

The Roots of the Industrial Revolution: Political Institutions or (Socially Embedded) Know-How?

Scott F Abramson* & Carles Boix†

February 7, 2013

Abstract

In this paper we reassess the literature of growth by looking at the evolution of the European economy from around 1200 to 1900. Employing a comprehensive dataset for the European continent that includes geographic and climate features (1200-1800), urbanization data (1200-1800), per capita income data in the second half of the 19th century, location of proto-industrial centers (textile and metal sectors from 1300 to the Industrial Revolution), political borders and political institutions, we estimate the geographic, economic and political covariates of urbanization (commonly used as a proxy for per capita income) and 19th-century per capita income. We show that the process of economic take-off (and of a growing economic divergence across the European continent) was caused by the early emergence and growth of cities and urban clusters in an European north-south corridor that broadly runs from southern England to northern Italy. In contrast to previous findings in the institutionalist literature, we then show that the fortunes of parliamentary institutions in early modern Europe played a small part in the success of the industrial revolution and the distribution of income across the continent in late 19th century. Rather, industrialization took place in those territories that had a strong proto-industrial base, often regardless of the absence of executive constraints (in the two centuries preceding the industrial revolution).

*Department of Politics, Princeton University. sabramso@princeton.edu

†Department of Politics & WWS, Princeton University. cboix@princeton.edu

The economic history of the last 1,000 years is one of growing income divergence across the world. According to admittedly crude estimates, GDP per capita ranged from a maximum of \$600 to a minimum of \$400 (in 1990 International Geary-Khamis dollars) across the world and from \$468 to \$400 in Europe around the year 1,000 (Maddison 2003). Half a millennium later, that range had widened in Europe from a minimum of \$433 in Albania to a maximum of \$1,100 in Italy.¹ By contrast, it had hardly changed outside of Europe. By 1850, with the industrial revolution in full swing in England, the maximum European per capita income had experienced a five-fold increase to almost \$2,500 (Figure 1). With the exception of the English settler colonies, per capita incomes in the rest of the world, which had not experienced any significant growth, were similar to the poorest areas of Europe.

There is no dearth of theories to explain how (parts of) Europe managed to escape from a pre-industrial, Malthusian world through a long period of sustained growth. Endogenous growth theories see economic development as the outcome of a self-generating process of learning-by-doing and of technological innovation, arguably shaped by geographical factors (Sachs and Warner 1997) and/or population size (Kremer 1993). Most of the current literature, however, traces growth back to a particular configuration of institutions ranging from the presence of formal structures such as constitutional checks and balances and constraints on the executive (North and Weingast 1989) to social norms of cooperation and culture (North 1990, Putnam 1993).

Both strands of the literature are beset by key theoretical gaps. Institutionalists do not endogenize the emergence of institutions, mostly tracing them back to some unidentified critical historical juncture.² Endogenous growth models do not specify with much precision the mechanisms through which economic growth takes place. In addition, the existing empirical evidence on the causes of modern growth should be seen as tentative at best. Most historical work on modern growth has been conducted by economic historians who have emphasized the analysis of a particular case or

¹Urbanization rates, which have regularly been employed in the literature to proxy for economic development tell a similar story. Chanda and Putterman (2007) estimate GDP per capita as a function of urbanization for a cross-country sample in 1500. Employing their point estimates and applying them to urbanization rates, the highest per capita income in Europe was around \$600 in 1000 and around \$780 in 1300.

²Two partial exceptions are Engerman and Sokoloff (2002) and Stasavage (2003) who related specific institutional frameworks to the social coalition in power.

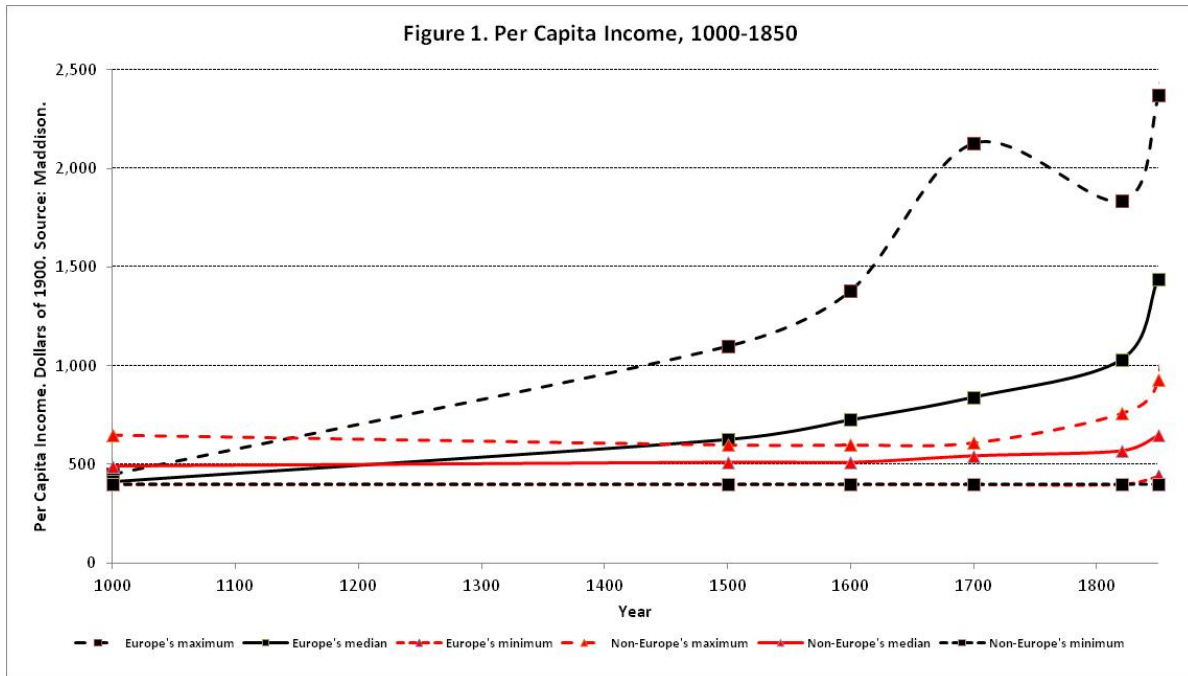


Figure 1

the comparison of a few cases with Britain normally included as a benchmark (Pomeranz 2001, Voigtländer and Voth 2006, Clark 2008). Econometric studies, generally based on cross-sectional models, have been less common (De Long and Shleifer 1993, Acemoglu et al. 2002; 2004, Hibbs and Olsson 2004, Chanda and Putterman 2007). These studies have made important empirical contributions. But their measurement of economic data and institutions has been incomplete, limited to the modern historical period starting in 1500 or 1800 (missing a process of economic divergence that started at some earlier point in time) and at times conceptually defective. Moreover, they have not modeled the choice of institutions (and their potentially endogenous relationship to the economy).

In this paper we reassess the literature of growth by looking at the evolution of the European economy from around 1200 (or the onset of the commercial and technological innovations that transformed Europe) to 1900 (a time at which the industrial revolution had been in place for about a century) using a new and wide-ranging dataset. Employing geospatial data coding techniques, we construct a comprehensive dataset for the European continent that includes: geographic and climate

features (1200-1800), urbanization data (1200-1800), per capita income data in the second half of the 19th century, location of proto-industrial centers (textile and metal sectors) and knowledge-producing institutions (universities) from 1300 to the Industrial Revolution, political borders and political institutions. All the data are calculated approximately at $225 \text{ km} \times 225 \text{ km}$ grid-square units as well as for independent and semi-independent political units (such as Genoa, Venice, France or Sicily). We then estimate the geographic, economic and political covariates of urbanization (commonly used as a proxy for per capita income) and 19th-century per capita income. Moreover, we assess causal relationships between urbanization and our political-economic outcomes of interest with an instrumental variables approach, in part exploiting random climatic variation across time and space in the propensity of territory to support large, urban, populations.

We claim that the process of economic take-off (and of a growing economic divergence across the European continent) was related to the emergence and growth of cities and urban clusters in an European north-south corridor that broadly runs from southern England to northern Italy. As shown in Figure 2, which plots the bivariate relationship between the total urban population in each geographical quadrant in 1200 or 1500 and the total urban population at a later time, their location exhibited strong continuities throughout the late medieval and modern era.

Once a “long period of migration, invasion, and conquest (Strayer 1973; p. 16)” spanning from about 400 to around 900 ended and the European continent gradually stabilized, cities flourished in highly productive agricultural lands as well as on those regions that had cheap communication and transportation waterways. Those areas that could sustain a growing non-farming population specialized, in line with endogenous growth models, in the development of a set of proto-manufacturing sectors. Indeed, the correlation matrix in Figure 3 shows that European urban areas were dense in textile and metal production centers. It also attests to the strong relationship between urban life and the emergence of knowledge-production institutions such as universities. All these features set European cities apart from non-European large towns, which were the seat of imperial bureaucracies and rent-seeking elites. Moreover, Benefiting from increasing returns to scale due to sector- and location-specific positive agglomeration externalities as modeled by recent economic geography research (Krugman 1991), more heavily urbanized areas at the beginning of the period tended to add

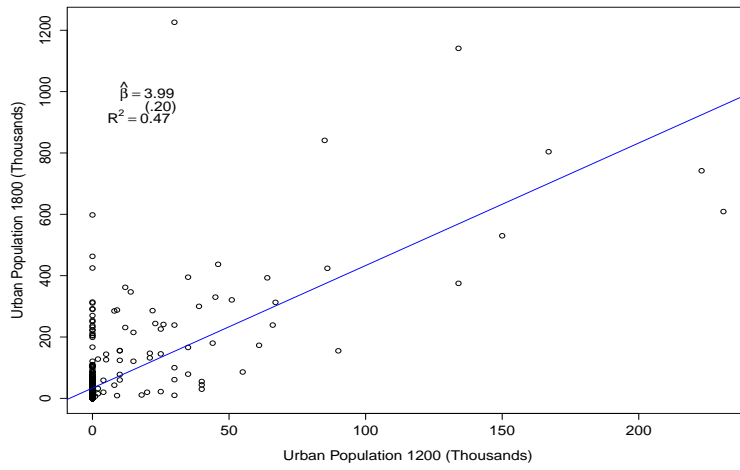
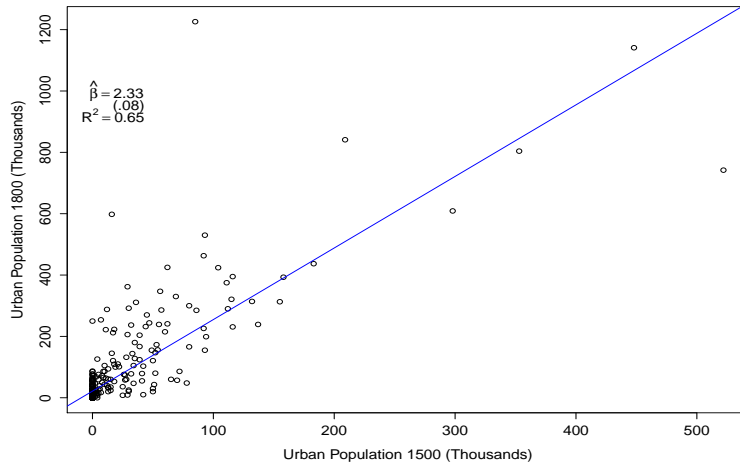
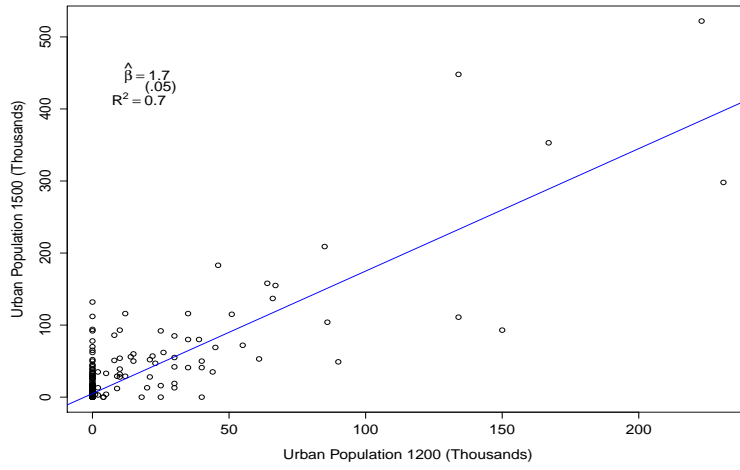


Figure 2: Bivariate relationship between urban population and future urban population across time.

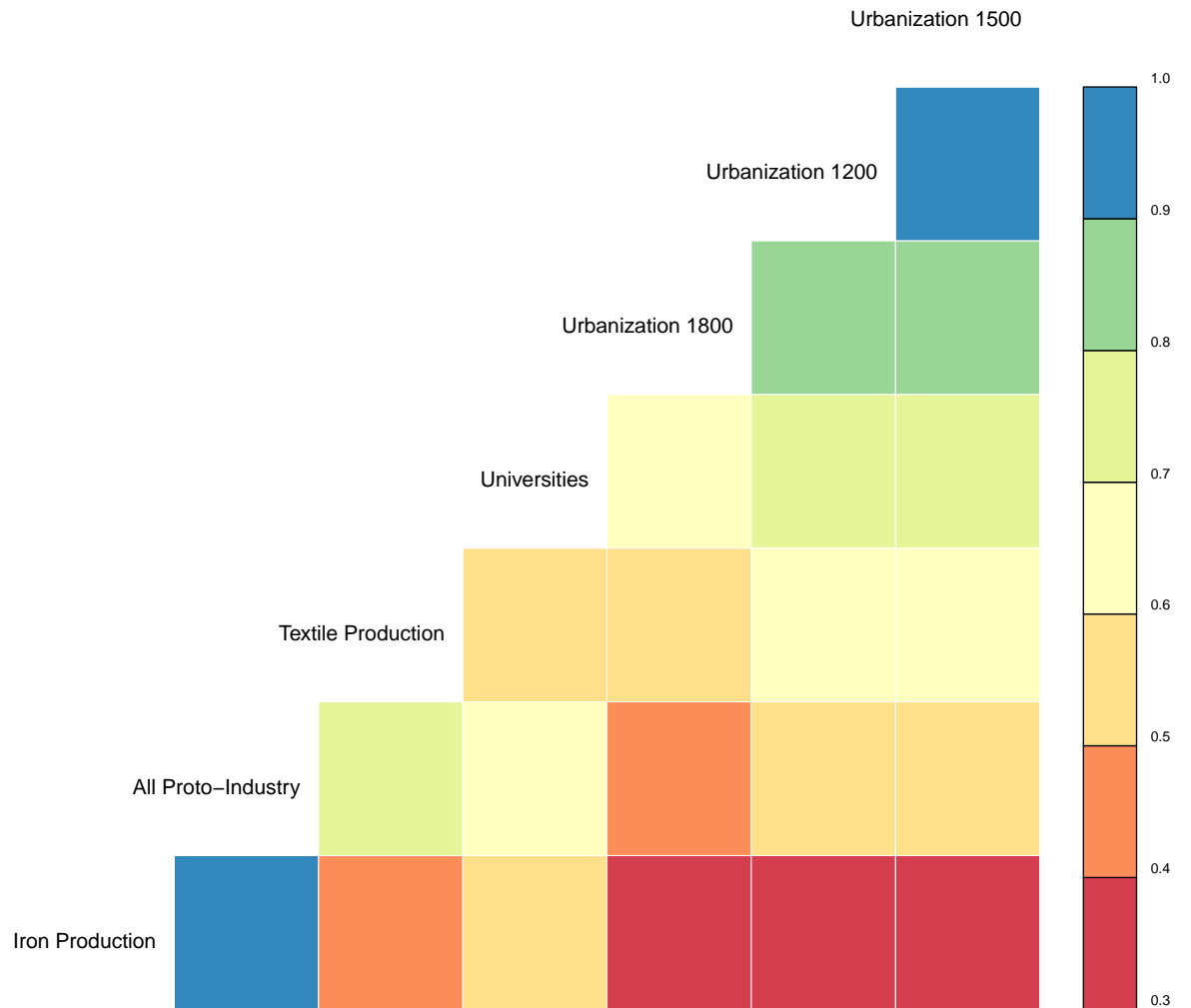


Figure 3: Correlation matrix of urbanization in 1200, 1500, and 1800 along with proto-industrial production centers and universities.

