Working Paper in Economics No. 722

Roman Roads to Prosperity: Persistence and Non-Persistence of Public Goods Provision

Carl-Johan Dalgaard, Nicolai Kaarsen, Ola Olsson
Pablo Selaya

Department of Economics, February 2018

ISSN 1403-2473 (Print)
ISSN 1403-2465 (Online)

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Roman Roads to Prosperity: Persistence and Non-Persistence of Public Goods Provision*

Carl-Johan Dalgaard†
Nicolai Kaarsen‡
Ola Olsson§
Pablo Selaya¶

February 16, 2018

Abstract

How persistent is public goods provision in a comparative perspective? We explore the link between infrastructure investments made during antiquity and the presence of infrastructure today, as well as the link between early infrastructure and economic activity both in the past and in the present, across the entire area under dominion of the Roman Empire at the zenith of its geographical extension (117 CE). We find a remarkable pattern of persistence showing that greater Roman road density goes along with (a) greater modern road density, (b) greater settlement formation in 500 CE, and (c) greater economic activity in 2010. Interestingly, however, the degree of persistence in road density and the link between early road density and contemporary economic development is weakened to the point of insignificance in areas where the use of wheeled vehicles was abandoned from the first millennium CE until the late modern period. Taken at face value, our results suggest that infrastructure may be one important channel through which persistence in comparative development comes about.

Keywords: Roman roads, Roman Empire, public goods, infrastructure, persistence.
JEL classification codes: H41, O40.

*We have received useful comments on various versions of the present study from seminar participants in Belfast, Bergen, Copenhagen, Gothenburg, London School of Economics, Manchester, Tokyo, Sheffield and the workshop "Deep-Rooted Factors in Comparative Development" at Brown University. The usual disclaimer applies.

†University of Copenhagen, CAGE (Warwick) and CEPR (London). carl.johan.dalgaard@econ.ku.dk
‡Danish Economic Council. nik@dors.dk
§University of Gothenburg. ola.olsson@economics.gu.se
¶University of Copenhagen. pablo.selaya@econ.ku.dk

1 Introduction

One of the most remarkable facts in the literature on comparative development is just how persistent relative development differences are over time. On average, societies that were relatively economically developed during the pre-industrial era tend to be comparatively successful today.[1] As an illustration of such persistence, Figure 1 depicts the conditional relationship between the intensity of Roman settlements in 500 CE and contemporary population density, within the area under Roman control ca. 117 CE, after controlling for country fixed effects. This insight has led researchers to search for the origins of comparative development in (geographic) initial conditions, or in historical processes that have shaped cultural traits and the institutional infrastructure of individuals societies.[2]

Meanwhile much less attention has been devoted to the study of persistence in *proximate* determinants of growth, factors that could be the channel or the transmission mechanism connecting fundamentals to economic development in the past and in the present. Yet a deeper understanding of the channels through which persistence in comparative development comes about may leave important clues as to which fundamentals are important, and how to potentially stimulate development in situations where important fundamentals are lacking.

Figure 1

The present study explores the persistence of physical infrastructure across time and space, starting in antiquity with the establishment of the Roman road network. As discussed below, Roman road construction did not follow the rules of infrastructure planning in the contemporary era: The roads were build chiefly with a military purpose in mind, and geographic obstacles in the landscape were usually surmounted rather than evaded. Despite this, our analysis uncovers a remarkable degree of persistence in road density across time and space: areas that attained greater road density during antiquity are characterized by a significantly higher road density today. Moreover, the Roman roads were strongly linked to economic activity by the end of

[1]See, for example, Olsson and Hibbs (2005); Comin et al. (2010); Chanda et al. (2015); Maloney and Valencia (2016).

[2]See Spolaore and Wacziarg (2013); Nunn (2014) and Ashraf and Galor (forthcoming) for recent reviews.

antiquity, and they remain a strong positive correlate with prosperity today. Overall, our analysis provides evidence that public goods provision is an important channel through which persistence in economic development, as depicted in Figure 1, may arise.

In our analysis we confine attention to localities, grid cells measuring one degree latitude by one degree longitude, that were part of the Roman Empire by the second century CE and were treated by at least one Roman road. By omitting areas that fell outside the Empire, as well as those completely unconnected to the network within the Empire, we hope to disentagle the influence of the physical infrastructure on economic outcomes from the legacy of Roman rule more broadly.[3] To further limit the risk of confounding the impact of the Roman roads with Roman influence more broadly, and to filter out the impact of modern day institutions and (sub-national) differences in cultural values on economic outcomes of interest, we control for country fixed effects as well as language fixed effects throughout our entire econometric analysis. We believe this strategy makes it unlikely that our results are driven by the legacy of Roman rule in the broadest terms. However, a natural concern is that areas receiving more Roman roads may differ in various geographic dimensions that by themselves may have influenced comparative economic development. We attempt to surmount this challenge in several different ways.

