Ireland. Using a threshold of 10,000 inhabitants, de Vries (1984, Appendix 3) puts the urban population of England and Wales at 80,000 in 1500, with Scotland at 13,000, and Ireland at 0. Bairoch, Batou and Chevre (p.39) give a total Irish urban population of 19,000. For 1500, McEvedy and Jones (p.43) give the population of England and Wales as 3.75m, Scotland as 500,000, and Ireland as 800,000. The de Vries urban population numbers would therefore imply urbanization of 2.1% in England and Wales and 2.6% in Scotland. With their lower threshold, Bairoch, Batou and Chevre (1988, p.259) give total urban population for the United Kingdom as 230,000; assuming this includes Ireland implies urbanization of 4.6%. Their implied Irish urbanization rate is 2.4%, but only Dublin had more than 5,000 inhabitants.

[10] De Vries 1984 p.270 does not report any towns with a population over 10,000 between 1500 and 1800. Bairoch, Batou and Chevre report a total urban population in 1500 of only 2,000 (in Turku). McEvedy and Jones (p.53) put the total Finnish population in 1500 at 100,000, implying urbanization of 2%.

[11] Bairoch, Batou and Chevre (1988, p.259) report urbanization for "Russie d'Europe" as 5.4% in 1500.

[12] Bairoch, Batou and Chevre (1988, p.259) provide an urbanization estimate of 4.8% for the aggregate of "Autriche-Hongrie-Tchéc."

[13] Bairoch, Batou and Chevre (1988, p.259) put total urbanization in the Balkans at 11.3%. This average estimate is puzzling and it is not clear to which countries they are referring. By their own numbers, total urbanization in Greece, Bulgaria, Albania and future Yugoslavia was 249,000. McEvedy and Jones (p.113) give total population in the Balkans as 4.5m (excluding 300,000 people in Turkey-in-Europe), which gives urbanization of 5.5%.

The Americas in 1500

There is some doubt about the exact level of urbanization in the Americas before the Europeans arrived, but the relative distribution of urbanization across North and South America is clear. To the north of Mexico, there was relatively little urbanization. In Mexico, other parts of central America and northern/western South America, urbanization was relatively high (probably average urbanization was just a little lower than in Asia at this time, but we cannot be sure).

For Latin America, Bairoch (1988, Table 24.1, p.389) reports the distribution of cities with populations of 20,000 or more. In 1500, there were 20 such cities in the Northern Andes (Bolivia, Colombia, Equador and Peru), and 10 in Mexico. There were no cities above this threshold size in the Caribbean, Brazil or the "Temperate Regions" (Argentina, Chile, and Uruguay). He further estimates that there were 4-7 cities with populations of more than 50,000 and "25-30 (perhaps even 40)" with populations of 20,000-50,000 (Bairoch 1988, p.66). There were two main concentrations of urban population: the Aztecs, Tairona, Tarascans, and Zapotecs (roughly in the area now Mexico); and the Incas, Mayas, Chibchas, and Cakchiquels (Central and north-western South America.) Urbanization may have been as high as 10-13% in some places, but Bairoch's assessment is that it was more likely around 7% on average. For the Chimu region of the Andes, Bairoch cites favorably an estimate as high as 14% (Bairoch 1988, p.66). For the Andean region as a whole, Bairoch estimates the low end for urbanization was 2-3%; and the high end was 10-13%.

For Colombia in 1500, Eggimann's estimates imply urbanization of 2%. As a Bairoch-equivalent urbanization measure, this would be 7.94%. We use this in our baseline data.

For Ecuador in 1500, Eggimann estimates urbanization as 15%. Part of this urban population should be ascribed to the area occupied by modern-day Bolivia. Ecuador and Bolivia both had 6% in Eggimann's numbers, if we ascribed urbanization evenly (probably there was more in Ecuador). We use an estimate of 10.62% urbanization for both in our baseline data.

For Peru in 1500, Eggimann estimates urbanization was 5.8%. This is converted to a Bairoch-equivalent urbanization level of 10.49%. We use this in our baseline data.

For Guatemala in 1500, Eggimann estimates 3.8% (deflating the urban population using whole population of Central America). Clearly part of this should be ascribed to other parts of Central America. Our baseline data has an estimate of 9.2%, which seems quite reasonable for Costa Rica, Honduras, Panama, Nicaragua, Belize, El Salvador, and Guatemala on average (it might have been lower in Costa Rica and Panama and higher in Guatemala, but there is no way to be sure.)

For Mexico in 1500, Eggimann estimates urbanization was 12.3%. Our baseline data uses the Bairoch-equivalent urbanization estimate of 14.84%. According to Bairoch, p.389, Mexico had half the number of cities with at least 20,000 population as the northern Andes (10 vs. 20) in 1500. The population in both areas was about the same according to McEvedy and Jones, around 5m. When Europeans arrived in 1519, Tenochtitlan (now Mexico City) probably had a population in the range of 150-200,000 and lay in the middle of a heavily urbanized region with up to 400,000 people living in smaller cities. "During this same period, there were only four European cities this size; and on the Iberian peninsula, the largest cities at that time, Granada and Lisbon, had around 70,000 inhabitants each" (p.63).

Despite his extensive discussion of urbanization in South America, Bairoch makes no mention of pre-European urbanization in Argentina or Brazil or other parts of South America. For Brazil in 1500, Eggimann estimates 0.1% urbanization (just 1,000 people in a town). We assign a value of zero in our baseline data. For other countries in South America with low levels of urbanization we assume the same value. Thus we use 0% for Argentina, Chile, Uruguay, Paraguay and Guyana.

While we also do not know the precise level of urbanization in North America, Bairoch (1988) is clear that this was much lower than in Central and South America. He refers to the overall situation as "Culture without Cities" (p.68). There was probably some urbanization, for example the Anasazi people may have had cities with populations of 10,000 or more between the eleventh and fourteenth centuries, but most of the southeastern and eastern United States and Canada was much less urbanized (with only the Huron-Iroquois having even slash-and-burn agriculture). He estimates that population density north of the Rio Grande before the Europeans was only 0.2 people per square kilometer or 0.4 people per square kilometer if we exclude all "truly uninhabitable" areas (p.69). "In any event, the displacement of villages caused by the practice of slash-and-burn cultivation considerably reduced the likelihood of the creation of urban centers by preventing or impeding any permanent large-scale settlements (p.69)." In his assessment, there were no "genuine cities" (p.69).

For the Americas, Chandler uses a minimum city size of 20,000 inhabitants. According to Chandler (1987), in 1500 urbanization was 6.5% in Mexico, 2.5% in Peru, 5% in Ecuador, 2% in Colombia. His urbanization estimate for Guatemala is 19.6%, but this is presumably case where the urban population should be divided by a total population that is larger than the number of people who were living then within the borders of Guatemala now. Chandler does not report separate data for Bolivia, but again this population was almost certainly part of the urban areas in Ecuador. He does not report any urbanization for Canada, Brazil, Paraguay, Uruguay, Argentina or Chile.[14]

We have very little information on urbanization in the Caribbean. Cuba, the Dominican Republic, Haiti and Jamaica all probably had some urbanization, although clearly they had less urbanization than Central America. We bring these countries into the analysis only when we use population density data.

The Pacific Region in 1500

For Australia, Bairoch's assessment is that "agriculture and urbanization emerged only in the nineteenth century, brought by the European immigrants" (p.69). The pre-European population existed as hunter-gatherers with no towns. We use a value of 0% in our baseline data.

There probably was some agriculture in Western Polynesia (p.70). In Bairoch's assessment, this level of agriculture generally supports a minimal level of urbanization at around 3%. Bairoch does not discuss pre-European New Zealand explicitly, but other sources suggest that the culture was similar to Western Polynesia. We therefore assign an urbanization level of 3%.

Africa in 1500

There is reasonably good urbanization data for North Africa. For Tunisia in 1900, Eggimann estimates urban population as 305,000 out of 1.5m; urbanization of 20.3%. In 1500, Eggimann estimates urban population as 65,000 out of 800,000; urbanization of 8.1%. We use the Bairoch-equivalent urbanization estimate of 12.3% in our baseline data.

For Libya in 1900, Eggimann estimates urban population as 57,000 out of 800,000; urbanization of 7.1%. In 1500, Eggimann estimates urban population as 16,000 out of 500,000; urbanization of 3.2%. We use the Bairoch-equivalent urbanization estimate of 8.74% in our baseline data.

For Morocco in 1900, Eggimann estimates urban population as 433,000 out of 5m; urbanization of 8.7%. In 1500, Eggimann estimates urban population as 250,000 out of 1.5m; urbanization of 16.7%. We use the Bairoch-equivalent urbanization estimate of 17.79% in our baseline data.

For Algeria in 1900, Eggimann estimates urban population as 499,000 out of 5m; urbanization of 10%. In 1500, Eggimann estimates urbanization as 165,000 out of 1.5m; urbanization of 11%. We use the Bairoch-equivalent urbanization estimate of 13.97% in our baseline data.

For Egypt in 1900, Eggimann estimates total urban population was 1.603m, while McEvedy and Jones put total population at 10m. For 1900, Eggimann's estimate is 16% and Bairoch's estimate is 19%, giving a ratio of 1.19. Urbanization in 1500, according to Eggimann, was 475,000, while population was 4m. In 1500, Eggimann's estimate is 11.9%, implying a Bairoch-equivalent estimate of 14.2%. We use the Bairoch-equivalent urbanization estimate of 14.57% in our baseline data.