new population at a faster rate than areas that were mainly rural in the Middle Ages. To give a sense of the magnitude of that divergence, Figure 4 reports the estimated difference in logged urban population between two areas from 1200 until 1800 that had an initial urban population of 1,000 and 24,000 respectively, a one standard deviation change for the year 1200. That figure is calculated by estimating are derived from the dynamic system GMM estimates of Model 1 (reported in Table 2), employing the coefficient on the lagged value to obtain an estimate for each period and then taking the difference of these estimates. Whereas in 1200 the difference was 23,000, six hundred years later the estimated difference is predicted to become just less than 468,000.³

Once the urbanization and proto-industrialization waves were well under way, pluralistic political institutions tended to spread, as permanent bodies of governance (in the form of city councils or territorial assemblies with stronger urban representation), across Europe. With the intensification of war competition in Europe after 1500 (and the consolidation of a few great powers), those institutions generally collapsed. They only remained in place, if at all, in the most proto-capitalist enclaves of modern Europe, although usually representing a much narrower oligarchy than in the late medieval period. However, contrary to the existing institutionalist literature, we find that the fortunes of parliamentary institutions in late modern Europe played a small part in the success of the industrial revolution and the distribution of income across the continent in late 19th century. Industrialization took place in those territories that had a strong proto-industrial base, often regardless of the absence of executive constraints (in the two centuries preceding the industrial revolution). In short, economic growth was only possible when there was a population of craftsmen who embodied a given stock of technological know-how that enabled them to take advantage of the technological breakthroughs of the 18th-century.

The plan of the paper is as follows. Section 1 reviews the two dominant explanations of development (endogenous growth models and neoinstitutionalism) and relates them to our explanation of the sources of Europe's development. Section 2 describes the data employed in the paper. Section 3 reports our empirical results: it shows the impact of early urbanization on subsequent urban-

³This also fits with recent evidence documenting a process of divergence in living standards between northwest Europe and eastern and southern Europe since the Middle Ages (Allen 2001).

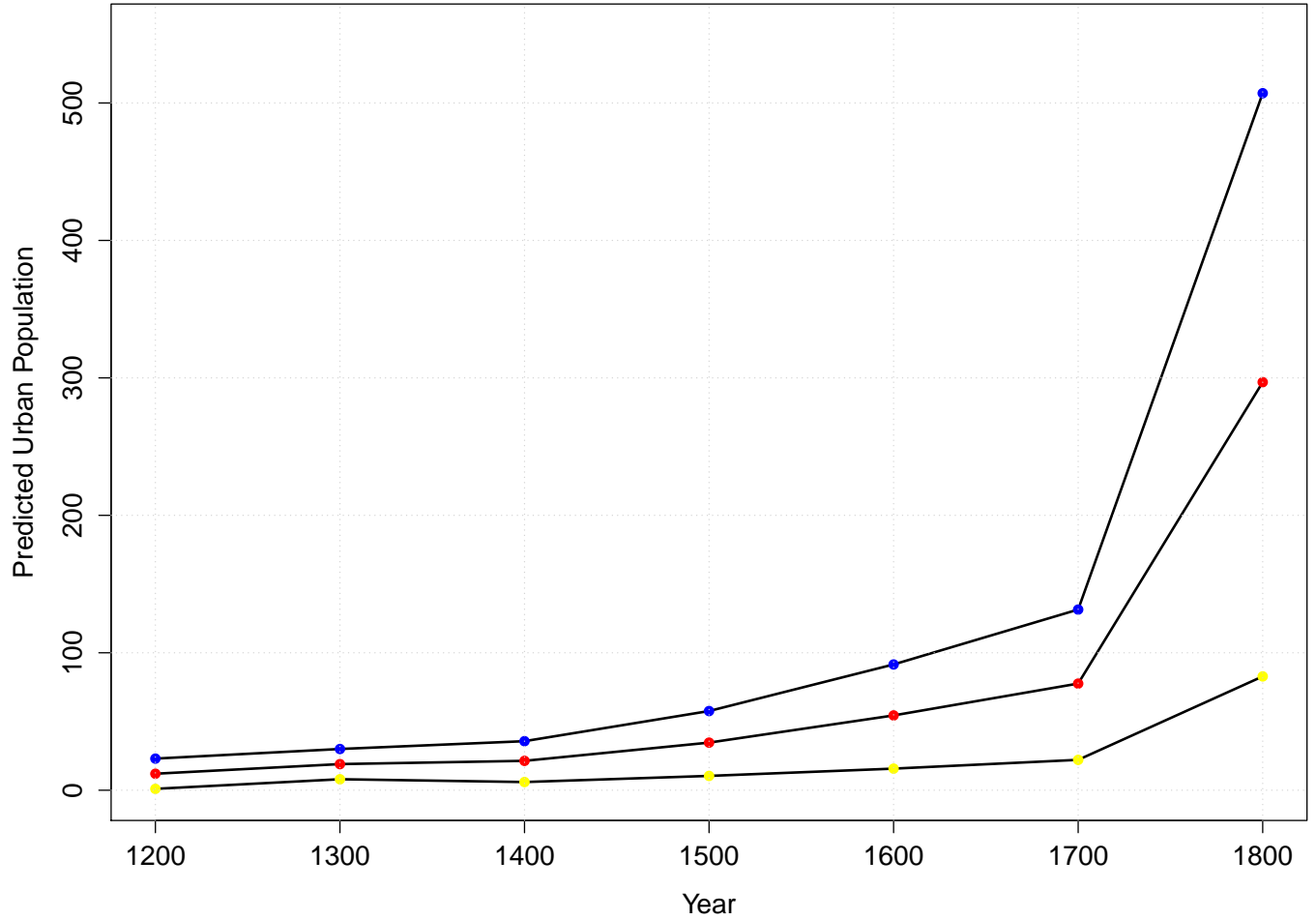


Figure 4: Estimates of the difference in urban population across time derived from the dynamic system GMM estimates of Model 1 with two lags. The predicted values are constructed using starting urban populations for two units of 1,000, 12,000 and 24,000 in the year 1200, approximately a one half, and one standard deviation difference for this period. The first lag is then constructed by increasing each unit by seven thousand, the mean increase across all units between 1200 and 1300. All successive estimates are derived from the estimates of the dynamic model

ization; it traces urbanization rates to a set of optimal climate conditions yielding an agricultural surplus that could support higher population densities in the core of Europe; it shows how high urban populations spurred a process of proto-industrialization in the textile and metal sectors; it then estimates the independent (and positive) effect of proto-industrialization on subsequent urban growth; it models the growth of parliaments; and it finally rejects the hypothesis that the presence of parliamentary institutions was a precondition for the development of Europe. Section 4 concludes by offering a theoretical interpretation of our results.

1 Theory

Besides standard accounting models of growth (Solow 1956), the scholarly literature offers two main alternative explanations of economic development: (1) institutionalist theories of growth and (2) endogenous growth models.

According to institutionalism, institutions – defined as “the rules of the game in a society or, more formally, the humanly devised constraints that shape human interactions – (North 1990; p. 3)– affect economic growth in two ways. In the first place, by defining property rights, they determine the private rate of return and hence investment decisions and the level of effort that individuals exert at technological innovation. More broadly, institutions affect transaction costs and therefore the extent to which individuals may be willing to engage in trade. As institutions reduce the costs of measurement and monitoring that are associated with any exchange, individuals have a higher incentive to specialize and trade with each other. In a Smithian economic framework, the process of economic specialization results in productivity gains and long-run growth (North 1990, Smith 1937 (1776)). The institutionalist theory of development has suggested three main institutional configurations leading to growth: (1) the existence of a stable political order guaranteed by the state (Olson 1993; 2000); (2) constitutional checks and balances to constrain the state and curb its incentives to exploit individual agents (North and Weingast 1989, De Long and Shleifer 1993); and (3) a stable set of norms of cooperation and “thick” trust, i.e. social capital, which reduces the incentives of individuals to take advantage of each other and empowers them to control state

institutions (Putnam 1993).

Endogenizing the formal and informal “rules of the game” (states, constitutions and social expectations over cooperation) is one of the crucial theoretical and empirical problems faced by the institutionalist literature. Depending on how we model the relationship between economic change and institutional change, institutional theories can be classified into “soft” and “hard” schools of thought. According to “soft” institutionalists, institutions adapt over time to technological change and changing relative prices. In other words, they are secondary to economic parameters and should be mainly thought of as intervening variables. North and Thomas (1973) offer a neoclassical rendition of this approach. Both modernization theory (Inkeles 1969, Fukuyama 1989) and classical Marxism fall also into this line of thought. Although most of modernization theory leaves the process of change unspecified, those scholars that attempt to model it tend to embrace some kind of functionalist theory of change. In Marx technological change and its carrying agents (embodied in the forces of production) assert themselves over the old relations of productions through a (normally violent) process of political conflict.

Since “soft” institutionalists cannot explain very prolonged instances of economic stagnation – for example, why Guatemala has been unable to catch up with Switzerland – “hard” institutionalists see institutions as exogenous parameters to the economy. Good institutions lead to growth. Bad institutions account for underdevelopment. In turn, the persistence of bad institutions is explained by the benefits they confer on a section of society that is able to block, through political means, any meaningful, pro-growth reform (Kuznets 1968, North et al. 1981, Mokyr 1990, Parente and Prescott 1994; 1999). Yet hard institutionalist theories run into their own set of problems. Given that the whole world was underdeveloped before the modern European breakthrough and since pre-industrial political and economic elites had an incentive to block change, how that change happened is hard to imagine. Accordingly, hard institutionalists do not model the origins of good institutions and tend to trace institutional change back either to the underlying distribution of endowments (Engerman and Sokoloff 1997; 2002) or to a set of not well defined critical junctures happening a few centuries ago (Anderson 1974, Acemoglu et al. 2002, Brenner 2003, Acemoglu et al. 2004).

The incompleteness of institutionalist theories of growth forces us to take a step back to rely on the endogenous growth literature, where development is generated by an internal feature of the economy. The latter is generally represented as having two sectors, a goods-producing one (in which capital and labor have constant returns to scale) and a knowledge or R&D sector where technological innovation takes place (and where the quantities of labor and capital employed to generate knowledge may experience increasing returns). Technological change and the long-run growth rate of output per worker are an increasing function of the size of the population and of the rate of population growth since, as the number of persons engaged in the R&D sector model, the number of technological innovations also increases (Rivera-Batiz and Romer 1991, Romer 1996, Kremer 1993).⁵

Endogenous growth models remain mostly silent about the determinants of population growth.⁶ However, a parallel strand of economic research has linked specific biological and geographic conditions to the rate of population growth and therefore to the rates of technological innovation and output growth (Sachs and Warner 1997). Naturally, biogeography features are unlikely to be the only conditions shaping the size and growth rate of population. Specific political and military shocks must do too.

Several waves of wars and massive migrations that started with the penetration of Barbarian populations in the late Roman empire and lasted until the Hungarian invasion and the Viking raids of the ninth and tenth centuries resulted in the collapse of urban life and the absence of any significant economic growth (Pirenne 1936, Randsborg 1991).⁷ Yet, as soon as those political shocks disap-

⁴According to Anderson (1974; p. 97-431) the modern breakthrough happened after urban and parliamentary institutions combined with a revival of the Roman law and “the reappropriation of virtually the whole cultural inheritance of the classical world” (426). For Brenner (2003) it was the post-1500 expansion of overseas trade that strengthened Europe’s emerging capitalist bourgeoisie. Taking a different position, Jones (1981) concludes that European growth resulted from chance (aided by some biogeographical factors) in a context of political fragmentation of the continent.

⁵In a variant of this general approach, technological innovation is sometimes modeled not as an outcome of making deliberate efforts at generating new ideas and tools but rather as a by-product of the production process itself. New knowledge accumulates as a result of some learning-by-doing: while producing goods and delivering services, economic agents come up with new ways to improve their production (Arrow 1962).

⁶For an exception, see Becker et al. (1990) who offers a growth model that endogenizes population choices

⁷The Arab expansion of the early Middle Ages pushed the European center of commercial activity northward and

peared and the European continent stabilized, European growth took place, as tested later, through an endogenous process of population growth and self-sustaining technological progress. European regions endowed with rich soils and optimal temperatures generated a large crop yield per hectare that allowed them to support high population densities and the formation of urban agglomerations. These conditions fostered, in turn, an environment where technological change proceeded at a faster pace than in less productive and less populated territories. Mineral endowments may have had similar effects. In addition, geography itself contributed to accelerate the process of urban agglomeration and technological innovation in the European core: proximity to cheap transportation means such as waterways augmented trade. As trade flourished, both the size of markets and the volume of population engaged in commercial and intellectual exchanges increased, leading to a higher rate of technological progress. In the presence of increasing returns to scale to knowledge and positive agglomeration externalities, that initial (and probably modest) variation in soil fertility and transportation costs across European regions resulted in much faster growth in the better-endowed territories and in a growing process of economic divergence between the central European corridor running from England to Northern Italy and the rest of the continent.

Although the European feudal order had already led to the formation of territorial assemblies representing the three medieval social orders (Duby 1968, Hintze 1975), the formation of towns spawned the emergence of pluralistic institutions mainly in the European core. During the late medieval and early modern period, the new commercial urban class played a key role in checking or at least funding the government through its participation in estates, diets and city councils. Yet, with the exception of a few enclaves across Europe, those proto-parliamentary institutions declined in power after the emergence of several large continental monarchies and a new international order after the treaty of Westphalia in 1648. According to our empirical results, parliamentary institutions had, contrary to the institutionalist literature, a small role in economic growth and the success of the industrial revolution. Economic development happened in heavily urbanized territories, rich in proto-manufacturing clusters – regardless of whether executive constraints were in place or not in approximately the two centuries that preceded the industrial revolution.

away from the eastern Mediterranean basin. Similarly, the fall of Constantinople in 1453 and the subsequent closing of the land trading routes to China shrank the size of markets in the Mediterranean basin.

