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

November 18, 2015

Abstract

This paper reassesses the causes and nature of economic growth in Europe from 1200 to 1900. Employing a comprehensive dataset that includes biogeographic features and urbanization data (1200-1800), per capita income data in the second half of the 19th century, location of proto-industrial centers and political institutions, it shows that the initial economic take-off (and of a growing economic divergence across Europe), sparked by new agrarian technologies interacting with suitable biogeographic conditions, led to the emergence and growth of cities and urban clusters in the European corridor broadly running from southern England to northern Italy. In contrast to the institutionalist literature, it shows that the fortunes of parliamentary institutions in early modern Europe played a small part in the success of the industrial revolution. The paper claims, instead, that development followed a process of endogenous growth, often regardless of the absence of executive constraints or access to energy resources.

Two broad features define the economic history of Europe in the last millennium: first, sustained growth, slowly at first and then at an accelerating rate after 1750; second, increasing income divergence within Europe. According to admittedly crude estimates, GDP per capita ranged from a maximum of $468 to a minimum of $400 (in 1990 International Geary-Khamis dollars) in Europe around the year 1,000 (Maddison 2003). Half a millennium later, that range had widened from $433 in Albania to $1,100 in Italy.[1] By 1850, with the industrial revolution in full swing in England, the maximum European per capita income had risen to almost $2,500 while GDP per capita had remained stagnant in the poorest areas of the continent.

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 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, Tabellini 2010).

Both strands of the literature are beset by important theoretical and empirical 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 why self-sustaining economic growth took place in some regions but not in others. In turn, most empirical work on modern growth has been conducted by economic historians either focusing on a particular case or comparing 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, Johnson and Robinson 2002; 2004, Hibbs and Olsson 2004, Chanda and Putterman 2007). Although these

[1] Urbanization rates, which have regularly 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.

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

studies have made important contributions, their measurement of economic data and institutions has been incomplete, limited to periods 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) claimed to matter for growth.

In this paper we reassess the literature on European growth by looking at the evolution of that continent 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 (textile and metal) centers and of coal mines, political borders, and political institutions. All the data are calculated at 225 km $\times$ 225 km grid-square units as well as for sovereign and semi-sovereign 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 using both a sensitivity analysis procedure and an instrumental variables approach that exploits random climatic variation across time and space in the propensity of territory to support urban populations.

Contrary to currently dominant institutionalist theories of growth, we show that the economic take-off and development of parts of Europe has to be understood as a process of endogenous growth and endogenous institutional development characterized by four key traits. First, the end of the "long period of migration, invasion, and conquest" (Strayer 1973; 16) spanning from about 400 to around 900 and a set of new agrarian technologies such as the heavy plow and the three-field rotation (White 1964) led to faster population growth and the emergence of cities in highly productive agricultural lands as well as in those regions that had cheap communication and transportation waterways. Second, in those areas, mostly clustered in the European north-south corridor that broadly runs from southern England to northern Italy, most urban settlements specialized in the development of a set of an artisan craft manufactures that, mostly through a process of learning-by-doing, fostered incremental technological innovations. Those urbanized, proto-industrial, regions benefited from increasing returns to scale due to sector- and location-specific positive agglomeration externalities (Krugman 1991): the areas that urbanized earlier in time grew into much larger towns than the areas that were mainly rural in the Middle Ages. Third, the processes of urbanization and proto-industrialization spurred or at least coincided with the diffusion of pluralistic political institutions (in the form of city councils or territorial assemblies with stronger urban representation) as permanent bodies of governance across Europe. Finally, and contrary to the existing institutionalist literature, 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 and where there was a sufficiently strong commercial and artisanal class that had the stock of technological know-how that enabled those areas to take advantage of the technological breakthroughs of the 18th and 19th centuries.

The paper is organized as follows. Section 1 reviews dominant explanations of development and relates them to our explanation of the sources of Europe's development. Section 2 describes the data employed in the paper. The following four sections examine our main empirical implications. Section 3 characterizes economic growth in Europe as following an endogenous process – with early urbanization leading to subsequent urbanization, reinforced by agglomeration effects. Sections 4 and 5 assess the underlying engines behind that process. Section 4 discards institutionalist explanations: after showing that parliaments multiplied with urban growth, it rejects the hypothesis that they were a precondition for the development of Europe. Section 5 examines the endogenous model of economic growth: it shows how high urban populations spurred a process of proto-industrialization in the textile and metal sectors and identifies how proto-industrialization was the central economic channel through which urban growth took place. Section 6 reassesses and finds little empirical support for two central non-institutional sources of growth put forward by the literature: easy access to energy, and Atlantic trade, alone and in interaction with institutions.. Section 7 concludes by offering a theoretical interpretation of our results. This article focuses on growth within Europe, both for theoretical reasons (such as the existence of a long, important scholarly debate on the ”rise” of the continent or its fundamental biogeographical homogeneity (Jones 1981)) and methodological motives (the availability of very detailed data at the local level combined with its institutional heterogeneity). However, we believe that continent offers enough variance in outcomes (some parts remained completely underdeveloped until the 20th century) to provide key insights on why the rest of the world did not experience the kind of breakthrough that took place in northwestern Europe.

1 Theory

Besides standard accounting growth models (Solow 1956), the current literature offers three alternative explanations of economic development: institutionalism, biogeography, and endogenous growth.

According to institutionalism, which has become the predominant explanation of development, institutions – defined as “the rules of the game in a society or, more formally, the humanly devised constraints that shape human interactions” (North 1990; 3) affect the economy in two ways. By structuring property rights, they determine the private rate of return, investment decisions and the level of effort individuals exert at technological innovation. By shaping transaction costs, they define the incentives of economic agents to specialize and trade with each other, and, in a Smithian economic framework, productivity gains and long-run growth (Smith 1937, North 1990). Institutionalists have suggested three main institutional configurations leading to growth: a stable political order guaranteed by the state (Olson 1993; 2000); 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 a stable set of norms of cooperation and “thick” trust, i.e. social capital, reducing the incentives of individuals to take advantage of each other and empowering them to control state institutions (Putnam 1993).

As pointed out before, 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 institutionalism. This literature explains the persistence of bad institutions (and therefore of underdevelopment) as the result of the benefits they confer on a section of society, which is then able to block, through political means, successful pro-growth reforms (Kuznets 1968, North et al. 1981, Mokyr 1990, Parente and Prescott 1994; 1999). However, the emergence of growth-promoting institutions (through a process exogenous to the economy) is not modeled. At most, it is traced back to a set of poorly defined critical junctures happening a few centuries ago (Anderson 1974, Acemoglu, Johnson and Robinson 2002, Brenner 2003, Acemoglu, Johnson and Robinson 2004).[3]

Given the limits of institutionalism, we propose, instead, to interpret the transformation of (parts of) Europe as the outcome of a process of endogenous economic change. In endogenous growth models, growth is triggered by technological innovation, itself the result of either some investment on a knowledge or R&D sector or, more simply, the by-product of the production process (Arrow 1962, Rivera-Batiz and Romer 1991, Kremer 1993, Romer 1996). In turn, the rate of technological change (and the long-run growth rate of output per worker) is an increasing function of the size and growth rate of the population. Variation in the latter is shaped, following the literature on the effect of biogeographical factors on population and economic growth (Sachs and Warner 1997), by the quality of soils and their effect on food availability.[4] In short, in this theoretical framework it is technological change and economic growth that lead to the emergence of new institutions and not the other way around.[5]

In the context of European development this story of endogenous economic and institutional change took place as follows. With the end of the waves of wars and massive migrations that had started with the penetration of Barbarian populations in the late Roman empire and lasted until the

[3] According to Anderson (1974; p. 97-431) the European breakthrough happened after urban and parliamentarian 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), in a way followed by Acemoglu, Johnson and Robinson (2004), 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.

[4] Its variation was also probably shaped by political and military shocks. For growth models that endogenize population choices, see Becker, Murphy and Tamura (1990) and Galor, Moav and Vollrath (2009).

[5] North and Thomas (1973), for example, model institutions as adaptative responses to technological change and changing relative prices. In Marx institutions also adjust to the economy over the long run: technological change creates a new class (e.g. capitalists) that asserts itself over the old relations of production through violence and political conflict. Among sociologists, modernization theorists see institutions as a response to the effects and challenges of social and economic development (Lipset 1959, Inkeles 1969).

Hungarian invasion and the Viking raids of the 9th and 10th centuries and that had resulted in the collapse of urban life and the absence of any significant interregional trade, the continent stabilized politically and growth resumed (Pirenne 1936, Randsborg 1991). Across Europe, at that point completely rural and autarkic, the introduction of new agricultural techniques such as the heavy mould plough and the three-field rotation system boosted yields and population growth (Lynn 1964, Andersen, Jensen, and Skovsgaard 2013). The latter varied, however, with biogeographical conditions. European regions endowed with rich soils and optimal temperatures generated a large crop yield per hectare, which allowed them to support high population densities and the formation of urban agglomerations. The latter were then conducive to the formation of a class of traders, artisans and craftsmen, which, through a process of learning-by-doing, led to a relatively faster rate of technological change than in less urbanized territories. The initial advantage of early urbanizers gave them a persistent lead over time. In the presence of increasing returns to scale to knowledge and positive agglomeration externalities, the 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 European corridor running from England to Northern Italy and the rest of the continent.

That process of economic development triggered (or at least co-evolved with) key institutional transformations and sustained them. Once towns had grown in size and wealth, their dwellers had the numbers and money to defeat the heavy cavalry of the old feudal class and introduced pluralistic institutions in autonomous or semi-autonomous city-states in the 13th and 14th centuries (Tilly 1990, Abramson 2015). The use of gunpowder and the intensification of war competition around 1500 led to the emergence of several large continental monarchies and to a generalized decline of most proto-parliamentary institutions. The latter only remained in place, if at all, in the most commercially dynamic enclaves of northwestern Europe. In other words, the strength of parliamentary institutions reflected (rather than generated) the particular economic and social structure of each territory, playing, 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 during the approximately two centuries that preceded the industrial revolution.

2 Data

We explore the covariates associated with economic growth by employing two types of units as our observations: 225 km-by-225 km grid-scale units or quadrants that have some mass of land;[6] and political units that were either sovereign or semi-sovereign polities. Sovereign units were fully independent territories with their own executive (monarchical or not). Semi-sovereign units were those territories that, although under the control of a different state, retained some measure of political autonomy (defined by the existence of their own governing institutions or 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 Catalan Crown in 1444) or Valencia (member of the Catalan confederation and later of Spain) 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 endowments (Tilly 1990, Abramson 2015).[7]

Our data coverage for political institutions is broader than existing studies in two ways. First, we include Scandinavia and most of Eastern Europe. Second, we code our observations going back to 1200 whereas most current studies employ instead historical panels that start at a moment in time when economic divergence had already taken place.

2.1 Economic Development

Following the current literature (Acemoglu, Johnson and Robinson 2002, Chanda and Putterman 2007), we first rely on urbanization data to proxy for economic development. Employing Bairoch,

[6] In the appendix we reproduce all of our results with units half this size.

[7] Our measures of historical state boundaries are taken from Abramson (2015) which allows the size of units to change over time. Therefore, when measuring urban density we account for both changes in the size of states as well as the addition of urban population via expansion.

Batou and Pierre (1988), who provide a comprehensive dataset with information on about 2,200 towns that had 5,000 or more inhabitants at some time between 800 and 1800, 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.

Figure 1 represents 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 maps also include the grid we use to define our observations. 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, Batou and Pierre (1988) the number of Europeans living in towns grew from 8.4 million in 1300 to 23 million in 1800. Urban growth did not simply track total population growth. It resulted in a higher proportion of the population living in cities. In Western Europe, the urbanization rate went up from 2.1% in 1000 to 8.1% in 1500 and 21.2% in 1800 (Boix 2015). Urbanization rates also varied across countries – for example, in 1500 they ranged from 29.5% in the Netherlands to 2.2.% in Scandinavia.