For starters it is important to observe that the risk of confounding the influence of roads with geography may not be as great as one might think, as alluded to above and elaborated upon below. According to the historical literature, it is conventional wisdom that major Roman roads were built to facilitate the movement of troops across the empire, rather than with the objective of enhancing economic development. Moreover, the roads were arguably only to a limited extent dictated by geographic circumstances. In our analysis we examine these arguments statistically. The role of many geographic characteristics does in fact seem limited. In the end, however, we find a claim of orthogonality between geography and Roman road location to be untenable, for which reason we control for a rich set of geographical characteristics when exploring the persistence of early road infrastructure over time, and its predictive power vis-a-vis eco-

[3]See Landes (1998) on the legacy of Roman rule. More recent research has shown that areas under Roman influence arguably developed different institutions from areas outside Roman direct influence (e.g. Glaeser and Shleifer, 2002).

nomic development.

Still, doubts may linger whether our control strategy is fully sufficient. Therefore, our second strategy to assess the importance of potentially omitted geographic charactaristics consists in exploiting the remarkable timespan of abandonment of wheeled transportation in North Africa and the Middle East. According to the landmark study by William Bulliet (1990 [1975]), wheeled transport disappeared in North Africa and the Middle East somewhere between the fourth and sixth century CE. Eventually, wheeled transport vehicles had to be reintroduced with the ascent of the automobile.[4] Consequently, following the fall of the Western part of the Roman Empire, the roads fell into disrepair, and ultimately went out of use in North Africa and the Middle East. In contrast, Roman roads continued to be maintained and in use in Europe after the fall of the Western Roman Empire (Glick, 2005 [1979]; Hitchener, 2012).

From the point of view of the present study, this natural experiment has two important implications. First, as the ancient roads fall into disuse and thus are left unmaintained, it becomes much less likely that modern roads are built in their place. Consequently, one would expect to see a far weaker link between Roman road density and modern road density outside the European part of the Empire. Second, if the roads create persistence in economic development, one would expect to find a far weaker link between ancient infrastructure and modern-day economic activity within the regions where roads temporarily lost relevance. Consistent with these conjectures, we find that there is no significant link between ancient infrastructure and modern infrastructure within North Africa and the Middle East. Moreover, within these two regions the ancient infrastructure is not a significant predictor of economic activity today. In contrast, in Europe – the region where roads continued to be used and therefore maintained – ancient roads predict modern roads as well as prosperity.

This differentiated effect of Roman road density is revealing from the point of view of identifying channels of influence. Naturally, the fundamental principles governing the construction of the Roman roads were the same throughout the Empire. If our baseline results are tainted by omitted variable bias, for example due to missing geographic characteristics that matter both for road location and subsequent economic activity, one would expect to see evidence of an apparent persistence of infrastructure

[4] See also Chaves et al. (2014) on the absence of wheeled transport in Sub-Saharan Africa and its introduction during the early colonial period.

density as well as a persistent impact from ancient infrastructure on modern economic activity *throughout* the Empire. Accordingly, in light of our findings, the "abandonment of the wheel"-experiment provides fairly compelling evidence of the mechanism under scuntiny: persistence in public goods provision leads to persistence in economic activity.

The present paper is related to several strands of literature. First, it is related to the literature on long-run persistence in economic development summarized in Spolaore and Wacziarg (2013); Nunn (2014) and Ashraf and Galor (forthcoming). This literature has largely been concerned with the influence and origins of fundamental determinants of productivity that ultimately can explain persistence in comparative development.[5] In contrast, the present paper focuses on channels of persistence, or the persistence of proximate sources of growth. In particular, the focus is on whether public goods provision appears to be persistent across long periods of time. From this perspective our work is related to Chen, Kung and Ma (2017), who document within China a remarkable persistence of another proximate source of growth: education. The authors argue persuasively that the observed persistence is (in part, at least) explained by the emergence of a pro-education culture, prompted by early educational investments. The argument is importantly supported by the fact that persistence in schooling weakens markedly in areas particularly exposed to the anti-intellectual Cultural Revolution. Our study documents that the provision of public goods – road infrastructure – is similarly persistent over very long periods of time, except in areas where roads lose economic value early on. By carefully controlling for the influence of (countrywide) formal institutions and (within-country) informal institutions, our findings indicate that this persistence is unlikely to be mediated by the emergence of cultural values or institutions. Instead, our analysis suggests that later roads likely were built on top of older roads, thus creating persistence in road density.[6] Of course, in places where the signs of the early roads disappear, modern roads would be less likely to be located along the ancient trajectories.