2 Data

To examine the covariates of economic growth we employ as our observations two types of units: first, 225km-by-225km grid-scale units or quadrants that have some mass of land; and, second, political units that are either sovereign or semi-sovereign polities. Sovereign units are fully independent territories with their own executive (monarchical or not). Semi-sovereign units are those territories which, although they are under the control of a different state, retain some measure of political autonomy (defined by the existence of their own governing institutions or a special “colonial” institutions such as having a permanent viceroy). Examples of sovereign units are Portugal before 1580 and after 1640 or Venice until 1798. Examples of semi-sovereign units are Naples (after passing to the Crown of Aragon-Catalonia in 1444) or Valencia (member of the Catalan confederation and the Spanish kingdom) until 1707. Using semi-sovereign units allows us to employ smaller territories and more fine-grained data. More generally, coding our data at either the quadrant level or according to old borders minimizes a fundamental problem in studies that employ current sovereign countries as their main unit of analysis: the fact that political boundaries are endogenous to territorial economic conditions and factor endowment (Tilly 1990, Abramson 2012).

Our data coverage is broader than existing studies in two ways. Spatially we include Scandinavia and most of Eastern Europe. Temporarily we code our observations going back to 1200 approximately: most current studies instead employ historical panels that start at a moment in time when economic divergence has already taken place.

For each of these units we estimate the covariates of interest described below.

2.1 Economic Development

We first employ urbanization data to measure economic development. Bairoch et al. (1988) provide a comprehensive dataset with information on about 2,200 towns which had 5,000 or more inhabitants at some time between 800 and 1800. Urbanization rates have been usually taken as a reasonable proxy for economic development in the literature (Acemoglu et al. 2002, Chanda and

Putterman 2007). We construct two measures of urbanization. The first one is the number of cities with more than a given number of inhabitants (1,000, 5,000, 10,000 and 20,000 inhabitants) in each unit. The second one is the ratio of urban population over geographical size of the unit.⁸ When we employ polities as our observational units we only use the second measure of urbanization.

Figures 5, 6, and 7 represent the location of all the cities in the Bairoch dataset for 1200, 1500 and 1800 respectively. The diameter of each dot is proportional to population size. The three maps capture a continuous process of urban expansion over time. By 1200 an urbanized axis had emerged in the old Lotharingian kingdom, with cities mostly clustered in today's Benelux and in Northern Italy. The map also records the existence of a set of (by that time declining) towns in the southern half of the Iberian Peninsula. Three hundred years later the urban population had grown quite rapidly. According to Bairoch et al. (1988) 8.4 million Europeans lived in towns in 1300 or about 9.1% of total population. In 1500 Europe's urbanization rate was 10.3% in 1500, ranging from 29.5% in the Netherlands to 2.2% in Scandinavia. In 1800 the number of Europeans living in cities had more than doubled to about 23 million people and the urbanization rate had edged up slightly to 11.9%.

Besides Bairoch's urbanization data, we employ regional per capita income in 1870 across Europe. To construct this measure at the regional level, we rely on a growing number of new estimations of GDP and GDP per capita done at the subnational level by several economic historians, harmonized across countries using Maddison's per capita income data at the national level as a benchmark. Since most of the data is reported at the NUTS-2 level, we apply geospatial smoothing procedures to attribute the data to observational units. Appendix B describes the sources and procedure employed.

2.2 Urbanization and proto industrialization

Towns may embody a process of economic specialization, technological innovation and higher income. However, they may just be urban agglomerations where a rent-seeking class (served by a

⁸This second measure proxies the standard urbanization rate (urban population over total population), which cannot be estimated at a subnational level for lack of data on total population.

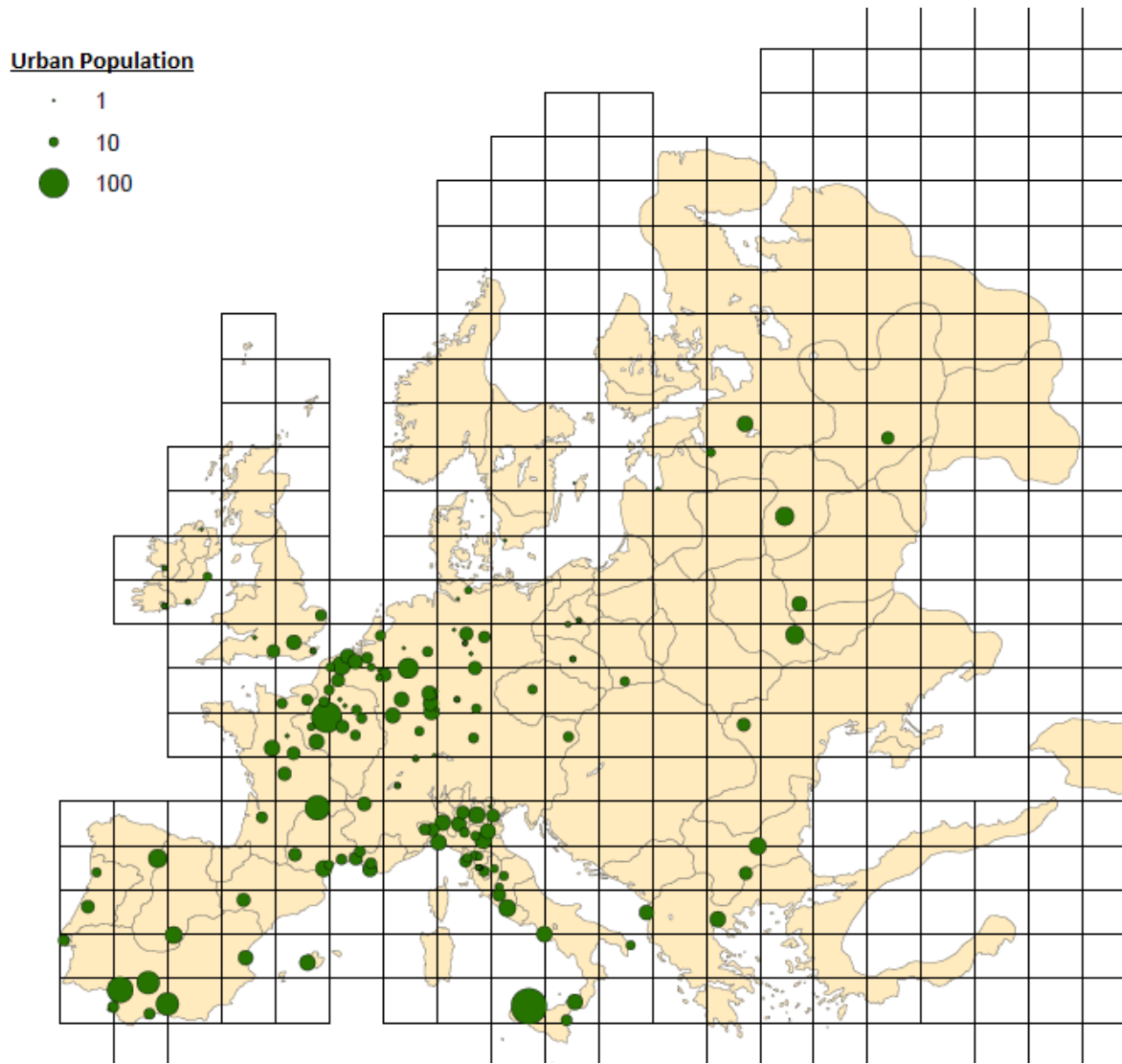


Figure 5: Urban populations for the year 1200 (in thousands).

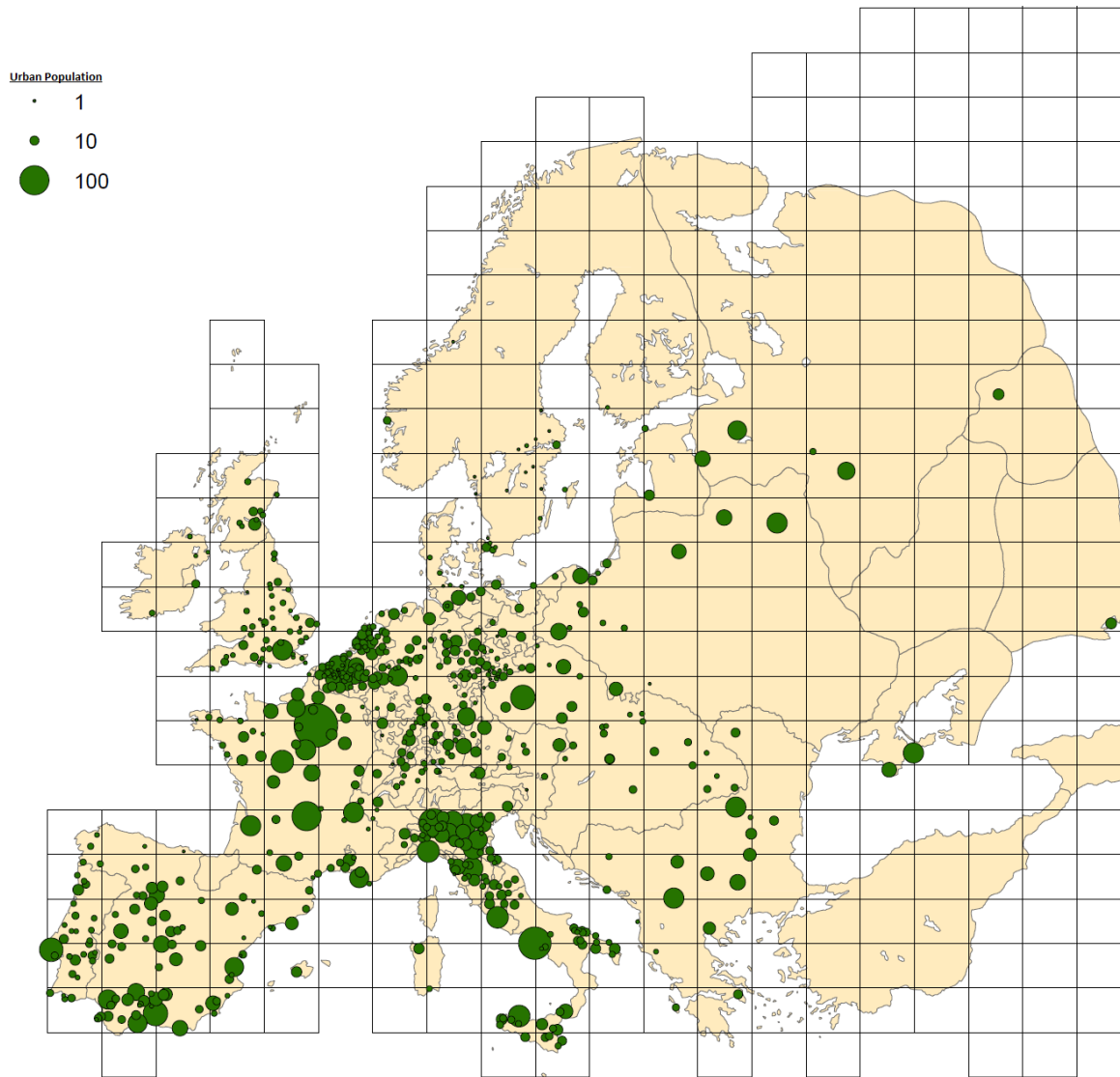


Figure 6: Urban populations for the year 1500 (in thousands).

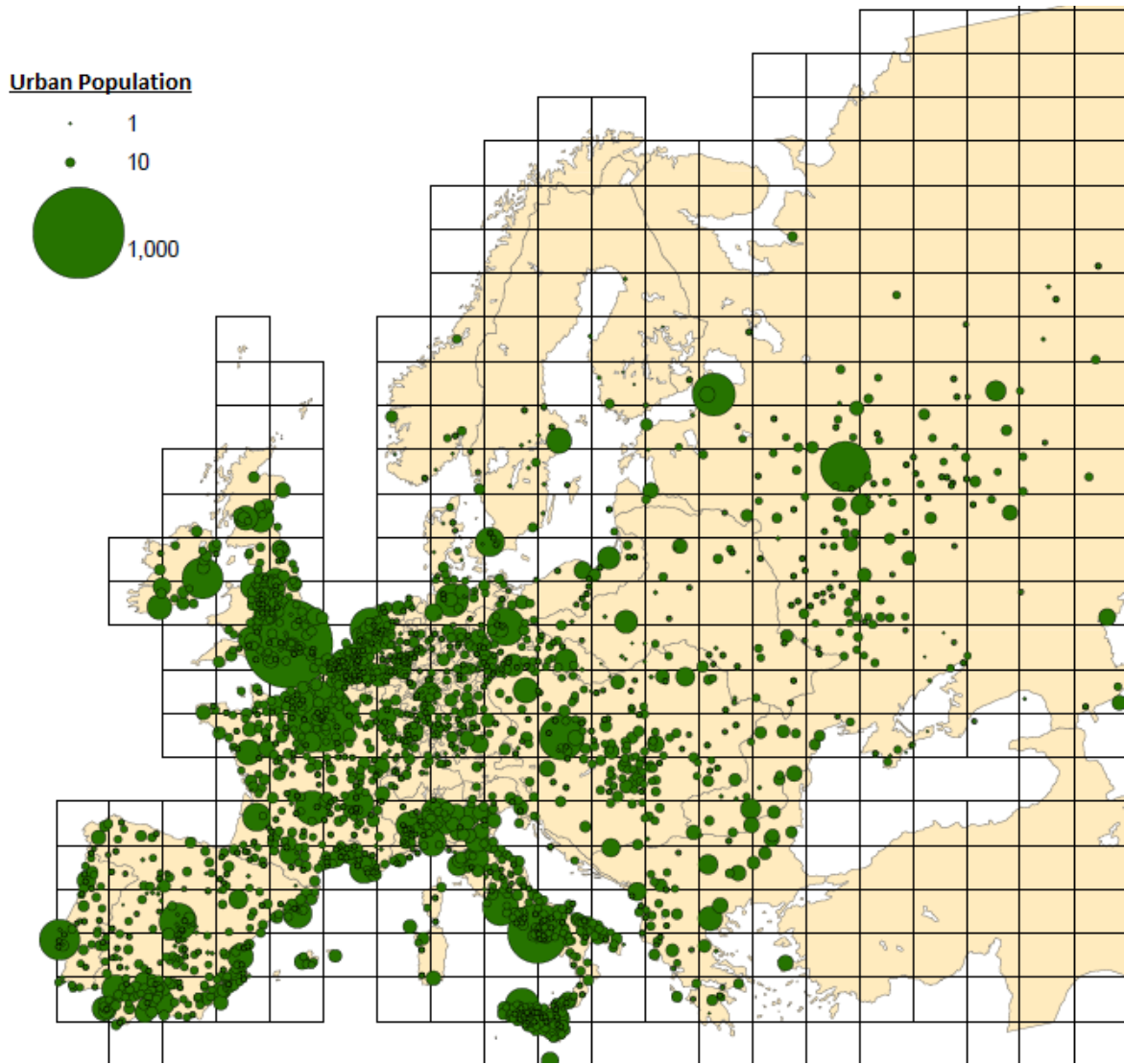


Figure 7: Urban populations for the year 1800 (in thousands).

class of servants) lives out of the surplus it has extracted from its particularly productive agricultural hinterland. Aware of this possibility, Weber (1968; p. 1212 ff.) already made a clear-cut distinction between European cities featuring a core of craftsmen, tradesmen and financiers and the imperial cities of Asia (and sections of Europe) built around a royalty, its court and its tax and military bureaucracy. Both cities may be located in agricultural rich lands. But only the former fostered the kind of technological innovation that ended up breeding the industrial breakthrough of the 18th and 19th centuries.

