City seeds: geography and the origins of the European city system

Maarten Bosker and Eltjo Buringh[*+]

< THIS VERSION, MARCH 2012 >

Abstract

Cities are the focal points of the world economy. This paper aims to better understand their origins. Using a new dataset covering over 250,000 randomly selected potential city locations, and all actual actual cities during the period 800 – 1800, we disentangle the different roles of geography in shaping today's European city system. We find that a location's physical, *first nature*, geography characteristics are the dominant determinant of city location. A location's relative position to already-existing cities (its *second nature geography*) only becomes important during the later centuries. Interestingly, we show that it does so in a way corresponding closely to predictions from new economic geography theory.

Keywords: city origins, economic geography, Europe

JEL codes: N93, O18, R10

---

[*] Bosker is affiliated to the Erasmus University of Rotterdam, Utrecht University, the Tinbergen Institute and CEPR. Buringh is affiliated to Utrecht University. Please send any correspondence to: Maarten Bosker, P.O. Box 1738, 3000DR, Rotterdam, The Netherlands; or email to bosker@ese.eur.nl.

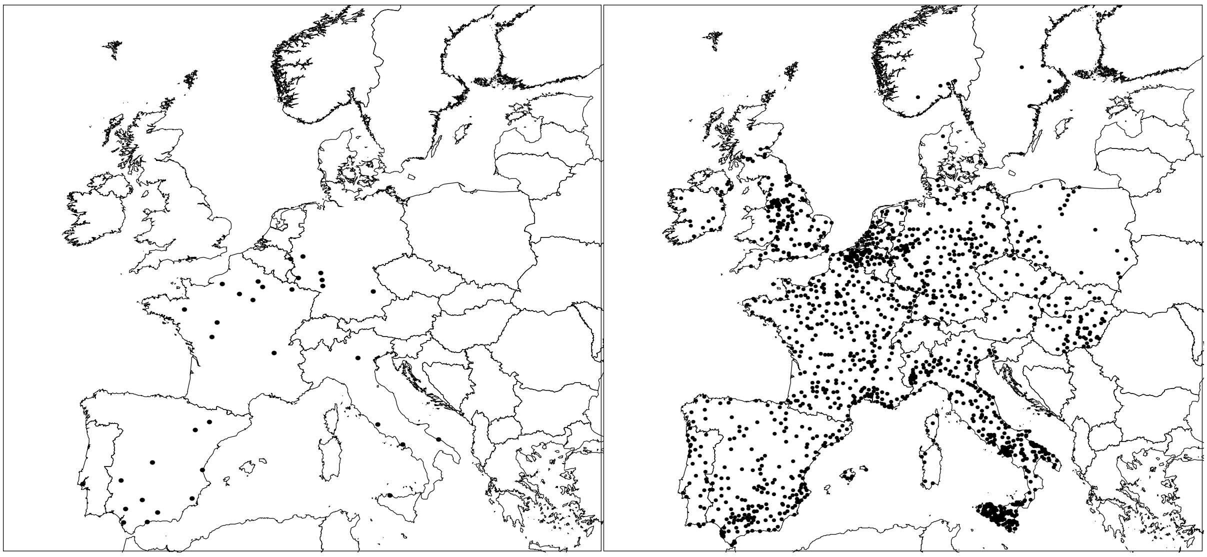
[+] We thank Rob Alessie, Bob Allen, Bas van Bavel, Francesco Billari, Christiaan van Bochove, Marius Brülhart, Gilles Duranton, Jessica Dijkman, Rui Esteves, Giovanni Facchini, Oliver Falck, Ewout Frankema, Harry Garretsen, Knick Harley, Laura Hering, Tommy Murphy, Nathan Nunn, Rick van der Ploeg, Maarten Prak, Joppe de Ree, Roger Smeets, Tony Venables, Wouter Vermeulen, Joachim Voth, Bastian Westbrock, Nikolaus Wolf, Jan Luiten van Zanden, and seminar participants in Oxford, Cambridge, Milan, Barcelona, Berlin, Rotterdam, Tilburg, Groningen, Utrecht, the Netherlands Bureau for Economic Policy Analysis, and at the 2009 World Economic History Conference, the 2009 North American Regional Science Conference, the 2010 Econometric Society World Conference, and the 2012 AEA meetings for very useful comments and/or suggestions. Special thanks go to Giovanni Zambotti at Harvard's Center for Geographic Analysis for helping us a lot in obtaining the necessary information for all our randomly drawn coordinate pairs.

"In a more advanced era, when better methods would permit man to conquer Nature […], it would doubtless have been possible to build towns anywhere the spirit of enterprise and the quest of gain might suggest a site. But it was quite another matter in a period when society had not yet acquired enough vigor to rise above the physical conditions in the midst of which it developed. […] the towns of the Middle Ages were a phenomenon determined as much by physical surroundings as the course of rivers is determined by the conformation of the mountains and the direction of the valleys." (Henri Pirenne, 1925 p.138/39).

1 Introduction

Today the European landscape is dotted with cities. Historically this was quite different. In the early medieval period Europe only knew a handful of cities. Over the next millennium this changed dramatically. Cities started to appear on an unprecedented scale, and virtually everywhere on the continent. Figure 1 shows that in 800, we only find a few scattered cities in mainly Spain, Italy, France and Germany. One thousand years later, cities have appeared all over the continent[1].

Figure 1. The European city system in 800 and 1800



a) 800 b) 1800

Notes: cities are denoted by black dots [see section 3.2 for more detail on the city definition used]. In 800 there are 34 cities, in 1800 this number has increased to over 1,450.

The rise of the city in the European landscape is important for several reasons. Throughout history, cities have been the important loci for technological innovation, institutional progress, (international) trade, political power, and culture (Pirenne, 1925; Glaeser, 2011). Also, cities are generally more productive places. The concentration of many people e.g. allows for a greater degree of specialization, carries positive externalities such as knowledge spillovers,

[1] Figure A1 and Table A1 in Appendix A further illustrate the rise of the city in the European landscape. Over our sample period, Europe's urbanization rate increased from only 3% in 800 to 15% in 1800. Urban population increased 30-fold from 0.7 to 21 million, whereas total population increased 6-fold from 23 to 137 million. A full, century-by-century, visualization of the formation of the European city system is available upon request.

and facilitates a more efficient provision of public goods (Lampard, 1955; Marshall, 1890). It may therefore not be surprising that cities are argued to have played a very important role in Europe's economic 'take-off' during the late Medieval and Early Modern period. Economic development and urbanisation often go hand in hand (Acemoglu et al., 2005; De Vries, 1984; Galor, 2005). Today, an estimated 75% of world production takes place in cities (World Bank, 2009).

The importance of cities in the development process makes understanding their origins of great interest. Cities do not develop everywhere. The question '*why do cities form in some locations, and not, or only much later, in others?*' lies at the heart of this paper. In particular, we empirically uncover the role(s) of geography, widely viewed as the most important determinant of a location's urban chances, in 'sowing the seeds' of the European city system.

The narrative urban (economic) history, the economic geography, as well as the more recent urban economic and new economic geography literatures, stress two important, but very different, roles for geography in the origins of an urban system[2].

First, geography determines a location's physical, or *1st nature geography*, characteristics. A location's agricultural potential, its transportation possibilities and its defensive advantages, have all been noted as important city seeds. The second role for geography, although already stressed by e.g. Christaller (1935) and Lösch (1938; 1940)[3], has received renewed attention in the economics literature following Krugman (1991; 1993b). While not denying an important role of 1st nature geography, this line of literature stresses the importance of a location's position relative to the rest of the (already existing) urban system, its *2nd nature geography*, for its urban prospects. As already acknowledged by Pirenne (1925, p.145), some locations may be well suited for urban development based on their own characteristics, but *"situated too far from the great highways of communication,* […] *they remained sterile, like seed fallen upon stony ground."*

The debate on the relevance of the two different roles of geography in determining cities' origins has up to now largely taken place without using rigorous empirical evidence[4]. Instead, it relies on historical narratives, largely descriptive accounts of European

---

[2] Influential contributions in these literatures are among others Pirenne (1925), De Vries (1984) or Bairoch (1988) [narrative urban history], Christaller (1935), Lösch (1940), Ullman (1941) or Lampard (1955) [economic geography], and Krugman (1993a), Fujita and Mori (1997) or Behrens (2007) [urban economics / geographical economics].

[3] An even earlier contribution focussing on 2nd nature geography is von Thünen (1826). He however considered the evolution of only one isolated city in relation to its hinterland, instead of the evolution of a system of cities.

[4] Several papers do look at the relative importance of 1st and 2nd nature geography for the evolution of a city system after its initial establishment, e.g. looking for evidence of path-dependence in urban development (see Bleakley and Lin, 2010; Davis and Weinstein, 2002; Bosker et al., 2008, Redding and Sturm, 2008).

urbanization, or detailed case studies looking at one particular city or region only. This paper fills this gap. Using a large, newly collected, data set on both actual, as well as, potential city locations in Europe, we empirically establish the (relative) importance of 1st and 2nd nature geography in determining city location over the period 800 – 1800[5].

The European case in our view provides an ideal testing ground. Historical data availability on the size and characteristics of individual cities in Europe is better than that of other continents in terms of both spatial and temporal coverage. This is largely due to the work of Bairoch et al. (1988) and De Vries (1984). They have constructed comprehensive data sets providing population estimates for many cities in Europe starting as early as the year 800[6]. We build on this data in two ways. First we extend its coverage by also considering over 250,000 randomly selected potential city locations: locations that in principle could have become a city but never did. Second, we complement the existing datasets with detailed information on each location's 1st and 2nd nature geography characteristics.

All this data is available for the period, 800 – 1800, during which one can forcefully argue that the seeds for the eventual European city system were sown. Following the eclipse of the Roman empire, cities in Europe withered (Pirenne, 1925; Greif, 1992). In 800, only about 30 cities can be found. But, over the next millennium Europe witnessed an unprecedented revival of urban activity and the establishment of cities on a scale not seen before (Davis, 1955 p.432). In 1800 their number had increased to over 1,450 (see Figure 1).

Using our data set, we quantify the role of 1st and 2nd nature geography in conditioning the spread of cities across the European continent. We explicitly base our empirical analysis on the main theoretical insights regarding the role of 1st and 2nd nature geography in sowing the seeds of cities. These insights come from the economic and urban history literature on the

---

[5] We focus on the 800 – 1800 period for the following reasons. We start in 800 as it is the first year for which comprehensive data on city population exists for Europe, i.e. Bairoch et al. (1988). We stop in 1800 because not doing so would add the Industrial Revolution to our sample (see e.g. Ashton, 1948). The substantial changes during that period in terms of transportation (railroads, steamships), production (both industrial and agricultural), and the importance of different natural resources (coal), turned many locations that previously had little chance of becoming a city into viable city sites (e.g. many locations in the coal-rich areas of Germany, Sweden, north-east England, and the Limburg provinces of both Belgium and The Netherlands). Including the Industrial Revolution in our view requires a detailed account of its effects, which lies beyond the scope of this paper.

[6] This data has up to now been used either to provide descriptive accounts of urban expansion (Bairoch et al., 1988; De Vries, 1984), or to uncover the major drivers of a city's size *once a city is established* (Acemoglu, Johnson and Robinson, 2005; De Long and Shleifer, 1993; Bosker et al., 2008; Kim, 2000; or Bosker et al, 2012). By looking at city size *conditional* on a city's existence, although very interesting in itself, these papers effectively take cities' location as given and refrain from shedding empirical light on the question why these cities were formed at their particular locations in the first place. They do not answer the question why other, often a priori equally viable, locations never became a city or only did so at a much later stage.

one hand, and from the more recent new economic geography literature on the other hand. They serve as the theoretical underpinning of our empirical analysis, guiding the selection of 1^{st} and 2^{nd} nature geography variables. In case of 2^{nd} nature geography, we develop a novel, flexible, way to quantify the effect that an already established city exerts on the urban chances of its surroundings, that allows us to test the main predictions from new economic geography theory.

We find that both 1^{st} and 2^{nd} nature geography played an important role in the origins of the European city system. Their (relative) importance however shows substantial changes over time. First nature geography dominates in the early stages of the formation of the European city system. But, as trade costs fall, economies of scale increase, and the overall European population increases, 2^{nd} nature geography also starts to be an important determinant of city location from around the sixteenth/seventeenth century onwards. Interestingly, the effect that an already existing city exerts on the urban chances of its surroundings corresponds closely to the predictions made by economic geography theory.

2 *Insights from theory*

2.1 *Economic and urban history*

Traditionally, the debate on cities' origins was conducted within the realm of the, largely narrative, economic and urban history literature (see e.g. Pirenne, 1925; Weber, 1922; Bairoch, 1988; De Vries, 1984). This literature stresses a priori differences between locations as the main reason why some locations are more likely than others to develop into a city. Such spatial inhomogeneities, what we call *1^{st} nature geography*, arise most notably from differences between locations in their resource abundance or transportation possibilities.

Attractive city locations were those close to natural resources (fertile plains, mineral deposits, thermal springs, etc.) and locations with good access to the main trade routes. Given that the city relies on exchange with its hinterland (most notably for the feeding of its population), location on a navigable river, an overland transport route, or at sea offers substantial advantages in terms of transportation possibilities (a recent paper by Bleakley and Lin (2010) aptly illustrates this for portage sites in the US).

A location's defensive possibilities and its climatic conditions are also often mentioned (see Pirenne, 1925 pp.72-76; or Hohenberg and Lees, 1995 p.30). A favourable climate for agricultural production, or a strategic location at a river crossing, the foot of a mountain pass or a hill overlooking the countryside, makes locations more attractive city sites.

A location's transportation possibilities however, are mostly viewed to overshadow these other motives. As put by Bairoch (1988, p.143) *"The critical role played by transport in the location of cities does not rule out exceptions, but statistically speaking these are in the minority."*

2.2 Economic geography

Spatial inhomogeneities also feature prominently in the economic (geography) literature on city creation (Duranton, 1999; Anas, Arnott, Small, 1998; Fujita and Mori, 1996; Krugman, 1993a; Behrens, 2007; Konishi, 2000). Although this literature does not deny that soil quality or climate are important determinants of city location, the 1st nature geography characteristic that receives most attention is preferential location on the main trade routes. Trade costs[7], together with scale economies, are viewed as the crucial elements in the process of city formation. Trade costs are vital to a city given that it relies entirely on exchange with its hinterland to meet its own demand for agricultural produce. When the cost of transporting these agricultural goods (or the goods the city produces in exchange for them) are very high, this results in the so-called tyranny of distance and cities only form in locations offering good 1st nature geography conditions so that sufficient food can be imported from nearby (see e.g. Duranton, 1999, p.2173).

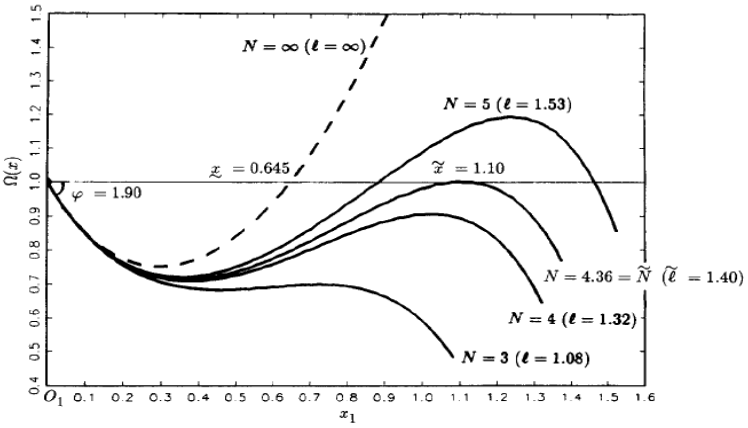
However, when trade costs diminish due to e.g. improvements in transportation technology or lower trade barriers (decreased tariffs, safer roads, improved freight insurance, etc), the tyranny of distance is alleviated and the (relative) importance of 1st nature geography diminishes. Since agricultural products can now be shipped over longer distances at lower costs, it becomes possible to establish cities at locations that, given their lack of 1st nature geography advantages, were previously unviable to host a city.