Second, our paper also contributes to a small literature on the economics of the Ro-

[5]Recent contributions include Galor and Özak (2016), Gorodnichenko and Roland (2017) on culture and development; Angelucci et al (2017) on institutions and Andersen et et al. (2016) on geography.

[6]Our finding of a persistency in road investments over time is also consistent with the presence of local *increasing returns* to scale in infrastructure investments, as emphasized in a literature inspired by Krugman (1991).

man period (Finley, 1973; Temin, 2006; Bowman and Wilson, 2009; Michaels and Rauch, 2016). A striking feature of the classic work by Finley (1973), for instance, is its relative neglect of the general importance of roads. A common theme is that road transport was inferior to shipping in terms of efficiency and hence of less importance. Recent research however has started to re-examine the influence of the Roman road network on long-run develeopment. Bosker et al. (2013), studying determinants of city growth between 800 CE and 1800, document that cities located at intersection points between Roman roads grew bigger in Europe, but not in North Africa and the Middle East.[7] The present study focuses on the persistence of infrastructure over time, which is not discussed in Bosker et al. Moreover, we study the impact of ancient infrastructure on 21st century economic activity rather than urban population size at the eve of the late Modern period. Also related is the work of Wahl (2017), who investigates the Roman *limes* inside Germany and finds that contemporary development is more advanced on the old Roman side of the border using a regression discontinuity design. Wahl identifies the road system as an important explanatory factor, which is consistent with our findings that pertain to the European part of the Empire more generally. At the same time, our analysis provides evidence on how shocks in the past, such a the abandonment of the wheel, can importantly perturb development trajectories with long-run implications for comparative development.

Third, in recent years, a surge of interest in the economic effects of infrastructure has resulted in a number of important studies. The results indicate that infrastructure investments often have a strong positive influence on population growth and economic activity.[8] A major difference with the present study obviously consists of the time horizon over which the consequences of infrastructure investments are assessed; with a two millennia perspective we believe the present study has the longest observation window hitherto explored. At the same time it is important to stress that the objective of the present study is not to estimate the productivity gains from infrastructure per se. In fact, our approach does not allow us to distinguish whether areas that received more infrastructure investments outgrew areas with less infrastructure because of productivity benefits from public goods, or if more public goods in a particular location

[7]On the link between city location and Roman roads within Europe, see also Bosker and Buringh (2017).

[8]On roads see Fernald (1999), Michaels (2008), and Bird and Straub (2015). On railways see Banerjee et al (2012), Donaldson (2012), Jedwab and Moradi (2015), and Hornung (2015).

simply attracted activity from other locations. As should be clear, the main focus of the present study is rather to assess the persistence of public goods provision and thereby whether public goods provision has a significant role to play in the observed pattern of persistence of comparative development.

The paper proceeds as follows. In section 2, we present our data on Roman roads, and our outcome variables. In section 3, we outline an historical background on the assignment of Roman roads as well as formal tests of the geographical determinants of Roman road density. In section 4, we present and discuss the main empirical results. Section 5 concludes.

2 Data

In this section we describe the central independent variable in the regressions to follow, the Roman roads variable, as well as the main dependent variables that appear below. The appendix contains a description of remaining (control) variables, as well as summary statistics.

2.1 Independent variable: Roman Roads

The raw data for the Roman roads come from a digitized map version of the road network illustrated and documented in Talbert's (2000) "Barrington Atlas of the Greek and Roman World".[9] We focus on all roads identified in the digitized map as being of major importance, and drop all roads identified as minor because of the difficulties involved in getting precise traces of small ancient roads.[10] In terms of timing and geographic coverage, we concentrate on roads within the borders of the Roman Empire in the year 117 CE, which is around the time when the empire attained its maximum

[9] The digitization was done by McCormick et al (2013), as part of the Digital Atlas of Roman and Medieval Civilization (DARMC) project of the Center for Geographic Analysis at Harvard University, http://darmc.harvard.edu.