To measure the commercial and industrial dimension of cities, we have collected data on the geographical location of textile and metal production centers before 1500 in Europe. For the textile industry, we plot the location of wool, linen and silk manufacturing centers reported in Gutmann (1988), who in turn follows Carus-Wilson (1966). For the metal industry we employ the exhaustive data set built by R. Sprandel on the location of iron forges between 1200 and 1500 (Sprandel 1968; p. 93-220). Figures 8 and 9 depict the distribution of textile and metallurgic centers respectively. In addition to metal and textile centers, we have gathered spatially referenced data on the presence of premodern universities across Europe from two sources (Verger 1992, Frijhoff 1996). They are displayed in Figure 10.

Textile and metal proto-industries were clustered in the European core, indeed matching the distribution of urban population. As reported in Figure 3, the correlation coefficient between those economic sectors and urbanization levels fluctuated around 0.5. This correlation reinforces one of the main theoretical points of this paper: most European urban centers were not simply political or administrative centers but embodied a cluster of proto-industrial economic activities.

2.3 Political Institutions

We examine the impact of political institutions by looking at the presence of parliamentary institutions, which are seen in the literature as a main guarantor of property rights and as the foundation of the rule of law (North and Weingast 1989, North 1990). Our index of parliamentary strength, coded at the level of political sovereign (and semi-sovereign) units, is the fraction of years with

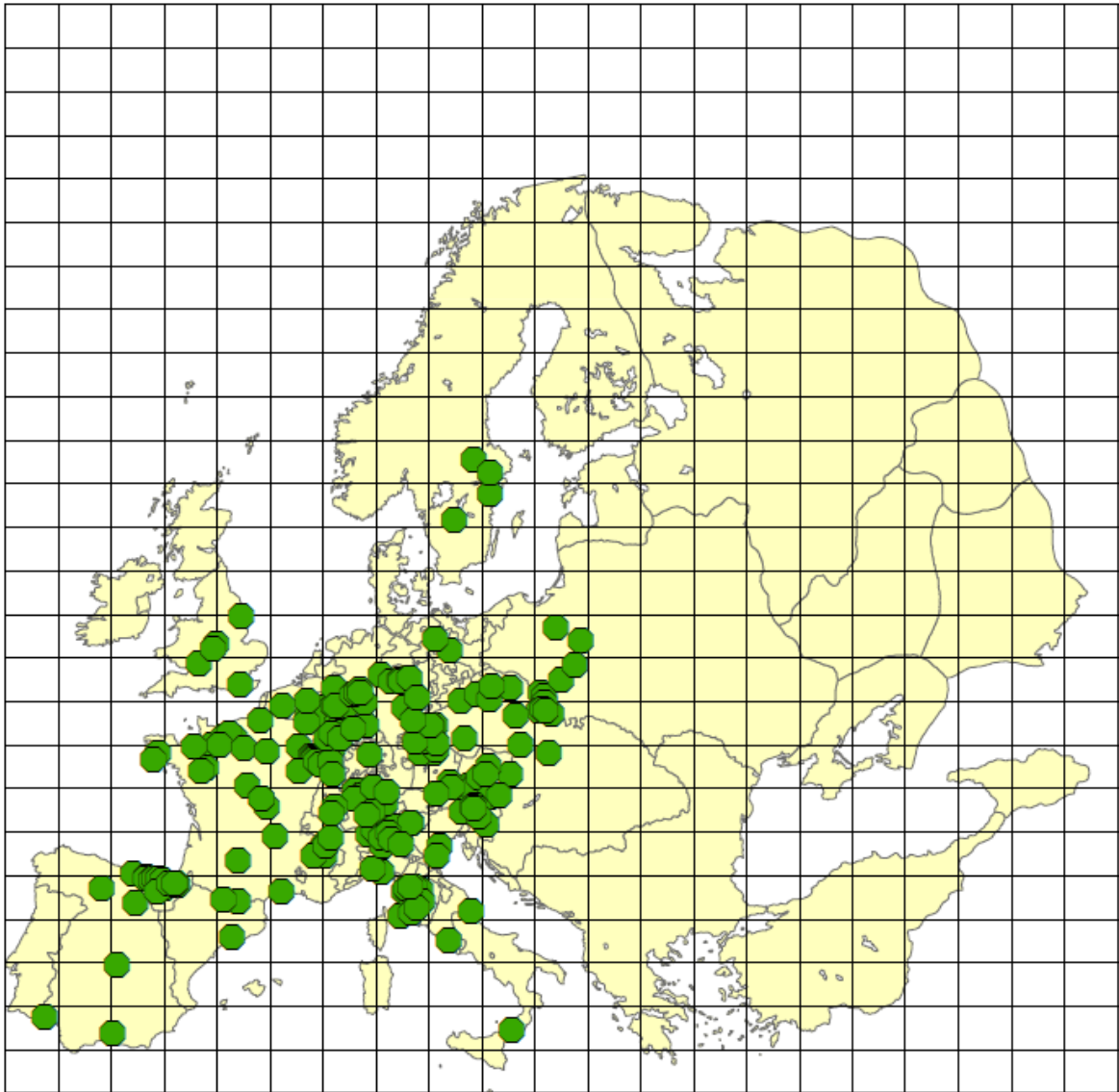


Figure 8: Iron Production Centers before 1500.

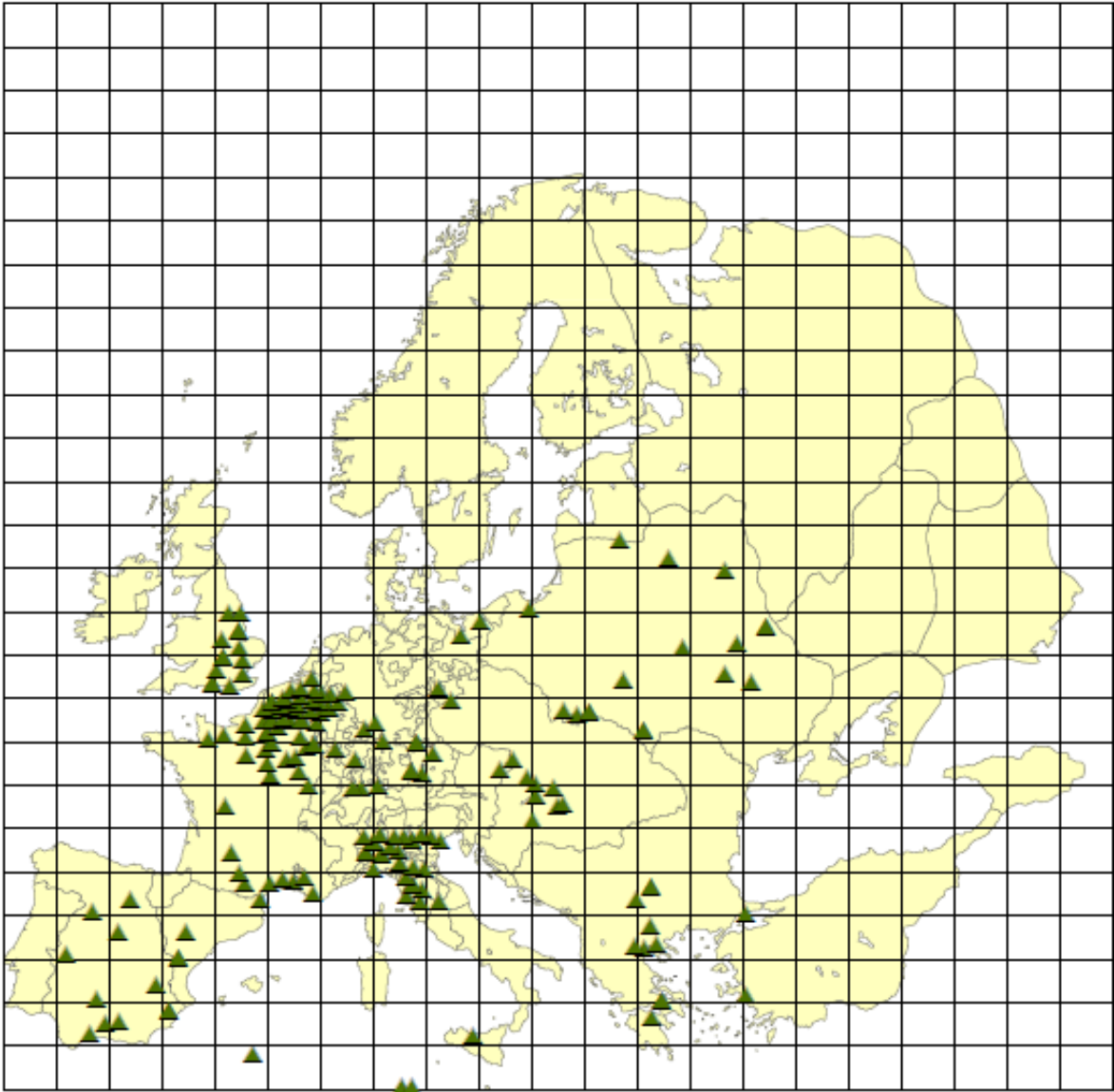


Figure 9: Textile Centers before 1500.

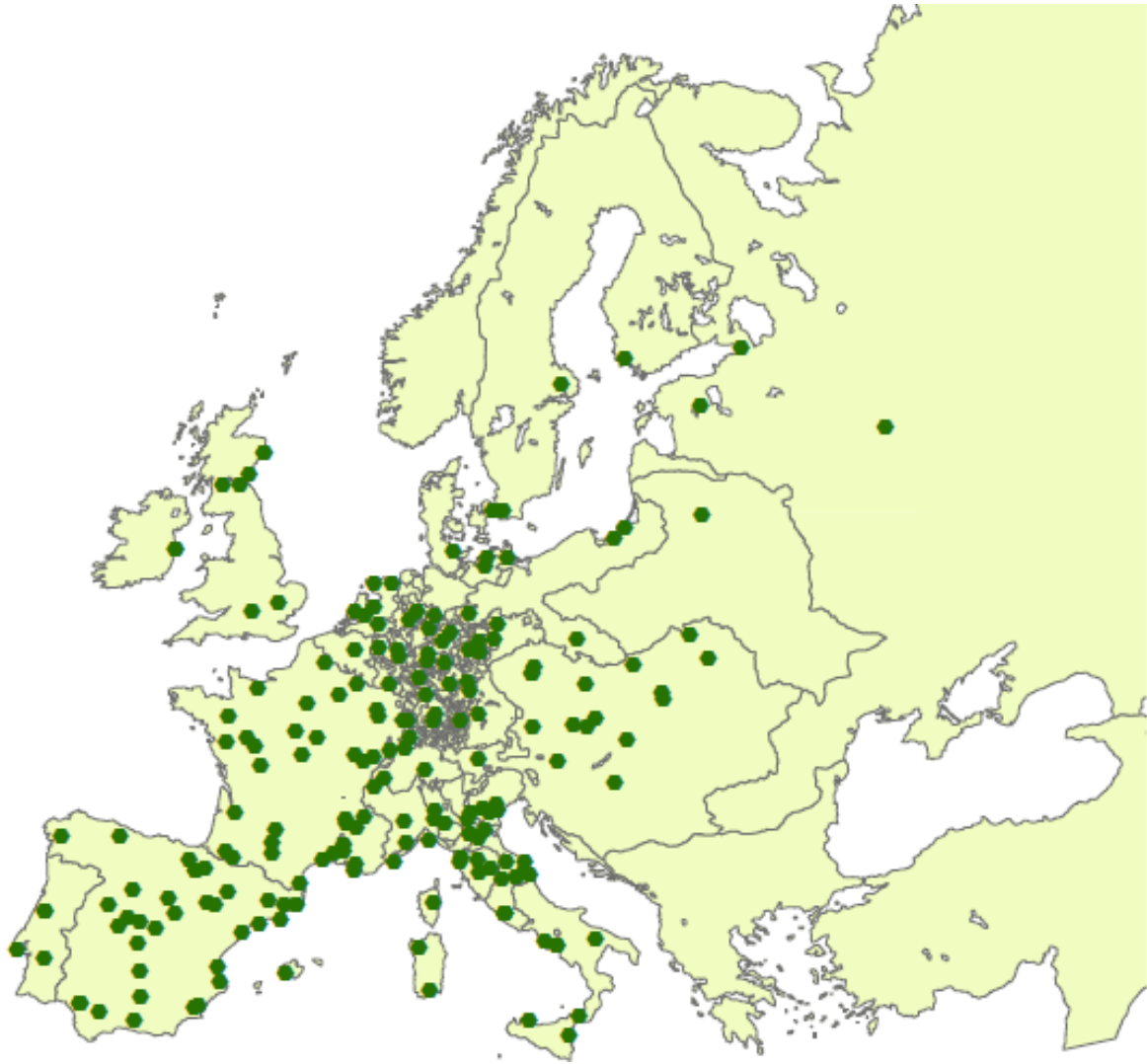


Figure 10: Universities Founded Before 1800

parliamentary meetings in each given century.

Parliamentary bodies include traditional territorial assemblies (like the British parliament, the French General Estates or the Catalan Corts) and permanent local councils (like Genoa's Maggiore Consiglio or Florence's executive committee). The coding, done annually, is then converted to century averages that range from 0 (Spain in the second half of the 18th century) to 1 (with a meeting every year, like Venice through 1798). To be defined as having a "parliament, the political unit under analysis has to have a non-executive body (i.e. a body that fulfills legislative and sometimes judicial functions as opposed to or in addition to strict executive tasks) formed by a plurality of members. This non-executive body must be chosen through procedures (elections or lottery) not directly controlled by the executive.⁹

The coding partly follows the data bases collected by van Zanden et al. (2010) and Stasavage (2011), corrected and complemented using secondary sources and historical collections of parliamentary sessions. However, our data base differs from previous studies in two ways. In the first place, we also code as parliamentary bodies those parliaments that did not include third estate representatives. Requiring urban representatives to code legislative bodies as parliaments conflates a purely institutional effect (i.e. a body capable of constraining the executive) with the presence of a particular social sector which is in fact endogenous to (proto-industrial) growth. In the second place, our data is more exhaustive than the existing data sets: it includes parliaments from territories that were members of political confederations (such as the Aragonese, Catalan or Valencian Corts, that were fully autonomous till early 18th century) and imperial structures (such as the parliaments of Naples, Sicily or Sardinia, which continue to meet under Catalan and later Spanish control); it also incorporates data on the governance structures of city-republics (as well as small duchies and principalities) such as Genoa, Lucca, Modena, Verona, etc. As a result, institutions are coded at a much lower level of aggregation than previous studies, which tend to use contemporary borders and throw away key regional variation.

⁹A council directly appointed by the executive (generally a monarch, prince or lord) is not counted as a parliament. Directly appointed councils range from early medieval curiae to advisory bodies set in place by absolutist kings. Multimember committees renewed through pure co-optation are not counted as "parliamentary bodies" unless they are the direct executive (generally in urban settings).

Parliamentary institutions were present in a broad swath of Europe at the end of the Middle Ages. At the peak of the Enlightenment era, however, they had receded to a fraction of the European territory: the Netherlands, Switzerland, Britain and parts of Germany. At the onset of the industrial revolution, the territorial extension of parliamentary institutions was much smaller than the areas that were heavily urbanized and that would become highly developed by the end of the 19th century.