Still, even with a diminishing importance of 1st nature geography, not all locations become equally viable future city sites. This crucially depends on their *2nd nature geography characteristics*, i.e. their position relative to the urban system already in place. Earlier contributions by e.g. Christaller (1935), Lösch (1938; 1940), or Ullman (1941) already stressed that *"no city is ever an island existing in and of itself"* (Lampard, 1955). Yet, it was only recently that several papers explicitly focus on the *where-do-cities-form* question in a theoretical framework of endogenous city location that formalizes the idea that already

[7] All costs associated with moving goods from one location to another, including not only transportation costs but also tolls, tariffs and less tangible costs associated with differences in e.g. language, institutions or culture.

existing cities influence the urban chances of their surroundings[8]. These papers (e.g. Fujita and Mori, 1996 and 1997; Fujita, et al., 1999 and Behrens, 2007) not only establish theoretically, using fully specified general equilibrium models, under what conditions a city (or subsequent cities) will form, they also make clear predictions about which locations are more likely to become a city than others.

Figure 2. The 2[nd] nature geography effect of an existing city



Notes: This figure is taken from Fujita and Mori (1996, p.108). The x-axis (x_1) indicates the distance from the already established city, which is located at the origin. The y-axis depicts the value of the so-called market potential function[9]: locations where the value of the market potential curve exceeds 1 (the solid straight line in the figures) are locations where a new city is viable. *N* denotes overall population.

Figure 2 (taken from Fujita and Mori, 1996) illustrates how an already existing city affects the urban chances of other locations[10]. It depicts so-called market potential curves that can be interpreted as indicating the likelihood of a location, located at a distance *x* from an already existing city at the origin, to become a city too. Whenever a location's market potential exceeds 1, it is in principle a viable new city location. Whether or not this is the case depends first and foremost on a location's distance to the already existing urban center:

---

[8] Earlier urban economic theories relying on scale economies and transport costs remain silent on the *where do cities from*-question. A city's relative location either bears no consequences for its further development (e.g. Henderson, 1974 or Black and Henderson, 1999), or, often despite assuming no differences in 1[st] nature geography characteristics between locations (i.e. a continuous homogenous plain), the (relative) position of a discrete number of possible city locations is a priori assumed (see e.g. von Thünen, 1826; Christaller, 1935; Lösch, 1940). Moreover, a drawback of these latter models is that the final structure of the urban system does not follow endogenously from a set of assumptions concerning the behavior of firms and consumers (see Ottaviano and Thisse, 2005 for an extensive and very useful overview of the history of location analysis in urban economic and economic geography theory).

[9] See Appendix B and D in Fujita and Mori (1997) for the analytical details of these market potential functions. Also, see section 4.2 in their paper for a more thorough discussion of the market potential curve.

[10] Figure 2 depicts the case when no potential city location has an a priori advantage in terms of their 1[st] nature geography. Fujita and Mori (1996) and Behrens (2007) further generalize this and show that locations with a 1[st] nature geography advantage in terms of their transportation possibilities (hubs) produce sharp positive kinks in the market potential function, making them more likely future city candidates (see Figure A2 in the Appendix).

Locations too close to an already existing city face too strong competition with that city, both for agricultural produce and for inhabitants[11]. On the other hand, locations too far from an already existing city can not take full advantage of the trading possibilities with the already existing city. This leaves locations at medium range from existing cities as preferred new city locations: they offer relatively cheap trading possibilities with the already existing cities compared to locations further off, as well as only limited competition with these same existing cities compared to locations at too close range.

The strength, and spatial reach, of this 2nd nature geography effect depends on the important model parameters. Most importantly, when trade costs are too high, and/or the productivity advantages of co-locating in a city are too low (relative to the disadvantages of co-locating in a city), 2nd nature geography plays no role in determining the location of new cities. Also, when transportation costs are extremely low, overall population too small to sustain multiple cities, or productivity advantages of co-location very high, the models predict that only one city will emerge. Only at intermediate values of trade costs and scale economies, and given a sufficiently large overall population, does the above-described non-linear 2nd nature geography effect come into play[12].

By introducing an important role for the current state of the urban system in determining its future development, 2nd nature geography offers a substantially different and more dynamic answer to the *where-do-cities-form* question than the more static[13] explanation offered by 1st nature geography hinging on a priori spatial differences between locations. This makes establishing their (relative) importance the more interesting. In the remainder of this paper we do just that. We construct a new dataset on the basis of which we can empirically identify the (relative) importance of both 1st and 2nd nature geography in 'sowing the seeds' of the European city system.

[11] Not uncommon in medieval times, the existing city may even use force to prevent a competitor city forming in its immediate backyard. Or, less violently, put severe restrictions on any economic activity in its immediate vicinity The German 'Bannmeile' is a good example (see Ennen, 1972).

[12] Our exposition is admittedly a bit too stylized and does not do entire justice to the richness of the models, where the relevance of the discussed 2nd nature geography effect depends delicately on the interaction between trade costs (and the relative size of those for agricultural and non-agricultural goods respectively), (dis)economies of scale, the share of agricultural consumption in overall consumption, and overall population size (for particular configurations of these model parameters, it can even be the case that only one city, or even no city, emerges). However, the effect of an existing city is always negative at close and at large distances from an already existing city. It is the positive effect at medium range (and the extent of this range) that depends delicately on the model parameters. We take this non-linear effect exerted by an already existing city as the main insight from theory that we take to the data in our empirical sections.

[13] Not completely static however. The importance of particular spatial inhomogeneities, or the inhomogeneities themselves, may change over time. A good example is cities formed for defensive purposes only. Located at impregnable locations, these offer limited possibilities for expansion in more peaceful times. Another example is location near natural resources. These locations lose their attractiveness once the resource is depleted or becomes obsolete. In section 5.2 we explicitly allow the importance of 1st nature geography to change over the centuries.

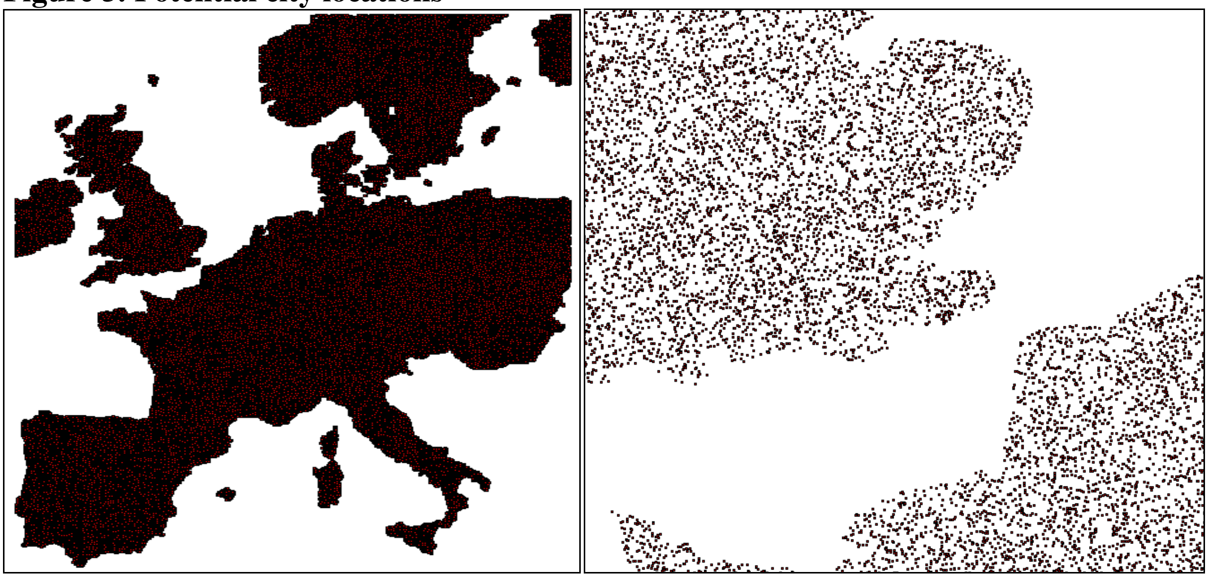
3 Data and descriptives

We focus in turn on our choice of potential city locations, the city-definition that we employ, and on the 1st and 2nd nature geography variables we collected[14]. We discuss in particular detail how we incorporate 2nd nature geography into the analysis. We propose a novel way to construct our 2nd nature geography variables that corresponds closely to the main theoretical insights presented in section 2.2.

3.1 Potential city locations

In order to empirically study the rise of cities in Europe[15], we consider 259,776 randomly drawn coordinate pairs[16] as our baseline sample. This effectively implies that we, in principle, consider each and every coordinate pair as a potential location for a city. Figure 3 below shows our sample.

Figure 3. Potential city locations



a. full sample **b. zoomed in: the English Channel**

Notes: each dot represents a potential city location.

Given the large number of randomly drawn locations, plotting all of them (Figure 3a) basically reproduces the entire map of Europe. It clearly shows the density as well as the boundaries of our sample. Figure 3b, zooming in on the English Channel, shows that Figure

[14] Tables A1 – A3 in Appendix A provide descriptive statistics on all variables discussed in this section.

[15] We define Europe as roughly everything west of the line Trieste – St. Petersburg. This line is well known from the literature on the European Marriage Pattern (see Hajnal, 1965): it coincides with the border of the Catholic Church during the Middle Ages. See also De Vries (1984), Findlay and O'Rourke (2007) or Bosker et al. (2012). Europe thus defined comprises current-day Norway, Sweden, Finland, Poland, Germany, Czech Republic, Slovakia, Austria, Hungary, Belgium, Luxembourg, the Netherlands, France, Great Britain, Ireland, Switzerland, Italy, Spain and Portugal.

[16] We actually randomly drew 400,000 coordinate pairs. Of them, 259,776 were located on land.

3a is indeed made up of the above-mentioned 259,776 randomly drawn coordinate pairs. To put some numbers on the spatial detail of our sample; the median [average] distance between a randomly drawn coordinate pairs and its closest neighbour is only 1.67km [1.79km][17].

3.2 Actual cities – definition and location

Next, we need information on where, and when, actual cities appear. First, this requires us to define what we mean by a city. In all our baseline analyses, we define a city as an agglomeration of at least 5,000 inhabitants. In doing so, we basically adopt the definition proposed by both De Vries (1984) and Bairoch (1988)[18]. An important reason for using this city definition is that *"a population of 5,000 is* […] *a criterion that may be questionable in certain respects but which nevertheless remains for all that the most adequate and especially the most operational."* (Bairoch, 1988 p.494).

Of course, using an absolute size criterion of 5,000 inhabitants may in certain cases be too low and thus wrongly ascribe an urban role to a location (see e.g. Malanima (1998) on Sicilian agrotowns). On the other hand, the opposite, i.e. the cutoff being too high, has also been argued, especially for the early medieval period (see e.g. Dyer, 1995). Both Bairoch (1988, pp 137/138) and De Vries (1984, pp. 53/54 or 21/22) view the use of a population cutoff of 5,000 inhabitants as providing a 'best of both worlds'.

The alternative would be to define cities on the basis of more criteria than population size only (e.g. having city rights or certain economic, religious or institutional features). In the words of Bairoch (1988) this would however be "*much less operational*" (p.494). Not only would it constitute a very time consuming exercise; to agree on what features a location needs to have in order to qualify as a city would be subject to much debate. Are city rights sufficient, or should it also have a fair, a market and/or a mint to qualify as a city? And, if so, should these fairs or markets be of a certain size, or of regional importance, before a location qualifies as a city? Even if we were to agree on which features to include in this city definition (and data on all these features would be readily available), the substantial institutional, political and religious differences between the different societies in Europe further complicates the task of consistently applying this definition (e.g., city rights in one part of Europe are not necessarily directly comparable to those in other parts).

[17] Table A2 in Appendix A provides additional detail on the geographical distribution of our (potential) city locations, indicating how many of them are found in each of the (current-day) European countries in our sample.
[18] Also in archaeology, it is common practice to define cities as population centres with more than 5,000 inhabitants. See for example Fagan (1997, p.27) or Bahn (1996, p.57).

An absolute population cutoff to define a city avoids these issues of comparability, it makes the city definition less subjective, more transparent, and much more up to scrutiny as one can easily compare the results using different, even possibly time-varying, population cutoffs (in Table A6 in Appendix A we do just that).

With our city definition in hand, we use Bairoch et al. (1988) as our main source of the location of actual cities. They provide centennial population estimates for all places in Europe that in some century over the 800 – 1800 period have more than 5,000 inhabitants[19]. In total, this gives us 1,588 actual city locations. Next, we match these cities to our random sample of coordinate pairs maintaining a margin of error of 2.5km[20]. That is, we replace a randomly selected coordinate pair with an actual city if the random draw lies within a range of at most 2.5km to that city. This results in 1,150 matches (72.4% of our 1,588 actual cities). In robustness checks we also use a stricter 1.5km matching criterion. This results in using only 624 (or 39.3%) of all actual cities, but leaves our main results basically unaffected[21].

In sum, our baseline sample consists of 259,776 potential city locations, of which 1,150 (or 0.4%) actually develop into a city at some point during our sample period.).

3.3 Explanatory variables determining city location

3.3.1 1st nature geography

To capture a location's opportunities for water- and land-based transportation, we construct dummy variables that indicate whether or not it has direct access to the sea, to a navigable waterway, or to the former Roman road network. For all our potential city locations, the information on their location with respect to sea, lakes and major rivers is obtained using GIS maps of Harvard's Center for Geographic Analysis. The information on the presence of a Roman road comes from Talbert (2000). This information is digitized by the DARMC project at Harvard[22]. Besides documenting whether or not a location has direct access to a Roman road, we also classify locations where two (or more) Roman roads crossed as hub locations. We use location on a Roman road instead of on an actual road for two reasons. First, the Roman road network is argued to have played an important role in trade long after the

---

[19] There are no population estimates for 1100. For this century we linearly interpolated the reported 1000 and 1200 population estimates. All our results are fully robust to excluding these interpolated 1100 numbers from the analysis. Results available upon request.

[20] The median [average] distance of a randomly drawn coordinate pair to its nearest already existing city is 1.79km [2.11km].

[21] Making the distance criterion even stricter, e.g. 1 or even 0.5km, the number of actual cities matched decreases to 308 (19.4%) or 82 (5.2%) respectively. The loss of variation that results from using an ever smaller number of actual cities, makes it increasingly difficult to find significant effects.

[22] See http://darmc.harvard.edu.

withering of the empire itself[23]. Roman roads constructed using similar methods and adhering to uniform quality standards can be found throughout the formerly Roman parts of Europe. Second, using location on a Roman road or a hub of Roman road avoids some of the reverse causality issues that could arise when using actual roads (i.e. roads being built to future city locations, instead of a road increasing the urban chances of locations along this road).

Besides these transportation related 1st nature geography variables, we collected information on each potential city location's elevation [in meters] and on its ruggedness [calculated as the standard deviation of the elevation of the terrain within 10km of a potential city location]. This data is taken from the (1 x 1km) GLOBE database made available by the US National Geophysical Data Center[24]. Both serve as a proxy of a location's accessibility[25].