Besides Bairoch's urbanization data, we employ regional per capita income in 1870 and 1900 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.[9]

[8]This 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.

[9]For the sources and procedure employed to build per capita incomes at the regional level, see appendix online.

2.2 Urbanization and proto industrialization

Towns may embody a process of economic specialization and technological innovation leading to higher incomes. However, they may just be urban agglomerations where a rent-seeking clique (served by a class of servants) lives out of the surplus it extracts from its particularly productive agricultural hinterland. Aware of this possibility, Weber (1968; 1212 ff.) distinguished between towns featuring a core of craftsmen, tradesmen and financiers and cities 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 could have 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 Rolf Sprandel on the location of iron forges between 1200 and 1500 (Sprandel 1968; 93-220).

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 politically sovereign (and semi-sovereign) units, is the fraction of years with parliamentary meetings in each given century. The frequency of parliamentary meetings is an indirect but plausible measure of institutional strength. The history of the conflict between parliamentary forces and absolutist monarchs in modern Europe revolved around the capacity of the latter to first domesticate and then suppress parliaments (Anderson 1979, Williams 1970). Recent literature on dictatorships and semi-democratic regimes shows that the presence of working institutions (such as legislatures) is fundamental to the preservation of power-sharing agreements among governing elites (Gandhi and Przeworski 2007, Svolik 2009, Boix and Svolik 2013).

Parliamentary bodies include traditional territorial assemblies (like the British parliament or the French General Estates) and permanent local councils (like Genoa's Maggiore Consiglio or Florence's executive committee). More precisely, 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.[10] 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).

The coding partly follows the data bases collected by van Zanden, Buringh and Bosker (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 that was 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 Catalan or Valencian Corts, which were fully autonomous untill the early 18th century) and imperial structures (such as the parliament of Sicily, which continued to meet under Catalan, Spanish and Austrian 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 by using contemporary borders throw away key regional variation. The number of political units coded reaches over three hundred.

[10] A 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 also control executive powers directly.

2.4 Controls: Climate, Agricultural Suitability, and Urban Population

As pointed out in Section 1, the growth of cities and proto-industrial centers and the development of quasi-representative political institutions was likely to have been endogenous to a self-sustained process of population growth and technological innovation through learning-by-doing. The literature on premodern city growth (De Vries 1984, Bairoch and Braider 1991) highlights the fact that urban centers required an agricultural surplus to sustain themselves. As Nicholas (1997) points out, "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 disentangle the causal relationship between economic development, proto-industrialization, and parliamentary constraints we take two approaches. First, because the bio-geographical conditions that promote early urban development may also affect the later economic and political outcomes we are interested in, we control for a large number of possible confounders and, then, using the sensitivity analysis procedure discussed in Oster (2013), we show that our estimates of the impact of early urban development on subsequent patterns of economic growth are robust to substantively large violations of the assumption of exogeneity conditional on observable covariates.[11] Second, we employ an instrumental variables approach where we exploit climatic shocks to the ability of some areas to grow cereals as a cause of urban growth that arguably had no direct effect on our variable of interest in later periods.

The list of our controls includes, in the first place, the rain-fed suitability of our units of interest to produce agricultural output. This variable, which measures the capacity for a given piece of territory to produce agricultural output without extensive irrigation, is derived from the FAO's GAEZ combined land suitability dataset (FAO 2000). In the second place, we control for how mountainous an area is using the spatial data on terrain ruggedness collected by Shaver, Carter

[11]For examples of recent empirical work in political science and economics using this method see Satyanath, Voigtländer and Voth (2013), Alesina, Harnoss and Rapoport (2013), Laitin and Ramachandran (2014), and Cagé and Rueda (2014)

and Shawa (2013). Third, since the ability to trade may have affected both the development of cities and our outcomes of interest, we account for access to trade routes by controlling for river density, distance to coasts, and the total length of coastline. Fourth, we include measures of latitude and longitude for the centroid of each unit. Finally, we add unit fixed effects whenever we have repeated observations over time allowing us to control for qualities specific to each territory and identifying any effects through within-unit variation.

Our results rely upon the standard assumption that conditional upon observable variables that we control for, our independent variable of interest, urban density, is exogenous. We assess the validity of this assumption with the test developed by Oster (2013). In order to place bounds on the bias of a treatment effect estimate caused by the presence of unobservables, i.e., omitted variables, this method uses information from changes in both point estimates and R^2 values derived from comparing the unconditional estimated causal impact of our main independent variable of interest, early urban density, to this variable's estimated effect after conditioning on all other observable covariates. The procedure allows us to evaluate the degree to which unobservable factors are likely to bias our results. Oster (2013) considers to be robust those results that survive the presence of hypothetical unobservables explaining variation in the outcome of interest equal to 1.3 times the R^2 associated with the regression containing the full set of observed controls.[12] As we detail later, all of our results relating the presence of early urban clusters to proto-industrial skills, future urban density, and future incomes survive at or beyond this level.

Again, recognizing the presence of cities were contingent upon the capacity to feed large populations, we also employ an instrumental variable approach where we use climatic perturbations in the capacity of some places to produce cereals like wheat. We do this for two reasons. First, the European diet of the premodern era was centered around the consumption of complex carbohydrates derived from cereals across all social classes (Lopez 1976, Duby, Clarke and Becker 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

[12] Oster (2013) finds that 90% of a random sample of randomized control trials (N=65) published in the *American Economic Review, Journal of Political Economy, Quarterly Journal of Economics, Econometrica* and *American Economic Journal: Applied Economics* would survive this threshold.

because they 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.

In order to use agricultural suitability (measured as deviation from optimal temperature) as an instrumental variable for urban population, several assumptions must be met. First, deviations from this temperature must be a strong encouragement of urban growth. Throughout, these shocks prove to satisfy all tests against weak instrumentation. Second, the instrument should meet the requirement of being randomly assigned: we understand it does since, at least until the 19th century, there was no direct human effect on climate. Finally, our instrument must satisfy the exclusion restriction: climate shocks to the ability to sustain large populations in period t should have no effect on political or economic outcomes like the development of proto-industry or parliaments in period t+1 other than through its effect on urban populations at time. Using the sensitivity analysis proposed by (Conley, Hansen and Rossi 2012), we show that it would take a substantively large violation of the exclusion restriction to nullify the causal interpretation of our findings. In the appendix we fully detail and present results from the instrumental variables strategy.

3 Endogenous Growth and the Persistence of Initial Advantages

3.1 Economic Development

Figure 2 plots the bivariate relationship between total urban population in each geographical quadrant in 1200 and 1500 and total urban population at a later time. It also 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. This effect is smaller in the first half of the series than in the second.[13] Total urban population on a given unit increased 1.7 times between 1200 and 1500 and then approximately 2.3 times in the following three centuries. This differential rate of growth suggests a widening gap between early and laggard urbanizers.

Since we have data covering more than three points in time we can exploit the full series to estimate the dynamic effect of past urban population (both from the immediately preceding century as well as from more distant times). Accordingly, we begin by estimating autoregressive models of the following form:

$$\mu_{i,t} = \alpha + \phi_{t-1}\mu_{i,t-1} + \phi_{t-2}\mu_{i,t-2} + ... + \phi_{t-k}\mu_{i,t-k} + \delta_t + \eta_i + \epsilon_{it} \qquad (1)$$

Where μ_{it} is a measure of urban population - its total or its logged value - on a given geographical unit 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 unit-specific effect η_i captures the existence of other determinants of a geographical unit's steady state. The period-specific effects, δ_t, capture common shocks affecting urban populations across the continent such as the plague of the 14th century.

Table 1 reports estimates of ϕ. We present models that include one, two and three lags sequentially. Columns 1 through 3 present pooled OLS estimates not accounting for unit specific heterogeneity. Since, as shown by Nickell (1981), estimating equation 1 in a standard fixed effects framework will yield biased parameter estimates, we follow a now conventional approach and reports in Columns 4 to 6 a system GMM estimator to consistently and efficiently identify equation 1 (Arellano and Bond 1991, Arellano and Bover 1995, Blundell and Bond 1998).[14] The estimates of ϕ_{t-1} in the first six columns of of Table 1 are close to one, indicating that the panel has a unit-root and that the data generating process contains an exploding trend across time. Recognizing this, in Columns 7-12 we conduct the same exercise, estimating the same set of models but with the data log-transformed. Once this transformation is taken into account all estimates of ϕ_{t-1} fall between -1 and 1. However, when second order lags are included the sum of their coefficients, $\phi_{t-1} + \phi_{t-2}$,

[13] In the appendix we provide further evidence that the inter-temporal relationship is robust to the inclusion of controls and, using the method of Oster (2013) is insensitive to the presence of unobservables. This and all of the subsequent results are robust to successive changes in the specification of urban population (as the population living in towns larger than 5,000, 10,000 and 20,000 inhabitants)

[14] For an example of this approach applied to growth outcomes see Caselli, Esquivel and Lefort (1996).

either exceed the bounds of stationarity or come very close to doing so.

In order to determine if the time-series component of urban population, either in logs or levels, is non-stationary, we conduct two unit-root tests, the results of which are presented in the lower panel of Table 1. The first, proposed by Breitung (2000), takes as the null hypothesis that all panels contain unit-roots. In both logs and levels, using first test we are unable to reject the null that geographical units in all panels have a unit root. The second test, developed by Hadri (2000), takes as the null hypothesis that all panels are stationary. In this case we can reject the null hypothesis that all panels are stationary with a high degree of confidence. In short, both tests suggest that the development of urban population was a non-stationary process.

From a substantive point of view, those results indicate that very early differences in urban population had a persistent effect on present outcomes greater than those in later periods. In other words, the "great divergence" between the European core and its peripheries cannot be pinned down to a structural break (at a given point in time) but was rather the result of a slow and continuous effect of early advantages: those places that urbanized early in time continued to be so, growing faster than places that were not urbanized early on due, again, to the persistent and cumulative effects of past advantages.[15]

To give a sense of the magnitude of that divergence, Figure 3 plots the estimated difference in logged urban population between three areas from 1200 until 1800 that had an initial urban population of 1,000, 12,000 and 24,000 respectively. The 23,000 difference between the two extreme values represents approximately one standard deviation for the year 1200. This figure is derived from the dynamic system GMM estimates of Model 1 (reported in Table 1), employing the coefficient on the lagged value to obtain an estimate for each period and then taking the difference of these estimates.[16] Figure 3 makes apparent that an initial advantage has a cumulative effect over time. Whereas in 1200 the maximum difference was of 23,000 individuals, six hundred years later the

[15]To see this, take as an example a non-stationary AR(1) process where $\mu_{it} = \phi\mu_{it-1} + \epsilon_{it}$. Iteratively substituting in for the lagged value yields

$$\mu_{it} = \epsilon_{it} + \phi\epsilon_{it-1} + \phi^2\epsilon_{it-2} + ... + \phi^k\epsilon_{it-k}.... \quad (2)$$

Since the series is non-stationary,$\phi > 1$, it implies that temporally distant shocks have a *greater* effect on the present than those which are closer in time. In simple terms, the effect of the past is not only persistent but compounding.

[16]Because the first and second lags are needed to simulate this model, we add the mean increase between 1200 and 1300 of seven thousand to each of these values. For the subsequent five periods we simulate the predicted urban populations using the estimates from this model.

estimated difference is predicted to become about 470,000.[17]

3.2 From Urbanization to Per Capita Income

Because urbanization is only a proxy for development we proceed now to regress per capita income in 1870 and 1900, that is, at height of the industrial revolution, on urban density in 1800 (i.e., right before the process of takeoff occurred). The unit of analysis is the current NUTS-2 region (as defined by the European Union). The data covers eleven countries of western and central Europe.[18] Results are reported in the upper panel of Table 2.