2.4 Climate, Agricultural Suitability, and Urban Population

As pointed in Section 2, the growth of cities and proto-industrial centers and the development of quasi-representative political institutions may have been endogenous to a self-sustained process of population growth and technological innovation through learning-by-doing. The historical literature on premodern city growth (De Vries 1984, Bairoch and Braider 1991) always noted that urban centers required an agricultural surplus to sustain themselves appears systematically in the literature on premodern city-growth. In a succinct description of this logic, Nicholas (1997) writes that “cities could not develop until the rural economy could feed a large number of people who, instead of growing their own food, compensated the farmer by reconsigning his products and later by manufacturing items that the more prosperous peasants desired. The ‘takeoff’ of the European economy in the central Middle Ages is closely linked to changes in the rural economy that created an agricultural surplus that could feed large cities” (p. 104).

To measure agricultural surplus and a territory’s carrying capacity (in terms of population), we focus on the ability of some places to produce cereals like wheat for two reasons. First, the European diet of the premodern era was centered around the consumption of complex carbohydrates derived from cereals. So important to the typical European’s life were foods derived from these plants that economic historian Robert Lopez notes that “ in the form of bread, porridge, or mush, cereals were almost everywhere the basis of human alimentation...(Lopez 1976).” Not only were cereals central to diets across European geography but across classes as well, being integral to both the consumption patterns of the aristocracy (Duby et al. 1974).

Second, the ability to grow cereals has been directly linked to the support of large populations. Cereals like wheat, unlike other plants, are most capable of feeding large populations with minimal effort. This is because cereal crops, unlike fruits, pulses, or nuts, are extremely fast growing, high in calories from carbohydrates, and have extremely high yields per hectare (Diamond 1997). Moreover, unlike other crops, cereals can be stored for long periods of time enabling communities to smooth consumption over extended periods. To summarize, the ability to feed large populations was key to the development of cities. Since in pre-modern Europe the principle component of diets were cereals like wheat, foods that are particularly good at supporting large populations, climatic variation across time and space in the ability to grow these crops serves as a good encouragement for urban growth.

We construct the measure of agricultural suitability in two steps:

1. We take spatially referenced temperature data from two paleo-climatological sources, both measured at half-degree by half-degree latitude/longitude intervals. The first measure from Mann et al. (2009) records temperature anomalies for the past 1500 years. A temperature anomaly captures the deviation at each point from the 1961-2000 mean temperature. We then construct a measure of absolute temperature by adding back the 1961-2000 baseline mean temperature as calculated from Jones et al. (1999)'s twentieth century data. This yields a half degree by half degree grid of temperatures for every year over the past 1500 years. Hundred year averages of these yearly measures are then taken.
2. Next, using tension weighted splines we take these estimates, measured at fixed intervals, and construct a smoothed measure of temperature for the entire continent. From this continuous measure the average for each grid-square is taken yielding an estimate of temperature across our fixed but arbitrary pieces of geography. When the unit of observation is the semi-sovereign state instead of taking the average of this smoothed measure at the level of the grid square we measure it by political unit. All of these operations are taken using the interpolation and zonal averaging tools found in ArcGIS 10.

To assess the validity of this temperature proxy for wheat propensity, we employ two data sets

Variable	FAO Wheat Suitability	Avg Wheat Yield
Intercept	9.89 (20.1)	36901 (138.7)
Growth Temperature _{it}	-.61 (-12.5)	-1612.7 (6.3)
R^2	.55	.30
N	119	94

Table 1: The relationship between temperature and the suitability to produce wheat. The first column regresses the FAO wheat suitability index against the absolute average distance from 10.5 degrees Celsius between 1960 and 2000. The second column regresses the average wheat yield on this measure. T statistics in parentheses

from the FAO. The first one, “Agro-climatically attainable yield for rain fed wheat,” is from the Global Agro-ecological Assessment for Agriculture in the 21st century and captures the ability of land to produce wheat absent modern irrigation techniques. We then estimate the optimal climate to grow wheat (at around 10.5 degrees Celsius) by fitting a parabolic relationship between average annual temperature between 1960 and 2000 and this FAO measure. Regressing this FAO measure on the absolute deviation from 10.5 degrees, the correlation is strong. The R-squared is .55 and the coefficient is -.61 - a large unit effect size since the FAO measure is on a fourteen point scale (Table 1, Model 1). The second measure is FAO data on actual annual wheat yields, measured in tons per hectare. Regressing annual wheat yields on deviation from the optimal growing temperature again shows a similarly robust relationship. A one degree deviation from this optimal temperature has a large effect on annual wheat yields – a decline of approximately 1600 hectograms per hectare (Table 1, Model 2).

3 Empirical Analysis

3.1 Economic Development

Figure 2 in the previous section reports bivariate regressions looking at the relationship between urban population in 1200, 1500 and 1800. The units of analysis are 225km-by-225km quadrants and urban population is defined as population living in cities of 1,000 inhabitants or more. It shows that

there is a strong, persistent, and statistically significant relationship between early urban densities in 1200 and later urban densities in 1500 and 1800, respectively. For every thousand individuals living on 225 km \times 225 km grid in 1200, approximately four times this number are expected to be living there six centuries later, implying a century on century effect of approximately 1.26. As indicated by the bivariate relationship between 1200 and 1500 and that between 1500 and 1800, this effect is smaller in the first half of the series than it is in the second. We find an increase of 1.7 times the total urban population on a given unit between 1200 and 1500 and an increase of approximately 2.33 times the value between 1500 and 1800. This is not unexpected as there was a substantial decline in urban population between 1300 and 1400, in large part caused by the onset of the Bubonic Plague. All of the subsequent results are robust to successive changes in the specification of urban population, that is, to defining urban population as the population living in towns larger than 5,000, 10,000 and 20,000 inhabitants.

We have, however, data covering more than three points in time. Exploiting the full series, we can better estimate the dynamic, century-on-century, effects of past urban population. However, we are not merely interested in the impact of the previous century's levels of urbanization. If knowledge and skills have independent effects across extended periods of time there should remain effects even after controlling for the immediate past. We explore these temporal dynamics by estimating equations of the following form:

$$\mu_{i,t} = \alpha + \rho_{t-1}\mu_{i,t-1} + \rho_{t-2}\mu_{i,t-2} + \dots + \rho_{t-k}\mu_{i,t-k} + \delta_t + \eta_i + \epsilon_{it} \quad (1)$$

Where μ_{it} is a measure of urban population - its total or its logged value - on a given piece of geography i in period t , η_i is a country specific effect, δ_t is a period-specific constant, and ϵ_{it} is an error term. The country-specific effects η_i captures the existence of other determinants of a country's steady state and the period-specific effects, δ_t , capture common shocks affecting urban populations across the continent, for example the plague of the fourteenth century.

As shown by Nickel (1981) estimating Equation 1 will yield biased parameter estimates. So, fol-

lowing Arellano and Bond (1991), Arellano and Bover (1995), and Blundell and Bond (1998), we follow a now conventional approach and use an Arellano-Bond type system GMM estimator to consistently and efficiently identify Equation 1.¹⁰

Table 2 summarizes estimates of ρ . The first three columns give the estimate derived from pooled OLS, not accounting for unit specific heterogeneity, the second three columns gives results from a “within” fixed effects estimator, and the last three columns give results from the system GMM estimator. Consistent with the known downward bias of the dynamic fixed effects models, Model 2 gives a smaller estimate than the other models. Because the estimates of ρ_{t-1} in the upper panel of Table 2 are close to one, it indicates that the panel has a unit-root. That is, it is evidence of that the data generating process contains a exploding trend across time. Recognizing this, the lower panel presents the same set of models but with the data log-transformed. Once this transformation is taken into account all estimates of ρ_{t-1} fall between -1 and 1.

Across specifications - except for the biased fixed effects estimator - we see a that even after accounting for the first lagged value, the second lagged value remains positive and statistically significant. That is, urban population at $t - 2$ has an effect on urban population at t even after controlling for the value at $t - 1$, evidence of a strong and persistent effect of the past. Moreover, these results demonstrate that a “great divergence” between Eastern and Western Europe can be explained not as a structural break but as a slow and continuous compounding effect of early advantages. That is, those places that were early urbanized, continued to be so and, moreover, grew faster than places that were not urbanized early on. Doing so, not because one was exposed to some suddenly new conditions but because of the persistent and growing effects of past advantages.

To better see the geographic persistent and continued effect of early differences we plot the predicted values derived from Model 8. in Panel B. of Table 2 for units which had urban populations of 1,000, and 12,000, and 24,000 in the year 1200. The 23,000 difference between the maximal and minimal of these values represents an approximately one standard deviation for the year 1200. Then, because the first lag is needed to simulate this model, we add the mean increase between 1200 and 1300 of

¹⁰For example of this approach applied to growth outcomes, Caselli, Esquivel, Lefort (1996).

	Pooled OLS			Within Estimate			System GMM		
	(1.)	(2.)	(3.)	(4.)	(5.)	(6.)	(7.)	(8.)	(9.)
Urban Population $_{t-1}$	1.29 (24.12)	1.12 (6.35)	1.26 (5.72)	1.15 (9.19)	.98 (5.64)	.93 (4.64)	1.27 (14.15)	.94 (11.80)	.95 (4.27)
Urban Population $_{t-2}$.23 (1.18)	.17 (1.06)		.54 (6.29)	.43 (4.31)		.57 (6.29)	.41 (3.65)
Urban Population $_{t-3}$			-.10 (-.56)			.39 (3.13)			.32 (2.22)
R ²	.85	.86	.87	.85	.86	.85	.95	.70	.09
m_2									
B.									
log(Urban Population $_{t-1}$)	.89 (66.99)	.64 (18.24)	.62 (14.00)	.18 (4.76)	.19 (3.91)	.10 (13.88)	.69 (15.52)	.78 (16.26)	.71 (12.52)
log(Urban Population $_{t-2}$)		.34 (9.15)	.28 (5.54)		.15 (4.31)	-.15 (-3.07)		.24 (3.78)	.38 (4.80)
log(Urban Population $_{t-3}$)			.09 (2.67)			-.06 (-2.18)			.14 (2.67)
R ²	.69	.76	.78	.69	.75	.032	.33	.18	.26
m_2									
N	468	468	468	468	468	468	468	468	468
T	6	5	4	6	5	4	6	5	4

Table 2: Two-Step GMM estimates. All possible lags ($t-2$) included as instruments in the differenced equation. Δ_{t-1} included as an instrument in the levels equation. All estimates include time effects. T statistics derived from heteroskedasticity robust standard errors in parentheses.

seven thousand to each of these values. For the subsequent five periods we simulate the predicted urban populations using the estimates from this model. Plotted in Figure 4 it is apparent that this difference is increasing across time, the gap between the largest and the smallest of these hypothetical units growing from an initial 23,000 to a predicted difference of just under 468,000 by 1800.

3.2 Protoindustrial Activity

We now turn to the task of identifying the effect of urban growth on the development of knowledge regimes as proxied by the existence of protoindustrial centers of medieval universities. To do so, we estimate the following model:

$$\text{Proto}_i = \mathbf{f}(\alpha + \beta \cdot \text{Urban Population}_i + \epsilon_i) \quad (2)$$

Where the outcome Proto_i is a measure of proto-industry, either the existence or the count of textile or metallurgic centers on unit i , is treated as a function of level of urban population on the same unit. Table 3 examines these relationships, regressing metal and textiles centers (separately and jointly) on urban population. Models 1 and 2 report OLS and negative binomial estimations regressing the number of proto-industrial centers in each geographical quadrant on total urban population. Model 3 reports two-stage least squares estimates. In all models the estimated effects are positive and statistically significant. Again, all these results are robust to successive changes in the specification of the independent variable, that is, restricting urban population to be above towns larger than 5,000, 10,000 and 20,000 inhabitants as well as to the dichotomization of the independent variable into similarly categorized binary treatments.

Using the coefficient estimates in Table 11, Panel C. Models 3 and 4, in Figure 11 we plot the predicted change in the number of proto-industrial centers located on a given grid-square after a one standard deviation change in the size of the urban population in the year 1200. These results suggest that European urbanization was intertwined with a process of proto-industrialization. Based upon

The Relationship Between Urbanization in 1200 and The Development of Proto-Industry

	(1.)	(2.)	(3.)
N: 468	OLS	N. Binomial	2SLS
A. Textiles Production Centers			
Urban Population	.03 (3.49)	.05 (20.85)	.07 (4.73)
log(Urban Population)	.19 (4.71)	.30 (8.31)	.31 (5.36)
R ²	.28	.25	-.27
ϕ		.28	.72
B. Iron Production Centers			
Urban Population	.03 (2.58)	.04 (11.35)	.13 (4.17)
log(Urban Population)	.26 (4.72)	.22 3.89	.54 (5.85)
R ²	.11	.15	- 0.76
ϕ		.06	.07
C. Total Proto-Industry			
Urban Population	.06 (3.13)	.05 (13.71)	.20 (4.89)
log(Urban Population)	.45 (5.71)	.25 (4.43)	.85 (7.02)
R ²	.23	.25	-.80
ϕ		.12	.14

Table 3: OLS, Negative binomial, and 2SLS estimates of the number of proto-industrial centers in a given unit. In the two-stage least squares regression we use the optimal growth temperature of cereals as an instrument for urban population in the year 1200. T-statistics computed from heteroskedasticity robust standard errors in parentheses.

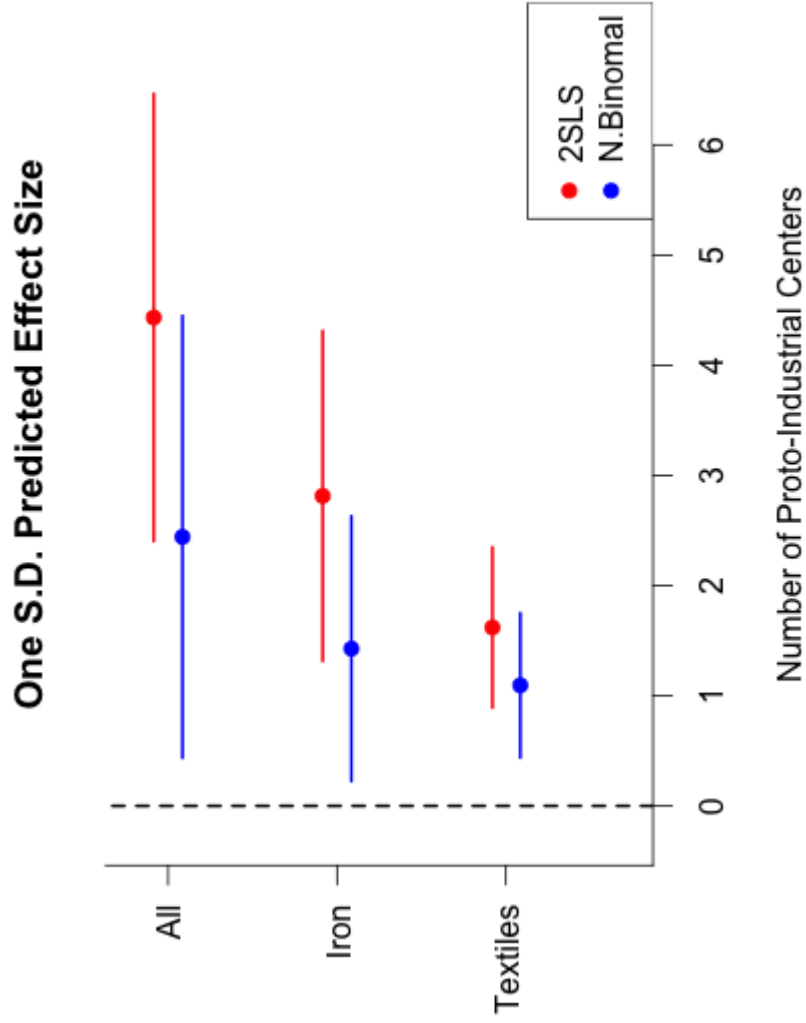


Figure 11: Predicted change in the number of proto-industrial centers on a given unit after a one standard deviation change in the number of people living in cities in the year 1200. The dashed line represent 95% confidence bands derived from heteroskedasticity robust standard errors

First Stage Results		
	Urban Population	log(Urban Population)
Growth Temp	-1.01 (-4.17)	-.43 (-11.24)
Intercept	(11.31) (4.99)	(-4.23) (-13.29)
F test on Excluded Instrument	15.19	126.36

Table 4: First stage regressions from the instrumental variables estimates presented in Column 4. of Table 3. T statistic computed with heteroskedasticity robust standard errors in parentheses.

the two-stage least squares results, for a one standard deviation increase in early urban population between four and five additional proto-industrial centers are predicted to have formed. Since we have a valid instrument for urban population, we interpret the results as pointing to a casual effect of urban population on the emergence of proto-industrial centers.