Finally, we collected information on each location's agricultural conditions from Ramankutty et al. (2002). That study combines information on climatic conditions (surface air temperature, precipitation and potential sunshine hours) and soil quality (total organic content [carbon density], availability of nutrients [pH] and water holding capacity) into one index that gives the probability that a certain location will be cultivated. This data is available in gridded form at a resolution of 0.5 degrees latitude-longitude (in case of our sample this corresponds to a grid of on average 56 km by 39 km). We match each potential city location to this data on the basis of its coordinates. Locations falling within the same grid cell have the same cultivation probability.

The Ramankutty et al. (2002) data provides a time-invariant indication of a location's agricultural possibilities. It it not unlikely that a location's agricultural conditions (and most notably its climatic conditions) varied over the centuries. To our knowledge however, historical climate data is not available at a sufficiently disaggregated scale to be useful for our purposes. To overcome this difficulty we capture the possibly time-varying agricultural conditions at a somewhat more aggregated spatial scale by including *country-century* fixed effects in all our baseline model specifications. Besides controlling for time-varying agricultural conditions that possibly differ between European countries, these *country-century* fixed effects also capture any country-specific[26] institutional, political, demographic or

---

[23] Glick (1979, p.23) gives several examples of policies by medieval Spanish states and cities to maintain the system of Roman roads. See also Bairoch (1988, p.110) or Lopez (1956). The latter offers a much more critical view on the importance of Roman roads in the centuries after the demise of the Roman Empire.
[24] See http://www.ngdc.noaa.gov/mgg/topo/globe.html.
[25] Although they can also be argued to be related to a location's agricultural possibilities.
[26] Countries are defined using current country boundaries. This arguably does not do full justice to the actual political, or institutional situation during our sample period. However, we think it serves as a good proxy (see also Acemoglu et al., 2005 (footnote 8); or De Long and Shleifer, 1993).

economic developments that may have left their mark on locations' urban chances[27]. Notably, they control for the general increase in overall population that European countries experienced (each at a different rate) during our sample period (McEvedy and Jones, 1979).

In robustness checks we also use two other fixed effects specifications. First, we provide results that allow for time-varying, geographically clustered, unobserved effects by including *block-century* fixed effects, with locations grouped in geographically clustered blocks on the basis of their coordinates. And, second, we show results when also controlling for unobserved time-*invariant* location-specific fixed effects.

3.3.2 2nd nature geography

We propose a novel way to uncover the effect(s) of 2nd nature geography. The most commonly used measure of a location's 2nd nature geography is its market or urban potential (see e.g. Stewart, 1947; De Vries, 1984; Black and Henderson, 2003; or Bosker et al., 2012). This measure is the distance weighted sum of the population of all other already existing cities. In each century *t*, city *i*'s urban potential (*UP*) is calculated as follows:

$$UP_{it} = \sum_{j=1, j \neq i}^{N} \frac{pop_{jt}}{D_{ijt}} \qquad (1)$$

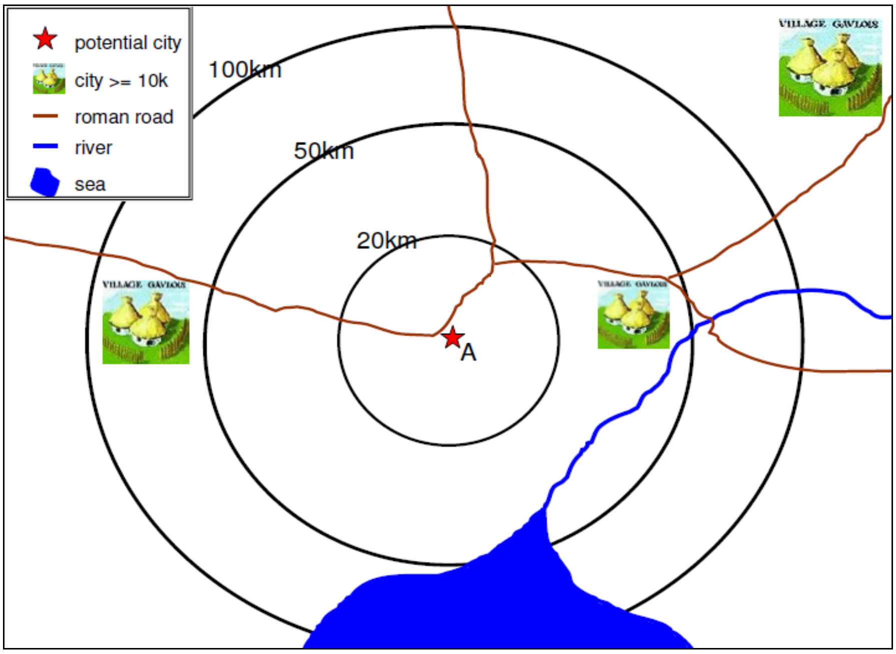
We argue that such *UP*-type measures do not do justice to theory when looking at the establishment of new cities. The way *UP* is constructed implies that the impact of 2nd nature geography diminishes *linearly* with the size of, and distance to, other already existing cities. Moreover, it implicitly assumes that the impact of an already existing city on a location's own urban chances is either always negative or always positive (depending on the sign of the estimated coefficient on *UP*).

This is clearly a too strong restriction when looking at Figure 2. An existing urban centre exerts an urban shadow at close range, prohibiting the formation of new cities in its immediate neighborhood. At the same time, potential locations that are too far removed from an already existing city also have little chance of becoming a city. It are the locations at medium distance from an already existing city that have the best urban chances. Theory thus predicts that an existing city exerts a *non-linear* effect on its surroundings. *UP*-type measures fail to adequately capture this.

---

[27] Allowing for unobserved time-varying but geographically clustered heterogeneity is particularly important to get accurate estimates of the effect of 2nd nature geography: A location that is located in a country that is, for unobserved reasons, a good seedbed for city development, will have a high probability of becoming a city. But, so do other locations in that country. As a result, this location is also more likely to be surrounded by some already existing cities. When not adequately controlling for geographically clustered unobserved heterogeneity, one can thus easily, and wrongly so, ascribe an important role to 2nd nature geography.

Figure 4. Constructing dummy variables to capture 2nd nature geography



To do more justice to theory, we adopt the following dummy variable approach that does not a priori restrict the effect of existing cities to be positive or negative at all distances[28]. We first draw three concentric circles around each potential city location at ever further distance[29], i.e. 20km, 50km and 100km respectively. These distances roughly correspond to a 1 day, 2.5 days, and 5 days round-trip during most of our sample period[30]. Next, we construct three dummy variables that indicate whether or not we find at least one already existing city of at least 10,000 inhabitants[31] within each respective distance band.

Figure 4 illustrates in more detail how we construct these dummy variables in case of a hypothetical potential city location A. For this location, the dummy variables indicating the presence of an established urban centre are only 1 in case of the 20-50km and the 50-100km distance band (there are no already existing urban centres within 20km of A). In further extensions (see section 6.1) we also consider more elaborately specified dummy variables indicating e.g. the presence of more than one already existing city, or an already existing city

---

[28] It does constrain the effect to be the same within each distance band. But, one can experiment with different distance bands (see Table A7). We also show results of abandoning our distance band-approach altogether, and fitting a sixth order polynomial in distance to the nearest city to the data instead.

[29] We calculate great circle distances between all locations in our data set on the basis of their coordinates.

[30] Roughly because this depends on mode of transportation, travel on horseback or donkey was generally faster than travel by foot, cart or water.

[31] We construct the dummy variables on the basis of existing cities larger than 10,000 inhabitants instead of 5,000 inhabitants to limit possible reverse causality (simultaneity) issues from including a spatially lagged variable (similar to Hanson, 2005). We further limit simultaneity issues by considering these dummy variables lagged one century (see section 4). In robustness checks in section 6.1 we also show results when constructing these dummy variables on the basis of a larger and/or smaller population threshold for existing cities.

with more than 5,000 or 25,000 inhabitants, within each of our three specified distance bands.

4 *Empirical framework*

With our data in hand, we empirically quantify the effect of a location's 1st and 2nd nature geography characteristics on its chances of developing into a city, using the following simple empirical model:

$$P\left(c_{ict}=1 \mid c_{ict-1}=0, X_{ict-1}, X_i, \alpha_{ict}\right)=F\left(X_{it-1}\beta_1+X_i\beta_2+\beta_3+\alpha_{ict}\right) \quad (2)$$

, where c_{ict} is a dummy variable indicating whether or not location i in country c is a city at period t, X_i are our time-invariant 1st nature geography variables measured at the location level, and X_{it-1} are the three 'already existing city'- dummy variables capturing a location's 2nd nature geography. We include all time-varying variables lagged one century to limit potential endogeneity issues resulting from reverse causality. Moreover, in combination with our assumption of no serial correlation in the error term, it also ensures that our inclusion of spatially lagged variables (i.e. all 2nd nature geography variables) does not result in inconsistent estimates (see also footnote 31). The α_{ict} capture any unobserved effects at the city, country or century level. In our main specification we specify these unobserved effects to be *country-century-specific* fixed effects: $\alpha_{ict}=\alpha_{ct}$, but we also show results using various different specifications (e.g. location-specific fixed effects). In most of the paper F denotes the CDF of the standard normal distribution, Φ (i.e. we estimate a probit model), but in robustness checks (see Table A5 in Appendix A) we also use the logistic function (a logit model) or simply the identity function (a linear probability model[32]).

Our main empirical specification is essentially a (restricted) dynamic probit model. Therefore, we need to assume that we have no serial correlation in the error term in order to obtain consistent estimates of our parameters of interest using standard probit techniques. Note that this assumption precludes us from basing our inference on clustered standard errors. Although these can be calculated, their use would be internally inconsistent with the necessary assumption of no serial correlation in the error term that underlies our estimations.

The β's are our parameters of interest. They reveal the sign, size (after calculating Average Partial Effects, see footnote 33), and, together with their estimated standard error, significance of the included 1st and 2nd nature geography variables.

[32] Given that the identity function is not a distribution function, an error term would be added to (2) in this case.

5 *Results*

Table 1 shows our baseline results. Unless noted explicitly, all tables in our paper do not report the estimated coefficients of (2) but Average Partial Effects (APEs) instead. Contrary to the estimated coefficients in (2), APEs also reveal the magnitude of each included variable's effect[33]. When interpreting our findings, it is useful to keep in mind that the average unconditional probability of becoming a city in one of the centuries in our sample is about 0.05%. This puts perspective on the generally small magnitude of the estimated APEs.

Table 1. Baseline results

		1	2	3	4
	P(city t | no city t-1)	**BASELINE**	coordinate block / century FE	potential city location FE (CRE)	matched 1.5km
X_i	sea	0.007***	0.006***	0.002***	0.004***
		[0.00]	[0.00]	[0.00]	[0.00]
	river	0.012***	0.013***	0.002***	0.008***
		[0.00]	[0.00]	[0.00]	[0.00]
	hub	0.007***	0.006***	0.002***	0.005***
		[0.00]	[0.00]	[0.00]	[0.00]
	road	0.003***	0.003***	0.001***	0.002***
		[0.00]	[0.00]	[0.00]	[0.00]
	ln elevation	-0.0001**	-0.00004*	-0.00002***	-0.00003
		[0.01]	[0.09]	[0.00]	[0.11]
	ruggedness	0.0001***	0.00003	0.00003***	0.0001***
		[0.00]	[0.26]	[0.00]	[0.01]
	P(cultivation)	0.001***	0.001***	0.0001***	0.0003***
		[0.00]	[0.00]	[0.00]	[0.00]
X_{it-1}	city >= 10k? (t-1)				
	0 – 20 km	-0.0002**	-0.0001	-0.000002	-0.0002***
		[0.02]	[0.12]	[0.95]	[0.01]
	20 – 50 km	0.0001**	0.0002***	0.00003**	0.0001
		[0.05]	[0.00]	[0.04]	[0.28]
	50 - 100 km	0.0001***	0.0002***	0.00004**	0.0001**
		[0.01]	[0.00]	[0.02]	[0.02]
	P(city) unconditional	0.0005	0.0005	0.0005	0.0003
	country/century FE	yes	block/century	yes	yes
	nr observations	1840091	1998901	1894008	1670472
	ln pseudo likelihood	-7171	-7076.5	-6673	-4185

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. The p-values are based on the estimated coefficients and their standard errors since these do not depend on the particular calculation of the average partial (or marginal) effects. Result are the same when using the p-values based on the calculated standard errors of the average partial effects. Only in this table, we add X_i and X_{it-1} to the left of our variable names to clarify the correspondence between equation (2) and our results.

[33] Average Partial Effects are an estimate of the derivative of the expected value of the independent variable with respect to the included variables of interest (see e.g. Wooldridge, 2005). In case of our model specified in (2), the APE of X_{it-1} is for example calculated as:

$$\hat{\beta}_1 \frac{1}{NT}\sum_{it} F'\left(X_{it-1}\hat{\beta}_1 + X_i\hat{\beta}_2 + X_{ct-1}\hat{\beta}_3 + \hat{\alpha}_{ict}\right) \qquad (3)$$

Using a linear probability model, with F the identity function, this would simply be $\hat{\beta}_1$. When F is a nonlinear function (Φ in our baseline probit case), this is no longer so.

Column 1 shows our baseline findings. *1st nature geography* is very important in determining a location's urban chances. Especially preferential location for water-based transportation substantially increases a location's probability to become a city (by 1.2ppt (rivers) and 0.7ppt (sea) respectively). Location at overland transport routes also carries positive effects, but they are smaller than being located close to navigable water (0.7ppt (hubs) and 0.3ppt (roads) respectively).

Besides a location's position on the main transportation networks, a favorable location in terms of agricultural possibilities, at lower altitudes, or in more rugged terrain also contributes positively to a location's urban chances. These latter three effects are however much smaller than that of favourable location for river-, sea- or road-transportation, hereby corroborating claims by e.g. Bairoch (1988, p.143) – see also our discussion in section 2.1 – about the dominant role of transportation in determining city-location. Moreover, the effects we find of elevation and ruggedness are the least robust of our findings so that we are reluctant to stress their relevance based on our baseline findings (see e.g. column 2 or 4; column 4 in Table 3a; column 1 in Table 3b; or Tables A5 and A8).

Compared to 1st nature geography, and the effect of location on the main transportation networks in particular, 2nd nature geography's impact on a location's urban chances is smaller. What is very interesting however, is that our flexible modelling strategy uncovers almost the exact prediction made by new economic geography theory (see section 2.2). The effect of an already existing city on another location's urban chances depends non-linearly on a location's distance to that city. Only locations at medium distance (20-100km) from an already existing city have significantly better urban chances. They have about a 0.01 ppt higher probability to become a city than locations located further away. Competition instead dominates at close range: an already existing city within the nearest 20km significantly diminishes a location's own urban chances by about 0.02 ppt.

5.1 Robustness

In this section we show that the above-discussed baseline results hold up to a wide variety of robustness checks. These checks result in one important refinement that shows the changing importance of 1st and of 2nd nature geography in determining city location over the centuries. Besides the robustness checks discussed in the main text, Tables A5, A6 and A7 in Appendix A further verify the sensitivity of our baseline results to the estimation strategy used, the absolute population cutoff of 5,000 inhabitants that we use to define a city, and to the particular distance bands we use to construct our 2nd nature geography variables.

5.1 *Unobserved heterogeneity, estimation strategy and match potential – actual cities*

In our baseline results we control for any unobserved *country-century-specific* variables that could leave their effect on a location's urban chances. Columns 2 and 3 verify whether or not our findings critically hinge upon this country-century specification. In column 2, we use block-century fixed effects instead that are based on a division of Europe into 25 geographically clustered *blocks*, using the 20th, 40th, 60th and 80th quantile of the distribution of all locations' latitude and longitude as boundaries. Doing this, does not substantially change our baseline findings. The only variable that loses its significance is a location's ruggedness.