The relationship is both statistically significant and strong from a substantial point of view. Taking the model from the first column of Table 2 and manipulating urban density across its interquartile range we get predicted incomes of $1,714 and $2,213 in 1870. These are extremely close to the true interquartile values of income per capita in 1870 of $1,312 and $2,429. These results are robust to the log-transformation of income, the inclusion of country fixed effects, and the addition of geographic controls. Moreover, a sensitivity analysis (following the procedure proposed by Oster (2013)) indicates that these results are, indeed, robust to the presence of unobservable factors. Across specifications the magnitude of relationship between urban density in 1800 and incomes in the 1870 and 1900 is roughly unchanged at her proposed "rule-of thumb" level of significance.

The relationship between urban density in 1800 and income per capita persisted into the 20th century. The lower panel of Table 2 now treats per capita income for all NUTS-2 regions in 2008 as the outcome and shows a positive and statistically significant relationship with urban density in 1800. Again, the size of the point estimate is substantial - a one hundred percent change in urban density in 1800 is predicted to yield between a $2583 and $3060 increase in per capita income in 2008. To make the results directly comparable to the analysis for the 19th century, columns 7 through 12 in the lower panel exclude regions not employed in the upper panel. The results remain qualitatively unchanged. When we conduct Oster (2013)'s sensitivity analysis, across specification, the relationship between urban density in 1800 and incomes in the 19th century and early 20th

[17]This divergence 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).

[18]The United Kingdom, France, Germany, Spain, Austria, Italy, Sweden, the Netherlands, Denmark, Switzerland, and Belgium.

centuries is robust to the presence of unobservables.

To sum up our results so far, very early random shocks to levels of early urban development explain later differences in urbanization rates across Europe. Moreover, the growth of cities before the industrialization revolution was a non-stationary process where very early differences across location compounded upon each other, leading to the wide divergence in the urbanization rates observed in 1800. Finally, those differences in urbanization just prior to the industrial revolution were correlated with both late 19th- and early 20th-century incomes.

4 Urbanization and Political Institutions

Were parliaments, i.e. institutions imposing checks and balances on rulers, related to development? And if so, in what ways? Did they lead to the development of commercial groups and cause economic growth or did they just reflect the distribution of power across social groups - typically urban, commercial, elites versus landed interests? We answer this question in two steps. First, we estimate the impact of urban growth on parliamentary life, exploiting century on century within-unit changes in urban population to assess how changes in urban density were related with the frequency of parliamentary meetings. Second, we take a similar approach and estimate the relationship between past parliamentary meetings and future urban development. Throughout this section the units of analysis are sovereign and semi-sovereign territories.

The first four columns of Table 3 provide estimates of the impact of urban development on future parliamentary constraints. Column 1 reports pooled OLS estimates, regressing our index of parliamentary institutions (the fraction of years with parliamentary meetings in a given century) on the logged value of urban density (total urban population divided by square kilometers of a given political unit) measured at the beginning of the century.[19] In Column 2 we introduce region and year fixed effects and the result remains substantively unchanged.[20] These models, where we are making comparisons across the entire pooled sample, demonstrate a statistically significant and

[19] In the appendix we replicate all of our results treating the log of the parliamentary meeting index as the outcome. Our results remain substantively unchanged.

[20] We define the following regions: The British Islands, Galliae (Contemporary France and Historical Burgundy), The Holy Roman Empire, Eastern Europe, Scandinavia, Iberia, and Italy

positive relationship. However, when we successively introduce political unit fixed effects (e.g., fixed effects for each sovereign or semi-sovereign state) and year fixed effects in Columns 3 and 4, we find that the relationship between urban growth and the frequency of future parliamentary meetings disappears.

The next five columns test institutionalist theories of growth by regressing urban density on the frequency of parliamentary meetings (conditional on lagged urban density) for all states between 1200 and 1800.[21] In Columns 5 to 7 we report pooled OLS estimates. In Column 5 we report the unconditional relationship between urban density and past parliamentary life, showing a positive and statistically effect of the latter on urban development. This effect remains nearly identical when we include region and year effects (Column 6). However, once we control for past levels of urban density (Column 7), the magnitude of the estimate falls by over two-thirds. Once we include political unit fixed effects (Column 8), i.e. once we identify the relationship between city growth and parliamentary institutions from within political unit changes, the relationship becomes negative and statistically significant. Finally, the effect becomes null after introducing time fixed effects (Column 9) and this null result persists after controlling for past value of urban density as well as unit and time effects (Column 10). In short, parliamentary constraints had a negative or null effect on urban growth.

In Table 4 we explore the causes of the robust cross-sectional relationship but null within-unit relationship described in Table 3. Initial economic conditions (urban density in 1200) are a consistently strong predictor of parliamentary meeting frequency (in each century) even after we control for overall changes in urban population across each period. We control for changes in urban population density to the regressions in two ways. In columns 1 through 5 we included only the change over the entire interval between the initial period and the period of observation. In columns 6 through 9 we control for urban density change in all previous centuries. Urbanization growth does not enter significantly into the regressions in columns 1 through 5 except for urban density change during the fourteenth century. When we include all the lagged values of urban density change, urban growth has a shifting relationship with parliamentary growth: higher urban growth is associated

[21] In the appendix we replicate all of our results treating the log of the parliamentary meeting index as the independent variable. Our results remain substantively unchanged.

with more parliamentary meetings in the five-, three-, and one-century differences and with stronger parliaments in four- and two-century century differences. To sum up, the main finding of Table 4 is that initial conditions shaped the distribution of parliamentary institutions in the medium and – in the same way that, according to the results in Table 1, the initial urbanization patterns of 1200 determined the subsequent urbanization path across Europe until 1800 (and beyond).

5 Endogenous Growth: The Effect of Know-How or Natural Resources?

5.1 The Emergence of a Proto-Industrial Core

Having tested the impact of institutions on urban development, we proceed now to examine our alternative explanation where economic growth is a function of technological innovation, embodied in a class or sector of society, namely artisans and craftsmen, that invest in either the generation of new technologies or the incremental transformation of existing techniques.

To measure this process, we employ our data on the existence of protoindustrial centers, which, as Figures 4 and 5 make apparent, matched the distribution of European urban population. We model that process in two steps. Table 5 regresses the number of textile or metallurgic centers on each geographical quadrant between 1200 and 1500 on of level of urban population on the same unit in the year 1200. Next, Table 6 examines the effect of having a proto-industrial center in the middle ages on urbanization growth after 1500 (controlling for past urbanization).

Table 5 reports estimates first treating the number of textile (columns 1-4) and iron centers (columns 5-8) separately as the outcome and next taking the sum of both types of protoindustry as the outcome (columns 9-12). The unit of observation is the geographical quadrant. Employing OLS (the first two columns for each dependent variable) and negative binomial estimates (the last two columns), across specifications early urban density is positively associated with the presence of proto-industry. Moreover, the magnitude of this relationship is substantively large: the OLS estimates indicate that a one-hundred percent change in urban population in the year 1200 would result in between .28 and .45 of a new industrial center. Our findings survive the inclusion of the full set of controls and are robust well beyond rule-of-thumb levels of significance proposed in Oster (2013)'s method for detecting bias based on the presence of confounding unobservables.[22]

In the appendix we provide further evidence that the relationship between early urban density and the development of proto-industrial skills is causal. There, we exploit climatic shocks in the ability to feed large populations in order to identify this effect in an instrumental variables framework. The 2SLS estimates are larger, indicating a .85 predicted increase following a 100% change in initial urban population. Again, we assess the robustness of these instrumental variables estimates by conducting the sensitivity analysis proposed by Conley, Hansen and Rossi (2012). Using the more conservative union-of-confidence-intervals approach, we find that there would need to be a substantively large direct effect of our instrument (between 76% and 81% the magnitude of each of our estimated effects) to violate the exclusion restriction and make our results statistically insignificant.

Table 6 turns to assess the impact of proto-industrial centers (in place before 1500) on urbanization in 1500 controlling for the effect of urbanization in 1200. In Table 6 we regress urban population in 1500 on the number of proto-industrial center in a particular unit. The relationship between proto-industrial activity (measured through number of proto-industrial centers in a particular unit before 1500) and urban density in 1500 is positive and statistically significant. The introduction of a full set of geographical controls (in the odd columns) does not substantively alter the relationship between the presence of these skills and density in 1500 even though it reduces the magnitude of the independent effect of proto-industry in some cases.

Measuring the outcome in logs, the addition of a single center of proto-industrial activity is estimated to be between over one-quarter and about one half of the magnitude associated with past urban development. For example, depending on the set of controls included, the addition of a single industrial center before 1500 is estimated to yield between a 14% and 39% increase in urban density in 1500. In comparison, the effect of a one-hundred percent change in urban population in 1200 is predicted to yield between a 44% and 65% change over the same period (Columns 11 and

[22] In the appendix we also show that these results are robust to successive changes in the specification of the independent variable, 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.

12).

The results of Table 6 can be interpreted as corroborating endogenous growth theories (Romer 1990) as well as geographic concentration models (Krugman 1991), where urban or economic clusters foster, due to 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 their biogeographical endowments experienced ever-faster growth rates.

6 Coal and Trade

In addition (or independently) of a proto-industrial base, easy access to cheap and abundant energy sources (coal) has been seen as a crucial condition behind the growth of Europe and, particularly, the location of the industrial revolution (Landes 1965, Pomeranz 2002). To test this hypothesis, following Findlay and O'Rourke (2014) we digitized the Les Houilléres Européennes map from Chatel and Dollfus (1931), which records the location of 124 major nineteenth century coal fields in Europe, and measure the distance for each of our units - either arbitrary grid squares or regions - to the closest coal field.

To explore the relationship between access to coal and urban development in 1800 we estimate the following model:

$$\mu_{i,t} = \alpha + \beta_1 \mu_{i,t-1} + \beta_2 D_i + \beta_3 D_i \times \mu_{i,t-1} + \epsilon_i \tag{3}$$

where $\mu_{i,t}$ is the logged urban population of quadrant i in the year 1800, $\mu_{i,t-1}$ is logged urban population in the same quadrant in an earlier period - either 1500 or 1700 - and D_i is the logged distance of quadrant i to the nearest coal field. Again, ϵ_i is a mean zero random disturbance. The parameter β_1 tells us the direct effect of past urban density, β_2 the direct effect of distance to coal, and β_3 the effect of urban density as it varies by distance to coal. The marginal effect of distance to coal is given by $\beta_2 + \beta_3 \mu_{i,t-1}$.

Estimates of these effects are given in the top panel of Table 7. Both the effect of past urban density and distance to coal are significant and in the expected direction in models where we exclude the interaction term (Columns 1-2 and Columns 5-6). The effect of urban population is substantively larger than that associated with distance to coal. According to model 1, for example, a 100% change in distance to coal reduced urban population in 1800 by 23% whereas the effect of a 100% change in urban population in 1500 is associated with 53% change. However, once we interact past values of urban population with the distance to the nearest coal field (Columns 3 and 4 of Table 7), we find that the negative effect of distance to coal declines in past levels of urban agglomeration. In fact, once we condition on the full set of controls (Columns 4 and 8) the relationship between distance to coal and urban density in 1800 is null.

The lower two panels of Table 7 examines both urban density in 1800 and distance to coal on regional per capita incomes in 1870 and 1900, respectively.[23] Corroborating the results from Table 2, a 100% change in urban density is associated with an increase in per capita income of between \$346 and \$442 in 1870. The relationship between a 100% change in distance to coal mines and income is 1870 is between $\$-297$ and -\$407. However, once we include country fixed effects, the relationship between coal and incomes becomes statistically indistinguishable from zero. Moreover, when we interact urban density in 1800 with distance to nineteenth century coal mines, the effect of distance to coal is not statistically significant in any model with control variables.