Including Egypt in 1900 the population of North Africa was 22.3m. Urban population from Eggimann was 2.9m. Implied urbanization for all of North Africa was therefore 13%. Bairoch (p.429) estimates North African urbanization in 1900 as 16%. In 1900 North Africa without Egypt had a

[14] Chandler suggests there was some urbanization in the area now occupied by the United States; he reports an urban population of 20,000 in Nanih Waiya. With a total population of 800,000 (from McEvedy and Jones), this implies urbanization of 2.5%. However, this seems to have been an isolated number. He does not report any urbanization for 1400 or 1600 or any other date before 1750.

population of 12.3m. Urban population for this area according to Eggimann was 1.294m, which implies an average urbanization level of 10.5%. In 1500, the population of the same area was 4.3m, while urban population according Eggimann was 0.496m. This would imply an urbanization rate of 11.5%.

In 1500, the total population of North Africa was 8.3m (McEvedy and Jones p.221, p.224 and p.227). Urban population according to Eggimann was 0.971m, implying an urbanization for all of North Africa of 11.7%.

While the data on sub-Saharan Africa is worse than for any other region, it is clear that urbanization before 1500 was at a higher level than North America or Australia. By 1500 urbanization was "well-established" in sub-Saharan Africa (Bairoch 1988, p.55). In fact, there are much older records from Arab explorers regarding African cities. "In summary, there were sizable cities in Black Africa by 1000 B.C., if not earlier" (Bairoch 1988, p.56).

Leading examples of urbanization around the time of 1500 are to be found in the kingdoms of Ghana, Songhai, Benin, Congo, Zimbabwe and the Yoruba states (of which Benin was one) (pp.56-60). Around 1500, the city of Benin had a population in the range of 60-70,000 and was a "well-ordered urban center with a system of water conduits and a sizable artisanry working at an advanced technical level" (p.58). The same kingdom had about 10 other cities at this time. At the time the Portuguese first made contact with the kingdom of the Congo (in 1484), the capital city probably had a population in the range of 40-60,000 (p.58).

There were also a number of cities that were closely connected with or even founded by Islamic traders from further north. Gao, Gober, Jenné, Kano, Kazargamu, Timbuktu and Zaria all fall into this category (p.61). "Around the beginning of the sixteenth century, these cities had populations ranging from between twenty-five thousand and seventy thousand, with an average somewhere on the order of forty thousand" (p.61).

Bairoch estimates that urbanization on average in sub-Saharan Africa was 3% in 1900 (p.430). For sub-Saharan Africa in 1800, Bairoch (1988, p.393) estimates the total urban population was 3-4 million, of which Europeans were between 30,000 and 120,000. For 1900, he puts urbanization in the range of 2-5% (p. 413-414, in particular Table 26.1). His discussion makes it clear that there was substantial variation within sub-Saharan Africa. West Africa, West-Central Africa (i.e., around the Congo) and East Africa (around Zanzibar and Kenya) were clearly more urbanized than Southern Africa (with the exception of Zimbabwe, but the decline of this civilization can be placed around 1450).

Given the weakness and incompleteness of data for sub-Saharan Africa, we do not include any estimates in our baseline dataset.

Asia in 1500

Bairoch estimates that average urbanization in Asia in 1900 was 9%, but he argues that this was higher in the eighteenth century before European colonization (p.430). He also argues there was more homogeneity of urbanization within Asia than within Africa or America (pp.430-431).

Bairoch (1988, table 25.1, p.407) reports urbanization in Asia in 1900 as follows.
China, 7-9%; India, 9-11%; Indonesia, 5-8%; Philippines, 14-16%; Vietnam, 6-9%; Iran, 12-15%; Korea, 8-11%; Thailand, 8-10%; Turkey, 15-18%; and Asia as a whole, 8-10%. Bairoch expresses the general view that urbanization had not changed much in the years before 1900. Bairoch (1988, p.406) estimates that "Around 1700, as well as around 1930, 10-12% of the population of Asia lived in cities." He further estimates that because urbanization fell in China, the level of urbanization had dropped to 8-10% by the start of the nineteenth century.

Bairoch emphasizes that there is a long-standing history of urbanization throughout Asia. In fact, prior to 1500, Asia was almost consistently among the most urbanized places in the world. By 1900 this had changed. For example, Bairoch 1988, p.350, states, "Between 100 B.C. and A.D. 1200 there were always one or two Asian cities among the three largest in the world. And between 1200 and 1820 there were invariably six or seven Asian cities (even excluding the Middle East) among the ten largest in the world. By contrast, there were only three around 1875, and none at all in 1900."

Overall, "Around 1500 the world appears to have had some fifty to sixty cities with populations or more than 100,000, and all but four lay in the regions destined to become the Third World of today" (p.436). Most of those cities were in Asia, with perhaps 2-3 in the Americas, and 3-4 in North Africa. In contrast, there were only 4 cities in Europe with more than 100,000 in population in 1500" (Bairoch 1988, Table 27.3, p.437).[15]

For China in 1900, Eggimann provides an urbanization estimate of 2.95% and Bairoch directly estimates 7.75%. For 1500, Eggimann gives 2.18%, and we use a Bairoch-equivalent urbanization estimate of 8.06% in our baseline data. This is a little low, as Bairoch says, p.403, that urbanization in China in the sixteenth century was 11-14% (falling to 6-7.5% by 1850 but rising slightly to 7-8.5% by 1900.) According to Chandler, urbanization in China in 1500 was 2%.

For China, Bairoch favors an estimate of urbanization by the twelfth century around 10-13%, taking the population threshold of 5,000 for a city (p.353). This was 1-2 points higher than Europe (excluding Russia) at the same time and "close to the maximum reached in Europe (minus Russia) at any time before the Industrial Revolution." He justifies this level of urbanization partly by reference to the relatively high levels of agricultural productivity in China at that time (p.354). At the beginning of the sixteenth century, the level of urbanization in China was probably 11-14% (Bairoch 1988, p.356). By the middle of the nineteenth century, urbanization was probably 6-7.5% (Bairoch 1988, p.357.)

Regarding India, Bairoch (1988, p.350) writes, "By 1300, probably five cities, and perhaps six to ten, had surpassed [100,000 population]. It may be recalled that during this same period, with a population of some 80-100 million, just about the same as India's, Europe also had five cities with populations of more than 100,000." According to Bairoch, urbanization in India in 1300 was roughly at the same level as in Europe. By the early sixteenth century the largest city in India, Vijayanagar, had a population of 500,000 while Paris, the largest city in Europe, had a population of only 200-250,000.[16]

For 1500, Eggimann's estimate for India, Pakistan and Bangladesh combined is urban population 2.323m. With a population of 80.6m, this implies urbanization was 2.9%. We convert to our baseline Bairoch-equivalent urbanization estimate of 8.54% for India. We use this also for Pakistan, Bangladesh and Sri Lanka.[17]

For India in 1900, Eggimann's estimate is 6.9% and Bairoch's estimate is 10%. But Bairoch and Eggimann overlap with an estimate of urban population closer to 1500. In 1700, Eggimann reports an urban population of 2.879m; McEvedy and Jones report a total population of India of, implying urbanization of 2.2%. For India, Bairoch cites favorably the estimate of urbanization around 15% in 1600 (but indicates this may be on the high end) (p.351).

The midpoint of Bairoch's estimated urbanization in this year is 12%. Bairoch's assessment is that urbanization in India declined after 1500.

Bairoch reports direct estimates of urbanization in India from census figures as follows (p.400): in 1700, 11-13%; in 1800, 9-11.5%; in 1850, 7.5-9.5; in 1900, 9-10; in 1940, 14-16%. He argues forcefully that the evidence of deindustrialization in the nineteenth century points to a decline of urbanization.

According to Bairoch's assessment, China and India probably had about the same level of urbanization in 1500, although India was at the top of an urbanization cycle and China was at the bottom**.** According to Chandler, urbanization in India in 1500 was 1.8%.

At the start of the eighteenth century, Bairoch argues that Japan had urbanization of 11-14%. Bairoch, pp.45-46: urbanization in Japan dates only from about 800 AD, but large cities developed fast; largest city was larger than in Europe at this time. In the sixteenth century, Bairoch p.46 says urbanization was "certainly less than 10% and probably on the order of 5-8%". We use a rate of 6.5% in our baseline estimates. Eggimann does not have an estimate for Japan as it is not in the "Third World."

---

[15] Bairoch (1988) doesn't name those four European cities. According to Chandler, there were five: Paris had a population of 185,000, Venice had 115,000, Naples had 114,000, and Adrianople had 127,000. The largest city in Europe was Constantinople, with 200,000 inhabitants.

[16] Chandler's estimate for Vijayanagar is XX and for Paris is YY…

[17] We have to estimate the population of India from McEvedy and Jones's estimate of population on the subcontinent. We use the population ratios of 1945 *(check)*, when India was 80.6% of the total subcontinent population. 160x0.806=128.96m

According to Chandler, urbanization in Japan was as high as 5% in 800 but fell to a low of 0.2% in 1500; it was 4.8% in 1600.