In column 3, we additionally control for any unobserved time-*invariant* factors at the *potential city location level*.[34] Allowing for such location-specific unobserved heterogeneity when employing non-linear panel data techniques is not as straightforward as in linear panel data models where one can simply include a dummy variable for each location and obtain consistent estimates of the parameters of interest (see e.g. Heckman, 1981; Wooldridge, 2005; or Carro, 2007)[35]. In our baseline probit case including such dummy variables results in inconsistent estimates of the parameters of interest (the incidental parameters problem).

To get around this problem, we use the Conditional Random Effects (CRE) strategy proposed by Wooldridge (2005), and specify the distribution of the unobserved *location-specific* effects conditional on the individual specific mean, $\overline{X}_{i\cdot}$, of the included time-varying variables, the country-century specific fixed effects, α_{ct}, and a location's initial city status in 800[36], c_{i800}, i.e.: $\alpha_{ict} = \alpha_{ct} + c_{i800}\zeta + \overline{X}_{i\cdot}\xi + \eta_i$, with $\eta_i \mid \alpha_{ct}, c_{i800}, \overline{X}_{i\cdot} \sim N(0, \sigma_\eta^2)$. Under this assumption for the location-specific unobserved heterogeneity we can employ random effect probit techniques to get consistent estimates of the parameters of all our second nature geography variables. We also get estimates of the parameters on the 1st nature geography variables, but we do not discuss these in any detail as we can not separately identify each 1st nature geography variable's effect on a location's urban chances from its partial correlation with the location-specific unobserved effects (see Wooldridge, 2005).[37]

---

[34] Our baseline results are only valid under the assumption of no time-invariant location-specific heterogeneity other than that controlled for by our included 1st nature geography variables (even when this heterogeneity is uncorrelated with the variables of interest, i.e. random effects, we would get incorrect estimates of our parameters of interest given the implicit dynamic nature of our model).

[35] See Table A5 column 4 for the results of estimating a linear probability model including location fixed effects.

[36] Given the dynamic nature of the model, the presence of any (random or fixed) location-specific time-invariant unobserved heterogeneity requires the inclusion of this initial value to address the 'initial condition problem'.

[37] If one is willing to assume that this partial correlation is zero, the shown APE's on our 1st nature geography variables can directly be interpreted. In that case, all our baseline findings, come through.

The column 3 results show that our main 2nd nature results also hold up to controlling for location-specific unobserved heterogeneity. Although the negative competition effect at short range (0-20km) is no longer significant, we still find the significant positive effect of being located at medium distance to an already existing city.

Finally, in column 4, we show results of using a much stricter criterion to match actual cities to our randomly drawn coordinate pairs (see section 3.2 for more detail), i.e. 1.5km instead of 2.5km in our baseline estimates. Doing this, we only match 624 (or 39.3%) of all 1,588 actual cities. This amounts to a substantial loss of variation (in our baseline sample, we match 1,150 (or 72.4%) of all actual cities), yet still our main results come through. Of our 1st nature geography variable only elevation loses its significance. We also still find the non-linear 2nd nature geography effect, predicted by theory (although the positive effect at 20-50km is no longer significant[38]).

5.1.2 Additional variables and sample composition

Table 2 shows a second set of robustness checks. They concern the inclusion of additional non-geography related control variables to our baseline model, as well as a check to establish whether our results are primarily driven by developments in earlier or later centuries only.

Columns 1 – 2 add additional variables to our baseline specification. Reassuringly, and with only few exceptions, all our main baseline results come through. The results on the extra included variables are of interest by themselves however.

Column 1 controls for a potential city location's religious, political and educational status in period *t-1*. We know for each location in each century whether it was an archbishopric, the capital of a large political entity, and whether it had a university or not. These data are taken from Bosker et al. (2012). We find that having an important religious [archbishopric] or political [capital] function substantially increases a location's urban chances. Note however, that the results on these non-geography related variables should be taken with some care. The extra information on the included political and religious variables at the lower-right bottom of Table 2 shows that only 28 (or 0.001%) of all our potential city locations are a capital or have a university, and only 75 (or 0.003%) are an archbishopric, before becoming a city. Although these characteristics significantly improve a location's urban chances, such locations are major exceptions.

[38] The 20-50km and 50-100km already existing dummies are jointly significant [p-value: 0.04]. Also, note that our final time-varying results are even more robust to the use of a stricter matching criterion (see Table 3b, column 2).

Table 2. Robustness (variables included, sample composition)

	1	2	**3**	**4**
P(city t | no city t-1)	political and religious function	ever a city before?	**800 - 1500**	**1600 - 1800**
sea	0.007***	0.006***	0.004***	0.009***
	[0.00]	[0.00]	[0.00]	[0.00]
river	0.012***	0.010***	0.010***	0.013***
	[0.00]	[0.00]	[0.00]	[0.00]
hub	0.005***	0.005***	0.005***	0.008***
	[0.00]	[0.00]	[0.00]	[0.00]
road	0.003***	0.003***	0.003***	0.003***
	[0.00]	[0.00]	[0.00]	[0.00]
ln elevation	-0.0001***	-0.0001***	-0.0001**	-0.0001*
	[0.01]	[0.01]	[0.03]	[0.09]
ruggedness	0.0001***	0.0001***	0.00002	0.0002***
	[0.00]	[0.00]	[0.43]	[0.00]
P(cultivation)	0.0005***	0.0004***	0.001***	0.0004***
	[0.00]	[0.00]	[0.00]	[0.00]
city >= 10k? (t-1)				
0 - 20 km	-0.0002**	-0.0001*	-0.0004***	-0.00004
	[0.04]	[0.08]	[0.00]	[0.67]
20 - 50 km	0.0001*	0.0001**	-0.0001	0.0002***
	[0.06]	[0.02]	[0.18]	[0.00]
50 - 100 km	0.0001***	0.0001***	-0.0001	0.0003***
	[0.00]	[0.01]	[0.14]	[0.00]
P(city) unconditional	0.0005	0.0005	0.0002	0.0012
country/century FE	yes	yes	yes	yes
nr observations	1840091	1840091	1139055	701036

extra included variables	1	2
capital t-1	0.154***	-
	[0.00]	-
archbishop t-1	0.041***	-
	[0.00]	-
university t-1	0.061***	-
	[0.00]	-
ever a city before?	-	0.064***
	-	[0.00]

extra info on political and religious variables	# (%) no city t-1, but
archbishop t-1	75 (0.003%)
capital t-1	28 (0.001%)
university t-1	28 (0.001%)

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. The p-values are however based on the estimated coefficients and their standard errors.

Column 2 instead controls for a location's (urban) population history. We include a dummy variable indicating whether or not a location had ever been a city before. This is done to control for the presence of cities (0.02% of the sample) that at some point pass our 5,000 inhabitants criterion, subsequently fall back below this number, to pass it again in a later century[39]. These cities would – so to speak – be counted double in our sample, which could leave an effect on the results. This is however not the case, but the results do show that locations that once already qualified as a city, but subsequently lost their city status, have a 6.4 ppt higher probability to regain city status.

[39] The Black Death (the 14th century plague epidemic) is responsible for many of these 'city-disappearances'. 40% of the existing cities in 1300 'disappeared' in 1400, i.e. fell back below the 5,000 inhabitants thresehold.

Contrary to all other robustness checks, columns 3 and 4 reveal that our results do change markedly when we separately consider the earlier (800-1500) or later (1600-1800) centuries in our sample period. In Table A5 in the Appendix we also show results when using a finer decomposition of the sample along century lines. The patterns shown in these results are very similar to those using a pre- and post-1600 split of the sample. For parsimony reasons we decided to show the pre- and post 1600 results in the main text[40].

The (relative) importance of the various 1st nature characteristics is quite similar in the pre- and post-1600 period. A good position on the main transportation arteries, and next to navigable water in particular, is most important in both periods. The relative importance of water- over land-based transportation slightly increases in the later centuries, corresponding to narrative accounts by (economic or urban) historians that document an increased dominance of water- over land-based transport in late Medieval and pre-modern Europe (Lopez, 1956; Hohenberg, 2004)[41]. Also, the importance of a location's agricultural possibilites somewhat diminishes over time. This is not only consistent with gradually improving agricultural methods diminishing the relative advantage of being located near highly productive lands, but also with the notion that in the later centuries food could be transported over greater distances at lower costs due to improvements in transportation technology (Duranton, 1999, p.2173)[42]. Finally, the positive effect of being located in rugged terrain only shows up in the later centuries. As mentioned before however, this finding is not very robust (see e.g. column 4 of Table 3a, Table A4, but also column 5 of Table A6).

Results change much more substantially when considering the relevance of 2nd nature geography. In the pre-1600 period, we find a strong competition effect at close distance, but no evidence of any significant positive effects of being located at medium distance from an already existing city. This changes markedly in the later centuries. The competition effect at close distance is much weaker in the post-1600 period; instead, we find strong evidence that the presence of an existing city at medium distance increases a location's urban chances.

---

[40] We stress at this point that we not claim in any way that 1600 is the exact year in which these changes occurred. What we do want to stress is that the (relative) importance of 1st and 2nd nature geography in determining city location changed significantly over the centuries. Since our data come at 100 year intervals, taking the year 1600 to be some kind of a crucial 'breakpoint' year would in our view be unwarranted.

[41] We do not want to emphasize this too much however, given that we focus on the Roman road network only. Lopez (1956) e.g. argues that the importance of the Roman road network diminished during the Middle Ages. One important reason was that the Roman road system was planned mostly for military purposes that did not always correspond to the most economical route. As a result (Lopez, 1956 p.21): *"in the later Middle Ages […] little by little a new network of roads was put into effective operation, different totally in structure and methods from the ancient one […] The routing reflected the needs of commerce rather than the convenience of soldiers and civil servants."*

[42] In north-western Europe e.g. the grain trade with eastern Europe became increasingly important (Hybel, 2002).

5.2 *The changing importance of 1st and 2nd nature geography over time*

The difference in results when considering the earlier or later centuries of our sample period is the most important refinement to our baseline results. In this section we further explore this finding.

Table 3a. Pre- and post-1600 results

P(city t | no city t-1)	1 MAIN RESULTS		2 CRE		3 no UK		4 no 1800	
period:	800-1500	1600-1800	800-1500	1600-1800	800-1500	1600-1800	800-1500	1600-1700
sea	0.009***	0.015***	0.001***	0.002***	0.006***	0.010***	0.039***	0.067***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
river	0.023***	0.022***	0.003***	0.003***	0.015***	0.014***	0.075***	0.074***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
hub	0.011***	0.014***	0.002***	0.002***	0.006***	0.008***	0.046***	0.042***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
road	0.008***	0.007***	0.001***	0.001***	0.004***	0.004***	0.038***	0.034***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
ln elevation	-0.0002**	-0.0001*	-0.00004***	-0.00002**	-0.0006*	-0.0005*	-0.001**	-0.002**
	[0.03]	[0.09]	[0.00]	[0.01]	[0.09]	[0.07]	[0.03]	[0.05]
ruggedness	0.0001	0.0004***	0.00002	0.00004***	0.0003	0.001***	0.0005	0.0002
	[0.43]	[0.00]	[0.12]	[0.00]	[0.47]	[0.00]	[0.43]	[0.77]
P(cultivation)	0.002***	0.001***	0.0002***	0.0001***	0.001***	0.0003***	0.011***	0.004
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]	[0.00]	[0.17]
city >= 10k? (t-1)								
0 – 20 km	-0.0015***	-0.0001	-0.0001**	0.00002	-0.001***	-0.0001	-0.010***	-0.002
	[0.00]	[0.67]	[0.03]	[0.49]	[0.00]	[0.5]	[0.00]	[0.49]
20 – 50 km	-0.0003	0.0005***	-0.00001	0.0001***	-0.0001	0.0003***	-0.002	0.004***
	[0.18]	[0.00]	[0.85]	[0.01]	[0.33]	[0.00]	[0.18]	[0.01]
50 - 100 km	-0.0003	0.0008***	-0.00003	0.0001***	-0.0001	0.0003***	-0.002	0.006***
	[0.14]	[0.00]	[0.18]	[0.00]	[0.15]	[0.00]	[0.14]	[0.00]
P(city) unconditional	0.0002	0.0012	0.0002	0.0012	0.0002	0.0012	0.0002	0.0006
country/century FE	yes		yes		yes		Yes	
nr observations	1840091		1894008		1695121		1585283	
ln pseudo likelihood	-7145		-6656		-6551		-3787	
p-value H0: pre 1600 = post 1600								
sea	[0.03]**		[0.00]***		[0.05]**		[0.00]***	
river	[0.67]		[0.62]		[0.34]		[0.60]	
hub	[0.36]		[0.54]		[0.49]		[0.64]	
road	[0.90]		[0.23]		[0.54]		[0.84]	
ln elevation	[0.43]		[0.15]		[0.73]		[0.00]***	
ruggedness	[0.04]**		[0.24]		[0.02]**		[0.79]	
P(cultivation)	[0.13]		[0.09]*		[0.03]**		[0.06]*	
city >= 10k? (t-1)								
0 – 20 km	[0.00]***		[0.01]***		[0.00]***		[0.00]***	
20 – 50 km	[0.01]***		[0.05]*		[0.00]***		[0.00]***	
50 - 100 km	[0.00]***		[0.00]***		[0.00]***		[0.00]***	

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. Whenever the effect of a variables is split in a pre- and post-1600 effect, the average partial effect is calculated using only the observation in the pre- or post-1600 period only. The p-values are however based on the estimated coefficients and their standard errors.

Instead of simply splitting the sample, we estimate (2) allowing all variables to have a pre- and post-1600 specific effect by interacting each variable with a pre- and a post-1600 dummy.

This has the advantage that we can formally test the equivalence of the pre- and post-1600 effect of each of the included variables. The p-values corresponding to tests for the equality of each respective variable's effect in the pre- and post-1600 period are shown at the bottom of Table 3a.

Column 1 of Table 3a shows our main results. They largely confirm our earlier findings in column 3 and 4 of Table 2. In case of 1st nature geography however, we find that most of the changes in the pre- and post-1600 effect are not significant. Those that are significant are location at sea and in rugged terrain, both becoming more important in the later centuries. But, we do not want to stress these findings as they do not hold up to one (or more) robustness checks. In these same robustness checks, we also, contrary to the results in column 1, very often find a significant decrease in the importance of location right next to productive fields (see e.g. column 3 and 4 in Table 3a, or column 1 in Table 3b).

Instead, the changing importance of 2nd nature geography is strongly confirmed. Only in the later centuries we find that locations at medium range (20 – 100km) from an already existing city have significantly higher urban chances than locations located close or further away. This significantly differs from the earlier centuries, when we do not find this effect. During this earlier period, we instead find evidence of significantly lower urban chances for locations at too close range (0 – 20km) to an already existing city. Over time, this urban shadow effect significantly decreases, turning insignificant in the post-1600 period.

5.2.1 Robustness of our time-varying results

Before discussing the implications of our findings, we first verify the robustness of our time-varying results[43].

The first three robustness checks are shown alongside our main results in Table 3a. Column 2 shows that our main results hold up when also allowing for any time-invariant location-specific unobserved heterogeneity (in addition to any country-century specific unobserved heterogeneity)[44].

Next, column 3 shows the insensitivity of our main results to excluding the UK, the earliest industrializing country, from the sample. The number of cities in the UK more than

---

[43] Again, see also the Appendix for additional robustness checks. Appendix A.3 in particular addresses possible concerns that our 2nd nature geography results may arise by construction due to the increased density of the European city system over the centuries.

[44] We only show results of estimating a conditional random effects probit model. Results when estimating a linear probability model including location-specific fixed effects instead are very similar and available upon request. Also, using different specifications to capture possible time-varying unobserved heterogeneity (see e.g. column 2 in Table 1) does not change our main results.

triples in the 18th century[45]. Relatedly, in column 4 we omit the eighteenth century from the sample. As shown in Figure A1 and Table A1 in Appendix A, the eighteenth century saw an unprecedented increase in the number of cities. Column 4 however shows that it is not only this episode that drives our results (except for our 'ruggedness'-results).