Easy access to transportation means, such as the sea, has been associated with the rise of trade, the expansion of urban life, and growth (Bairoch 1988, Tracy 1990, Jones 1991, Braudel 1995). Within this general interpretation of the effects of geography on the economy, several authors link the rise of incomes in the European northwest to the rise of the Atlantic trade (and the closing of Mediterranean routes after the fall of Constantinople) (Davis 1973, Acemoglu et al. 2005).

To examine the effect of having access to both the Atlantic and the Mediterranean, we estimate the following model:

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

[23] In the online appendix we report our estimations on the impact of coal mines location on regional per capita incomes in 2008. The results with respect to coal access and urban density are qualitatively similar.

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 parameters β_t and γ_t capture the time varying effect of access to the Atlantic and Mediterranean seas respectively, in interaction with the set of time dummies, δ_t.

We operationalize access to the sea in two ways: as a dummy for whether or not a given grid-square contains Atlantic coast (where Atlantic coast is defined following Acemoglu, Johnson and Robinson (2004)) or the Mediterranean coast; and employing distance in kilometers from the geometrical center of the quadrant to the coast. To compare the change in urban growth associated with Mediterranean versus Atlantic coasts, we test the restriction that Atlantic-exposed units grew at the same rate as those on the Mediterranean for each period ($\beta_t - \gamma_t = 0$). Results are presented in the top panel of Table 8. While access to both the Atlantic and Mediterranean were associated with increases in urban population, we cannot reject the null that the access to the Mediterranean gave the same advantage as access to the Atlantic for any period.

The bottom panel of Table 8 examines the impact of having access to the Atlantic conditioned by level of urbanization. The estimation includes a lagged dependent variable and allows the effect of the Atlantic to vary by period. In the specification that employs a dichotomous measure of access, the relationship is negative for 1300 and statistically insignificant for the years 1400 to 1700. It only becomes positive and statistically significant for 1800. When we use distance to the Atlantic as our measure of access, territories closer to the Atlantic were, on average, less developed than those far away.

6.1 Trade and Parliaments

Moving beyond a standard story stressing the unconditional effect of trade access on growth, Acemoglu, Johnson and Robinson (2004) claim that the rise of Western Europe after 1500 can be traced back to the combination of constraining political institutions, e.g. parliaments, and access to the Atlantic trade. We revisit their analysis here using our political unit time-varying measures of parliamentary constraints - instead of their time-invariant measure (for the year 1415) coded at a much higher level of spatial aggregation.

To begin, we follow Acemoglu, Johnson and Robinson (2004) in estimating the following baseline model:

$$\mu_{it} = \alpha_i + \sum_{t \geq 1500}^{T} \beta_{1t} \times \delta_t \times \text{Atlantic}_i + \beta_{2t} \times \delta_t \times \text{Atlantic}_i \times \text{P-Index}_{it-1} + \sum_{t \geq 1500}^{T} \gamma_t \times \delta_t \times \text{W. Europe}_i + \delta_t + \theta \times \text{P-Index}_{it-1} + \epsilon_{it} \quad (5)$$

where β_{1t} captures the effect of access to the Atlantic in period t, β_{2t} captures how this effect varies with the frequency of parliamentary constraints, θ captures the direct effect of parliamentary constraints, and δ_t are a set of time dummies. As in Acemoglu, Johnson and Robinson (2004) we estimate these parameters after having controlled for the broader trend of urban growth in Western Europe, given by the parameters γ_t, and unit fixed effects, α_i. Our unit of observation is political unit as defined in each century.

Table 9 presents our results. In Columns 1 through 4 we use a dichotomous, time invariant, measure of potential for Atlantic trade. Column 1 reproduces Acemoglu, Johnson and Robinson (2004)'s main result and confirms that after the 17th century access to the Atlantic was positively associated with changes in urban development. However, in Column 2, where we condition on the previous century's level of urban density, the relationship between access to the Atlantic and urban growth is null except for 1800 or 300 years after the discovery of the New World. The next two columns estimate the interactive relationship between access to the Atlantic and the existence of parliamentary constraints. Column 3 interacts our measure of parliamentary constraint with the Atlantic trade access dummy. Column 4 interacts the parliamentary constraint measure with the full set of Atlantic access and post 15th century time dummies. In both models there is no evidence of a statistically significant relationship between parliaments, trade, and growth.

In Columns 5 through 9 of Table 9, we use Acemoglu, Johnson and Robinson (2004)'s second measure of potential for Atlantic trade: the ratio of Atlantic coast-line to the total area of the state. Instead of using the boundaries of 20th century states to measure access to the Atlantic, we measure the ratio contemporaneously with urban density and parliamentary constraints, giving us a time-varying measure of coast access. Because of this we can include it directly as a covariate instead of only estimating changes in its century-on-century effect via interactions with a series of time dummies, yielding the total effect of access to the Atlantic across time rather than just how access to the Atlantic changed across time.

Column 5 estimates the average effect across time and finds a statistically significant relationship between Atlantic access and urban growth. The next two columns repeat the same exercise as in columns 1 and 2, estimating the effect of access to the Atlantic across time. The time varying components are each statistically significant but the direct effect is null when we condition on past values of urban population. The last two columns, which report the interactive effect of access to the Atlantic and parliaments, find no significant relationship between them and urban density. To sum up, while the time varying effect of the Atlantic trade as estimated by Acemoglu, Johnson and Robinson (2004) is significant and positive, the total effect of access to the Atlantic, which in their models is absorbed by unit-fixed effects, is indistinguishable from zero across time periods when we control for past levels of urbanization. Moreover, there is no positive interaction of Atlantic trade and institutional set-up on growth.

7 Conclusion

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, segmented in local, autarkic markets. Urban centers were small and far apart from each other. Excluding those cities under Arab or Byzantine control in the Mediterranean, the largest towns were Rome with 35,000 inhabitants and London and Laon (in France) with 25,000 people. Less than four hundred years later, however, areas such as Italy and the Low Countries had a flourishing urban economy. By 1850, the industrial revolution was under course in northwestern Europe.

Employing fine-grained geographic, economic and political data covering 700 years of history we show, in the first place, that the long-run process of European economic development was related to the formation and expansion of urban clusters across the continent. After political conditions stabilized around the turn of the millennium, the introduction of new techniques such as the heavy plow and the three-field rotation led to larger crop yields in those regions endowed with rich soils and suitable climate conditions. Those areas with a substantial cereal surplus sustained a growing non-farming population that joined in urban agglomerations and that specialized in a variety of artisanal and proto-industrial activities. In turn, those urban clusters made of traders and artisans fostered an incremental process of technological innovation and of capital accumulation.

Besides accounting for the very strong historical continuities in urban life across the continent from 1200 till today, these processes also explain why, in line with a growth model with increasing returns to scale and positive intra-sectoral externalities, urban growth exhibited a divergent pattern across the continent. Cities that were relatively larger at the beginning of the period kept adding population at a faster rate than smaller towns such that, by the end of modern era, Europe had 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. Urban densities remained much lower, instead, in its western and eastern peripheries.

That process of urban agglomeration and growth, strongly correlated with the formation of protoindustrial centers, appears to be at the roots of 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), λ-knowledge) needed to take advantage of the new general knowledge generated by the scientific revolution of the 17th and 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. Indeed, urban life and the distribution of proto-industries in medieval and early modern Europe predict cross-regional variation in per capita income in the late 19th century and at the turn of the 21st century quite strongly. By contrast, proximity to energy sources and access to the Atlantic did not seem to play any significant role in the rise of northwestern Europe.

Proto-parliamentary institutions and pluralistic governance structures followed the expansion of urban life but, contrary to current institutionalist theories, which make them "fundamental determinants" of development (North and Weingast 1989; Acemoglu, Gallego and Robinson 2014), played a negligible causal role (alone or in interaction access to the Atlantic trade) in fostering economic growth across the continent. The wealth and population of Italian and Flemish towns allowed them to defeat their ecclesiastical or feudal lords over the 12th and 13th centuries (Weber 1968, Pirenne 1969, Najemy 2006). Likewise, parliamentary institutions only remained in place in modern Europe in those proto-capitalist enclaves where a wealthy urban class had the means to oppose absolutism. Dutch cities joined in a military league and then a republic that eventually defeated Spain. In England the parliamentary forces and the pro-trade party won over the royal forces in 1640 and again in 1688. As Pincus (2009) writes in his landmark study of the Glorious Revolution, England in the second half of the 17th century was rapidly becoming a modern society with a booming economy, growing cities and expanding trade. Inspired by the Dutch example, the opponents of James II supported the principle of limited government, rejected James II's political-economic program based on land interests at home and territorial acquisition abroad and embraced urban culture, manufacturing and economic imperialism - understood as - commercial hegemony (ibid, 484). Political institutions were endogenous to the structure of economic and commercial life: it was only those places that had a sufficiently wealthy and cohesive class of "burghers" that could defeat the landed and monarchical elites and sustain the process of endogenous growth that eventually led to the industrial revolution.

References

Abramson, S. 2015. "The Economic Origins of the Territorial State." *PhD Dissertation* .

Acemoglu, D., S. Johnson and J.A. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4):1231–1294.

Acemoglu, D., S. Johnson and J.A. Robinson. 2004. "The Rise of Europe: Atlantic Trade, Institutional Change and Economic Growth." *American Economic Review* .

Alesina, Alberto, Johann Harnoss and Hillel Rapoport. 2013. Birthplace diversity and economic prosperity. Technical report National Bureau of Economic Research.

Allen, R.C. 2001. "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.

Anderson, P. 1974. *Lineages of the Absolutist State*. Verso Books.

Arellano, M. and O. Bover. 1995. "Another look at the instrumental variable estimation of error-components models." *Journal of econometrics* 68(1):29–51.

Arellano, M. and S. Bond. 1991. "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.

Arrow, K. 1962. "Economic Welfare and the Allocation of Resources for Invention.".

Bairoch, P. and C. Braider. 1991. *Cities and Economic Development: From the Dawn of History to the Present*. University of Chicago Press.

Bairoch, P., J. Batou and C. Pierre. 1988. *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)*. Librairie Droz.

Becker, G., K. Murphy and R. Tamura. 1990. "Economic Growth, Human Capital and Population Growth." *Journal of Political Economy* 98(5):S12–S137.

Blundell, R. and S. Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of econometrics* 87(1):115–143.

Boix, C. and M. Svolik. 2013. "The foundations of limited authoritarian government: Institutions and power-sharing in dictatorships." *The Journal of Politics* 75(2):300–316.

Boix, Carles. 2015. *Political Order and Inequality: Their Foundations and Their Consequences for Human Welfare*. Cambridge University Press.

Brenner, R. 2003. *The Boom and the Bubble: the US in the World Economy*. Verso Books.

Cagé, Julia and Valeria Rueda. 2014. "The long-term effects of the printing press in Sub-Saharan Africa." *Available at SSRN 2292660* .

Carus-Wilson, E.M. 1966. *The Woollen Industry. Cambridge Economic History of Europe. Volume II.* Cambridge University Press.

Caselli, Francesco, Gerardo Esquivel and Fernando Lefort. 1996. "Reopening the convergence debate: a new look at cross-country growth empirics." *Journal of economic growth* 1(3):363–389.

Chanda, A. and L. Putterman. 2007. "Early Starts, Reversals and Catch-up in the Process of Economic Development*." *The Scandinavian Journal of Economics* 109(2):387–413.

Clark, G. 2008. *A Farewell to Alms: A Brief History of the World*. Princeton University Press, Princeton.

Conley, Timothy G, Christian B Hansen and Peter E Rossi. 2012. "Plausibly exogenous." *Review of Economics and Statistics* 94(1):260–272.