Urbanization in Japan was likely 14-15% by 1800 and continued to rise (Bairoch 1988, p.360). Note that none of this increase was due to international trade, as Japan was cut off from the outside world at this time.

Urbanization in Byzantium (present-day Turkey) was "certainly more than 8-9%, and probably around 10-13%" for several hundred years after 1000 A.D. (Bairoch 1988, p.370). For Turkey in 1900, Eggimann's estimate or urbanization is 12.9 and Bairoch's estimate is 16.5%, giving a ratio of 1.28. For 1500, Eggimann's estimate is 7.4%, which implies a Bairoch-equivalent urbanization estimate of 11.56%. We use this in our baseline data. According to Chandler, in 1500 urbanization in Turkey was 5.9%.[18]

For Thailand in 1900, Eggimann's urbanization estimate is 6.25% and Bairoch's estimate is 9%. For 1500, Eggimann's estimate is 6.3, implying a Bairoch-equivalent estimate of 10.82%, which we use in our baseline estimates. According to Chandler, urbanization in Thailand in 1500 was 6.3%.

For Korea in 1900, Eggimann's urbanization estimate is 4.8% (note that we combine north and south Korea's urban population for this calculation) and Bairoch's estimate is 9.5. In 1500, Eggimann's estimate is 3.1%, implying a Bairoch-equivalent estimate of 8.68%. According to Chandler, urbanization in Korea in 1500 was 4.1%.

For Indonesia in 1900, Eggimann's urbanization estimate is 2.28 and Bairoch's estimate is 6.5, giving a ratio of 2.85. For 1500, Eggimann does not have an estimate. From Eggimann, the urban population in Malaysia in 1500 was 80,000. The population of Malaysia was only 400,000, which would imply 20% urbanization. However, making a distinction between Malaysia and Indonesia at that time is quite artificial.[19] The population of Indonesia was 7,750,000, giving a total population of 8.15m, and an Eggimann urbanization estimate of 1%. In our baseline data we use the Bairoch-equivalent estimate of 7.27% for both Malaysia and Indonesia. According to Chandler, urbanization in Indonesia and Malaysia (using combined urban populations and combined total populations) was 0.5% in 1500.

Eggimann reports the urban population in Laos in 1500 as 40,000. The population was 400,000, giving an urbanization rate of 10%. But it seems more reasonable to ascribe part of this urbanization also to Vietnam which, in 1500, had a population of 4m. This implies urbanization of 1%. Converting this to Bairoch-equivalent urbanization gives 7.27%, which we use in our baseline estimate. According to Chandler, urbanization in Laos was 10% in 1500, but he does not give any urban population after that date.

For Vietnam in 1900, Eggimann's estimate is 2.9% and Bairoch's estimate is 7.5. There is no data from Eggimann for 1500. According to Chandler, urbanization in Vietnam in 1500 was 2%.

According to Chandler, urbanization in Cambodia was 3.1% in 1400; he does not give an estimate for 1500.

According to Eggimann, in 1500 there was an urban population of 110,000 in Burma/Myanmar. With a population of 4m, this implies an urbanization rate of 2.8%. Converted into Bairoch-equivalent urbanization, this gives 8.48% and we use this in our baseline estimates. According to Chandler, urbanization in Burma was 3.5% in 1500.

We assign an urbanization value of 3% to the Philippines, based on the qualitative assessment by Bairoch (1988). On p.48: Bairoch says of the Philippines, "before the arrival of the Spanish colonizers, there were no genuine cities."

There is not much information on Singapore and Hong Kong. We know that neither area was developed and neither was even a minor port. Hong Kong had some small fishing villages. We assign 3% urbanization to Singapore and Hong Kong (Bairoch's estimate of urbanization when there is settled agriculture), but check our results carefully under alternative assumptions. Both Singapore and Hong

[18] The Ottoman Turks took Constantinople in 1453 and over the next one hundred years developed a major international city. By 1600 the population was 700,000 (up from 200,000 in 1500, according to Chandler) and according to Eggimann urbanization reached 12%.

[19] Cite material from A History of Malaysia here

Kong were built up by the Europeans, i.e., the Europeans did not take over an existing port but rather started from scratch.

The Middle East in 1500

The Middle East was one of the first regions of the world to be urbanized (see the section on urbanization before 1500 below.) Around the year 1000 A.D., the Muslim world had urbanization around 10-13%, compared with only 8-9% in Europe (excluding Russia and Spain) (Bairoch 1988, p.375). In 1500, however, the region was probably no more urbanized than Western Europe on average and most likely less urbanized than the most advanced parts of Europe and Asia.

For Iran in 1900, Eggimann's urbanization estimate is 10% and Bairoch's estimate is 13.5. For 1500, Eggimann's estimate is 11.95, giving a Bairoch-equivalent urbanization estimate for our baseline estimates of 14.61%. According to Chandler, urbanization in Iran was 11.1% in 1500.

According to Eggimann, the urban population in Iraq was only 12,000 in 1500. Total population was 1m (McEvedy and Jones, p.151), giving urbanization of 1.2%. We use the Bairoch-equivalent urbanization estimate of 7.4%. According to Chandler, urbanization in Iraq was 9% in 1400 and 5% in 1700, but he does not provide numbers for 1500 or 1600 (i.e., no city was above his minimum threshold).

According to Eggimann, the urban population in Saudi Arabia in 1500 was 34,000 and in 1600 was 80,000. McEvedy and Jones (p.147) put the population of "The Interior" at 2m in 1500 and perhaps 2.1m in 1600. Eggimann does not report any urbanization numbers for this year for the Gulf, Oman or Yemen (although he does for later years). In comparison, McEvedy and Jones put the population in 1500 at 100,000 on the Persian Gulf Coast, 2.25m in Yemen, and 200,000 in Oman. This implies an Eggimann urbanization estimate of 1.7% if we use just the population of the Interior, 1.5% if we use the population of the Interior plus the Gulf and Oman, and 0.7% if we include also the population of Yemen. We use 1.5% as our Eggimann estimate, which implies Bairoch-equivalent urbanization of 7.6%. According to Chandler, urbanization in Saudi Arabia in 1600 was 1%. He does not provide numbers for 1500.

According to Eggimann, in Syria the urban population was 137,000 in 1500. The population of Syria and the Lebanon combined was 1.5m in 1500 (McEvedy and Jones, p.139), most of which was in Syria. This implies an urbanization rate of 9.1%, which converts into a Bairoch-equivalent urbanization estimate of 12.7%. According to Chandler, urbanization in Syria/Palestine was 6.4% in 1500.

According to Chandler, urbanization in Central Asia was 4.9% in 1500 (using the population numbers for Turkestan from McEvedy and Jones.) This converts into a baseline Bairoch-equivalent urbanization estimate of 9.9%.

Eggimann does not report any urbanization numbers for Lebanon or Jordan in 1500.

3. Urbanization from 1500 to 2000

There has been a great deal of scholarly work on urbanization in Europe since 1500 (to be completed using de Vries, BBC). The pattern of urbanization in 1500 in Eastern Europe fits the discussion of the situation in 1800 from de Vries (1984), Appendix 4. This suggests there was persistence within Eastern Europe.

In Europe as a whole, urbanization did not increase from 1700 to 1800 (Bairoch 1988). For the Americas, Bairoch (1988) comments as follows "Around 1700 North America north of the Rio Grande was still very thinly populated, there being some 250,000 Europeans, some 30,000 blacks, and perhaps one to three million native Americans. But in contrast to what was the case south of the Rio Grande, the white newcomers did not confront indigenous urban cultures… [W]hile certain of the cultures of northern North America had attained a high degree of development in art and in a number of techniques, they remained on the whole essentially rural and even, in most cases, preagrarian." (p.221, emphasis added). For 1800, Bairoch (1988, p.222) estimates that urbanization in North America was only 5-6%, compared with 12% in Europe. Around the same time, urbanization in Latin America was around 12-14% (p.222).

By 1900 North America was more urbanized than Europe.

In 1800, Bairoch (1988, p.389) estimates that Latin America was the most highly urbanized part of the world, with urbanization around 13-16%, i.e., 2-3 points higher than Europe (excluding Russia). Bairoch 1998, Chapter 21 is about the link between urbanization and economic development during the Industrial Revolution.

According to Bairoch, table 27.1, p.428, urbanization in the Third World was 8.5% of total population in 1800, 9.1% in 1900, and 15.7% in 1950. In comparison, in Europe (without Russia) it was 12.1% in 1800 (down slightly from 12.3% in 1700), 37.9% in 1900, and 50.7% in 1950 (Table 13.2, p.216.)

Bairoch p.460-461: urbanization and growth in industrial employment proceeded together until at least 1900. On China, p.476, "Progress in agriculture stimulated urban growth, and declines in agriculture prompted declines, if not in the cities themselves, at least in the level of urbanization."