Table 3b. Pre- and post-1600 results

	1		2	
P(city t | no city t-1)	matched 1.5km		measurement error	
period:	800-1500	1600-1800	800-1500	1600-1800
sea	0.004***	0.004***	1	1
	[0.00]	[0.00]	< 1 >	< 1 >
river	0.010***	0.010***	1	1
	[0.00]	[0.00]	< 1 >	< 1 >
hub	0.006***	0.006***	1	1
	[0.00]	[0.00]	< 1 >	< 1 >
road	0.003***	0.003***	1	1
	[0.00]	[0.00]	< 1 >	< 1 >
ln elevation	-0.0001	-0.00004	0.915	0.211
	[0.24]	[0.19]	< 1 >	< 0.666 >
ruggedness	0.00002	0.0001***	0	1
	[0.52]	[0.01]	< 0 >	< 1 >
P(cultivation)	0.001***	0.0002**	1	1
	[0.00]	[0.01]	< 1 >	< 1 >
city >= 10k? (t-1)				
0 – 20 km	-0.0004***	-0.0001	1	0
	[0.01]	[0.13]	< 1 >	< 0 >
20 – 50 km	-0.0002**	0.0001**	0.006	1
	[0.04]	[0.01]	< 0.208 >	< 1 >
50 - 100 km	-0.0001	0.0002***	0	1
	[0.19]	[0.00]	< 0.026 >	< 1 >
>= 1k t-1?	-		-	
P(city) unconditional	0.0001	0.0007	-	-
country/century FE	yes		yes	
nr observations	1670472		-	
ln pseudo likelihood	-4170.5		-	
p-value H0: pre 1600 = post 1600				
sea	[0.85]		-	
river	[0.77]		-	
hub	[0.94]		-	
road	[0.75]		-	
ln elevation	[0.82]		-	
ruggedness	[0.26]		-	
P(cultivation)	[0.01]***		-	
city >= 10k? (t-1)				
0 – 20 km	[0.04]**		-	
20 – 50 km	[0.00]***		-	
50 - 100 km	[0.00]***		-	

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. Whenever the effect of a variables is split in a pre- and post-1600 effect, the average partial effect is calculated using only the observation in the pre- or post-1600 period only. The p-values are however based on the estimated coefficients and their standard errors. In column 2 we show the fraction of simulations that each respective variable is significant at the 5% or < 10% > level.

[45] Ashton (1948) dates the start of the Industrial Revolution in Britain in the late eighteenth century. In continental Europe it only gathered steam in the first half of the nineteenth century. Excluding also Belgium, the earliest industrializer on the continent, also leaves our results unaffected. Results available upon request.

The next two robustness checks are shown in Table 3b. In column 1, we show results of using a stricter 1.5km criterion to match actual cities to our randomly drawn coordinate pairs. Doing this turned the already existing city dummy within 20-50km insignificant in Table 1 (column 4). In the time-varying results this is no longer the case. The positive effect of an already existing city within 20-50km still comes through in the later centuries. By contrast, this variable is significantly negative in the pre-1600 period, suggesting an even stronger urban shadow effect than that implied by our baseline findings (also explaining this variable's nonsignificance when not distinguishing between the earlier and later part of our sample period – see Table 1 column 4). All other results come through when using this stricter matching criterion.

The second robustness check deals with the issue of measurement error. Bairoch et al. (1988) acknowledge that getting spot-on population estimates is sometimes difficult, especially for the smaller cities and for the earlier centuries. As we are using a nonlinear transformation of the city population data (by estimating a probit model), such measurement error, even if it were random, could leave its effect on our results (see e.g. Hausman, 2001). To verify the sensitivity of our results to measurement error, we adopt the following simulation strategy. We assume that each reported population estimate has a 40% probability of being misreported. Conditional upon being misreported, we subsequently assume that there is an equal, 25%, chance of being underestimated by 2,000 inhabitants, overestimated by 2,000 inhabitants, underestimated by 1,000 inhabitants, or overestimated by 1,000 inhabitants respectively. This structure for the measurement error implicitly assumes that Bairoch et al. (1988) made relatively bigger mistakes for smaller population numbers[46]. We generate 1,000 different population samples using this sampling strategy and estimate our baseline model using each of them separately. Column 2 reports the fraction of simulations that each variable is significant at a 5% and at a 10% level respectively. Under the assumption of measurement error, each of the 1,000 simulated samples is 'equally true'. If we find that a significant variable in our main results is less than 90% of the times significant at the 10% level, this would shed some doubts on the actual relevance of this variable. Reassuringly, column 2 shows that all our results hold up to this measurement error check.

5.3 The importance of our time-varying results

Overall, the time-varying impact of both 1st and 2nd nature geography is the most important refinement to our baseline results in Table 1. 1st nature geography, and preferential location

[46] Bairoch (1988, p.525) expects a margin of error of about 10% for overall European city population around 1300 and 1500, increasing to 15% in 1000 and even 20% in 800.

on the main transportation arteries in particular, is the dominant driver of city location throughout our sample period. There is some evidence of an increased importance of water- over land-based transportation[47], and a diminishing importance of being located right next to fields of high agricultural productivity[48], over the centuries, yet these findings do not hold up to all of our robustness checks.

By contrast, our results show strong evidence that the relevance of 2nd nature geography in determining city location changed substantially over time. In the early centuries, we find no evidence that an already existing city exerts a positive influence on other locations' urban chances at any distance. At too close range the already exisiting city even exerts a strong urban shadow: locations within a 20km range of an already existing city have significantly smaller urban chances. These effects change substantially when looking at the later centuries in our sample. On the one hand the urban shadow looses its importance, locations at close range from an already-exisiting city no longer have significantly smaller urban chances. On the other hand, we now find a positive effect of an already existing city at medium distance: locations at medium distance from already existing cities have significicantly higher urban chances than those at closer or further distance. They combine the advantage of cheaper trade with existing cities compared to locations at further distance, with that of weaker competition with existing cities compared to locations at closer distance.

Interestingly, this increased importance of 2nd nature geography over the centuries is consistent with predictions from theory (Behrens, 2007; Fujita and Mori, 1997; Duranton, 1999). As set out briefly in section 2.2, theory predicts that 2nd nature geography only is an important *positive* determinant of city location when overall trade costs are sufficiently low, the advantages of co-locating in a city are sufficiently large compared to its disadvantages, and overall population is large enough to sustain multiple urban centres.

Each of these three developments occurred over the 800 - 1800 period. Trade costs diminished substantially. Not only did transportation technology improve significantly, also the introduction of e.g. the bill of laden, insurance contracts (Greif, 2006 p.24), and other institutional and political changes that improved security and law[49] greatly reduced the costs

[47] This corresponds nicely to earlier accounts by for example Lopez (1956) or Pirenne (1925). Also, it concurs with the notion that improvements in shipping technology [not only in the size and speed of the vessels used, but also in e.g. navigation (van Zanden and van Tielhof, 2009) and canal building (Bairoch, 1988)] were larger than those in land transportation despite the fact that e.g. horseshoes, rigid tandem horse collars, and the use of explosives to build tunnels, did all significantly improve land-based transportation (see Lopez, 1956).

[48] A finding that is not only consistent with improving agricultural methods, but also with the notion that in later centuries food could be transported over greater distances at lower costs due to improvements in transportation.

[49] Lopez (1956, p.24): " *an English statute of 1285 ordered that along highways between market towns "there be no dyke, tree or bush whereby a man may lurk to do hurt within 200 feet of either side of the way"* "

of (long-distance) trade (Greif, 1992; Hohenberg, 2004 p.3025; Duranton, 1999 p.2177). Second, the advantage of co-locating in a city gradually increased due to improved non-agricultural production techniques (e.g. the blast furnace, finery forge, treadwheel crane, water- and windmills, and the printing press), while its disadvantages decreased (improved living conditions). And finally, overall European population increased markedly over our sample period, largely because of improvements in agricultural production (crop rotation, heavier plows, the introduction of new crops).

Our finding that 2nd nature geography only exerts a significant positive influence on a location's urban chances during the later centuries is consistent with these three developments. In earlier centuries trade costs were (too) high, and economies of scale and/or overall population (too) low, making 1st nature geography the only dominant determinant of city location. Only with the gradual increase in economies of scale, a growing overall population and decreasing trade costs, do we start to find the positive effect of location at medium distance from existing cities that corresponds closely to the predictions from economic geography theory.

6. Refining 2nd nature geography

In this section we show that further refining the impact of already existing cities[50], gives useful additional insights into the role of 2nd nature geography in determining city location[51]. It also gives further confidence in our baseline 2nd nature geography results[52]. Throughout this section, we always distinguish between the earlier and later centuries in our sample.

6.1 Refining the impact of already existing cities

In column 1 and 2 of Table 4, we replace our three 'already existing city'-dummy variables with more standard measures of 2nd nature geography: a location's urban potential in column 1 [see equation (1)], and the distance to, and population of, the nearest already existing city in

[50] We also experimented with various interaction terms of our 1st nature geography dummy variables. Doing this generally gives the result that having a favourable location for both land- and water-transport increases a location's urban chances compared to favourable location for only one of the two transport modes, but not always significantly so. Adding these interaction terms leaves the baseline results regarding 2nd nature geography unchanged. They are available upon request.

[51] An even more elaborate way to refine our 2nd nature geography variables would be to take account of e.g. actual road or river systems, or the ruggedness of the terrain, and come up with more detailed indicators of travel distance between locations than our great circle distances. Aside from the additional data requirements, note that such extensions require making assumptions on the relative importance of each of the additionally considered characteristics in determining overall travel distances. We leave such extensions for future work.

[52] If we would for example find that the presence of an existing city larger than 25,000 inhabitants does exert a positive influence on potential city locations' urban chance within a 0 – 20 km range (whereas an existing city larger than 10,000 does not [i.e. our baseline result]), this would shed some doubts on our main findings.

column 2. The results of using these more standard measures corroborate our claim (see section 3.3.2) that they are too restrictive to do justice to the patterns in the data.

Table 4. 2nd nature geography – some extensions

	1		2		3		4	
	(a) ln UP (cities >= 10k)		(a) ln pop near city >= 10k		(a) city >= 5k		(a) 1 city >= 10k	
	(b) -		(b) ln dist near city >= 10k		(b) city >= 10k		(b) >= 2 cities >= 10k	
	(c) -		(c) -		(c) city >= 25k		(c) -	
P(city t | no city t-1)	800-1500	1600-1800	800-1500	1600-1800	800-1500	1600-1800	800-1500	1600-1800
1st nature geography as in Table 3a								
	(a) t-1		(a) t-1		(a) t-1		(a) t-1	
0 - 20 km	-0.001***	-0.0002	-0.0001	0.00002	-0.001**	0.0006***	-0.0015***	-0.0002
	[0.00]	[0.15]	[0.16]	[0.89]	[0.03]	[0.00]	[0.00]	[0.31]
20 - 50 km	-	-	-	-	0.0002	0.001***	-0.0003	0.0004***
	-	-	-	-	[0.6]	[0.00]	[0.19]	[0.01]
50 - 100 km	-	-	-	-	0.0007*	0.0001	-0.0002	0.0006***
	-	-	-	-	[0.07]	[0.70]	[0.32]	[0.00]
	(b) t-1		(b) t-1		(b) t-1		(b) t-1	
0 - 20 km	-	-	0.0003***	-0.00003	-0.0007	-0.0003	-	0.001*
	-	-	[0.00]	[0.36]	[0.41]	[0.21]	-	[0.05]
20 - 50 km	-	-	-	-	-0.0004	-0.00003	0.0003	0.00002
	-	-	-	-	[0.3]	[0.88]	[0.55]	[0.92]
50 - 100 km	-	-	-	-	-0.0006*	0.0007***	-0.0002	0.0004**
	-	-	-	-	[0.05]	[0.00]	[0.44]	[0.02]
	(c) t-1		(c) t-1		(c) t-1		(c) t-1	
0 - 20 km	-	-	-	-	0.0007	-0.0005	-	-
	-	-	-	-	[0.5]	[0.12]	-	-
20 - 50 km	-	-	-	-	-0.0001	-0.0003	-	-
	-	-	-	-	[0.83]	[0.12]	-	-
50 - 100 km	-	-	-	-	-0.0002	-0.0002	-	-
	-	-	-	-	[0.50]	[0.19]	-	-
P(city) unconditional	0.0002	0.0012	0.0002	0.0012	0.0002	0.0012	0.0002	0.0012
country/century FE	yes		yes		yes		yes	
nr observations	1840091		1840090		1840091		1838523	
p-values tests					H0: $\beta_{city>=10}$ >0?		H0: $\beta_{2\ cities>=10}$ >0?	
0 - 20 km	-	-	-	-	[0.00]***	[0.40]	-	[0.11]
20 - 50 km	-	-	-	-	[0.40]	[0.00]***	[0.82]	[0.04]**
50 - 100 km	-	-	-	-	[0.82]	[0.00]***	[0.10]*	[0.00]
					H0: $\beta_{city>=25}$ >0?			
0 - 20 km	-	-	-	-	[0.01]***	[0.23]	-	-
20 - 50 km	-	-	-	-	[0.32]	[0.00]***	-	-
50 - 100 km	-	-	-	-	[0.33]	[0.01]***	-	-

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. The p-values are however based on the estimated coefficients and their standard errors. All regressions contain the same 1st nature geography variables as in column 2 of Table 1. The estimated parameters on these variables correspond closely to those reported in Table 3. They are available upon request. In column 4, the dummy variables indicating the presence of at least 2 cities within 0-20km perfectly predicts failure, so that we do not have any variation to pinpoint its effect (i.e. all locations located within 0-20km of two or more already existing cities never become a city themselves.)

For the earlier centuries one would still reach a similar conclusion as we do. The negative coefficient on urban potential in column 1, and the positive coefficient on distance to the nearest already existing city in column 2, indicate a negative effect of an already existing city on a location's urban chances that is larger the closer this location is to the already existing city. For the later centuries either of the measures in column 1 and 2 is insignificant. By assuming an always positive or always negative effect of already existing cities, these measures are unable to uncover the nonlinear effect that already existing cities exert on the urban chances of their surroundings. This is only revealed when using our more flexible dummy specification.

Columns 3 and 4 instead further specify the dummy variables included in our baseline estimations. Column 3 includes six additional dummy variables, three indicating the presence of *at least one already existing city larger than 25,000 inhabitants*, and three indicating the presence of *at least one already existing city larger than 5,000 inhabitants* in each of the three distance bands respectively. And column 4 includes three additional dummy variables, indicating the presence of *at least two cities larger than 10,000 inhabitants* in each of the three distance bands respectively.