De Long, J.B. and A. Shleifer. 1993. "Princes and Merchants'." *Journal of Law and Economics* 36.

De Vries, J. 1984. *European Urbanization, 1500-1800*. Methuen London.

Diamond, J. 1997. "Guns, Germs and Steel.".

Duby, G., H.B. Clarke and M.B. Becker. 1974. "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.

Engerman, S.L. and K.L. Sokoloff. 2002. Factor Endowments, Inequality, and Paths of Development Among New World Economics. Technical report National Bureau of Economic Research.

FAO, IIASA. 2000. "Global Agro-Ecological Zones (GAEZ).".

Galor, Oded, Omer Moav and Dietrich Vollrath. 2009. "Inequality in landownership, the emergence of human-capital promoting institutions, and the great divergence." *The Review of economic studies* 76(1):143–179.

Gandhi, J. and A. Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies* 40:1279–1300.

Gutmann, M.P. 1988. *Toward the Modern Economy: Early Industry in Europe, 1500-1800*. Temple University Press.

Hibbs, D.A. and O. Olsson. 2004. "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.

Inkeles, A. 1969. "Making Men Modern: On the Causes and Consequences of Individual Change in Six Developing Countries." *American Journal of Sociology* pp. 208–225.

Jones, EL. 1981. "The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia.".

Kremer, M. 1993. "Population Growth and Technological Change: One Million BC to 1990." *The Quarterly Journal of Economics* 108(3):681–716.

Krugman, P.R. 1991. *Geography and Trade*. The MIT Press.

Kuznets, S. 1968. *Toward a Theory of Economic Growth: With Reflections on the Economic Growth of Modern Nations*. Norton.

Laitin, David D and Rajesh Ramachandran. 2014. "Language Policy and Human Development.".

Lipset, Seymour Martin. 1959. "Some social requisites of democracy: Economic development and political legitimacy." *American political science review* 53(01):69–105.

Lopez, R.S. 1976. *The Commercial Revolution of the Middle Ages, 950-1350*. Cambridge Univ Pr.

Maddison, A. 2003. *The World Economy: Historical Statistics*. Organization for Economic Cooperation and Development.

Mokyr, J. 1990. "Punctuated equilibria and technological progress." *The American Economic Review* 80(2):350–354.

Mokyr, J. 2004. *The Gifts of Athena: Historical Origins of the Knowledge Economy.* Princeton Univ Pr.

Nicholas, D. 1997. "The Growth of the Medieval City: From Late Antiquity to the Early Fourteenth Century.".

Nickell, Stephen. 1981. "Biases in dynamic models with fixed effects." *Econometrica: Journal of the Econometric Society* pp. 1417–1426.

North, D.C. 1990. "Institutions, Institutional Change and Economic Performance.".

North, D.C. and B.R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-century England." *Journal of Economic History* 49(4):803–32.

North, D.C. and R.P Thomas. 1973. "The Rise of the Western World: A New Economic History.".

North, D.C. et al. 1981. *Structure and Change in Economic History.* Norton New York.

Olson, M. 1993. "Dictatorship, democracy, and development." *American Political Science Review* pp. 567–576.

Olson, M. 2000. *Power and Prosperity: Outgrowing Communist and Capitalist Dictatorships.* Basic books.

Oster, Emily. 2013. Unobservable selection and coefficient stability: Theory and validation. Technical report National Bureau of Economic Research.

Parente, S.L. and E.C. Prescott. 1994. "Barriers to technology adoption and development." *Journal of Political Economy* pp. 298–321.

Parente, S.L. and E.C. Prescott. 1999. "Monopoly Rights: A Barrier to Riches." *American Economic Review* pp. 1216–1233.

Pomeranz, K. 2001. *The Great Divergence: China, Europe, and the Making of the Modern World Economy.* Princeton Univ Pr.

Putnam, R.D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton Univ Pr.

Rivera-Batiz, L.A. and P.M. Romer. 1991. "Economic Integration and Endogenous Growth." *The Quarterly Journal of Economics* 106(2):531–55.

Romer, P.M. 1996. "Why, Indeed, in America? Theory, History, and the Origins of Modern Economic Growth." *American Economic Review* 86(2):202–206.

Sachs, J.D. and A.M. Warner. 1997. "Fundamental Sources of Long-run Growth." *The American Economic Review* 87(2):184–188.

Satyanath, Shanker, Nico Voigtländer and Hans-Joachim Voth. 2013. Bowling for fascism: social capital and the rise of the Nazi Party. Technical report National Bureau of Economic Research.

Shaver, Andrew C, David Carter and Tsering W Shawa. 2013. "Terrain Ruggedness and Land Cover: New Data." *Princeton University, unpublished manuscript* .

Smith, A. 1937. "The Wealth of Nations." *New York: Modern Library* p. 740.

Solow, R.M. 1956. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70(1):65–94.

Sprandel, R. 1968. *Das Eisengewerbe im Mittelalter*. A. Hiersemann.

Stasavage, D. 2003. *Public debt and the birth of the democratic state: France and Great Britain, 1688-1789*. Cambridge Univ Pr.

Stasavage, D. 2011. "When Distance Mattered: Geographic Scale and the Development of European Representative Assemblies." *American Political Science Review* .

Strayer, J.R. 1973. *On the Medieval Origins of the Modern State*. Princeton Univ Pr.

Svolik, M.W. 2009. "Power sharing and leadership dynamics in authoritarian regimes." *American Journal of Political Science* 53(2):477–494.

Tabellini, Guido. 2010. "Culture and institutions: economic development in the regions of Europe." *Journal of the European Economic Association* 8(4):677–716.

Tilly, C. 1990. *Coercion, capital, and European states, AD 990-1990*. Basil Blackwell.

van Zanden, J.L., E. Buringh and M. Bosker. 2010. "The Rise and Decline of European Parliaments, 1188-1789.".

Voigtländer, N. and H.J. Voth. 2006. "Why England? Demographic Factors, Structural Change and Physical Capital Accumulation during the Industrial Revolution." *Journal of Economic Growth* 11(4):319–361.

Weber, M. 1968. "Economy and Society (G. Roth & C. Wittich, Eds.)." *New York: Bedminster* .

Figures & Tables

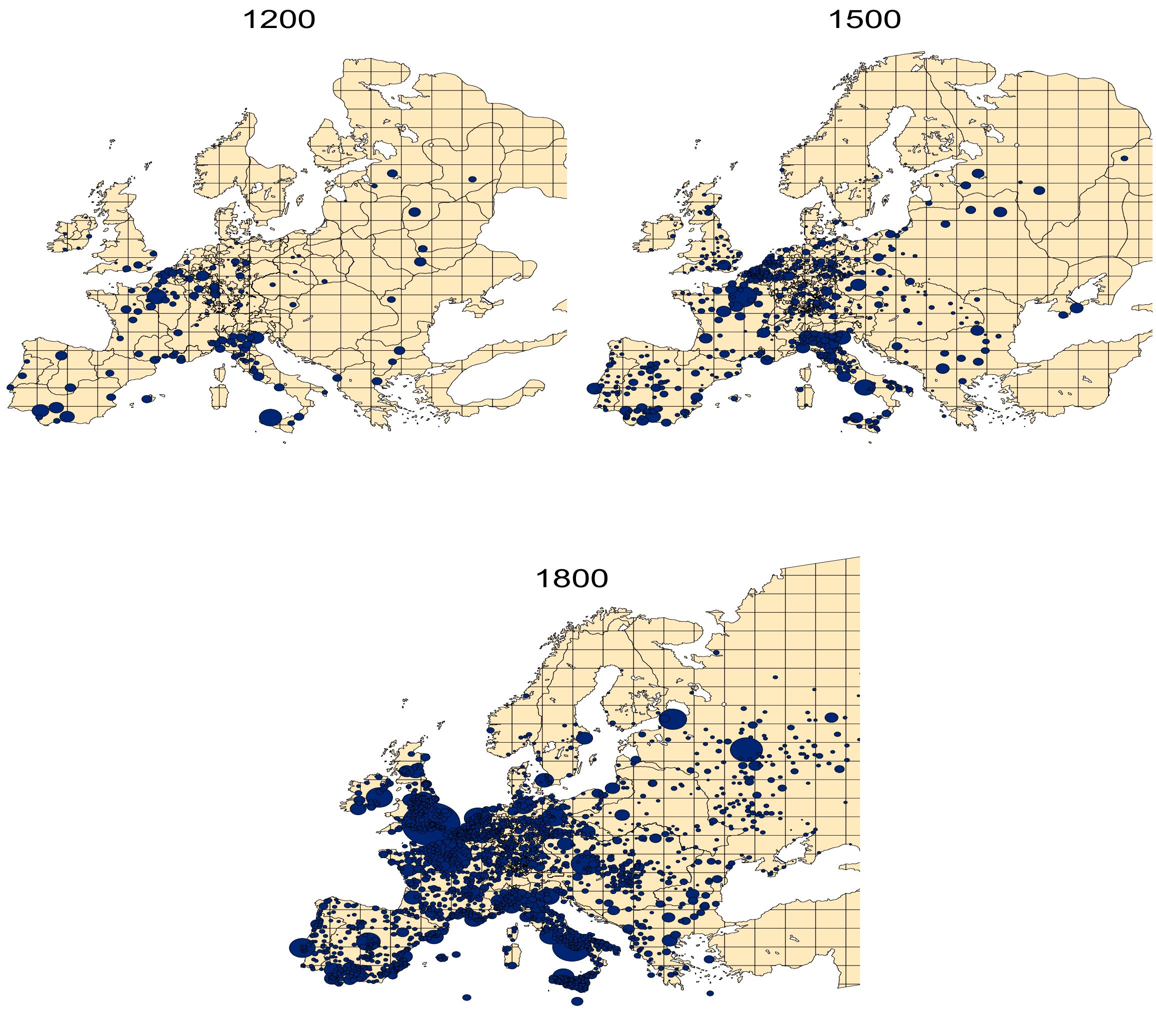

Figure 1: Urban populations in Europe in 1200, 1500 & 1800 (in thousands).