Urbanization in North America occurred relatively late and quite quickly. Chandler (1988) has the most detailed information through 1861 and also makes a comparison with 1775 for almost all countries in the Americas. In 1775 there was a total population of about 4m in what became the United States. The total urban population in cities above 10,000 inhabitants was 89,000 (34,000 in Philadelphia, 24,000 in New York, 19,000 in Boston and 12,000 in Charleston), implying an urbanization rate of 2.2% (Chandler 1987, p.52.)[20]

In 1800, according to Chandler, the urban population was 192,000 out of a total population of 6m, implying urbanization of 3.2%.[21] A gap was already opening up between the north and the south of the US. In the North (defined by the states that stayed with the union in the civil war), urbanization was 7.1% while urbanization in the South was around 0.8% by Chandler's figures.[22]

In 1850, total urbanization, using a threshold of 18,000 inhabitants from Chandler, was 19.8%. In 1861, urbanization in the North was up to 31.5% and while in the South it was no more than 5%.[23] At the same time, urbanization in Canada was 12.9%.

In contrast, in Central and South America, urbanization was relatively high in the mid-1700s but in most countries it either did not increase as rapidly as in the US or actually fell in the next 100 years. In 1775, Chandler (1987, p.52) reports that the urban population of Mexico was 351,000, implying an urbanization rate of 7%. In 1800, urbanization was still 7% and in 1850 it was only 7.5%. By 1861, urbanization was no more than 8.6%. In Peru, Chandler (p.52) reports an urban population in1775 of 165,000, implying an urbanization rate of 11%. By 1800, however, his numbers imply that urbanization was down to 9.3% and by 1850 it was down to 6.6%. In 1861, urbanization in Peru was no more than 8.7%. In Brazil, urbanization in 1775 was 7.3%, while in 1861 it was 7.4%. In Venezuela, urbanization was 5.2% in 1775 but only 7.1% in 1861. Urbanization in Guatemala and Haiti seems to have stayed around 2% throughout this period.

Urbanization fell in the former Inca empire. Urbanization in Bolivia was 9.1% in 1775 but no more than 7.7% in 1861. In Ecuador, urbanization was 15.8% in 1775, but no more than 10.4% in 1861. In Colombia, urbanization was 5.2% in 1775 but only 4.1% in 1861. In Cuba, urbanization in 1775 was 24.6%; in 1861 it was 23.7%. In Jamaica, urbanization was 9% in 1800 but only 8.3% in 1850. There was no appreciable increase in urbanization over the same time period in Martinique or Barbados.

The notable exceptions are in the southern cone, where urbanization was relatively low in the mid-eighteenth century but then grew rapidly. In Chile, urbanization was 2.5% in 1775, but up to 11.1%

[20] In 1700 Boston had a population of 6-7,000 (Bairoch 1988, p.296). In 1750 Boston was the only city in North America that had a population of 20,000 (Chandler 1987).

[21] Even with his lower threshold for urbanization (5,000 vs. 18,000 in this case for Chandler), Bairoch estimates that urbanization was just 5% (Bairoch 1988, p.221).

[22] For 1820, Bairoch (1988, p.297) reports that 6% of the US population lived in towns with at least 5,000 inhabitants. By Bairoch's figures, in 1850 urbanization was up to 13-14%. Bairoch emphasizes that urbanization grew particularly rapidly between 1820 and 1870.

[23] These North vs. South calculations use the population numbers of 1850.

in 1861. In Argentina, urbanization was 8% in 1775 and up to 12.5% in 1861. In 1861, urbanization in Paraguay was 9.6% and 51% in Uruguay.

Coatsworth, not using urbanization estimates, comes to the same conclusion that we do about when Latin America fell behind the USA.

4. Urbanization in 1000 AD

The most comprehensive cross-continent data on urbanization in 1000 AD is from Chandler (1987). Eggimann's (1999) data starts only in 1500; de Vries also starts only in 1500. Bairoch provides useful some qualitative assessments and quantitative estimates from 1300. Bairoch, Batou and Chevre (BBC, 1988) provide detailed European urban population estimates by city from 800 AD, although they put more reliance in these numbers from 1200/1300. As we want to compare urbanization in 1000 with urbanization in 1500 (on the Bairoch basis), we use the BBC numbers where they are available. The BBC estimates are consistently higher than those of Chandler. We convert urban population numbers into urbanization using total population from McEvedy and Jones (1978).

The other Chandler estimates we convert into BBC equivalents using coefficients from a regression of the 11 overlapping data points for 1000 AD (all European). The coefficient on the Chandler estimate is 1.13 (significant at the 1% level) and the constant is 1.03. Note that given the result of this regression, our adjustment is essentially the same as applying Davis-Zipf's law for low urbanization countries (although a smaller adjustment for high urbanization countries).

In this part of the appendix we focus on urbanization in 1000. However, to assess whether urbanization in that year was at all unusual, as well as to give a sense of persistence of urbanization over shorter periods, we also report urbanization for 800, 1200, 1300, and 1400 (if we do not report a particular year, this indicates that the data is missing.) We should emphasize that reporting these numbers (and using one decimal place) does not imply that we necessarily think these estimates are accurate to within a few percentage points. Rather we argue that the broad pattern of urbanization is clear from these estimates and we can discern (in conjunction with qualitative evidence) to what extent urbanization has changed over periods of time such as 500 years.

Europe in 1000

According to Chandler, France had an urban population of 95,000 in 1000, when total population was 5 million (McEvedy and Jones), implying an urbanization rate of 1.5%. Before and after this date urbanization was fairly constant, although showing a slow and steady rise towards 1500. This fits with what we know about the gradual spread of prosperity in North-West Europe at this time. According to Chandler, the French urbanization rate in 1.8% in 800, 1.5% in 1000, 4% in 1200, 4.2% in 1300 and 6.1% in 1400. BBC put total urban population in 1000 at 255,000, implying urbanization of 5%. We use the BBC estimate.

The total Italian urban population was 362,000 in 1000 (Chandler), implying urbanization of 7.2% (McEvedy and Jones give total population as 5 million). Italian urbanization was 5.7% in 800, 10.1% in 1200, 8.2% in 1300, and 10.9% in 1400. BBC put total urban population in 1000 at 623,000, implying an urbanization rate of 12%. We use the BBC estimate.

In Spain, urbanization peaked at 19.2% in 1000 using Chandler's data, primarily due to the large (but short-lived) size of Cordoba. Total urban population was 766,000, of which 450,000 are estimated to have lived in Cordoba (total population was 4 million, from McEvedy and Jones). Urbanization was 7.9% in 800, 8.8% in 1000, 7.2% in 1300 and 7.7% in 1400. BBC estimate total urban population in 1000 was 970,000, which gives urbanization of 24%. We use the BBC estimate.

In Portugal, the urban population (just Lisbon according to Chandler) was 15,000 in 1000. With total population estimated at 600,000, this implies urbanization was 2.5%. We have no data on urbanization in 800. Urbanization was 4.4% in 1200, 2.8% in 1300, and 7.8% in 1400. BBC estimate total urban population in 1000 was 15,000, giving the same estimate as Chandler for urbanization. We use the BBC estimate.

Chandler has no data on urbanization for the Netherlands before 1300. In 1300 urbanization was 2.5% and in 1400 it was 5.25%. BBC do not have any data for the urban population in 1000; for 1200 they estimate the urban population was 12,000. Our estimate is explained in the next paragraph.

For Belgium, the first Chandler data is for 1200, showing urbanization of 7.2%. In 1300 urbanization was 9.8%, and in 1400 it was 29.5%, with 70,000 people in Ghent and 60,000 people in Bruges (most of the prosperity was based on the production and export of woolen cloth). BBC put the urban population in 1000 at 34,000, implying urbanization of 9% (McEvedy and Jones, p.63, put total population at 400,000). Most of this urban population was on the border with the Netherlands, which had a total population of 300,000. Therefore if we apply this urban population to Belgium plus the Netherlands, it would give an average urbanization rate of 4.9%. We use this average BBC estimate for both countries.

Chandler's work implied that German urbanization was 2.8% in 1000 (111,000 urban population and total population, from McEvedy and Jones, of 4 million.) Urbanization was 0.6% in 800, 5.3% in 1200, 4% in 1300, and 4.3% in 1400. BBC put total urban population at 209,000, implying urbanization of 5.2%. We use the BBC estimate.

For Austria, Chandler provides no data before 1500. BBC also provides no data. We use an estimate of zero urbanization, which is probably too low.

BBC (who were based in Geneva) estimate the urban population of Switzerland in 1000 AD as just 1000 people. McEvedy and Jones, p.87, put the total population at 300,000 at this date, implying urbanization of 0.3%. We use this estimate. Chandler does not have any relevant data.

For Britain, including Ireland, the total urban population according to Chandler was 15,000. Using the population for the British Isles minus Ireland (2 million, from McEvedy and Jones p.49), gives urbanization of 0.8%. Urbanization was 1.7% in 1200, 1.9% in 1300, and 1.4% in 1400. BBC estimate the total urban population for the British Isles minus Ireland was 97,000 (total population for this area was 1.7 million), implying urbanization of 5.7%.

BBC estimate the urban population of Ireland was 19,000 in 1000 AD. McEvedy and Jones put total population in 1000 AD at 300,000, implying urbanization of 6.3%. We use this estimate. Chandler does not give any data for Ireland at this time.

For Bulgaria, Chandler estimates that total urban population was 80,000 in 1000. With total population of 800,000, urbanization was 10%. Urbanization is estimated to have been 6.8% in 800, 5.5% in 1200, 5.5% in 1300 and 1400 in 4.2%. BBC estimate the urban population was 70,000, implying urbanization of 8.8%.