The results in column 3 show an interesting pattern[53]. In the earlier centuries, the urban shadow effect does not depend on the size of the already existing city: a city larger than 5,000, 10,000 or 25,000 inhabitants always diminishes the urban chances of its nearby surroundings (0-20km). However, and by contrast to our baseline findings, we now also find a positive effect of an already existing city as well, but only so for a small already existing city (more than 5,000 but less than 10,000 inhabitants) that is sufficiently far away (50-100km). Larger already existing cities never carry positive effects for other locations' urban chances. This changes in the later centuries. Interestingly, we find that the larger the distance between an existing city and a potential city location, the larger the existing city has to be to exert a positive influence on that potential city location's urban chances. Put differently, the larger an already existing city, the larger its urban shadow (a finding that corresponds nicely with early observations by e.g. Lösch (1940, p.126) or Ullman (1941, p.856), and with predictions from economic geography theory[54]). A city larger than 5,000 inhabitants significantly increases the urban chances of locations within 0 – 50km, but it is too small to exert a positive influence beyond that distance. The effect of a city larger than 10,000 inhabitants instead is also positive at larger distances (50-100km), but, contrary to a smaller city, no longer carries significant

[53] Given the way the different dummy variables are specified (i.e. if there exists a city larger than 25,000 inhabitants within a certain distance band, not only the dummy variable indicating the presence of a city larger than 25,000 inhabitants will be 1, so will be the dummy variable indicating the presence of a city of at least 10,000 inhabitants), the p-values below the coefficients indicate whether or not the effect of a dummy variable is significantly different from the effect of having a smaller city within a distance band.
[54] See e.g. the discussion around figure 6 in Fujita et al., 1999.

positive effects at close range (0-20km). An even larger city of more than 25,000 inhabitants carries effects that are similar to those of a city larger than 10,000 (they are slightly smaller, but not significantly so).

A similar result follows from column 4. As in our baseline findings, the presence of only one city larger than 10,000 inhabitants exerts a positive influence on the urban chances of locations at 20 – 100km (in the post-1600 period). A second already existing cities does not further increase this effect, except when it is located sufficiently far away (at least 50km).

6.1.1 More flexible 2nd nature geography

Our final results show what happens when we completely abandon our 'three-distance-bands-based approach' (see also section A.3 in Appendix A for the results when using twentyfive 20km distance bands). Instead of using our baseline dummy variable approach, we now model the effect of 2nd nature geography by including as right hand side variables a sixth order polynomial in *distance to the nearest already existing city*. In order to facilitate interpretation of our results we estimate a linear probability model[55], so that (2) becomes:

$$P\left(c_{ict}=1 \mid c_{ict-1}=0, X_{ict-1}, X_i, \alpha_{ict}\right)=\sum_{m=1}^{6}\left(\left(D_i^{near}\right)^m \gamma_m\right)+X_i^{GEO1st}\beta_1+\alpha_{ict}+\varepsilon_{it} \quad (3)$$

, where D_i^{near} denotes the distance to the nearest already existing city from location *i*, and X_i^{GEO1st} the same 1st nature geography variables as in our baseline model (see Table 1). Based on the six estimated γ_m's we can verify whether, and if so how, the marginal effect of distance to the nearest already existing city varies at different distances. Moreover, we can calculate confidence intervals around these marginal effects. Figure 5 shows the results[56]. where we estimate (3) allowing all parameters to differ between the pre- and post-1600 period.

The patterns shown in the two panels of Figure 5 again confirm our findings[57]. In the pre-1600 period we find that being located further away from an already existing city

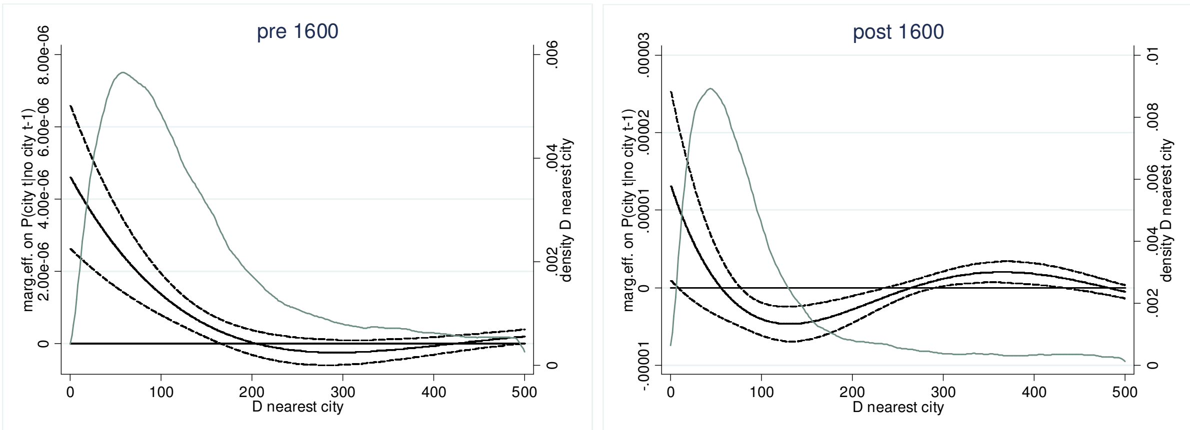
[55] Estimating a probit model gives the exact same results, however it is not very straightforward to calculate marginal effects (and corresponding confidence intervals).

[56] Again, we do not show the results on the 1st nature geography variables, they are very similar to those shown in column 1 of Table 3a.

[57] One crucial difference with our baseline distance band approach is that this 'polynomial-appraoch' only explicitly considers each location's nearest already existing city (i.e. *only one per location*), whereas the distance band approach is able to also take characteristics of the already existing urban system into account that go beyond each location's nearest already existing city. E.g. for a potential city location that is located at 15km to an already existing city, at 40km from another existing city, and at 95km to a third already existing city, all three of these already existing cities would be incorporated in the dummy-variable when using our distance band approach, whereas only the closest would be taken into account in the 'polynomial approach'. This difference can also explain why the distances at which the different effect are significant are generally larger for the

increases a location's urban chances up to a distance of about 200 kilometers (the effect is significant up unto 165 km), after which it becomes negative, but insignificant.

Figure 5. Estimated 6th order polynomial in distance to the nearest city



Notes: distances are in kilometres. The two distibutions shown depict the distribution of distance to the nearest city over all locations in our sample in the pre- and post-1600 period respectively. The 5%-, 25%-, 50%-, 75%-, and 95% quantile of these distributions are 25km, 67km, 131km, 381km and 1017km in the pre-1600, and 13.5km, 36km, 64km, 116km, and 446km in the post-1600 period respectively.

In the post-1600 period instead we find a picture that bears a striking resemblance to Figure 2. Now being located further away from an already existing city significantly increases a location's urban chances only up unto 8 km (the marginal effect remains positive up unto about 55km). This effect turns around at medium distance from an already existing city: at distances between 76 and 237km, we find a significant negative effect of being located further away from an already existing city. Finally, at furthest distance this effect disappears again (but note that only very few locations find their nearest already existing city at those distances [see the plotted distributions, and the notes to Figure 5]).

7. *Conclusions*

This paper empirically disentangles the different roles of geography in determining the location of European cities. We introduce a new data set covering over 250,000 randomly drawn potential city locations, as well as all actual European cities. This data is available during the 800 – 1800 period, when the foundations for today's European city system were laid. Using this data, we uncover the (relative) importance of physical - 1st nature – geography and of the urban system already in place (2nd nature geography) in determining locations'

---

polynomial compared to the distance band approach. The polynomial approach effectively ignores the effect of already existing cities at larger distance other than the nearest already existing city.

chances of becoming a city. To do this we use a novel, flexible way to empirically model the effect that an already established city exerts on the urban chances of its surroundings.

Our results, that hold up to a wide-variety of robustness checks, show that both 1st and 2nd nature geography played an important role in sowing the seeds of European cities, but very differently so. Most importantly, we find that their (relative) importance changes substantially during our sample period.

First nature geography is the dominant geographical force throughout the formation of the European city system. The (relative) importance of different 1st nature characteristics does not significantly change over time. A good position on the main transportation arteries is most important, with some (weak) evidence that the importance of favorable location on navigable water increased over the centuries (corresponding to the larger technological improvement and cost reductions in water- compared to land-based transportation). Also, we find some evidence that the importance of being located right next to fields of high agricultural potential diminishes in the later centuries. A development that is consistent not only with improved agricultural production techniques, but also with better (and cheaper) possibilities to transport food over larger distances.

The importance of 2nd nature geography instead changes markedly over our sample period. By virtue of our flexible modelling strategy, we moreover show that the way it matters corresponds closely to predictions from new economic geography theory. In the earlier centuries we only find evidence of an urban shadow effect, with already existing cities prohibiting the development of other cities in their immediate backyards (a range of about 20km or a one-day round trip). We do not find any positive effect of already existing cities in this period. This significantly changes in the later centuries in our sample. The strength of the urban shadow effect diminishes, and we start to find strong empirical evidence of a positive effect of being located at medium distance (20 – 100km) from an already existing city. This finding is consistent with predictions from economic geography theory. With trade costs falling, the advantages of co-locating in cities increasingly outweighing its disadvantages, and overall population increasing due to general improvements in agricultural productivity, locations at medium distance from existing cities become preferred city locations. They combine the advantage of cheaper trade with existing cities compared to locations at further distance, with that of weaker competition with existing cities compared to locations at closer distance.

Overall, our results show that geography played a crucial role in laying the foundations of the European city system as we know it today. First nature geography was a particularly dominant determinant of city location during the early formative stages of the European city system. Only from about the sixteenth century onwards, and as a result of falling trade costs, increasing net benefits of co-location and an increasing overall population, does 2nd nature geography become an important positive determinant of city location.

REFERENCES

Acemoglu, Daron, Simon Johnson and James Robinson. 2005. "The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth." *American Economic Review*, 95(3): 546-579.

Anas, Alex, Richard Arnold and Kenneth A. Small. 1998. "Urban Spatial Structure." *Journal of Economic Literature*, 26: 1426-1464.

Ashton, Thomas. 1948. *The Industrial Revolution (1760-1830)*. Oxford: Oxford University Press.

Bahn, Paul. 1996. *Archaeology: A Very Short Introduction*. Oxford: Oxford Universtiy Press.

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

Bairoch, Paul, Jean Batou and Pierre Chèvre. 1988. *La population des villes Européennes, 800-1850*. Genève: Librairie Droz.

Behrens, Kristian. 2007. "On the location and lock-in of cities: geography vs. transportation technology." *Regional Science and Urban Economics*, 37: 22-45.

Black, Duncan and Vernon Henderson. 1999. "A theory of urban growth." *Journal of Political Economy*, 107(2): 252-284.

Black, Duncan and Vernon Henderson. 2003. "Urban evolution in the USA." *Journal of Economic Geography*, 3(4): 343-372.

Bleakley, Hoyt and Jeffrey Lin. 2010. "Portage: path dependence and increasing returns in U.S. history." Working Paper, Federal Reserve Bank of Philadelphia.

Bosker, Maarten, Steven Brakman, Harry Garretsen, and Marc Schramm. 2008. "A century of shocks: the evolution of the German city size distribution 1925-1999". *Regional Science and Urban Economics*, 38(4): 330-347.

Bosker, Maarten, Steven Brakman, Harry Garretsen, Herman de Jong, and Marc Schramm. 2008, "Ports, Plagues and Politics: Explaining Italian city growth 1300-1861." *European Review of Economic History,* 12: 97-131.

Bosker, Maarten, Eltjo Buringh, Jan Luiten van Zanden. 2012. "From Bagdad to London. Unraveling urban development in Europe, the Middle East and North Africa, 800-1800." *Review of Economics and Statistics*, forthcoming.

Carro, Jesus. 2007. "Estimating dynamic panel data discrete choice models with fixed effects." *Journal of Econometrics*, 140(2): 503-528.

Christaller, Walter. 1935. *Die zentralen Orte in Suddeutschland*. Jena.

Davis, Kingsley. 1955. "The Origin and Growth of Urbanization in the World." *American Journal of Sociology*, 60: 429-437.

Davis, Donald and David Weinstein, 2002. "Bones, bombs and breakpoints: the geography of economic activity", *American Economic Review*, 92: 1269-89.

Duranton, Gilles. 1999. "Distance, land, and proximity: economic analysis and the evolution of cities." *Environment and Planning A*, 31: 2169 – 2188.

Duranton, Gilles and Henry Overman. 2005. "Testing for localization using micro-geographic data." *Review of Economic Studies*, 72: 1077-1106.

Dyer, Chris. 1995. “How urbanized was medieval England?” in J.M. Duvosquel and E. Thoen (eds.) *Peasants and Townsmen in medieval Europe*. Gent: Snoek Ducaju& Zoon, 169-183.

Ellison, Glenn and Edward E. Glaeser. 1997. “Geographic concentration in U.S. manufacturing industries: a dartboard approach,” *Journal of Political Economy,* 105(5): 889-927.

Ennen, Edith. 1972. *Die europäische Stadt des Mittelalters*. Göttingen: Vandenhoeck und Ruprecht.

Fagan, Brian M., 1997. *Archaeology: A Brief Introduction*. New York: Longman.

Findlay, Ronald and Kevin H. O'Rourke. 2007. *Power and Plenty: Trade, War, and the World Economy in the Second Millennium*. Princeton: Princeton University Press.

Fujita, Masahisa and Paul Krugman, 1995. “When is the economy monocentric: Von Thunen and Chamberlain unified.” *Regional Science and Urban Economics*, 24: 505-528.

Fujita, Masahisa, Paul Krugman and Tomoya Mori. 1999. “On the evolution of hierarchical urban systems.” *European Economic Review* 43(2): 209-51.

Fujita, Masahisa and Tomoya Mori. 1996. “The role of ports in the making of major cities: Self-agglomeration and hub-effect.” *Journal of Development Economics,* 49: 93-120.

Fujita, Masahisa and Tomoya Mori. 1997. “Structural stability and evolution of urban systems.” *Regional Science and Urban Economics*, 27(4-5): 399-442.

Galor, Oded. 2005. “From stagnation to growth: unified growth theory.” In: *Handbook of Economic Growth* eds. Phillipe Aghion and Steven N. Durlauf. Amsterdam: Elsevier.

Glick, Thomas F. 1979. *Islamic and Christian Spain in the early Middle Ages*. Princeton: Princeton University Press.

Greif, Avner. 1992. “Institutions and international trade: Lessons from the commercial revolution.” *American Economic Review - Papers and Proceedings*, 82(2): 128-133.

Greif, Avner. 2006. *Institutions and the path to the modern economy: lessons from Medieval trade*. Cambridge: Cambridge University Press.

Hajnal, J. 1965. “European marriage in perspective.” In: *Population in history* ed. David. V. Glass and David. E. C. Eversley. London: Arnold. 101-143.

Hanson, Gordon. 2005. “Market potential, increasing returns and geographic concentration.” *Journal of International Economics*, 67-1, pp.1-24.

Hausman, Jerry. 2001. “Mismeasured variables in econometric analysis: problems from the right and problems from the left.” *Journal of Economic Perspectives*, 15(4): 57-67.

Henderson, Vernon. 1974. “The sizes and types of cities.” *American Economic Review*, 64: 640-656.

Heckman, James. J. 1981. “The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process.” in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. F. Manski and D. McFadden. Cambridge, MA: MIT Press, 179–195.

Hohenberg, Paul M. 2004. “The historical geography of European cities: an interpretative essay.” In *Handbook of Regional and Urban Economics*, volume 4, eds. J.V. Henderson and J-F. Thisse, Amsterdam: North Holland. 3021-3052.

Hohenberg, Paul M., Lynn H. Lees. 1995. *The making of urban Europe: 1000-1994.* Cambridge: Harvard University Press.

Hybel, Nils. 2002. “The grain trade in northern Europe before 1350.” *Economic History Review*, 55(2): 219-427.

Kim, Sukkoo. 2000. “Urban development in the United States, 1690-1990.” *Southern Economic Journal*, 66(4): 855-880.

Konishi, Hideo. 2000. "Formation of hub cities: transportation cost advantage and population agglomeration." *Journal of Urban Economics*, 48(1): 1-28.

Krugman, Paul. 1991. "Increasing returns and economic geography." *Journal of Political Economy*, 99: 483-499.

Krugman, Paul, 1993a. "The hub effect: or, threeness in interregional trade." In: Wilfred. J. Ethier, Elhanan Helpman, J. Peter Neary (eds.). *Theory, Policy and Dynamics in International Trade*. Cambridge: Cambridge University Press: 29–37

Krugman, Paul. 1993b. "On the number and location of cities." *European Economic Review*, 37(2-3): 293-298.