The instrumental variables results have a substantive causal interpretation. In those regions where agriculture was highly productive, the existence of a food surplus supported the emergence of towns where a class of individuals could specialize in the development of non-agriculture economic activities and of new technologies of production. To see this, examine the first stage regression from the the instrumental variables estimates presented in Column 4. of Table 3. This result, presented in Table 4 demonstrates a substantial negative effect of deviation from the optimal growing temperature of cereals like wheat on the presence of urban populations in the initial period of 1200; a one degree change in this measure is predicted to lead to an approximately forty-three percent decline in urban population. In addition to being a substantively large effect, it is a strong encouragement in a purely statistical sense. That is, the F-test on the excluded instrument is well above conventional measures of strength, indicating that, indeed, we have estimated a local causal effect of early urbanization on the development of protoindustry.

Notice that, in line with endogenous growth theories (Romer 1990) as well as geographic concentration models (Krugman 1991), urban clusters fostered, due some increasing return-to-scale and positive externalities derived from the agglomeration of individuals, an endogenous process of economic specialization and technological innovation in which regions with an initial advantage in

Multi-Valued Treatment												
	(1.)	(2.)	(3.)	(4.)	(5.)	(6.)	(7.)	(8.)	(9.)	(10.)	(11.)	(12.)
	Log(Urban Population ₁₅₀₀)						Log(Urban Population ₁₈₀₀)					
Iron Production Center	0.86 (11.29)	.46 (6.24)					.75 (10.28)	.37 (5.52)				
Textile Production Center			1.40 (3.48)	.43 (2.29)					1.31 (3.30)	.41 (2.06)		
All Proto-Industry					.74 (11.72)	.36 (5.39)					.66 (10.22)	.31 (4.94)
Log(Urban Population ₁₂₀₀)		.72 (16.04)		0.75 (13.96)		.68 (13.07)		.69 (21.98)		0.70 (16.73)		.64 (18.16)
R ²	.19	.45	.17	.41	.25	.45	.11	.26	.10	.27	.14	.27
N	468	468	468	468	468	468	468	468	468	468	468	468
Binary Treatment												
	Prob(Urban Population ₁₅₀₀ ≥ 20,000)						Prob(Urban Population ₁₈₀₀ ≥ 100,000)					
Iron Production Center	3.11 (6.72)	2.40 (7.02)					3.23 (10.28)	2.49 (5.52)				
Textile Production Center			3.23 (7.42)	2.55 (4.81)					2.71 (6.29)	1.62 (3.16)		
All Proto-Industry					3.09 (7.51)	2.46 (5.67)					2.79 (7.22)	1.91 (4.26)
Log(Urban Population ₁₂₀₀)		1.18 (9.67)		0.75 (6.77)		.46 (7.67)		.41 (8.583)		.97 (8.05)		.39 (8.18)
AIC	332.91	229.46	297.12	215.9	293.73	214.64	292.4	221.72	305.08	238.22	292.47	227.67
N	468	468	468	468	468	468	468	468	468	468	468	468

Table 5: Panel A. presents OLS regressions of the log of urban population on the number of proto-industrial centers in a particular unit. Panel B. presents logistic regressions indicating the effect of a binary indicator of a proto-industrial center on the probability of having an urban population of greater than 20,000 in 1500 and 100,000 in 1800 respectively. T statistics and Z statistics computed from heteroskedasticity robust standard errors in parentheses

their biogeographical endowments experienced ever-faster growth rates. Indeed, Figure 4 already provided some evidence that early (more) urbanized areas diverged from rural and semi-rural territories over time. This process was likely to be embodied and fostered by the proto-industrial sectors that emerged in cities (and their hinterlands). Textile and metal craftsmen were relatively involved in making piecemeal technological changes through a process of learning-by-doing. Given that industrial sector had a much larger technological frontier than agriculture, the rate of technological innovation must have been higher in urban settings than in semi-urban or rural regions, first by a marginal fraction and then, after around 1750, by a large and ever increasing difference.¹¹ To test the impact of proto-industrial centers on urbanization (independently of the effect itself of past urbanization), Table 3 regresses the logged value of urban population (in 1500 and 1800) on late-medieval textile and iron production centers on economic development in the modern period (controlling for lagged urbanization). The first panel in Table 3 presents OLS regressions of the log of urban population (in 1500 and in 1800) on the number of proto-industrial center in a particular unit. The second panel reports OLS regressions of the log of urban population on a dummy that indicates whether the unit of analysis had a proto-industrial center or not. In both cases, the effect is statistically significant and quite large from a substantive point of view.

Figure 12 examines the effect size of having a proto-industrial center on future urban populations in two ways. First, to compare the effect size of changes in early urban population to changes in the number of proto-industrial centers we simulate the effect of a one standard deviation change, holding the other covariate at its mean value in 1200. Next we plot the predicted probability of having a urban center of 20,000 inhabitants or more in 1500 and of 100,000 inhabitants or more in 1800, again controlling for the mean value of urban population in 1200. Here we manipulate a binary “treatment” indicating the existence of at least one proto-industrial production center in the late medieval period.

In both instances the effects are substantial. A one standard deviation change in the number of proto-industrial centers is expected to increase urban population in 1500 and 1800 by 9,068

¹¹For a discussion of whether economic growth built up through a gradual process or took place through a sharp breakthrough, see Clark (2008), chapter 12.

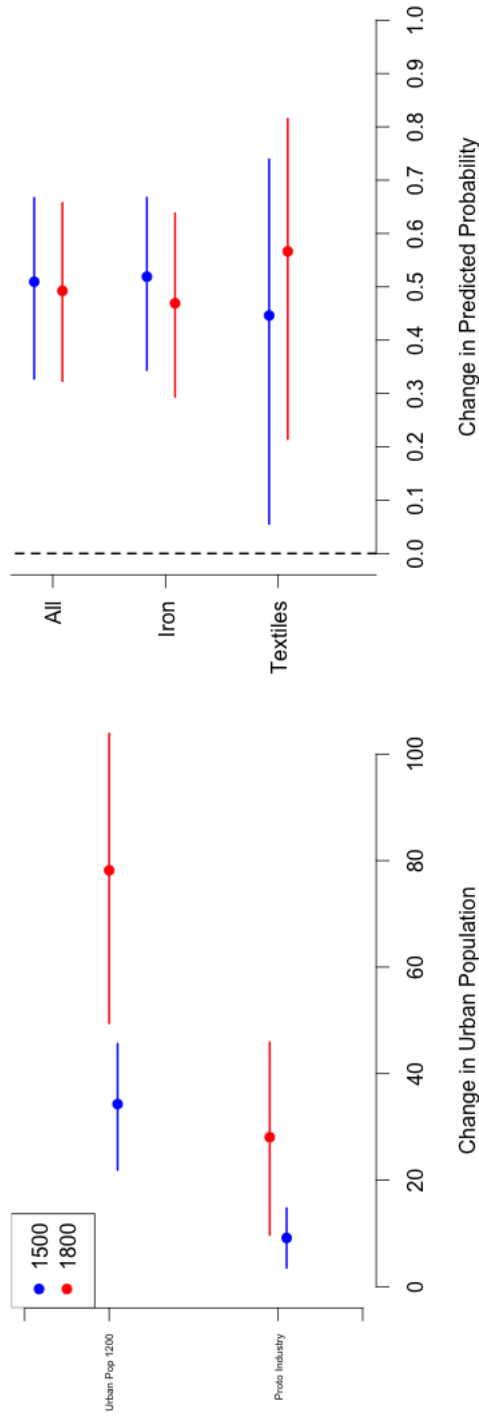


Figure 12: The lefthand figure plots the predicted effect of one standard deviation changes of both urban population in 1200 and of the number of protoindustrial centers. In each simulation we hold the other covariate at its observed mean in the year 1200 and the dependent variable is either urban population in 1500 or 1800. The right-hand plot gives the change in probability of a large urban population in either 1500 or 1800 as predicted by the existence of a proto-industrial center, controlling for urban population in 1200. We take operationalize "large" city as a population of 20,000 or greater in 1500 and 100,000 or greater in 1800. Estimates derived from the models presented in panel B. of Table 5. 95% confidence bands computed from heteroskedasticity robust standard errors.

and 28,059, respectively. Similarly, manipulating the presence of a single proto-industrial center, the probability of having an urban center greater than 20,000 inhabitants in 1500 is predicted to increase by 51%. Conducting the same exercise this time examining the predicted probability of a city greater than 100,000 in 1800, we see an increase of 49%.

3.2.1 Universities Before 1800.

Centers of learning, like centers of proto-industrial activity, have been linked to later increases in economic activity in the premodern world.¹² First, medieval universities may have contributed to economic activity by directly serving as locations of exchange and commerce. However, their larger effect on economic development was most likely the outcome of the human capital externalities they created. By producing individuals trained in logic, mathematics, and the sciences, the presence of institutes of higher learning engendered later innovation and commercial activity. Moreover, the individuals trained by these universities created networks composed of people trained with a generally shared curriculum, language (Latin), and most importantly a common understanding of canon and civil law. Specifically, the development of legal and educational institutions helped solve problems of exchange by resolving uncertainty in economic transactions and created a common language by which communication and business between distant traders could be conducted.¹³

We are interested in the effect of urban population on the number of universities on a given unit and estimate the following model:

$$\text{Universities}_i = \mathbf{f}(\alpha + \beta \cdot \text{Urban Population}_i + \epsilon_i) \quad (3)$$

Where Universities_i is a count of the number of universities present on geographic unit i in the eighteenth century and $\text{Urban Population}_i$ is the size of the urban population living on the same unit in the same time period.

Again, we see a strong and statistically significant relationship between urban population and

¹²See, for example, Cantoni and Yuchtman (2012)

¹³For a pair of references on premodern universities see Rashdall (1895) and de Ridder-Symoens (2003)

Dependent Variable:	Number of Universities			log(Number of Universities)	
Model:	OLS	Negative Binomial	2SLS	OLS	2SLS
log(Urban Population ₁₇₀₀)	.11 (5.5)	.56 (6.66)	.09 (2.25)	.31 (10.33)	.20 (1.81)
ϕ		.38			
R ²	.40		.40	.89	.89
F test on excluded instruments:			22.98		22.98
N	468	468	468	468	468

Table 6: The effect of urban population on the presence of universities in the eighteenth century. T statistic derived from heteroskedasticity robust standard errors in parentheses.

the presence of a university on a given piece of geography. Once more taking an instrumental variables approach, we can interpret this as a causal effect. These results, presented in Table 6, give a substantial predicted effect size. A one hundred percent change in urban population yields a predicted increase in the probability of another university by about ten percent.

3.3 Urbanization and Political Institutions

Figure 13 plots the evolution of parliamentary institutions (including multi-member councils in city-states), measured as a fraction of years in which they met in each century, and of urban population across the European continent from 1300 to 1800. Broadly speaking, parliamentary assemblies and councils peaked around 1500. Italian republican institutions, which were widely spread in the late medieval period, declined precipitously after the emergence of France and Spain as great continental powers. By the 18th century, the only Italian territories that maintained parliamentary institutions were Genoa, Lucca, the mark of Ancona (in the Papal States), Sicily and Venice. Parliamentary institutions became weaker in France and the Iberian Peninsula after 1500 and disappeared altogether after 1714. In central and eastern Europe they remained generally stable at their mid-16th century levels, with a brief rise in the Netherlands. Britain was the only country where parliament became systematically stronger over time. By the end of the 18th century it was meeting annually.

Table 7 examines the impact of urban development on political institutions. Models 1 and 2 regress

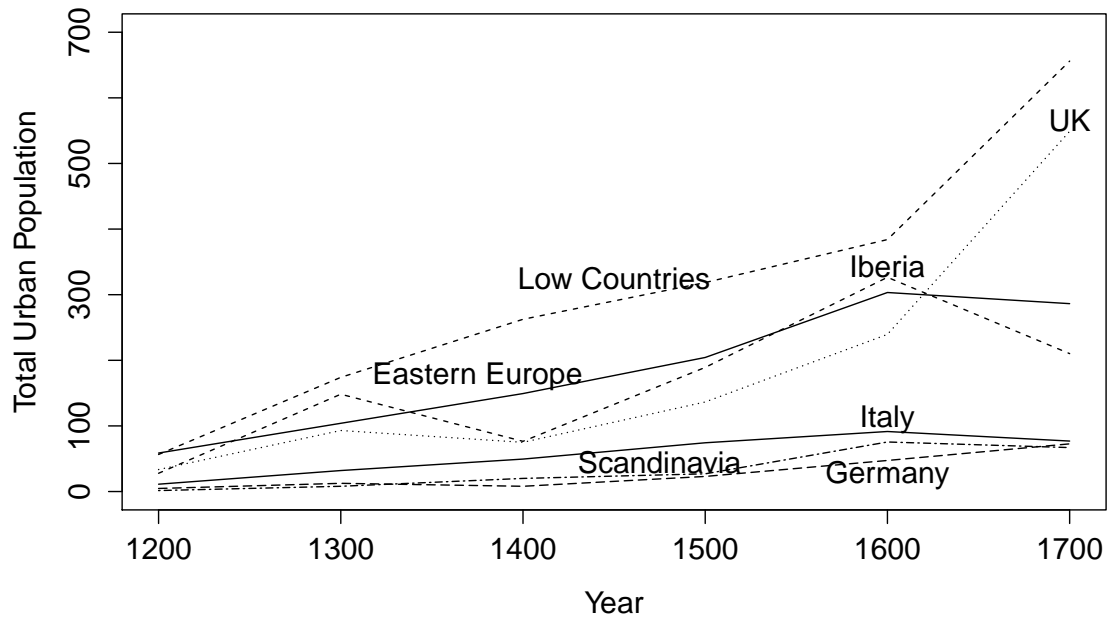
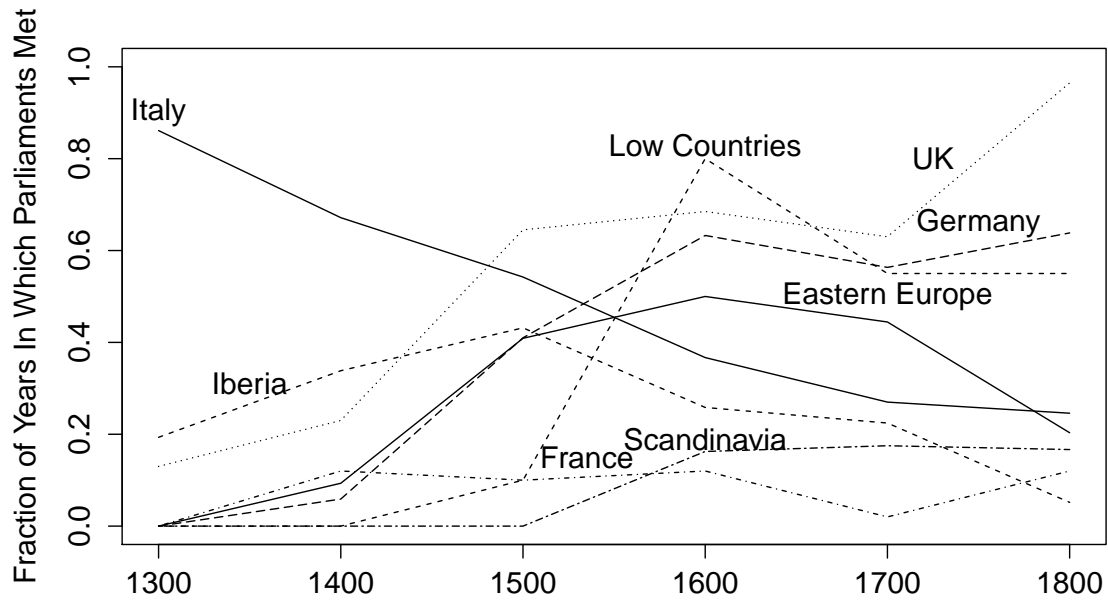