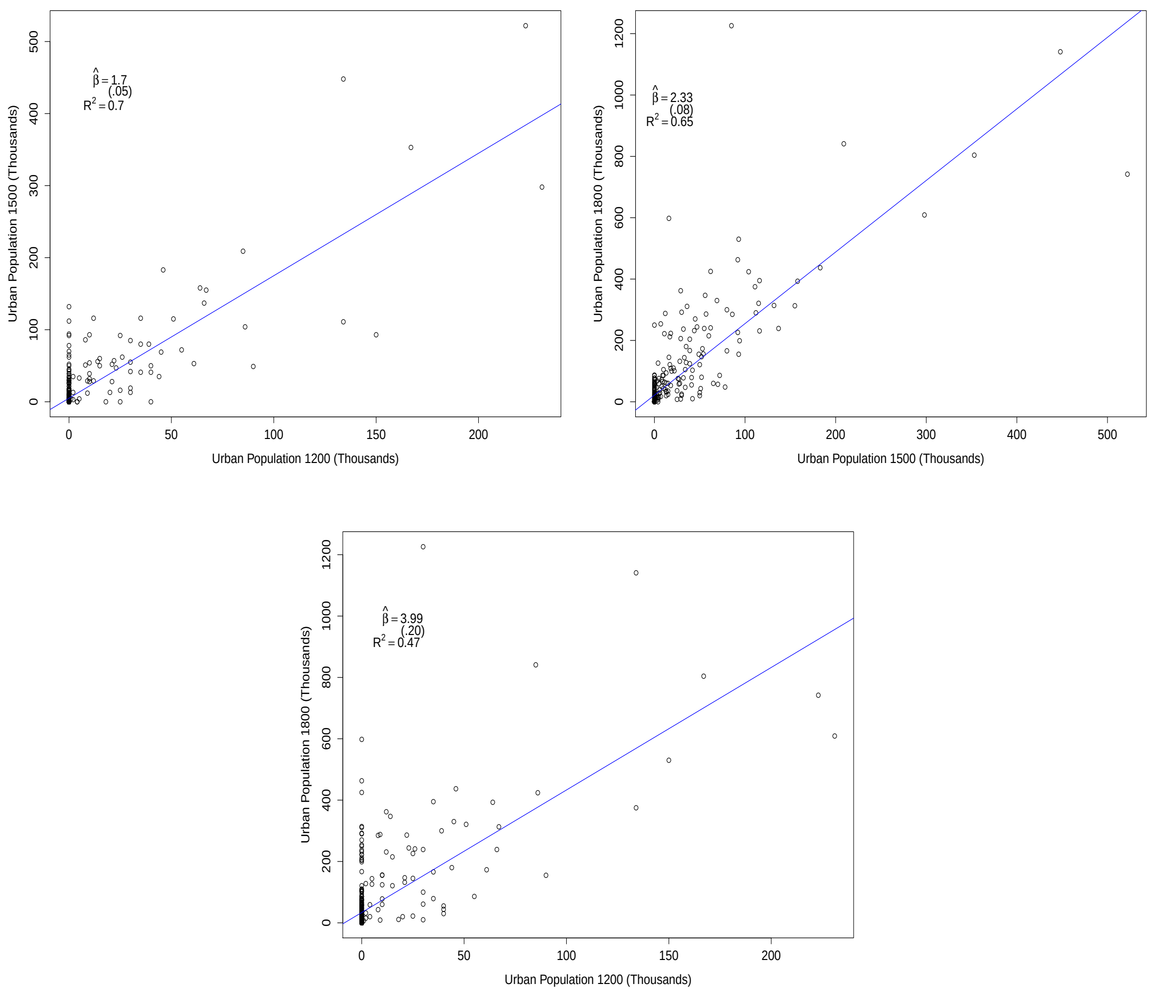

Figure 2: Bivariate relationship between urban population and future urban population across time.

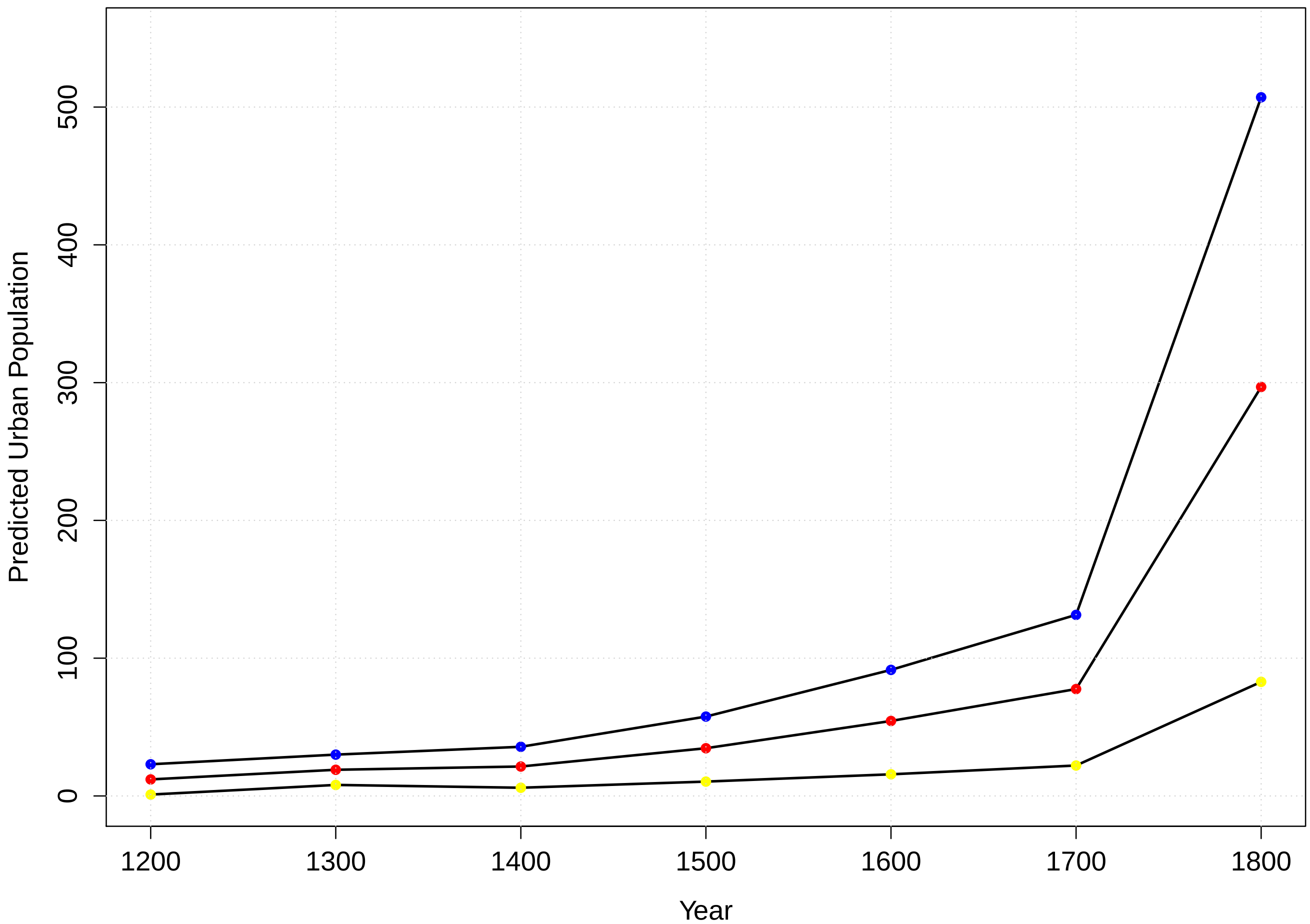

Figure 3: Estimates of the difference in urban population across time derived from the dynamic system GMM estimates of Model 1 with two lags. The predicted values are constructed using starting urban populations for two units of $1,000$, $12,000$ and $24,000$ in the year 1200, approximately a one half, 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

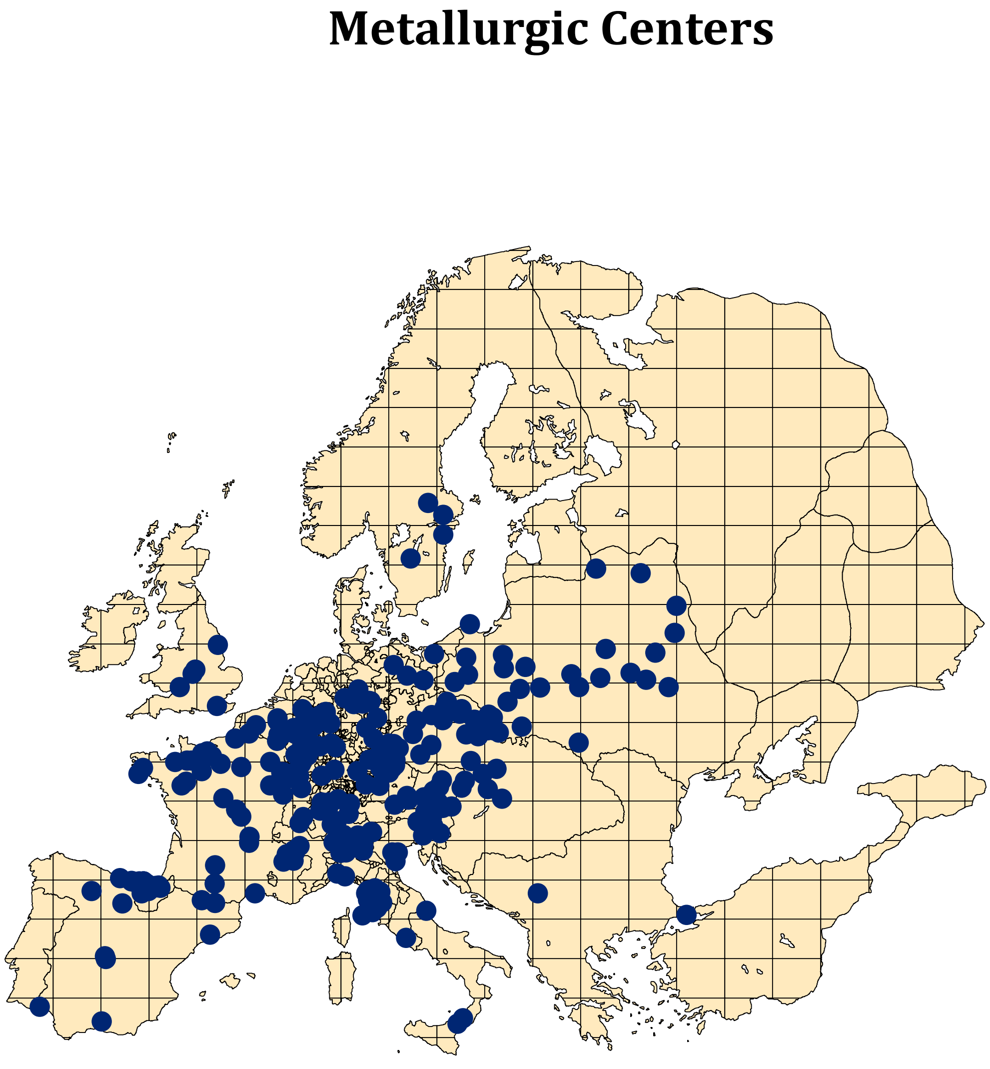

Figure 4: Iron Production Centers before 1500.

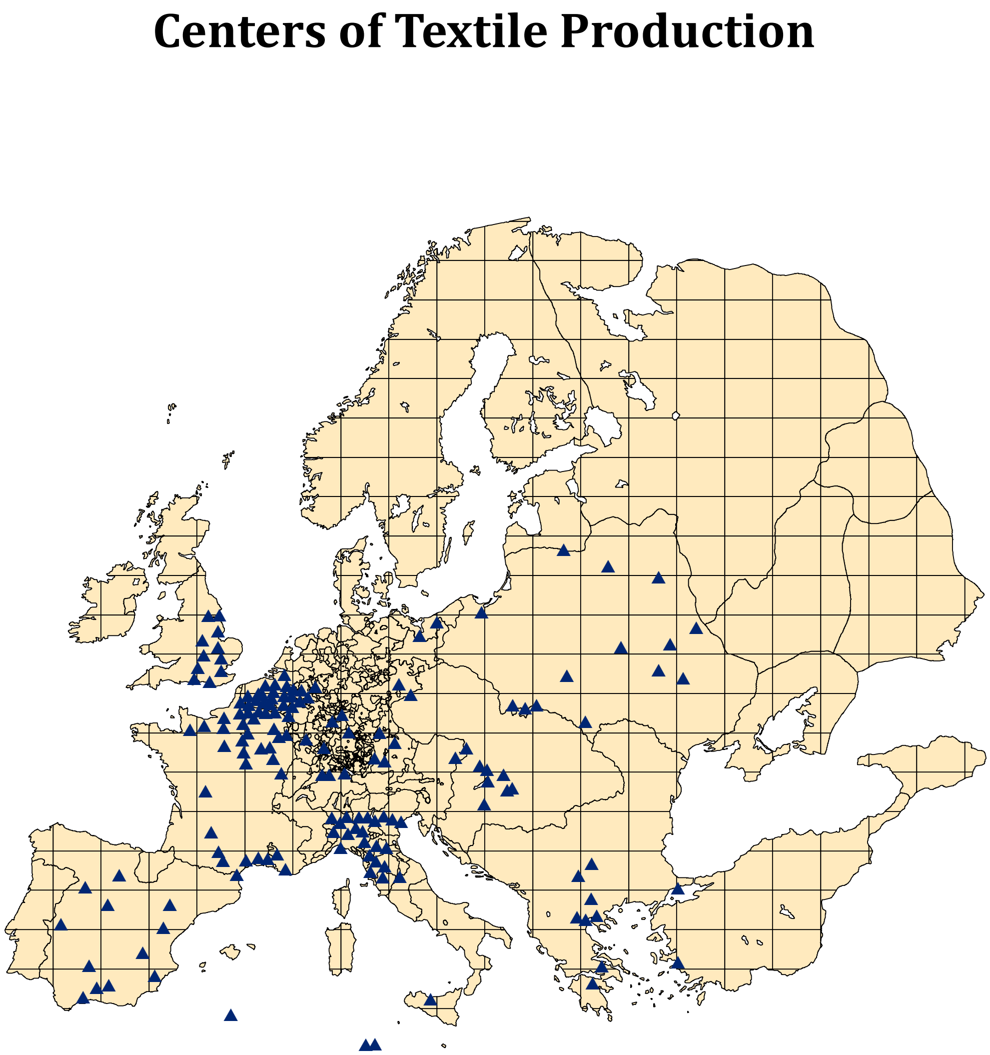

Figure 5: Textile Centers before 1500.