Lampard, Eric. 1955. "The History of Cities in the Economically Advanced Areas." *Economic Development and Cultural Change*, 3(2): 81-136

Lopez, R.S., 1956. "The evolution of land transport in the middle ages." *Past and Present*, 9: 17-29.

Lösch, August, 1938. "The nature of Economic Regions." *Southern Economic Journal*, 5: 71-78.

Lösch, August, 1940. *Die räumliche Ordnung der Wirtschaft* (Gustav Fischer, Jena); English translation: The Economics of Location (New Haven: Yale University Press, 1954)

Malanima, Paolo. 1998a. "Italian Cities 1300-1800. A quantitative approach." *Rivista di Storia Economica*, pp. 91-126.

Marshall, Alfred. 1890. *Principles of economics*, 8th ed. Macmillan, London

McEvedy, Colin, and Richard Jones. 1979. *Atlas of world population history.* London: Allen Lane.

Ottaviano, Gianmarco and Jacques Thisse. 2005. "New Economic Geography: what about the N?" *Environment and Planning A*, 37: 1707-1725.

Pirenne, Henri. 1925. *Medieval cities, their origins and the revival of trade*. Princeton, Princeton University Press.

Ramankutty, N., J.A. Foley, J. Norman, and K. McSweeney. 2002. "The global distribution of cultivable lands: current patterns and sensitivity to possible climate change." *Global Ecology and Biogeography*, 11 (5): 377-392.

Redding, Stephen and Daniel Sturm. 2008. "The costs of remoteness: evidence from German division and reunification." *American Economic Review*, 98(5): 1766-1797.

Stewart, John Q. 1947. "Empirical Mathematical Rules concerning the Distribution and Equilibrium of Population." *Geographical Review*, 37: 461-485

Talbert, Richard J.A. ed. 2000. *Barrington atlas of the Greek and Roman world*. Princeton: Princeton University Press.

Thünen, Johann Heinrich Von. 1826. *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalokonomie*. Hamburg: Perthes.

Ullman, Eric. 1941. "A theory of location for cities." *American Journal of Sociology*, 46(6): 853-864.

Vries, Jan de. 1984. *European urbanization, 1500-1800*. London: Methuen.

Weber, Max. 1922. *The City.* Translation and edited by Don Martindale and Gertrud Neuwirth. New York: The Free Press.

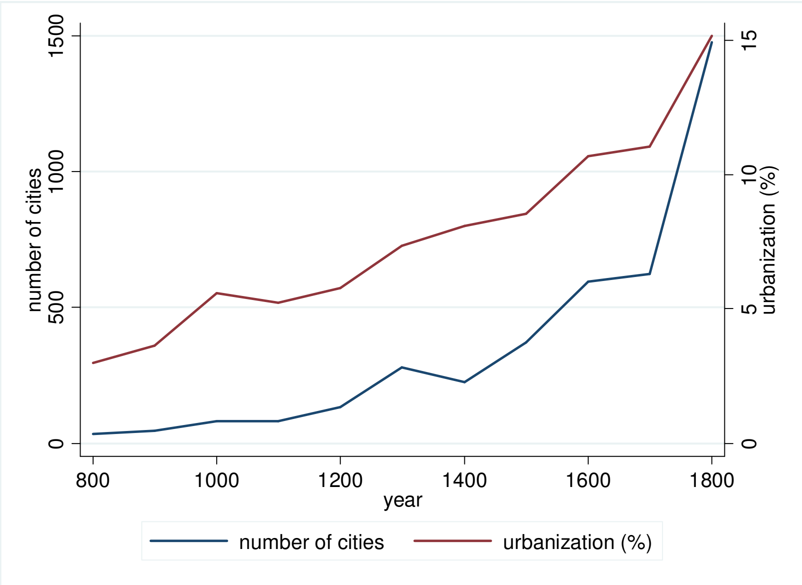
World Bank. 2009. *Reshaping Economic Geography*. Washington D.C.: World Development Report, World Bank.

Woolridge, Jeffrey. 2005. "Simple solutions for the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of Applied Econometrics*, 20: 39-54.

Zanden, Jan Luiten van and Milja van Tielhof. 2009. "Roles of growth and productivity change in Dutch shipping industry, 1500-1800." *Explorations in Economic History* 46(4): 389-403.

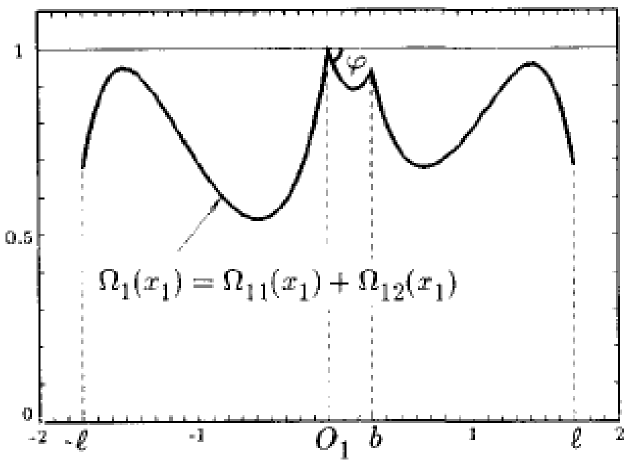
Appendix A. Additional results

Figure A1. Urbanization and the number of cities in Europe, 800 – 1800



Notes: Both the number of cities and the urbanization rate are based on defining cities as population centres with at least 5,000 inhabitants [see section 3.2 for more detail on this definition]. The urbanization rate is calculated by dividing total urban population (i.e. the total number of people living in cities with at least 5,000 inhabitants) by total population. Total population figures are taken from McEvedy and Jones (1979).

Figure A2. Market potential curves with 1st nature geography



Notes: The figure is taken from Fujita and Mori (1996, p.109). The location at *b* has a first nature geography advantage in the ease of transporting goods (i.e. it is a hub location).

Table A1. Century specific probability of being a city

year	city yes/no? no	yes	unconditional P(city t | no city t-1) (%)
800	259,750	26	
900	259,743	33	0.003
1000	259,720	56	0.01
1100	259,720	56	0.002
1200	259,682	94	0.01
1300	259,576	200	0.04
1400	259,616	160	0.02
1500	259,512	264	0.05
1600	259,353	423	0.07
1700	259,338	438	0.05
1800	258,715	1,061	0.25

Table A2. Potential city locations

country	# potential city locations (% total sample)	# becoming city (% total # potential city locations in country)
Andorra	37 (0.01)	0 (0)
Austria	6093 (2.35)	12 (0.2)
Belgium	2530 (0.97)	47 (1.86)
Czech Republic	5955 (2.29)	13 (0.22)
Denmark	3623 (1.39)	5 (0.14)
Finland	3575 (1.38)	1 (0.03)
France	40532 (15.6)	216 (0.53)
Germany	28581 (11)	161 (0.56)
Hungary	6698 (2.58)	39 (0.58)
Ireland	4937 (1.9)	6 (0.12)
Italy	19824 (7.63)	239 (1.21)
Liechtenstein	17 (0.01)	0 (0)
Luxembourg	190 (0.07)	0 (0)
Monaco	1 (0.00)	0 (0)
Netherlands	2888 (1.11)	28 (0.97)
Norway	19029 (7.33)	6 (0.03)
Poland	23964 (9.22)	40 (0.17)
Portugal	5857 (2.25)	25 (0.43)
San Marino	6 (0.00)	1 (16.67)
Slovakia	3718 (1.43)	11 (0.3)
Spain	33110 (12.75)	171 (0.52)
Sweden	24711 (9.51)	7 (0.03)
Switzerland	3139 (1.21)	11 (0.35)
United Kingdom	20761 (7.99)	111 (0.53)
total	259776	1150 (0.44)

Notes: The numbers in the third column are based on the city definition explained in section 3.2, i.e. population centres with at least 5,000 inhabitants.

Table A3. Descriptives

1st or 2nd nature characteristic	mean	sd	min	max	mean	sd	min	max
	all locations (259776)				locations ever >= 5,000 (1150)			
sea	0.006	0.08	0	1	0.09	0.28	0	1
river	0.02	0.13	0	1	0.48	0.50	0	1
hub	0.001	0.03	0	1	0.13	0.34	0	1
rroad	0.02	0.16	0	1	0.36	0.48	0	1
elevation (m)	382	429	-15	4356	232	242	-3	1218
ruggedness	90.4	109.5	0	922.8	79.0	86.6	0	720.8
P(cultivation)	0.53	0.34	0.001	0.999	0.71	0.23	0.006	0.999
latitude	50.0	6.8	36.0	63.5	46.4	5.5	36.7	63.4
longitude	7.4	8.6	-9.4	23.3	6.2	7.7	-9.3	22.8
D near. city >= 10k	233	295	0	1701	98	117	0	1424
--------------		nr. cities >= 10k				nr. cities >= 10k		
0 – 20km	0.05	0.23	0	1	0.10	0.30	0	1
20 – 50km	0.20	0.40	0	1	0.36	0.48	0	1
50 – 100km	0.42	0.49	0	1	0.62	0.49	0	1

Table A4. A finer century decomposition

P(city t | no city t-1)			baseline		
years	sea	river	hub	road	elevation
900-1000	0.005***	0.012***	0.007***	0.006***	-0.0001
	[0.00]	[0.00]	[0.00]	[0.00]	[0.31]
1100-1200	0.002**	0.016***	0.004***	0.007***	-0.0002*
	[0.05]	[0.00]	[0.00]	[0.00]	[0.1]
1300-1400	0.006***	0.018***	0.004***	0.004***	-0.0001*
	[0.00]	[0.00]	[0.00]	[0.00]	[0.07]
1500-1600	0.009***	0.013***	0.007***	0.004***	0.0001***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
1700-1800	0.007***	0.014***	0.009***	0.004***	-0.0001***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
				city >= 10k? (t-1)	
	ruggedness	P(cultivation)	0 – 20 km	20 – 50 km	0 – 100 km
900-1000	-0.0001	0.002*	-	-0.0002	-0.0004
	[0.24]	[0.08]	-	[0.67]	[0.12]
1100-1200	0.0001	0.001***	-0.0006	0.0001	0.0002
	[0.39]	[0.01]	[0.21]	[0.85]	[0.56]
1300-1400	0.0001	0.0006**	-0.0005***	-0.0001	-0.0001
	[0.1]	[0.03]	[0.01]	[0.3]	[0.53]
1500-1600	-0.0001	0.0003	-0.0003**	0.0002*	0.0002*
	[0.15]	[0.13]	[0.03]	[0.05]	[0.06]
1700-1800	0.0002***	0.0005***	-0.00001	0.0001*	0.0002***
	[0.00]	[0.00]	[0.87]	[0.09]	[0.01]
country/century FE			yes		
nr observations			1834732		

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. The p-values are however based on the estimated coefficients and their standard errors. The 0 – 20km version of the ‘already existing city’-dummy variable perfectly predicts failure during the earliest centuries, so that we do not have any variation to pinpoint its effect (i.e. all locations located within 0-20km of two or more already existing cities never become a city themselves).

A.1 Robustness to choice of estimation technique

The additional robustness checks reported in Table A5 below all concern the estimation technique that we use to obtain our baseline results in Table 1.

Table A5. Robustness to estimation technique

	1	2	3	4	5
P(city t | no city t-1)	logit	LP	duration model (Cox)	FE LP	P(city t)
sea	0.006***	0.006***	8.10***	-	0.023***
	[0.00]	[0.00]	[0.00]	-	[0.00]
river	0.012***	0.012***	28.8***	-	0.041***
	[0.00]	[0.00]	[0.00]	-	[0.00]
hub	0.004***	0.069***	5.95***	-	0.032***
	[0.00]	[0.00]	[0.00]	-	[0.00]
road	0.003***	0.004***	6.80***	-	0.016***
	[0.00]	[0.00]	[0.00]	-	[0.00]
ln elevation	-0.0001**	-0.00004*	0.91**	-	-0.0006***
	[0.03]	[0.07]	[0.01]	-	[0.00]
ruggedness	0.0001***	0.0001***	1.13***	-	0.0005***
	[0.00]	[0.00]	[0.00]	-	[0.00]
P(cultivation)	0.001***	0.0002***	2.04***	-	0.003***
	[0.00]	[0.00]	[0.00]	-	[0.00]
city >= 10k? (t-1)					
0 – 20 km	-0.0002***	-0.0004***	0.75***	0.0001	-0.001***
	[0.00]	[0.00]	[0.01]	[0.62]	[0.00]
20 – 50 km	0.0001	0.0001**	1.14*	0.0002***	0.001***
	[0.25]	[0.02]	[0.06]	[0.01]	[0.00]
50 - 100 km	0.0001**	0.0001***	1.15*	0.0002***	0.001**
	[0.05]	[0.00]	[0.07]	[0.01]	[0.01]
P(city) unconditional	0.05	0.05	0.05	0.05	0.05
country/century FE	yes	yes	yes	yes	yes
nr observations	1840091	2588903	2588507	2588903	2062221
% of-the-chart predictions		67.8		19.1	

Notes: p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects in columns 1 and 5. The p-values are however based on the estimated coefficients and their standard errors. Column 3 reports hazard rates instead of estimated coefficients, i.e. a hazard rate larger (smaller) than 1 indicates that the corresponding characteristic increases (decreases) the probability to become a city.

Instead of assuming F to be the standard normal CDF, column 1 and 2 show the results when taking the logistic distribution function or the identity function instead, and estimating (2) using logit or OLS techniques respectively. All our main baseline results do not crucially depend on the assumption made on F[58].

Column 3 shows that our baseline results also come through when we completely change our modelling strategy and adopt a duration model (of the time until becoming a city) instead of the transition model (becoming a city conditional upon not being a city before) that

[58] In column 1 the already existing city dummy at (20-50km) is not significant, but we note that we do not reject the joint significance of the 20-50km and 50-100km dummy variable at the 10% level (p-value 0.078).

we employ throughout the paper. Using a duration model can be argued to take better account of any duration dependence in the probability of becoming a city (i.e. this probability may not be the same depending on the time a location has already not become a city). Although the inclusion of country-century fixed effects in all our baseline specifications can be argued to already go a long way in controlling for duration dependence [in duration terms: they allow the baseline hazard to (arbitrarily) change over the centuries in a moreover possibly different way across countries], it is reassuring that we basically find the same results in Column 3. Column 3 reports hazard ratios. A hazard ratio significantly larger than 1 indicates that the corresponding characteristic increases a location's baseline hazard to become a city. Similarly, a hazard ratio significantly smaller than 1 indicates that the corresponding characteristic decreases a location's baseline hazard to become a city. Column 3 e.g. shows that a location on a navigable river is 29 times more likely to become a city than otherwise similar locations not on a river, and that locations within 0-20km of an already existing city are 25% less likely to become a city than (otherwise similar) locations outside the immediate urban shadow of an already existing city).

Columns 4 again shows results using a linear probability model, but we now, in addition to controlling for unobserved *country-century* specific heterogeneity, also control for unobserved *time-invariant location-specific* heterogeneity, i.e. we include location specific fixed effects. As such, this column is readily comparable to column 2 in Table 1 that employs a CRE-probit estimation strategy. However, the linear probability model does not take account of the fact, as both probit does, that the dependent variable is restricted to the [0,1] interval. It can result in severe off-the-chart predictions (see the bottom of column 2 and 4). Still, column 4 shows that our baseline results, except for the negative effect of being located too close to an already existing city, come through[59].

Finally, column 5 shows that we find very similar results when instead of looking at the probability of *becoming* a city, $P(city_{it} = 1 \mid city_{it-1} = 0)$, we consider the probability of *being* a city, $P(city_{it} = 1)$, unconditional on its status in the previous century.