For Greece, Chandler estimates total urban population was 40,000 in 1000 AD. With total population of 1 million, this implies urbanization of 4%. Urbanization was 6.9% in 800, 2.5% in 1200, 6% in 1300 and 7.7% in 1400. BBC estimate the urban population was 55,000, implying urbanization of 5.5%.

For Denmark, Chandler estimates the urban population was 15,000. With total population estimated at 400,000 (McEvedy and Jones), the implied urbanization estimate is therefore 3.8%. Urbanization was 3.8% in 1200, 4.4% in 1300 and 5.2% in 1400. BBC do not have data for Denmark in 1000. We use the BBC-equivalent Chandler estimate, which is 5.3%.

We use the population for Czechoslovakia, which was 1.25 million according to McEvedy and Jones, p.85 (this may give lower urbanization numbers than is reasonable, because Slovakia had lower urbanization and was not so closely linked to the Czech Republic at this time.) Chandler's estimate of urban population for 1000 AD is 15,000 (all in Prague). With total population of 1.25m, this implies urbanization of 1.2%. Urbanization in 1200 was 1.1%, urbanization in 1300 was 1.3% and urbanization in 1400 was 3.8%. BBC report urban population for Czechoslovakia in 1000 AD of 10,000, giving urbanization of 0.8%. We use the BBC estimate for the Czech Republic and Slovakia.

Chandler estimates total urban population in Hungary as 15,000 in 1000 AD. McEvedy and Jones put total population at 500,000, implying urbanization of 3%.
Urbanization was 4.1% in 1200 and 1.2% in 1300. We do not have data for 1400. BBC put the total urban population at 2,000 in 1000AD, implying urbanization of 0.4%. We use the BBC estimate.

In Poland, Chandler estimates the total urban population was 30,000 in 1000 AD, implying urbanization of 2.4% (McEvedy and Jones estimate the total population was 1.25m). In 1200 urbanization was 1.3%, for 1300 we do not have data and in 1400 it was 0.7%. BBC have no data on urbanization in Poland in 1000. We use the BBC-equivalent Chandler estimate, which is 3.74%.

In the area now comprising Russia, Ukraine and Moldova, Chandler estimates the urban population was 88,000 (of which 45,000 in Kiev). We deflate this using the population of Russia-in-Europe (McEvedy and Jones, p.79), which was 4 million, giving a total urbanization level of 2.2%. Urbanization was 0.6% in 800, 1.4% in 1200, 3% in 1300 and 2.8% in 1400. BBC put urban population in 1000 AD for this area at 152,000, implying urbanization of 3.8%. We use this for Russia, Ukraine and Moldova.

Chandler does not have urbanization estimates for Switzerland, Estonia, Lithuania, Latvia, Romania, Serbia, Macedonia, other parts of Yugoslavia, Albania, Finland, Norway or Sweden in 1000 AD.

BBC does not have estimates for these countries, with one exception. BBC estimates urban population in the former Yugoslavia was 55,000. McEvedy and Jones, p.113, estimate that the total population living in this area was 1.25m, which implies that urbanization was 4.4%. We use the BBC estimate for all countries that emerged from the former Yugoslavia.

Where both Chandler and BBC do not give data, and there is no nearby urban population in neighboring countries, we assume urbanization was zero. The exception is the Baltics, which may or may not reasonably be considered part of Russia in 1000 AD. As these are small countries, we prefer to leave their data as missing.

We use an estimate of zero urbanization for Finland, Norway, Sweden, and Romania. This seems quite consistent with the evidence from a variety of sources, for example de Vries.

Asia in 1000

Chandler's estimates of urbanization in Asia are almost certainly underestimates because he uses a minimum threshold of 40,000 for 1000 AD. We use the same conversion factor (from the regression of BBC on Chandler) for our base estimates. All our estimates for Asia are BBC-equivalent urbanization rates formed by thus converting Chandler's data.

For the area now occupied by Turkey in 1000 (the empire of Byzantium), Chandler estimates total urban population was 455,000. From McEvedy and Jones, using data for Turkey-in-Asia and Turkey-in-Europe, we have total population of 7.3 million. This implies an urbanization rate of 6.2%. Urbanization from Chandler was 4.6% in 800, 4.5% in 1200, 3% in 1300, and 4.8% in 1400. We use a BBC-equivalent estimate of 8%.

For Iraq, Chandler estimates the total urban population of Iraq to have been 175,000 in 1000 AD. McEvedy and Jones estimate total population to have been 2m, implying urbanization of 8.8%. Urbanization was 36.8% in 800 (when Baghdad was the capital of a large empire), 6.7% in 1200, 4% in 1300 and 9% in 1400. We use a BBC-equivalent estimate of 11%.

For Iran, urban population, according to Chandler, in 1000 AD was 427,000. Total population, according to McEvedy and Jones, was 4.5m, implying urbanization of 9.5%. Urbanization in 800 was 4.5%, 5.8% in 1200, 8.6% in 1300, and 9.4% in 1400. We use a BBC-equivalent estimate of 11.8%.

Chandler estimates that the urban population of Saudi Arabia in 1000 was 110,000. Given McEvedy and Jones' population estimate of 5m, this implies urbanization of 2.4%. Urbanization was 1.6% in 800 and we do not have any other information before 1600. We use a BBC-equivalent estimate of 3.7%.

To assess urbanization in Central Asia we use the estimate of population for Russian Turkestan, from McEvedy and Jones, p.163 – 2.5 million. Chandler estimates the urban population of this area was 270,000 (Samarkand, Bokhara and Balasaghun), implying urbanization of 7.6%. Urbanization was 7.3% in 800, 5.9% in 1200, and 5.2% in 1400. Eggimann does not have data for Central Asia in 1500, so we use Chandler's data (urban population of 110,000 in a total population of 3.5 million), which implies 3.1% urbanization. We use a BBC-equivalent estimate of 9.6% for both Kazakhstan and Uzbekistan in 1000 and 4.5% in 1500.

The urban population of Syria, Lebanon and Palestine was 140,000 in 1000 AD, according to Chandler. The total population of this same area (McEvedy and Jones, p.139 and p.143) was 2m, implying urbanization of 7%. Urbanization was 4.5% in 800, 7.8% in 1200, 2.8% in 1300 and 7.1% in 1400. We use a BBC-equivalent estimate of 8.9%.

The urban population of Pakistan, India and Bangladesh was 627,000 in 1000 AD, while its total population was 77 million (McEvedy and Jones, p.185), giving urbanization of 0.8%. Urbanization was 0.8% in 800, 0.8% in 1200, 0.8% in 1300, and 1.2% in 1400. These estimates are almost certainly much too low. For example, they imply that urbanization in 1500 was only 1.8%, whereas the evidence suggests it was much higher – probably at least 8%. Chandler's method seems to be particularly biased for larger countries (see also the problems with China below). We use his numbers here, however, as we are primarily interested in the change between 1000 AD and 1500 AD using the same methodology.

To assess urbanization in Indonesia and Malaysia we use the estimate of the total population of the Malay Archipelago, from McEvedy and Jones, p.199. Chandler provides no data for 1000 AD. For 800 AD, he estimates total urban population was 60,000. With total population of 3.5 million, this implied urbanization of 1.7%.

In 1300, urbanization was 0.7% and in 1400 it was 1.1%. We use the 800 AD estimate and convert this to a BBC equivalent number of 3%.

Chandler estimates the urban population of Cambodia was 14,000 in 1000 AD. McEvedy and Jones estimate total population was 1 million in the same year, implying urbanization of 14%. Urbanization was 5% in 800 AD, 13.6% in 1200, 7.5% in 1300 and 3.1% in 1300. We use a BBC-equivalent estimate of 16.9.

The area occupied by "China Proper" (excluding Turkestan and Tibet) contained 60 million people in 1000 AD (McEvedy and Jones, p.171). Chandler estimates the total urban population in this area was 1.504 million, implying urbanization of 2.5%. Measured in this way, urbanization was 2.9% in 800, 1.4% in 1200, 2.4% in 1300, and 2.8% in 1400. We use a BBC-equivalent urbanization estimate of 3.9%.

Note that this is almost certainly an underestimate, as in the case of India. Given that the urban population of China in 1500 is underestimated for the same reason (both Eggimann and Chandler count only large cities), this should still allow us to check for persistence, i.e., if urbanization either fell or rose dramatically, it would probably be picked up in these numbers. If anything, the use of a relatively high threshold – like 40,000 inhabitants – should exaggerate the fluctuations in urban population. We should exercise care, however, when comparing large countries such as China and India with smaller countries.

Chandler estimates that the urban population of Korea was 110,000 in 1000 AD. McEvedy and Jones put total population on that date at 2.5 million, implying urbanization of 4.4%. Urbanization was 2% in 800 and 2.9% in 1400. We use a BBC-equivalent urbanization estimate of 6%.

The urban population of Japan in 1000 AD was, according to Chandler, 175,000. Total population from McEvedy and Jones, was 4.5 million, implying urbanization of 3.9%. Urbanization was 5% in 800, 3.7% in 1200, 2.5% in 1300, and 1.5% in 1400. Japan is one the few countries in which urbanization in 1500 (0.2%) was much lower than in 1000 AD – the size of this discrepancy was temporary and largely the result of civil war. We use a BBC-equivalent estimate of 5.4%.