Figure 13: The top panel plots the regional average of the fraction of years in which a parliament met during the preceding century. The bottom panel plots the regional average urban population.

	Parliamentary Index _{it+1}		log(Parliamentary Index _{it+1})	
	(1.)	(2.)	(3.)	(4.)
	OLS	2SLS	OLS	2SLS
log(Urban Density _{it})	.012 (.79)	0.10 (3.84)	.11 (1.56)	.70 (4.49)
R ²	0.007	0.007	0.02	0.03
N	269	269	269	269
F-test on Excluded Instruments	.	10.18	.	10.18

Table 7: Estimates of the effect of urban density on the number of years in the subsequent century that a parliament was called. All regressions include country fixed-effects. T-statistics derived from heteroskedasticity standard errors clustered by country in parentheses.

the index of parliamentary institutions (fraction of years with parliamentary meetings in a given century) on the log value of urban density, the total urban population divided by the size in square kilometers of a given political unit, at the beginning of the century. Models 3 and 4 employ the log value of the parliamentary index as the dependent variable. All regressions are based on 100-year panels between 1200 and 1800 and include country fixed-effects. Models 2 and 4 instrument urban density again using optimal temperature for wheat. The instrument is robust with a first stage F-statistic of 10.18. Urban density has a causal effect on the fraction of years in a 100 year panel in which a parliament has met. According to the instrumented model (Model 2), a 100% increase in urban density caused about an 10% increase in the number of years parliaments met in a given political unit in the subsequent century. The relationship between urbanization and parliaments is almost all driven by within unit variation: the cross sectional relationship in any given panel between urbanization and parliaments is null.

To examine whether the effect of urbanization on parliaments differs across time periods, that is, to establish whether it is null in the 18th century, we first interact the urbanization measure with a full set of time dummies. The main effect does not change and none of the interaction terms are significant. When we take a random effects approach and estimate a varying slopes model (allowing the effect of urbanization to vary across time periods). Once more, the effect does not vary across time periods.

	GDP Per Capita		log(GDP Per Capita)	
	1870	1900	1870	1900
log(Urban Density)	400.11 (6.25)	556.82 (7.28)	.19 (5.45)	.18 (5.95)
Intercept	3969.3 (12.05)	5424.13 (13.90)	8.44155 (47.88)	8.71824 (55.52)
R^2	.27	.34	.22	.25
N:	106	106	106	106

Table 8: The relationship between urban density in 1800 and incomes in 1870 and 1900. T statistics computed from heteroskedasticity robust standard errors in parentheses.

3.4 Urbanization and Income

Since urbanization is only a proxy for development and given that one hundred year lags in the Bairoch data only extend to 1800, before the industrial revolution emerged in full force, we proceed by regressing per capita income in 1870 and 1900 on urban density in 1800. The results are reported in Table 8. The unit of analysis here is the current NUTS-2 region and the data only covers western and central Europe (plus Sweden). So, within this present day group of regions, all relatively well off in terms of present day income, we are examining variation in per capita wealth during the height of the industrial revolution as a function of urbanization right before the process of takeoff occurred. These results are summarized in Table 8

The magnitude of this relationship is substantial. For example, taking the model from the first column of Table 8 and manipulating the independent variable, urban density, across its interquartile range we get predicted incomes of \$1714.28 and \$2213.22 in 1870. These are extremely close to the true interquartile values of income per-capita in 1870 of \$1312 and \$2429. To see more clearly the relationship between urban density in 1800 and incomes at the height of the industrial revolution, we have plotted the data and fitted regression lines in Figure 14.

3.5 An Alternative Explanation: The Atlantic Trade

Acemoglu et al. (2004) propose an alternative explanation for the divergence in incomes across

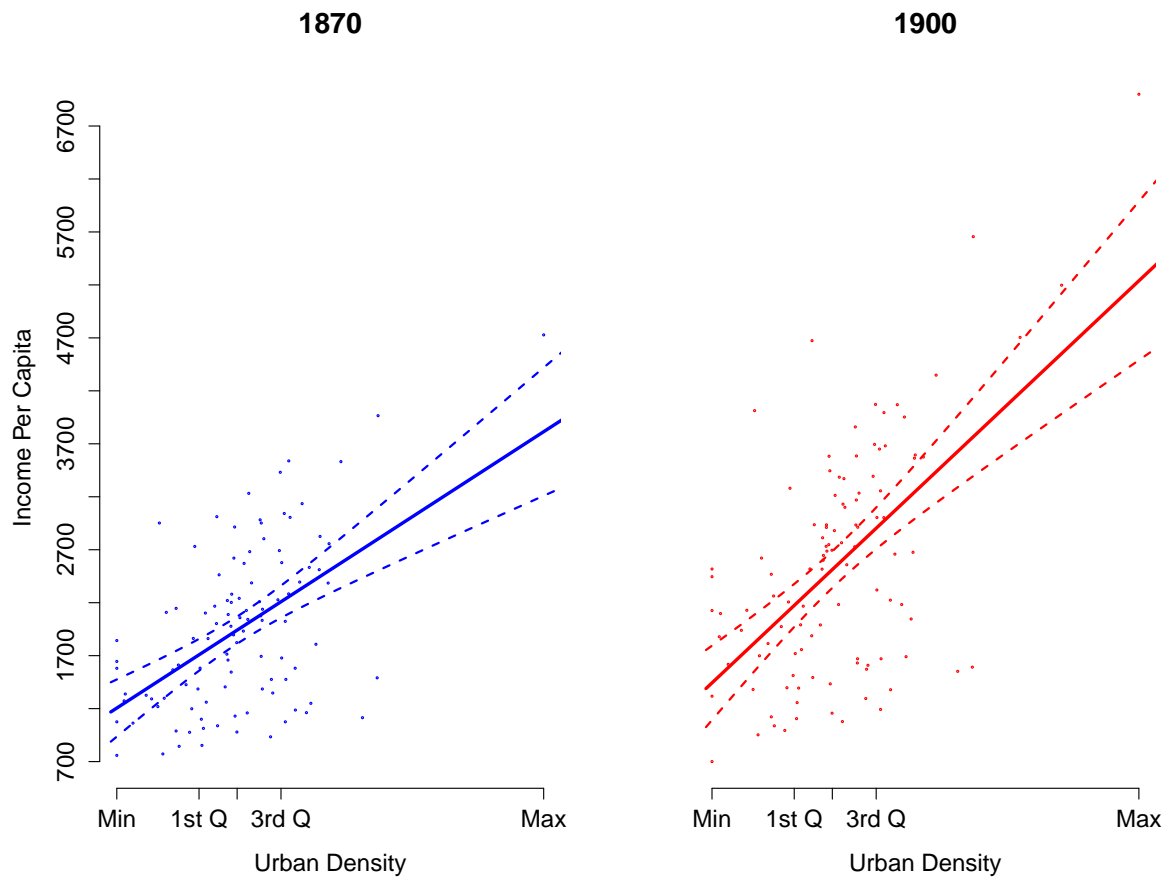


Figure 14: The bivariate relationships between urban density in 1800 (the x-axis in both plots) and per capita incomes in 1870 and 1900 respectively (the y-axis on each plot)

Europe beginning in 1500: the Atlantic trade. In their interpretation, access to the Atlantic and through this the ability to trade with the New World caused some countries to take-off and others to remain stagnant. In this section we examine the effect that access to the Atlantic had on urban populations using arbitrary pieces of geography. Following Acemoglu et al. (2004)'s empirical strategy we begin by estimating the following model.

$$\mu_{it} = \alpha + \sum_i^T \beta_t(\delta_t \times \text{Atlantic}_i) + \eta_i + \delta_t + \epsilon_{it} \quad (4)$$

Where μ_{it} is the total urban population living on grid square i in period t , η_i is an individual fixed effect, δ_t is a set of time effects, and ϵ_{it} an error term. The time-varying effect of access to the Atlantic trade is captured by the interaction of a set of time dummies, δ_t with a measure of access to the Atlantic, Atlantic_i . We operationalize this in two ways. The first, is simply a dummy for whether or not a given grid-square contains Atlantic coast as Atlantic coast is defined by Acemoglu et al. (2004). The second, captures distance in kilometers from the coast. The results are presented in Models 1 and 4 of Table 9. As predicted access to the Atlantic has a strong positive effect on the size of urban population.

However, comparing units with access to the Atlantic coast to all other units may not be the correct comparison. Rather, we should be comparing units with access to the Atlantic to units that had access to other trade routes. To do this we estimate the following model.

$$\mu_{it} = \alpha + \sum_i^T \beta_t(\delta_t \times \text{Atlantic}_i) + \gamma_t(\delta_t \times \text{Mediterranean}_i) + \eta_i + \delta_t + \epsilon_{it} \quad (5)$$

Where γ_t captures the time varying effect of access to the Mediterranean sea and all other parameters are defined as in Equation 4. Models 2 and 5 in Table 9 operationalize access to the Mediterranean as we do with the Atlantic, using instead a dummy or distance from the Western Mediterranean coast, here defined as all European coast west of Syracuse. Models 3 and 6 use all European Mediterranean coast. Using an F-test to test the restriction $\beta_t - \gamma_t = 0$, comparing estimates of the effect of access to Atlantic and Mediterranean coasts, we see that across specifications

Dependent Variable: Urban Population						
	1300	1400	1500	1600	1700	1800
1.)						
Atlantic Coast	8.9 (2.29)	11.0 (1.84)	22.0 (2.63)	37.89 (2.81)	64.06 (2.59)	129.07 (3.34)
R^2 : .11 N: 468 T: 7						
2.)						
Atlantic Coast	10.36 (2.62)	11.53 (1.95)	23.07 (2.78)	40.84 (3.17)	67.10 (2.72)	135.60 (3.54)
W. Mediterranean Coast	33.00 (2.91)	12.58 (1.40)	24.17 (2.16)	68.49 (3.39)	68.56 (3.43)	138.25 (4.30)
p value from F test that W. Mediterranean = Atlantic	.06	.92	.94	.23	.96	.97
R^2 : .20 N: 468 T: 7						
3.)						
Atlantic Coast	10.90 (2.80)	11.60 (1.96)	23.32 (2.81)	41.97 (3.21)	67.92 (2.76)	138.32 (3.61)
W. or E. Mediterranean Coast	23.07 (3.03)	7.29 (1.24)	15.23 (2.02)	48.079 (3.49)	44.544 (3.26)	101.17 (4.32)
p value from F test that W. or E. Mediterranean = Atlantic	.16	.61	.48	.74	.42	.67
R^2 : .18 N: 468 T: 7						
4.)						
Distance to Atlantic Coast	-.60 (-6.32)	-.42 (-4.28)	-.83 (-6.44)	-1.47 (-7.26)	-1.93 (-5.93)	-4.36 (-8.14)
5.)						
Distance to Atlantic Coast	-.16 (-1.82)	-.23 (-2.26)	-.42 (-3.09)	-.44 (-1.89)	-.92 (-2.44)	-2.16 (-3.09)
Distance to W. Mediterranean Coast	-.57 (-4.62)	-.25 (-2.20)	-.53 (-3.63)	-1.35 (-4.73)	-1.32 (-4.37)	-2.76 (-4.96)
p value from F test that W. Mediterranean = Atlantic	.03	.91	.64	.05	.50	.59
R^2 : .21 N: 468 T: 7						
6.)						
Distance to Atlantic Coast	-.35 (-4.64)	-.31 (-3.32)	-.60 (-4.83)	-.90 (-4.65)	-1.38 (-4.00)	-3.12 (-5.06)
Distance to min W. or E. Mediterranean Coast	-.44 (-4.28)	-.20 (-2.33)	-.42 (-3.70)	-1.02 (-4.68)	-.98 (-4.18)	-2.08 (-4.60)
p value from F test that Mediterranean = Atlantic	.45	.45	.34	.69	.40	.26
R^2 : .20 N: 468 T: 7						

Table 9: The relationship between Atlantic coastline and urban population across time. All models account for unit and period fixed effects. 95 % confidence intervals in brackets constructed from heterkedasticity robust standard errors. In models 2,3, 5 & 6 F tests on the restriction that the estimated parameter on Atlantic coastline is equivalent to the estimated parameter on Mediterranean coastline are presented

there is no statistically distinguishable difference between the two.

The second way we examine the effect of access to the Atlantic on urban growth is to estimate a dynamic model as in Equation 1 while including the time-varying effect of access to the Atlantic and Mediterranean coasts as in Equation 5. These results are summarized in Table 10. We see that only in the first model does access to the Atlantic have the predicted effect in any period, and here it does only in the final period. So, if this is the correct specification and the Atlantic trade mattered in the way Acemoglu and co-authors believe, it did so only three hundred years following the discovery of the New World. Moreover, in all of the models where access to the Atlantic is captured through distance to the coast, the coefficients all are significant but in the opposite direction one would expect. That is, increasing distance to the coast is predicted to *increase* urban population.