The Pre-Industrial Structure of Urban Population Growth (Autoregressive Models)

	Outcome: Urban Pop_t							Outcome: log(Urban Pop)$_t$					
	OLS			*GMM*				*OLS*			*GMM*		
Urban Pop_{t-1}	1.29***	1.12***	1.26***	1.27***	0.94***	0.94***	log(Urban Pop)$_{t-1}$	0.89***	0.65***	0.62***	0.68***	0.77***	0.67***
	(0.05)	(0.18)	(0.22)	(0.09)	(0.08)	(0.22)		(0.01)	(0.04)	(0.04)	(0.05)	(0.05)	(0.06)
Urban Pop_{t-2}		0.23	0.17		0.58***	0.42***	log(Urban Pop)$_{t-2}$		0.30***	0.27***		0.25***	0.31***
		(0.20)	(0.16)		(0.09)	(0.11)			(0.04)	(0.05)		(0.06)	(0.09)
Urban Pop_{t-3}			-0.10			0.32*	log(Urban Pop)$_{t-3}$			0.09**			0.07
			(0.17)			(0.14)			(0.03)			(0.04)	
R^2	0.85	0.86	0.87					0.70	0.77	0.77			
$m2$				0.94	0.70	0.08					.86	0.17	0.25
N	2664	2664	2220	2220	1776	1776		2664	2664	2220	2220	1776	1776
	Breitung Test (H_0: All Panels Contain Unit Roots)			Hadri LM Test (H_0:All Panels Are Stationary)				Breitung Test (H_0: All Panels Contain Unit Roots)			Hadri LM Test (H_0:All Panels Are Stationary)		
	8.51	24.33	28.85	22.00***	42.56 ***	16.21***		2.47	9.49	3.80	22.08***	16.35***	4.02***
	(0.99)	(0.99)	(0.99)	(0.00)	(0.00)	(0.00)		(0.99)	(0.99)	(0.99)	(0.00)	(0.00)	0.00)
Unit Effects		Y	Y		Y	Y			Y	Y		Y	Y
Time Trend			Y			Y				Y			Y

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1: This table presents the estimates of the autoregressive relationship between past and present urban development. The unit of observation is the 225 km x 225 km grid-square. The top panel measures total urban population and the lower takes the logarithm of this number. Heteroskedaticity robust standard errors clustered by unit in parentheses. p-value for the Arellano-Bond test of second order serial correlation in the errors denoted as $m2$.

The Relationship Between Urban Density in 1800 and Per Capita Income in the 19th and 20th Centuries

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
	GDP per capita 1870						**GDP per capita 1900**					
	GDP per capita			log(*GDP per capita*)			*GDP per capita*			log(*GDP per capita*)		
log(Urban Density$_{1800}$)	400.11*	329.61***	370.20***	0.19**	0.15**	0.16**	556.82**	509.65***	581.52***	0.18**	0.16***	0.18***
	(129.27)	(71.24)	(50.78)	(0.06)	(0.04)	(0.04)	(155.45)	(75.47)	(55.21)	(0.06)	(0.03)	(0.03)
Constant	3969.30***	3365.79***	37.64	8.44***	8.04***	5.85***	5424.13***	4828.27***	70.33	8.72***	8.37***	6.25***
	(804.35)	(420.85)	(429.47)	(0.35)	(0.24)	(0.34)	(979.86)	(445.83)	(759.07)	(0.33)	(0.18)	(0.23)
Controls	N	N	Y	N	N	Y	N	N	Y	N	N	Y
Country Fixed Effects	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
N	104	104	104	104	104	104	106	106	106	106	106	106
R^2	0.28	0.80	0.85	0.23	0.75	0.82	0.34	0.79	0.85	0.25	0.75	0.83
Effect at 1.3 × R^2 of Controlled Regression			362.16			0.16			588.74			0.18

GDP per Capita in 2008

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
	All Europe						**Western Europe**					
	GDP per capita			log(*GDP per capita*)			*GDP per capita*			log(*GDP per capita*)		
log(Urban Density$_{1800}$)	2583.81*	3060.22***	3379.21***	0.12*	0.10***	0.11***	1832.12*	3029.52**	3543.93***	0.04	0.07**	0.09***
	(978.03)	(478.38)	(349.97)	(0.05)	(0.02)	(0.02)	(867.93)	(614.13)	(419.44)	(0.02)	(0.02)	(0.02)
Constant	15411.75**	6794.65***	-2369.08	9.52***	9.12***	9.26***	21045.11***	32242.94***	-13408.05	10.07***	10.40***	8.74***
	(4609.22)	(516.70)	(26129.59)	(0.25)	(0.02)	(1.05)	(3990.74)	(1414.37)	(18083.10)	(0.11)	(0.04)	(0.67)
Controls	N	N	Y	N	N	Y	N	N	Y	N	N	Y
Country Fixed Effects	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
N	254	254	254	254	254	254	254	254	254	254	254	254
R^2	0.12	0.76	0.79	0.11	0.88	0.90	163.00	163.00	163.00	163.00	163.00	163.00
Effect at 1.3 × R^2 of Controlled Regression			3632.46			0.10			6162.70			0.15

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 2: The top panel of this table describes the relationship between urban density in 1800 and per capita income in 1870 and 1900, respectively. The lower of panel of this table describes the relationship between urban density in 1800 and per capita income in 2008. All observations are at the NUTS-2 level. In the lower panel the first six columns use all NUTS-2 regions for which there is income data. The last six columns use only those in Western Europe. Hetereoskedacticty robust standard errors clustered by country are in parentheses. Controls are: terrain ruggedness, agricultural suitability, distance to coast, coast length, latitude, and longitude. Following Oster (2013) we provide sensitivity estimates of the effects under the hypothetical condition when unobservables account for 1.3 × the R^2 from the controlled regressions.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
	Outcome: Frequency of Parliaments				Outcome: log(Urban Density$_{t+1}$)					
log(Urban Density$_t$)	0.099***	0.103***	0.003	0.001			0.717***			.300***
	(0.01)	(0.01)	(0.01)	(0.01)			(0.049)			(0.148)
Frequency of Parliaments					1.033***	1.079***	.386***	−0.242**	−0.135	-0.031
					(0.167)	(0.164)	(0.085)	(0.119)	(0.107)	(0.157)
Region Effects	N	Y	N	N	N	Y	Y	N	N	N
Year Effects	N	Y	Y	Y	N	Y	Y	Y	Y	Y
Country Effects	N	N	Y	Y	N	N	N	Y	Y	Y
System GMM	N	N	N	N	N	N	N	N	N	Y
R^2	0.094	0.109	0.000	0.042	0.093	0.175	0.583	0.093	0.141	
N × T	1790	1790	1790	1790	1793	1793	1793	1793	1793	1793
N = 309, T=6										
$m2$										0.04

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table 3: The first four columns of this table presets the estimated effect of urban density in period t on the frequency with which parliaments met in periods t to t+1. Columns 5-10 presets results of the relationship between the frequency of past parliamentary meetings and urban density in the subsequent century. Heteroscedasticity robust standard errors clustered by semi-sovereign unit in parentheses

Initial Urban Conditions and Parliamentary Life Across Time

	1.	2.	3.	4.	5.	6.	7.	8.	9.
	1700-1800	*1600-1700*	*1500-1600*	*1400-1500*	*1300-1400*	*1700-1800*	*1600-1700*	*1500-1600*	*1400-1500*
$\log(UrbanDensity)_{1200}$	0.127^{***}	0.129^{***}	0.129^{***}	0.137^{***}	0.160^{***}	0.155^{***}	0.163^{***}	0.165^{***}	0.178^{***}
	(0.024)	(0.020)	(0.019)	(0.019)	(0.020)	(0.031)	(0.021)	(0.019)	(0.022)
$\Delta_5 \log(UrbanDensity)$	-0.004					0.112^{***}			
	(0.002)					(0.020)			
$\Delta_4 \log(UrbanDensity)$		0.002				-0.543^{***}	-0.123^{***}		
		(0.004)				(0.096)	(0.020)		
$\Delta_3 \log(UrbanDensity)$			0.003			0.974^{***}	0.496^{***}	0.115^{***}	
			(0.009)			(0.191)	(0.073)	(0.020)	
$\Delta_2 \log(UrbanDensity)$				-0.024		-0.769^{***}	-0.689^{***}	-0.339^{***}	-0.119^{***}
				(0.016)		(0.194)	(0.112)	(0.057)	(0.021)
$\Delta_1 \log(UrbanDensity)$					0.090^{***}	0.183	0.352^{***}	0.271^{***}	0.219^{***}
					(0.017)	(0.094)	(0.076)	(0.054)	(0.036)
Intercept	1.158^{***}	1.231^{***}	1.104^{***}	1.136^{***}	1.169^{***}	1.344^{***}	1.434^{***}	1.337^{***}	1.406^{***}
	(0.193)	(0.203)	(0.156)	(0.203)	(0.147)	(0.231)	(0.209)	(0.153)	(0.212)
Region Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	288	287	286	289	290	288	287	286	289
R^2	0.096	0.101	0.100	0.090	0.135	0.183	0.200	0.183	0.171

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4: This table provides estimates of the relationship between initial urban density and parliamentary meeting frequency. Each column regresses the fraction of years in a given century on the logged value of urban density in the year 1200. We account for the overall change between any set of periods, such that δ_t represents the change in urban density over t centuries. Hetereskedacisity robust standard errors clustered by semi-sovereign unit in parentheses.

The Effect of Early Urbanization on the Development of Protoindustry by 1500

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
	Textile Production				**Iron Production**				**Total Protoindustry**			
Urban $Density_{1200}$	0.03^{***}	0.02^{***}	0.05^{***}	0.01^{***}	0.03^{**}	$0.02^{\dagger}$	0.04^{***}	0.01	0.06^{***}	0.04^{**}	0.05^{***}	0.01^{*}
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
R^2	0.28	0.39			0.11	0.35			0.23	0.47		
θ			3.50	0.71			15.41	2.89			8.33	1.43
Effect at $1.3 \times R^2$ of Controlled Regression		0.02				0.01				0.03		
log(Urban $Density_{1200}$)	0.19^{***}	0.14^{***}	0.30^{***}	0.15^{***}	0.26^{***}	0.13^{*}	0.22^{***}	0.11^{**}	0.45^{***}	0.28^{***}	0.24^{***}	0.11^{***}
	(0.04)	(0.03)	(0.03)	(0.02)	(0.05)	(0.06)	(0.03)	(0.04)	(0.08)	(0.08)	(0.03)	(0.03)
R^2	0.25	0.33			0.15	0.35			0.25	0.44		
θ			1.39	0.52			13.41	2.75			6.87	1.34
Effect at $1.3 \times R^2$ of Controlled Regression		0.07				0.07				0.16		
Model:	OLS	OLS	NegBin	NegBin	OLS	OLS	NegBin	NegBin	OLS	OLS	NegBin	NegBin
Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
N	444	444	444	444	444	444	444	444	444	444	444	444

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$

Table 5: This table presents estimates of the effect of early urban development (in the year 1200) on the number of proto-industrial centers in existence on a given 225 km × 225 km unit. Heteroscedasticity robust standard errors in parentheses. Controls are: terrain ruggedness, agricultural suitability, distance to coast, river length, coast length, latitude, and longitude. Following Oster (2013) we provide sensitivity estimates of the effects under the hypothetical condition when unobservables account for 1.3 × the R^2 from the controlled regressions.