The urban population of Burma/Myanmar in 1000 AD was 40,000 (Chandler). The total population was 2 million (McEvedy and Jones), implying urbanization of 2%. Urbanization was 7.3% in 1200 and 3.1% in 1400. We use a BBC-equivalent estimate of 3.3%. Note that while we have urbanization data on Myanmar in 1000 and 1500 but not data on GDP per capita today. This country therefore appears in graphs and regressions of urbanization in 1500 on urbanization in 1000 but not in our base sample.

Chandler estimates the urban population of Afghanistan was 60,000, while McEvedy and Jones put total population at 2.25 million in 1000 AD. This puts the "pure Chandler" urbanization estimate at 2.7% and we use a BBC-equivalent estimate of 4.1%. No other information on urbanization before 1500 is available.

Urbanization in Sri Lanka was 7.5% in 1200 – urban population of 75,000 from Chandler and total population of 1 million from McEvedy and Jones. No other data on urbanization before 1500 is available. We treat this as missing data for 1000 AD.

Urbanization in Georgia was 8% in 1200 (urban population of 80,000 and total population of 1 million). No other data is available on urbanization before 1500. We treat this as missing data for 1000 AD.

We have no data on urbanization in Vietnam until 1300. At that date, the urban population, according to Chandler was 40,000. McEvedy and Jones' total population estimate of 1 million implies an urbanization estimate of 4%. We treat this as missing data for 1000 AD.

We also have no data on urbanization in Thailand until 1300. At that date, the urban population (Chandler) was 40,000 and the total population was 1.5 million, implying urbanization of 2.7%. Urbanization was 2.3% in 1400. We treat this as missing data for 1000 AD.

We have no data on urbanization in Nepal or the Philippines on or around 1000 AD. We have no information about urbanization in Laos in or around 1000 AD. We treat this as missing data for 1000 AD.

We assume urbanization in Hong Kong and Singapore was 3% in 1000, following the same reasoning as in our previous appendix on 1500.

America in 1000

For Mexico, Chandler reports an urban population in 1000 AD of 170,000. McEvedy and Jones put total population at 3 million, implying urbanization of 5.7%. Urbanization was 6% in 800, 3.1% in 1200, 2.8% in 1300 and 4.9% in 1400. We convert Chandler's estimate to a BBC-equivalent estimate of 7.5% and use this.

For Peru, Chandler's urban population estimate in 1000 AD is 40,000. McEvedy and Jones put total population at 1.5 million, implying urbanization of 2.7%. Calculated in this way, urbanization was 1.6% in 800, 2.3% in 1200, 3.45 in 1300 and 5.3% in 1400. We use a BBC-equivalent estimate of 4.1%.

Chandler estimates that the urban population of Bolivia was 20,000 in 800 and 40,000 in 1200, but he does not give an estimate for 1000 AD. McEvedy and Jones put population at 500,000 in 800 and 800,000 in 1200, implying urbanization of 4% and 5% respectively. An estimate of 3% Chandler-equivalent urbanization in 1000 therefore seems reasonable. We use a BBC-equivalent estimate of 4.4%.

The earliest urbanization data we have for Ecuador is 1200, for which date Chandler estimates an urban population of 40,000. McEvedy and Jones suggest that total population was half a million, which would imply that urbanization was 8%. We use this estimate also for 1000 AD, converting it to a BBC-equivalent estimate of 10.1%. Urbanization was 16% (using Chandler's measure) in 1300, and 18% in 1400.

According to Chandler, the area now consisting of central America (particularly Guatemala) had an urban population of 100,000 in 800 (at the height of the Mayan civilization) and 45,000 in 1200. McEvedy and Jones put the population of Central America at 500,000 and 600,000 respectively at those dates, implying urbanization of 20% at the earlier date and 7.5% at the later date. Unfortunately, we do not have data on urbanization in 1000. To be on the safe side, given that we know the Mayan civilization collapsed suddenly before 1000, we urbanization from 1200. We convert this to a BBC-equivalent estimate of 9.5.

Chandler does not report urbanization data for Argentina, Brazil, Chile, Canada, Paraguay, Uruguay or the United States before 1500. For all these countries, it is safe to assume that urbanization was essentially zero in 1000. Chandler also does not report data for any part of the Caribbean, Colombia, Guyana, or Venezuela. We have reason to think that urbanization may have been above zero in some of these cases, but we assume zero urbanization for the sake of consistency (and because it goes against our hypothesis.)

Africa in 1000

As is the case for 1500, the data on North Africa is reasonable and probably not subject to greater error than other non-European data. We have information on urbanization in 1000 for all North African countries, except Libya.

For Egypt in 1000, Chandler reports an urban population of 285,000. McEvedy and Jones estimate the total population was 5 million at this date, implying urbanization of 5.7%. Urbanization, calculated from these sources, was 3.9% in 800, 10.6% in 1200, 13.5% in 1300, and 14.5% in 1400. We use a BBC-equivalent estimate of 7.5%.

For Tunisia, Chandler puts the urban population in 1000 at 60,000 and McEvedy and Jones estimate total population was 1 million. We convert this estimate of urbanization, 6%, into a BBC-equivalent estimate of 7.8%. Urbanization, in Chandler units, was 12.5% in 800, 6% in 1200, 7% in 1300 and 8.8% in 1400.

Chandler estimates the urban population of Algeria was 20,000 in 1000 AD. McEvedy and Jones put total population at 2 million, implying urbanization of 1%. Urbanization was 1.6% in 800, 2% in 1200, 7% in 1300 and 9.1% in 1400. We use a BBC-equivalent estimate of 2.2%.

For Morocco, McEvedy and Jones put total population 1000 at 2 million and Chandler puts the urban population at 100,000, implying urbanization of 5%. The same combination of sources gives urbanization of 3.25 in 800, 23% in 1200, 17.3% in 1300, and 14.8% in 1400. We use a BBC-equivalent estimate of 6.7%.

As previously discussed, the data on urbanization in sub-Saharan Africa is fragmentary, of dubious value and almost certainly underestimates with biases that are hard to assess. The data on 1000 AD is worse than for 1500. However, there is some relevant information on the persistence of urbanization (assuming similar biases in estimates for 1000 and 1500). All the information reported here comes from combining Chandler and McEvedy and Jones. As we do not calculate urbanization in 1500, we do not convert into BBC-equivalent measures (and this information on sub-Saharan Africa is not used in our graphs or regressions.)

For the Sudan, our sources imply urbanization of 0.8% in 800, 1.8% in 1000, 1.7% in 1200, 1.4% in 1300, 0.75 in 1400, and 0.5% in 1500. For Ethiopia, we have no data on urbanization in 1000 AD, but estimates of 2.5% for 800, 1.3% for 1200, 1.3% for 1300, 1.7% for 1400 and 2.7% for 1500. For Mali, urbanization was 4% in 800, 2% in 1000, 2.6% in 1200, 6.6% in 1300, 6.2% in 1400, and 7.5% in 1500. For Zimbabwe, urbanization is estimated to have been 5% in 800, 4% in 1000, 3.3% in 1200, 3.6% in 1300, 3.1% in 1400 and 2% in 1500. For coastal West Africa (using the population estimates for West Africa, from McEvedy and Jones, p.243), with most of the urbanization in and around Nigeria, we have urbanization in 800 and 1000 of 0.4%, 1.1% in 1200, 1.3% in 1300, 1.8% in 1400 and 2.65 in 1500.

We have no other urbanization data for African countries on or around 1000 AD.