3.6 Did political institutions matter?

We now test institutionalist theories directly by estimating the impact of parliamentary institutions on growth. All the analyses are based on 100-year panels from 1200 to 1800. With models containing a lagged dependent variable we again take a GMM approach to identifying the following model.

$$\mu_{it} = \alpha + \rho\mu_{it-1} + \gamma\text{Parliaments}_{t-1} + \eta_i + \delta_t + \epsilon_{it} \quad (6)$$

Where μ_{it} is a measure of urban density for or a given polity, i (such as Venice, Valencia or England), in time period t and Parliaments_{it} is again the fraction of years in century t that parliaments met in polity i . All other terms are defined as before.

Table ?? gives estimates the effect of lagged urban density and parliamentary meetings frequency on urban density for a given polity. Models 1, 2, and 3 estimate the effect of each independent variable separately. Again, we see that the lagged value of urban density takes on a value close to unity. Moreover, we see that parliamentary meeting frequency seems to have a large and statistically sig-

System GMM	Dependent Variable: Urban Population				
	1400	1500	1600	1700	1800
Urban Population _{it-1}	1.25 (14.25)				
Atlantic <i>m</i> ₂ : .88	-15.36 (-2.31)	.72 (.14)	5.52 (.85)	5.84 (.79)	46.58 (2.89)
Urban Population _{it-1}	1.37 (15.47)				
Atlantic Distance <i>m</i> ₂ : .86	.72 (4.61)	.49 (4.48)	.66 (4.72)	1.23 (6.53)	.17 (4.57)
Dependent Variable: log(Urban Population)					
log(Urban Population _{it-1})	.81 (16.01)				
Atlantic <i>m</i> ₂ : .25	-2.25 (-3.21)	.46 (.82)	.00 (.99)	-.49 (-1.11)	-.21 (-.33)
log(Urban Population _{it-1})	.80 (18.57)				
Atlantic Distance <i>m</i> ₂ : .24	.07 (8.25)	.02 (3.55)	.03 (5.86)	.05 (7.73)	.04 (5.38)

Table 10: The relationship between a one century lag of urban population current population. Each model allows the relationship between being coastal and urban population is allowed to vary by century. Models 2,3,5 & 6 account for unit specific confounders. T statistics computed from heteroskedasticity robust standard errors clustered by grid-square.

	Within Estimator		System GMM	
	Urban Density _{it}	log(Urban Density _{it})	Urban Density _{it}	log(Urban Density _{it})
Urban Density _{t-1}			1.00 (6.36)	.99 (6.13)
log(Urban Density _{t-1})				.20 (2.72)
Parliamentary Index _{t-1}	74.55 (2.16)	.76 (1.86)		33.04 (1.37)
				.21 (3.02)
				.26 (1.11)

Table 11: Two-Step GMM estimates. All possible lags (t-2) included as instruments in the differenced equation. Δ_{t-1} included as an instrument in the levels equation. All estimates include time effects. Heteroskedasticity robust standard errors in parentheses. p values for the Arellano-Bond test for serial-correlation are presented in the row denoted m_2 .

nificant effect on future urban density. Models 4 and 5. jointly estimates the effect of parliamentary frequency and lagged urban population. In these specifications lagged urban density continues to perform strongly, in line with the results in previous estimations. However, parliamentary institutions are not statistically significant. That is, that once past levels of economic development are accounted for the effect of parliamentary institutions on future economic development is null.

4 Interpretation

Around the year 1,000 Europe was an economic and political backwater. The last Carolingian attempt at unifying the continent had collapsed not long ago, leaving a myriad of small political units. The economy was strictly agrarian and organized in local, segmented markets. Urban centers were small and far apart from each other. The biggest town in France at that time, Laon, had 25,000 inhabitants. Around 25,000 people lived in London and 35,000 in Rome. The largest cities lay in the continental periphery, under either Arab or Byzantine control: Cordova with 468,000 people, Palermo with 350,000 and Constantinople between 300,000 and half a million persons. Europe was mired in poverty – the average daily income ranged between \$1 and \$2 (in dollars of 1990).

Five hundred years later, however, the Italian per capita income had risen to around \$1,100 or \$3.5 a day. By 1850 some areas of Europe boasted annual per capita income close to \$2,500. Why? Employing geographic, economic and institutional data that cover 700 years of history, we show that the long-run process of European economic development was related to the formation and expansion of urban clusters across the continent – specially in the European north-south corridor that runs from southern England through the Low Countries, the Rhine and Switzerland to northern and central Italy. Cities emerged in highly productive lands (measured through the deviation from the optimal temperature to grow cereals). Regions with a substantial cereal surplus could sustain a growing non-farming population that joined in urban agglomerations and that specialized in the development of two main proto-industrial sectors: textiles and metal.

A growing population and the emergence of urban clusters acted as an endogenous growth engine:

as the number of individuals increased, the rate of technological innovation rose, slowly at first and at an accelerating rate later in time. Indeed, as we show in the paper, there were strong historical continuities in the location of towns in Europe, the correlation coefficient between urban populations in 1200 and 1800 is over 0.6, and the level of urbanization in a given century was a strong predictor of urbanization at subsequent periods. At the same time, as predicted in a growth model with increasing returns to scale and positive intra-sectoral externalities, the rate of urban growth varied substantially across the continent. Broadly speaking, those cities that were relatively larger at the beginning of the period kept adding population at a faster rate than smaller towns. As a result, there were sharp economic divergences across Europe by the end of modern era with a highly urbanized core extending from Barcelona-Lyon-Naples in the south to Liverpool-Manchester in the northwest and Hamburg-Dresden-Prague in the east and gradually less densely populated toward the outer rings of the continent.

The European urban model had little to do with the political and economic nature of non-European towns. Asian, Middle Eastern and pre-Columbine American cities were the seat of rent-seeking elites and their religious or royal bureaucracies. Absolutist monarchs built large empires through the force of arms to extract an agricultural surplus that they then could spend in lavish palaces and founded vast cities that could cater to their needs. By contrast, European urban clusters were networks of textile and metal artisans. As a matter of fact, Italian cities were initially corporations based on the free associations of particular guilds (Weber (1968), Najemy (2006)).¹⁴ This type of production favored a growing process of economic specialization and fostered technological innovation. Independently of the initial size of a town, having a textile or metal production center in the 14th and 15th centuries had a powerful effect on urbanization throughout the modern period up until 1800. The fact that European towns were made (or at least included) a class of individuals that embodied and carried over a certain production know-how had important implications for the modern industrial breakthrough. European artisans were the only individuals who had the kind of “useful” or technical knowledge (or, in the terms of Mokyr (2004) the λ -knowledge) needed to take advantage of the new general knowledge generated by the scientific revolution of the 17th and

¹⁴Notice that if the differential economic nature of European cities is true, then urbanization measures mean very different things across continents and cannot be applied to study growth in a comparative manner.

the 18th centuries and to apply the latter to the production process. In other words, knowledge of the principles of Newtonian physics and Lavoisier's chemistry could travel quickly from Lisbon to Moscow and Athens. But their profitable application was only possible in those areas which had a proto-industrial tradition.

Parliamentarian institutions spread, as permanent formalized bodies of governance (either as territorial assemblies or as city councils), once the proto-industrialization wave was well under way. They peaked in the late 15th century, which has been identified by the historical literature as the era of estate representation (Strayer 1973). With the intensification of war competition in Europe after 1500 and the consolidation of France and Spain as continental hegemony, those institutions weakened across the continent, particularly in Italy. They only remained in place, if at all, in the most proto-capitalist enclaves of modern Europe, although usually representing a much narrower oligarchy than in the late medieval period. The process of urbanization had a positive impact on parliamentarism. However, contrary to the existing institutionalist literature, the fortunes of parliamentary institutions seemed to have played a small part in the success of the industrial revolution and in the level of development across the continent.

References

- S. Abramson. Production, predation, and the origins of the territorial state. *Mimeo*, 2012.
- D. Acemoglu, S. Johnson, and J.A. Robinson. Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *Quarterly Journal of Economics*, 117(4):1231–1294, 2002.
- D. Acemoglu, S. Johnson, and J.A. Robinson. The rise of europe: Atlantic trade, institutional change and economic growth. *American Economic Review*, 2004.
- R.C. Allen. The great divergence in european wages and prices from the middle ages to the first world war. *Explorations in Economic History*, 38(4):411–447, 2001.
- P. Anderson. *Lineages of the Absolutist State*. Verso Books, 1974.
- M. Arellano and S. Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2):277–297, 1991.
- M. Arellano and O. Bover. Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1):29–51, 1995.
- K. Arrow. Economic welfare and the allocation of resources for invention, 1962.
- P. Bairoch and C. Braider. *Cities and Economic Development: From the Dawn of History to the Present*. University of Chicago Press, 1991.
- P. Bairoch, J. Batou, and C. Pierre. *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)*. Librairie Droz, 1988.
- G. Becker, K. Murphy, and R. Tamura. Economic growth, human capital and population growth. *Journal of Political Economy*, 98(5):S12–S137, 1990.
- R. Blundell and S. Bond. Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, 87(1):115–143, 1998.
- R. Brenner. *The Boom and the Bubble: the US in the World Economy*. Verso Books, 2003.
- D. Cantoni and N. Yuchtman. Medieval universities, legal institutions, and the commercial revolution. Technical report, National Bureau of Economic Research, 2012.
- E.M. Carus-Wilson. *The Woollen Industry. Cambridge Economic History of Europe. Volume II*. Cambridge University Press, 1966.
- A. Chanda and L. Putterman. Early starts, reversals and catch-up in the process of economic development*. *The Scandinavian Journal of Economics*, 109(2):387–413, 2007.
- G. Clark. *A Farewell to Alms: A Brief History of the World*. Princeton University Press, Princeton, 2008.
- J.B. De Long and A. Shleifer. Princes and merchants?. *Journal of Law and Economics*, 36, 1993.
- H. de Ridder-Symoens. *A History of the University in Europe: Volume 1, Universities in the Middle Ages*. Cambridge University Press, 2003.

- J. De Vries. *European Urbanization, 1500-1800*. Methuen London, 1984.
- J. Diamond. *Guns, germs and steel*. 1997.
- G. Duby, H.B. Clarke, and M.B. Becker. The early growth of the european economy: warriors and peasants from the seventh to the twelfth century. *History: Reviews of New Books*, 2(8):208–208, 1974.
- S. Engerman and K. Sokoloff. Factor endowments, institutions and differential paths of growth among the new world economies,” in stephen haber, ed., *how latin america fell behind*, stanford: Stanford university press, 1997.
- S.L. Engerman and K.L. Sokoloff. Factor endowments, inequality, and paths of development among new world economics. Technical report, National Bureau of Economic Research, 2002.
- W. Frijhoff. Graduation and careers. *A History of the University in Europe*, 2:355–415, 1996.
- F. Fukuyama. The end of history? *The National Interest*, 1989.
- M.P. Gutmann. *Toward the Modern Economy: Early Industry in Europe, 1500-1800*. Temple University Press, 1988.
- D.A. Hibbs and O. Olsson. Geography, biogeography, and why some countries are rich and others are poor. *Proceedings of the National Academy of Sciences of the United States of America*, 101(10):3715, 2004.
- A. Inkeles. Making men modern: On the causes and consequences of individual change in six developing countries. *American Journal of Sociology*, pages 208–225, 1969.
- EL Jones. *The european miracle: Environments, economies, and geopolitics in the history of europe and asia*, 1981.
- PD Jones, M. New, DE Parker, S. Martin, and IG Rigor. Surface air temperature and its changes over the past 150 years. *Reviews of Geophysics*, 37(2):173–199, 1999.
- M. Kremer. Population growth and technological change: One million bc to 1990. *The Quarterly Journal of Economics*, 108(3):681–716, 1993.
- P.R. Krugman. *Geography and Trade*. The MIT Press, 1991.
- S. Kuznets. *Toward a Theory of Economic Growth: With Reflections on the Economic Growth of Modern Nations*. Norton, 1968.
- R.S. Lopez. *The Commercial Revolution of the Middle Ages, 950-1350*. Cambridge Univ Pr, 1976.
- A. Maddison. *The World Economy: Historical Statistics*. Organization for Economic Cooperation and Development, 2003.
- M.E. Mann, Z. Zhang, S. Rutherford, R.S. Bradley, M.K. Hughes, D. Shindell, C. Ammann, G. Faluvegi, and F. Ni. Global signatures and dynamical origins of the little ice age and medieval climate anomaly. *Science*, 326(5957):1256, 2009.
- J. Mokyr. Punctuated equilibria and technological progress. *The American Economic Review*, 80(2):350–354, 1990.
- J. Mokyr. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton Univ Pr, 2004.
- J.M. Najemy. *A History of Florence, 1200-1575*. Wiley-Blackwell, 2006.

- D. Nicholas. The growth of the medieval city: From late antiquity to the early fourteenth century, 1997.
- D.C. North. Institutions, institutional change and economic performance. 1990.
- D.C. North and R.P Thomas. The rise of the western world: A new economic history. 1973.
- D.C. North and B.R. Weingast. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century england. *Journal of Economic History*, 49(4):803–32, 1989.
- D.C. North et al. *Structure and Change in Economic History*. Norton New York, 1981.
- M. Olson. Dictatorship, democracy, and development. *American Political Science Review*, pages 567–576, 1993.
- M. Olson. *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships*. Basic books, 2000.
- S.L. Parente and E.C. Prescott. Barriers to technology adoption and development. *Journal of Political Economy*, pages 298–321, 1994.
- S.L. Parente and E.C. Prescott. Monopoly rights: A barrier to riches. *American Economic Review*, pages 1216–1233, 1999.
- K. Pomeranz. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton Univ Pr, 2001.
- R.D. Putnam. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton Univ Pr, 1993.
- H. Rashdall. *The Universities of Europe in the Middle Ages: Salerno. Bologna. Paris*, volume 1. Clarendon Press, 1895.
- L.A. Rivera-Batiz and P.M. Romer. Economic integration and endogenous growth. *The Quarterly Journal of Economics*, 106(2):531–55, 1991.
- P.M. Romer. Why, indeed, in america? theory, history, and the origins of modern economic growth. *American Economic Review*, 86(2):202–206, 1996.
- J.D. Sachs and A.M. Warner. Fundamental sources of long-run growth. *The American Economic Review*, 87(2):184–188, 1997.
- A. Smith. The wealth of nations (1776). *New York: Modern Library*, page 740, 1937 (1776).
- R.M. Solow. A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70(1):65–94, 1956.
- R. Sprandel. *Das Eisengewerbe im Mittelalter*. A. Hiersemann, 1968.
- D. Stasavage. *Public debt and the birth of the democratic state: France and Great Britain, 1688-1789*. Cambridge Univ Pr, 2003.
- D. Stasavage. When Distance Mattered: Geographic Scale and the Development of European Representative Assemblies. *American Political Science Review*, 2011.
- J.R. Strayer. *On the Medieval Origins of the Modern State*. Princeton Univ Pr, 1973.
- C. Tilly. *Coercion, capital, and European states, AD 990-1990*. Basil Blackwell, 1990. ISBN 155786067X.
- J.L. van Zanden, E. Buringh, and M. Bosker. The Rise and Decline of European Parliaments, 1188-1789. 2010.

J. Verger. Patterns. *A History of the University of Europe*, 1992.

N. Voigtländer and H.J. Voth. Why england? demographic factors, structural change and physical capital accumulation during the industrial revolution. *Journal of Economic Growth*, 11(4):319–361, 2006.

M. Weber. Economy and society (g. roth & c. wittich, eds.). *New York: Bedminster*, 1968.