The Effect of Proto-industry on Urban Development in 1500

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
	$UrbanDensity_{1500}$						$\log(UrbanDensity_{1500})$					
Textiles	8.25***	5.85**					0.45*	-0.00				
	(2.21)	(2.18)					(0.19)	(0.12)				
Iron Production			2.53*	0.75					0.48***	0.22**		
			(1.02)	(1.11)					(0.07)	(0.07)		
Total Proto-Industry					3.04***	1.74†					0.39***	0.14*
					(0.85)	(0.97)					(0.07)	(0.06)
$UrbanDensity_{1200}$	1.45***	1.40***	1.61***	1.53***	1.50***	1.47***						
	(0.21)	(0.21)	(0.24)	(0.24)	(0.23)	(0.23)						
$\log(UrbanDensity_{1200})$							0.74***	0.48***	0.70***	0.45***	0.65***	0.44***
							(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	(0.06)
R^2	0.75	0.77	0.72	0.75	0.74	0.76	0.40	0.59	0.44	0.60	0.44	0.59
N	444	444	444	444	444	444	444	444	444	444	444	444
Controls	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\dagger}p < 0.10$

Table 6: This table presents results giving the relationship between the existence of proto-industrial centers on future urban development in the years 1500 after conditioning on earlier levels of urban development (in the year 1200). Units are 225 km × 225 km grid-squares. Heteroscedasticity robust standard errors in parentheses. Controls are: terrain ruggedness, agricultural suitability, distance to coast, coast length, latitude, and longitude.

The Relationship Between Coal and Urban Density, and Incomes

	1.	2.	3.	4.	5.	6.	7.	8.
	Outcome: $\log(UrbanDensity_{1800})$							
$\log(UrbanDensity_{1500})$	0.53***	0.45***	0.33***	0.20*				
	(0.04)	(0.05)	(0.06)	(0.08)				
$\log(UrbanDensity_{1700})$					0.63***	0.55***	0.33***	0.20*
					(0.05)	(0.05)	(0.06)	(0.08)
Distance to Coal	−0.23***	−0.11**	−0.13***	0.00	−0.18***	−0.07*	−0.13***	0.00
	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Distance to Coal × $\log(UrbanDensity_{1500})$			0.02***	0.02***				
			(0.00)	(0.00)				
Distance to Coal × $\log(UrbanDensity_{1700})$							0.02***	0.02***
							(0.00)	(0.00)
Constant	5.43***	0.32	4.31***	−1.11	4.29***	−0.34	4.31***	−1.11
	(0.37)	(0.91)	(0.29)	(0.94)	(0.40)	(0.78)	(0.29)	(0.94)
R^2	0.65	0.73	0.65	0.74	0.71	0.78	0.65	0.74
N	444	444	444	444	444	444	444	444
Controls	N	Y	N	Y	N	Y	N	Y

	Outcome: GDP Per Capita 1870							
$\log(UrbanDensity_{1800})$	346.94**	442.20***	350.65***	369.47***	660.00***	416.33***	430.30***	389.44***
	(111.94)	(37.14)	(60.27)	(49.49)	(121.64)	(71.05)	(81.90)	(67.95)
Distance to Coal	-406.55*	-296.95*	-94.10	7.19	-1937.71***	-148.56	-552.42	-113.32
	(186.82)	(94.61)	(140.40)	(102.17)	(360.08)	(380.02)	(456.40)	(381.81)
Distance to Coal × $\log(UrbanDensity_{1800})$					-298.82**	28.52	-89.12	-23.13
					(73.08)	(74.22)	(70.58)	(65.59)
Constant	4006.70***	-16.94	3585.00***	27.81	5564.33***	-257.40	3975.43***	181.29
	(624.81)	(633.76)	(283.92)	(393.12)	(491.31)	(1057.09)	(469.40)	(798.32)
N	104	104	104	104	104	104	1104	104
R^2	0.43	0.78	0.80	0.85	0.53	0.78	0.80	0.85
Controls	N	Y	N	Y	N	Y	N	Y
Country Effects	N	N	Y	Y	N	N	Y	Y

→

	Outcome: GDP Per Capita 1900							
$\log(UrbanDensity_{1800})$	508.85**	630.61***	533.55***	585.60***	995.70***	695.07***	651.97***	634.59***
	(157.83)	(55.97)	(61.53)	(50.65)	(131.36)	(90.07)	(117.30)	(61.71)
Distance to Coal	-489.37	-312.66*	-112.90	-40.28	-2874.70***	-660.18	-779.37	-319.29
	(299.18)	(112.49)	(204.42)	(157.64)	(485.52)	(403.06)	(676.60)	(520.07)
Distance to Coal × $\log(UrbanDensity_{1800})$					-465.54**	-66.63	-129.97	-53.68
					(107.20)	(76.53)	(114.73)	(89.18)
Constant	5553.75***	286.83	5083.33***	124.62	7975.29***	862.33	5664.10***	484.20
	(871.14)	(504.29)	(318.83)	(825.50)	(487.68)	(903.21)	(643.68)	(1028.51)
N	106	106	106	106	106	106	106	106
R^2	0.45	0.80	0.79	0.85	0.58	0.80	0.80	0.85
Controls	N	Y	N	Y	N	Y	N	Y
Country Effects	N	N	Y	Y	N	N	Y	Y

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 7: The top panel of this table provides estimates of the effects of coal access and historical urban population on urban population in 1800. The unit of observation is the 225 km × 225 km grid-square. Heteroskedasticty robust standard errors in parentheses. The lower two panels provide estimates of the effect of urban density in 1800 and coal extraction in the nineteenth century at the regional level on incomes in the nineteenth century. The unit of analysis is the NUTS2 region. Heteroskedasticty robust standard errors clustered by country in parentheses. Controls are: terrain ruggedness, agricultural suitability, distance to coast, coast length, latitude, and longitude

	Effect of By Century					
	1300	**1400**	**1500**	**1600**	**1700**	**1800**
Atlantic Coast	10.71**	11.50	23.05***	41.59***	67.39***	138.00***
	(3.90)	(5.91)	(8.30)	(13.10)	(24.66)	(38.34)
Mediterranean Coast	22.88***	7.19	14.96***	47.70***	44.01***	100.85***
	(7.61)	(5.86)	(7.54)	(13.81)	(13.68)	(23.34)
p value from F test	.16	.61	.48	.74	.42	.67
H_0 Mediterranean = Atlantic						
R^2: .18 N: 444 T: 7						
Distance to Atlantic Coast	-.35***	-.32***	-.62***	-.91***	-1.44***	-3.14***
	(0.08)	(0.10)	(0.10)	(0.21)	(0.38)	(0.63)
Distance to Mediterranean Coast	-.45***	-.20***	-.41***	-1.03***	-.96 ***	-2.07***
	(0.11)	(0.09)	(0.12)	(0.23)	(0.24)	(0.46)
p value from F test	.48	.44	.30	.74	.37	.26
H_0 Mediterranean = Atlantic						
R^2: .20 N: 444 T: 7						
Urban Population$_{it-1}$	1.25***					
	(0.09)					
Atlantic Coast		-15.05**	.74	5.71	6.05	46.71***
m_2: .88		(6.71)	(5.23)	(6.51)	(7.44)	(16.11)
N: 444 T: 6						
Urban Population$_{it-1}$	1.37***					
	(15.47)					
Distance to Atlantic Coast		.76***	.51***	.68***	1.23***	.21***
m_2: .86		(0.15)	(0.09)	(0.11)	(0.15)	(0.03)
N: 444 T: 6						

Table 8: This table presents results first comparing access to the Atlantic to access to the Mediterainan. Then, estimates of the relationship between access to the Atlantic and urban population after controlling for past values of urban population. All models contain unit and time fixed effects. When the lagged dependent variable is included we use a system GMM estimator. Heteroskedasticty robust standard errors in parentheses. The unit of analysis is the 225 km × 225 km grid square.

The Effect of Atlantic Trade and Parliamentary Activity on Urban Density

	Atlantic Potential: Dummy				Atlantic Potential: Atlantic Coast/Area				
Atlantic Potential					17.283***	12.837***	164.371	13.777***	11.660***
					(1.08)	(1.43)	(106.81)	(2.08)	(3.28)
log(Urban Density)$_{t-1}$		0.318**					-0.038		
		(0.14)					(0.17)		
Atlantic Potential × 1500	0.152	0.142	0.077	0.071		3.300***	2.581**	1.745*	3.077**
	(0.11)	(0.11)	(0.10)	(0.12)		(0.70)	(1.25)	(0.98)	(1.47)
Atlantic Potential × 1600	0.183	0.163	0.103	0.085		4.653***	3.680**	2.702*	3.123
	(0.14)	(0.14)	(0.12)	(0.18)		(1.06)	(1.53)	(1.49)	(4.25)
Atlantic Potential × 1700	0.291*	0.092	0.222	-0.102		4.530***	3.767**	2.673	-0.519
	(0.16)	(0.16)	(0.16)	(0.25)		(1.29)	(1.66)	(1.67)	(2.78)
Atlantic Potential × 1800	0.319**	0.186	0.267	0.282		4.973***	4.771***	3.439**	7.828***
	(0.15)	(0.16)	(0.14)	(0.19)		(1.29)	(1.66)	(1.67)	(2.78)
P-Index$_{t-1}$			-0.159	-0.252				-0.158	-0.261*
			(0.12)	(0.16)				(0.11)	(0.15)
P-Index$_{t-1}$ × 1500				0.147					0.153
				(0.17)					(0.16)
P-Index$_{t-1}$ × 1600				0.222					0.216
				(0.16)					(0.15)
P-Index$_{t-1}$ × 1700				-0.085					-0.040
				(0.17)					(0.16)
P-Index$_{t-1}$ × 1800				0.332					0.350**
				(0.18)					(0.17)
Atlantic Potential × P-Index$_{t-1}$			0.285	0.003				2.228	3.098
			(0.33)	(0.38)				(2.23)	(2.26)
P-Index$_{t-1}$ × Atlantic Potential × 1500				0.024					-2.306
				(0.25)					(2.05)
P-Index$_{t-1}$ × Atlantic Potential × 1600				0.022					-1.232
				(0.31)					(4.53)
P-Index$_{t-1}$ × Atlantic Potential × 1700				0.650					3.294
				(0.38)					(8.16)
P-Index$_{t-1}$ × Atlantic Potential × 1800				-0.101					-6.934
				(0.35)					(3.94)
N	2106	1793	1793	1793	2106	2106	1793	1793	1793
R^2	0.182		0.144	0.150	0.181	0.186		0.147	0.153
m2	·	0.04	·	·	·	·	0.16	·	·

Table 9: This table estimates the relationship between access to the Atlantic and urban development as well as the interactive effect between Atlantic access and the existence of parliamentary constraints on the same outcome. The unit is the polity. Heteroskedasticty robust standard errors in parentheses. When the lagged dependent variable is included we use the system GMM estimator.