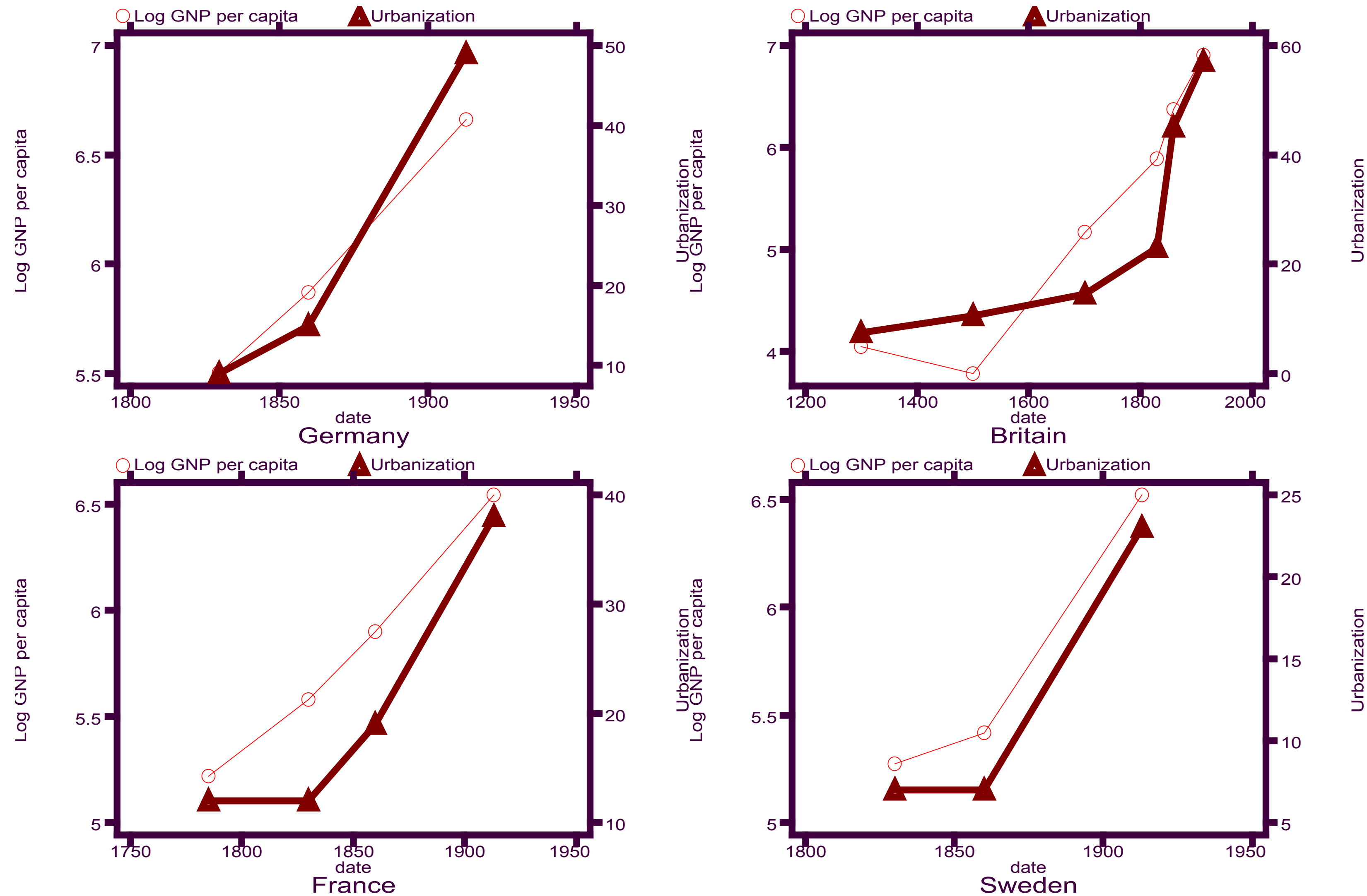
The Pacific Region in 1000

We assume zero urbanization in 1000 for Australia, New Zealand. This does not seem controversial. There is no sign of significant urban structures in Australia before 1800. In New Zealand, there may have been some very small scale urbanization, but even if we increased the urbanization estimate to 3% (as for 1500), this would not change any of our results.

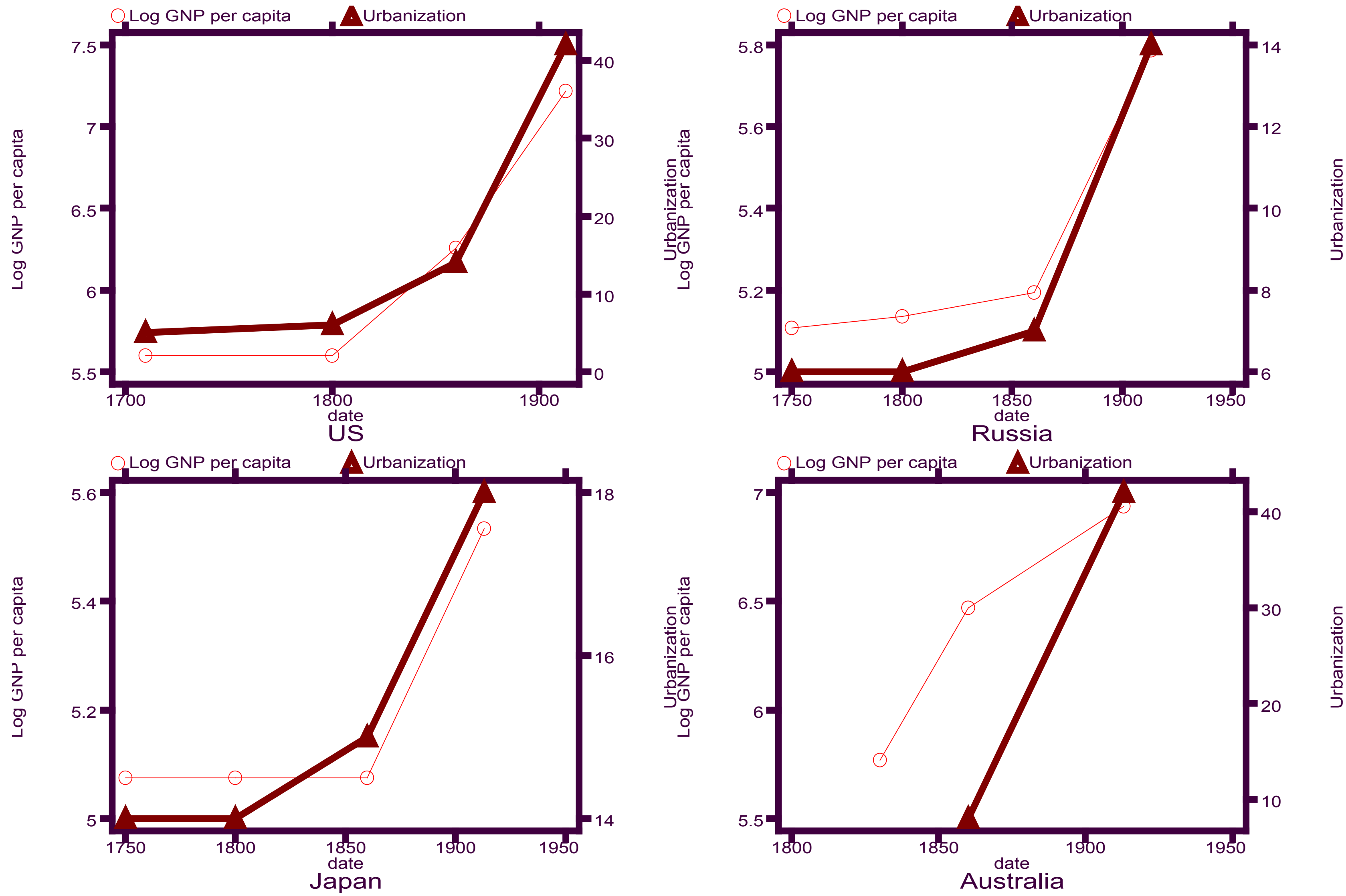
Missing Data for 1000

Given the state of knowledge about urbanization, we have less data on 1000 AD than we do on 1500 AD. In particular, reliable data about the Caribbean, Papua New Guinea and Philippines are missing. Sri Lanka, Vietnam, Laos, Libya are also missing data. However, for all of these countries, there is no reason for believe there was a big change in urbanization rates between 1000 and 1500.