A.2 Changing the city definition

It this section we assess the sensitivity of our results with respect to our city definition based on an *absolute* population cutoff of having at least 5,000 inhabitants. Table A6 shows the results when using a different absolute cutoff, or a time-varying population cutoff instead.

[59] This negative effect does show up in the earlier centuries when allowing the effect of each variable to differ between the pre- and post-1600 period (similar to Table 3a).

Table A6. Sensitivity to the choice of city definition

P(city t | no city t-1)	(1) >= 3000		(2) >= 4000		(3) >= 6000		(4) >= 10000		(5) step-wise	
period:	< 1600	>= 1600	< 1600	>= 1600	< 1600	>= 1600	< 1600	>= 1600	< 1600	>= 1600
sea	0.007***	0.008***	0.006***	0.008***	0.010***	0.020***	0.002***	0.004***	0.003***	0.007***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
river	0.017***	0.013***	0.015***	0.014***	0.027***	0.025***	0.006***	0.006***	0.010***	0.008***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
hub	0.006***	0.009***	0.006***	0.008***	0.012***	0.016***	0.002***	0.001***	0.004***	0.003***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
road	0.004***	0.004***	0.004***	0.004***	0.012***	0.010***	0.002***	0.002***	0.002***	0.002***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
ln elevation	-0.00003	-0.0001	-0.00004	-0.0001*	-0.0004**	-0.0002	-0.00004	-0.0001***	-0.0001**	-0.00002
	[0.42]	[0.10]	[0.29]	[0.07]	[0.04]	[0.16]	[0.13]	[0.00]	[0.03]	[0.55]
ruggedness	0.00004	0.0001***	0.00001	0.0002***	0.00004	0.0005***	-0.00001	0.0001***	0.00002	0.0001
	[0.38]	[0.00]	[0.77]	[0.00]	[0.79]	[0.00]	[0.83]	[0.01]	[0.43]	[0.1]
P(cultivation)	0.0006***	0.0004***	0.001***	0.0004***	0.003***	0.001***	0.0004***	0.0001*	0.0004***	0.0001
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.05]	[0.00]	[0.24]
city >= 10k? (t-1)										
0 – 20 km	-0.001***	-0.000004	-0.001***	-0.00002	-0.002***	-0.0001	-0.0002**	0.00001	-0.0004***	-0.00003
	[0.00]	[0.97]	[0.00]	[0.82]	[0.00]	[0.87]	[0.02]	[0.82]	[0.00]	[0.67]
20 – 50 km	-0.0001	0.0002***	-0.0001	0.0002***	-0.0005	0.0004	-0.00004	0.00003	-0.0001	0.0001
	[0.37]	[0.00]	[0.13]	[0.00]	[0.2]	[0.13]	[0.47]	[0.55]	[0.18]	[0.13]
50 - 100 km	-0.0001	0.0003***	-0.0001	0.0003***	-0.0001	0.001***	-0.0001	0.0001**	-0.0001	0.0002***
	[0.2]	[0.00]	[0.22]	[0.00]	[0.89]	[0.01]	[0.17]	[0.01]	[0.14]	[0.00]
P(city) unconditional	0.0003	0.0012	0.0002	0.0012	0.0002	0.0009	0.0001	0.0004	0.0002	0.0004
country/century FE	yes		yes		yes		yes		yes	
nr observations	1903819		1875244		1774380		1581112		1766214	
ln pseudo likelihood	-7753.4		-7452.6		-5610.8		-2831.9		-4031.1	
p-value H0: pre 1600 = post 1600										
sea	[0.34]		[0.16]		[0.00]***		[0.02]**		[0.01]**	
river	[0.07]*		[0.25]		[0.31]		[0.99]		[0.08]*	
hub	[0.14]		[0.37]		[0.20]		[0.35]		[0.20]	
road	[0.32]		[0.97]		[0.18]		[0.74]		[0.50]	
ln elevation	[0.70]		[0.81]		[0.38]		[0.55]		[0.32]	
ruggedness	[0.04]**		[0.00]***		[0.06]*		[0.05]**		[0.47]	
P(cultivation)	[0.21]		[0.34]		[0.11]		[0.07]*		[0.03]**	
city >= 10k? (t-1)										
0 – 20 km	[0.00]***		[0.00]***		[0.00]***		[0.02]**		[0.00]***	
20 – 50 km	[0.01]**		[0.00]***		[0.06]*		[0.35]		[0.04]**	
50 - 100 km	[0.00]***		[0.00]***		[0.06]*		[0.01]***		[0.00]***	

Notes: the last column shows the results when employing a step-wise city definition, i.e. from 800 – 1500 the size criterion is >= 5,000 inhabitants, from 1600 – 1700 it is >= 6,000, and in 1800 it is >= 10,000. p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. The p-values are however based on the estimated coefficients and their standard errors. Instead of the estimated coefficients in (2), the table reports average partial effects. Whenever the effect of a variables is split in a pre- and post-1600 effect, the average partial effect is calculated using only the observation in the pre- or post-1600 period only.

In columns 1 and 2, we lower our absolute population criterion to 3,000 and 4,000 inhabitants respectively. Bairoch et al. (1988) only provide population numbers smaller than 5,000 inhabitants for a very limited set of city locations, stressing (see their p.218) that on the on hand these numbers are subject to a much greater margin of error than those larger or equal than 5,000 inhabitants, and, on the other hand, that they did not systematically search for any numbers smaller than 5,000, so that these numbers are only very selectively available. Nevertheless, using these much less reliable population numbers, we find the same results as when using our preferred 5,000 inhabitants cutoff to define a city.

When instead raising the population cutoff, all but one of our main results come through. When increasing our population cutoff to 6,000 inhabitants, we find a slight change to our 2nd nature geography results that is further exacerbated when increasing the population cutoff to 10,000 inhabitants. In particular, we find that the positive effect of having an already existing city at 20 – 50km during the later centuries disappears when raising the criterion to 6,000 inhabitants, but it is still significantly different from its effect in the pre-1600 period. When raising the cutoff to 10,000 inhabitants, this effect is completely lost and it is also no longer significantly different from its pre-1600 effect[60].

However, this result does not necessarily invalidate our baseline results. In combination with our baseline findings, the results in columns 1 – 4 show a consistent pattern: the positive effect of an already existing city at medium distances gradually disappears when raising the absolute size criterion used to define a city. Having an existing city at 20-50km may significantly improve a location's probability of becoming a city of 5,000 or 6,000 inhabitants, it becomes increasingly difficult to grow larger, say 10,000 inhabitants, in the shadow of an already existing urban centre. An existing city, as it were, does only tolerate moderately sized new cities to appear in its immediate backyard.

Finally, column 5 shows results when using a time-varying population cutoff to define a city. We employ the following step-wise increasing population cutoff: 5,000 inhabitants before 1600, 6,000 in 1600 and 1700, and 10,000 in 1800. We choose this particular stepwise increase as it leaves the unconditional probability of becoming a city in any century around 0.04% in the period 1500 – 1800 (instead of increasing substantially over this period when using our absolute 5,000 inhabitants cutoff). Using such a time-varying definition is in itself not without difficulties. In particular, given that we condition on not already being a city in *t-1*, one has to choose which definition to use when constructing these variables (i.e. the 'new' definition in period *t* or the 'old' definition in period *t-1*).

In column 5 we use the city definition in period *t-1* as our conditioning variable (i.e. was there a city in period *t-1*). The results show that our main findings are again generally robust to using this time-varying city-definition. Again, only the effect of an already existing

[60] Note that a test for the joint significane of the 20-50km and 50-100km is always significant in the post-1600 period (corresponding p-values: 0.01 (column 3); 0.04 (column 4); 0.004 (column 5).

city within 20 - 50 km in the post-1600 period is sensitive to using this time-varying definition[61].

A.3 More than three distance bands

Table A7. Twenty-five 20km distance bands

P(city t | no city t-1)	pre 1600	post 1600	
	1st nature geography as in Table 3a		p-value < 1600 == >= 1600?
city >= 10k? (t-1)	APE [p-value]	APE [p-value]	
0 - 20 km (1 day)	-0.002*** [0.00]	-0.0003 [0.31]	[0.00]***
20 - 40 km (2 days)	-0.0004 [0.22]	0.0002 [0.23]	[0.10]*
40 - 60 km (3 days)	-0.0002 [0.62]	0.001*** [0.00]	[0.03]**
60 - 80 km (4 days)	0.0003 [0.34]	0.001** [0.01]	[0.65]
80 - 100 km (5 days)	-0.0004 [0.12]	0.0004* [0.05]	[0.02]**
100 - 120 km (6 days)	0.0003 [0.28]	0.0004** [0.02]	[0.76]
120 - 140 km (7 days)	-0.001** [0.03]	-0.0002 [0.20]	[0.21]
140 - 160 km (8 days)	-0.00002 [0.94]	-0.0004** [0.01]	[0.24]
160 - 180 km (9 days)	-0.0003 [0.36]	-0.0001 [0.77]	[0.51]
180 - 200 km (10 days)	0.0001 [0.78]	0.0003 [0.25]	[0.62]
200 - 220 km (11 days)	0.0003 [0.3]	-0.0002 [0.38]	[0.17]
220 - 240 km (12 days)	-0.00002 [0.95]	-0.0001 [0.77]	[0.91]
240 - 260 km (13 days)	0.0001 [0.78]	-0.0002 [0.38]	[0.47]
260 - 280 km (14 days)	-0.0004 [0.16]	0.0001 [0.71]	[0.17]
280 - 300 km (15 days)	0.0001 [0.84]	-0.00002 [0.94]	[0.83]
300 - 320 km (16 days)	-0.0003 [0.35]	-0.0003 [0.17]	[0.95]
320 - 340 km (17 days)	0.0001 [0.68]	-0.00003 [0.87]	[0.67]
340 - 360 km (18 days)	-0.00001 [0.98]	-0.0002 [0.29]	[0.54]
360 - 380 km (19 days)	-0.001** [0.03]	-0.001*** [0.00]	[0.88]
380 - 400 km (20 days)	0.0001 [0.67]	0.0001 [0.55]	[0.97]
400 - 420 km (21 days)	-0.0004 [0.22]	-0.0003 [0.14]	[0.96]
420 - 440 km (22 days)	-0.0002 [0.47]	0.00003 [0.91]	[0.53]
440 - 460 km (23 days)	-0.0004 [0.18]	-0.0002 [0.36]	[0.66]
460 - 480 km (24 days)	0.0004 [0.19]	-0.0001 [0.78]	[0.23]
480 - 500 km (25 days)	-0.0003 [0.23]	0.0001 [0.73]	[0.26]
P(city) unconditional	0.0002	0.0012	
country/century FE	yes		
nr observations	1840091		

Notes: behind each distance band, we denote, in brackets, the number of days needed to complete a round-trip from a potential city location to a city at this distance (20 kilometers roughly corresponds to a one day round-trip during most of our sample period (see also footnote 29). p-values, based on robust standard errors, between square brackets. *, **, *** denotes significance at the 10%, 5%, 1% respectively. Instead of the estimated coefficients in (2), the table reports average partial effects. The p-values are however based on the estimated coefficients and their standard errors. All regressions contain the same 1st nature geography variables as in column 1 of Table 3a. The estimated parameters on these variable correspond closely to those reported in column 1 of Table 3a. They are available upon request.

[61] Results are also robust to using the city definition in period *t* instead. Also, using a different 'step-wise' city definition (i.e. 5,000 before 1800 and 10,000 in 1800) all our baseline results come through. These results are available upon request.

Table A7 shows results when using twenty-five 20km distance bands instead of the three that we use throughout most of the paper. Doing this, we basically arrive at the exact same conclusions as when using our more parsimonious 'three-distance-bands-based approach'. In the earlier centuries, we again only find evidence for a significant competition effect at close range (0-20km), and no evidence for a significant positive effect of being located at medium distance from an already existing city. This positive effect only shows up significantly in later centuries. Moreover, the spatial reach at which we find this positive effect is very similar to that in our baseline findings (it does not extend beyond 120km).

Appendix A.4 A final robustness check: 2nd nature geography results by construction?

Given the steady increase in the number of cities over the centuries, one may be worried that especially our 2nd nature geography results could be obtained by construction. Europe's urban system becomes denser over the centuries. Is it simply this increased density that drives our finding of an increased importance of 2nd nature geography over the centuries?[62].

To assess this possibility we adopt the following *Dartboard Approach* in the spirit of Duranton and Overman (2005) and Ellison and Glaeser (1997) as a final robustness check. Using a simulation approach, we verify whether we would obtain the same results regarding our 2nd nature geography variables when cities appeared randomly at one of our potential locations instead of at the locations where they appeared in reality. If we do, this means we could be getting our results by construction, shedding doubts on our findings. This *Dartboard Approach* is operationalized as follows:

---

1. In each century t, randomly allocate n_t cities, the number of new cities actually appearing in century t, over the k_t available potential city locations in that century. We do this *conditional on each potential city location's 1st nature geography characteristics, i.e.:*

$$n_t \sim Binomial(k_t, p_t(X_i)), \text{ , where } p_t = \Phi(X_i b + a_{ct}) \qquad (4)$$

and b an a_{ct} are the estimated parameters on the 1st nature geography variables and the estimated country-century fixed effects respectively, obtained by estimating (2) including only these 1st nature geography variables and country-century dummies as explanatory variables.

---

[62] Note that this issue is different, yet related, to the possibility of dynamic selection bias. However, where dynamic selection bias concerns the dependent variable, the concern that we address here is that the increasing number of cities over time affects our 2nd nature geography regressors with possibly unwanted consequences for our results.

2. Using this hypothetical city configuration, we estimate our baseline model as in (2) and store the estimated parameters on each of the three 2nd nature geography variables.

3. Repeat the above-outlined procedure 1,000 times to obtain the empirical distribution of all three estimated 2nd nature geography coefficients. Next, for each respective 2nd nature geography variable we calculate the percentage of simulation runs that we *falsely not reject* that it is significant at the 1%, 5%, and 10% respectively when using the standard z-tests. Given that we allocate new cities randomly in each century, this percentage should be close to 1%, 5%, and 10% respectively to conclude that the standard tests perform well.

Table A8 shows the results of doing these simulations for our main specification allowing the effect of all variables to differ in the pre- and post-1600 period.

Table A8. Dartboard Approach – simulation results

	% falsely not rejected				% falsely not rejected		
800 - 1500	at 1%	at 5%	at 10%	**1600 - 1800**	at 1%	at 5%	at 10%
city >= 10k? (t-1)				city >= 10k? (t-1)			
0 - 20 km	1.2%	5.3%	10.1%	0 - 20 km	0.4%	5.5%	10.0%
20 - 50 km	1.4%	5.9%	10.6%	20 - 50 km	0.6%	5.5%	10.0%
50 - 100 km	1.0%	4.9%	10.9%	50 - 100 km	1.2%	4.2%	9.6%

Notes: All the '% falsely not rejected' are based on 1,000 simulation runs. The pre- and post-1600 results are based on 'dart-throws' conditional upon 1st nature geography, where 1st nature geography is also allowed to have a possibly different pre- and post-1600 effect.

The percentage of simulation runs in which we falsely reject the null hypothesis is always quite close to 1%, 5% or 10% respectively, providing strong confidence that we do not obtain results by construction as a consequence of the increased density of the urban system over the centuries. True, because of the denser urban system the number of locations becoming a city located closer to an already existing city increases over the centuries; but, so does the number of locations *not* becoming a city located closer to an already existing city. The estimated parameters on our, 2nd nature geography, dummy variables depend on the trade-off between these two. Our simulation results show that their significance in the pre-1600 period is not simply an artefact of increased density of the urban system in the later centuries.