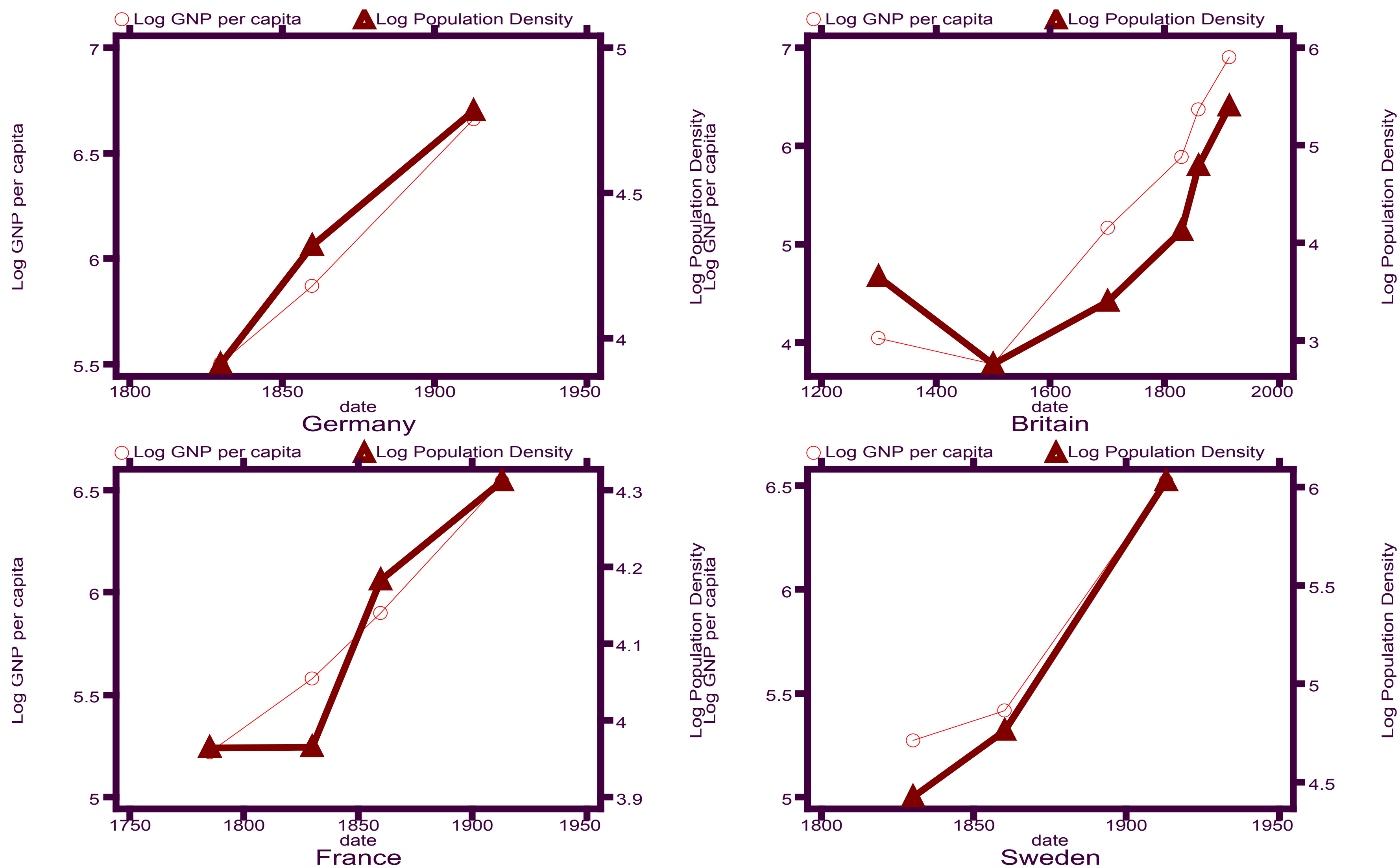
Appendix Figure A1



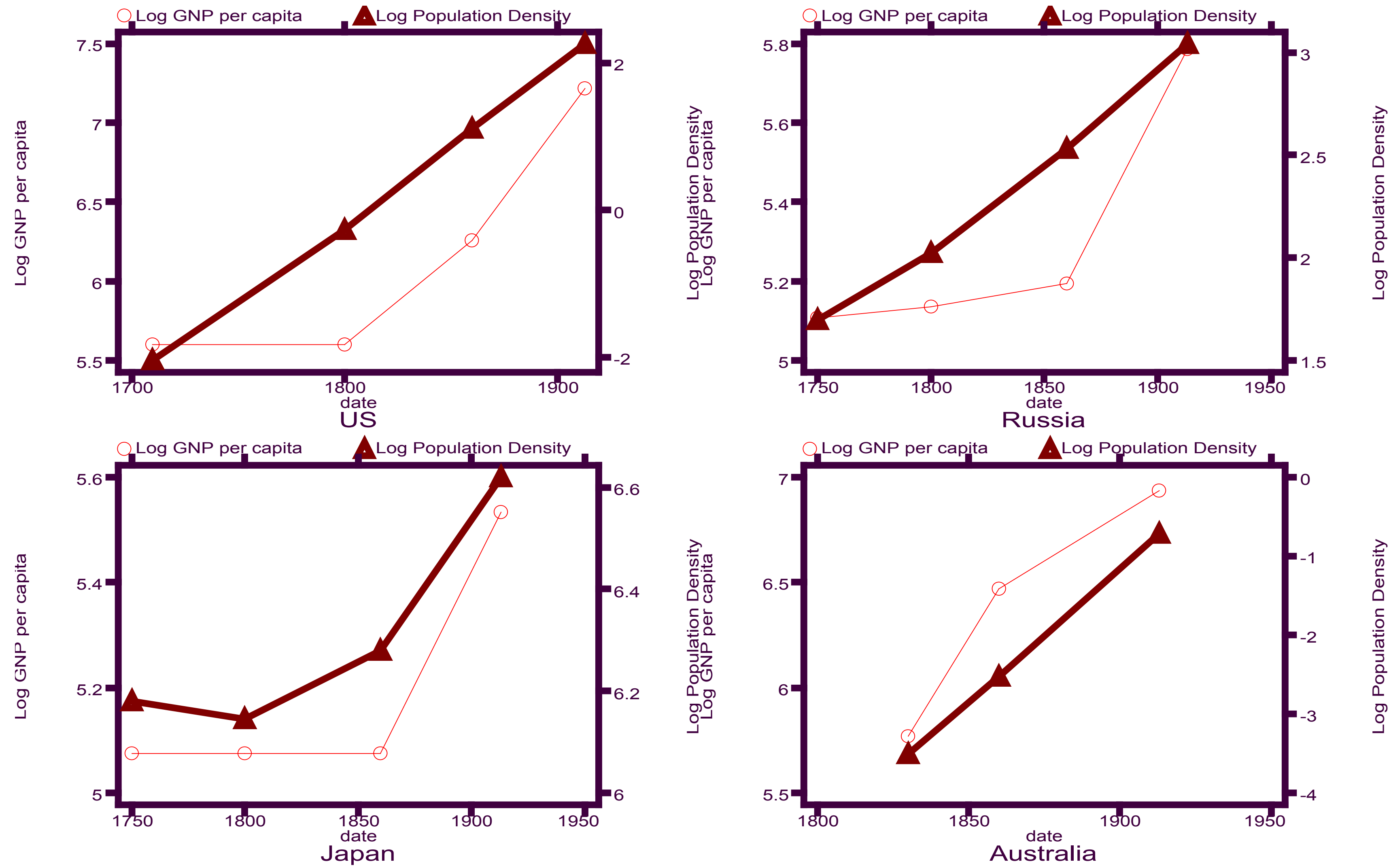
Appendix Figure A2



Appendix Figure A3



Appendix Figure A4



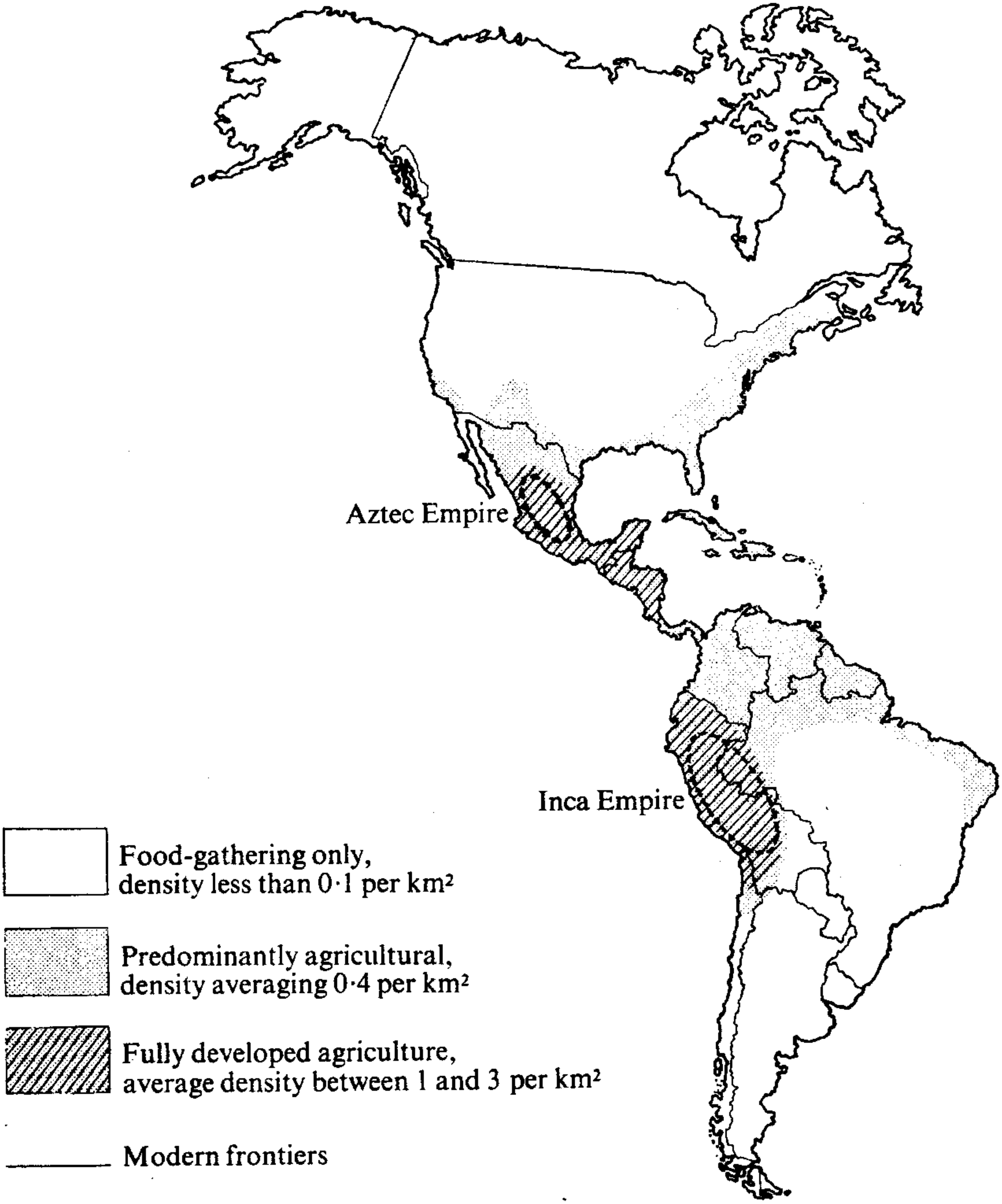


Fig. 4.3 *The Americas, agricultural development and population densities in* AD 1500