[10] The Barrington Atlas notices, in fact, that the roads system is perhaps "the most difficult element to map". Page 262 in the map-by-map directory that accompanies Talbert's (2000) Barrington Atlas describes, for example, the difficulty of mapping minor roads that had neither milestones nor paving in the outline of Augustan roads or in routes through the Alpine valleys (Map 18). Page 169 describes, as another example, the use of tree-ring dating methods to overcome the difficulties in tracing parts of the Via Claudia Augusta (Map 12). See http://press.princeton.edu/B_ATLAS/B_ATLAS.PDF for a detailed description of all roads and other features contained in the Barrington Atlas.

territorial expansion.[11]

To construct a measure of the influence that Roman roads have had, we draw areas (buffers) of 5 km around the trace of each road in the whole network, and compute the percentage of area of these buffers within the contours of *country-cells*, or the portion pixels of 1x1 degrees of latitude by longitude that lie within the borders of each modern country that the Roman Empire covered in 117 CE. As an illustration of our procedure to measure the degree of influence of Roman roads, Figure 2 shows the buffer around all major roads in the Roman Empire, and Figure 3 zooms in on the road system and the buffer around the main Roman roads crossing the area of Lutetia (contemporary Paris) in France.

Fig 2-3

At first sight, a 5 km buffer may seem large, and it is certainly possible to construct a smaller buffer. The choice is made so as to accomodate the fact that early Roman infrastructure was associated with a range of adjacent investments, as discussed below, including drainage and more. In order to be sure to envelope the total treatment resulting from the road construction, we use the said buffer size. This still allows a lot of variation since our unit of observation are areas of about 10,000 km2 (if measured at the equator).

As mentioned in the introduction, we confine our attention to country-cells that were treated by at least one Roman road. That is, we focus the analysis exclusively on the intensive margin of infrastructure investments. Also, in order to disentagle the influence of the physical infrastructure on economic outcomes from the legacy of Roman rule more broadly, we drop areas that fell outside the Empire or were unconnected to the road network within the Empire. Still, in the interest of completeness, we also report, in the appendix, results where we measure Roman roads along the extensive margin, within the borders of the Roman Empire.

2.2 Dependent variable I: Modern roads

Data for modern roads are taken from the Seamless Digital Chart of the World (SDCW) Base Map version 3.01, which is one of the most comprehensive global GIS databases

[11] Data on the extent of the Roman Empire in the year 117 AD are from the Ancient World Mapping Center at the University of North Carolina at Chapel Hill, http://awmc.unc.edu.

that is freely available.[12]

The road network data in the SDCW version 3.01 was published in 2000 and include various characteristics of roadways, which are classified according to whether they are operational or under construction, including a median (the central area reserved to separate opposing lanes of traffic in divided roads) or not, and covering primary, secondary or unknown (unexamined or not surveyed) routes.[13]

As the modern (2000) counterpart to the major Roman Roads in the Barrington Atlas, we select all primary and secondary modern roads, with or without a median, and with known (examined/surveyed) characteristics (that is, we drop all unexamined or not surveyed modern roads). Primary modern roads are defined as hard surface, all weather roads with two or more lanes in width, and maintained for automobile traffic. Secondary roads are defined as all other roads maintained for automobile traffic.[14]

Just as we do for the Roman roads, we build our indicator of the intensity of modern roads by constructing a buffer of 5 km around the network of modern roadways, and computing the percentage of the buffer area within each country-cell of 1x1 degrees of latitude and longitude within each country within the contours of the Roman Empire in the year 117.

2.3 Dependent variable II: Roman settlements

As a measure of economic development at the end of antiquity we use the number of Roman settlements in the year 500 CE. We take these data from the Digital Atlas of the Roman Empire (DARE) constructed by Åhlfeldt (2017).[15] The dataset was compiled using Talbert (2000) and other sources, and contains the location of settlements, mines,

[12]The SDCW dataset, built by Global Mapping International, is a collection of shapefile layers for a variety of geographic and population features. The underlying data are from the U.S. Government's Digital Chart of the World (DCW), which is a comprehensive and consistent cartographic global database at a scale of 1:1,000,000 (Langas 1995), developed and maintained by the US National Imagery and Mapping Agency. The underlying US Government's DCW data in SDCW version 3.01 is based on the Vector Map (VMap) Level 0, Edition 5 database published in September 2000 by the US National Imagery and Mapping Agency. See World GeoDatasets (2017) for full documentation of the SDCW database.

[13]Specifically, the road categories in te SDCW are: (1) road with a median (for primary, secondary or unknown routes), (2) primary route, operational (including all primary routes except those under construction and without a median), (3) primary route under construction (including all primary routes and under construction roads), (4) secondary route or unknown, operational, without a median, (5) secondary route or unknown, under construction or doubtful, (6) unexamined/unsurveyed.

[14]See Langas (1995) for details.

[15]The DARE data is continuously updated. Our version was downloaded from the Pleiades data site https://pleiades.stoa.org/home on August 16, 2017.

forts, villas and various other localities. The observed time span of existence is indicated for each locality. We compute the number of large settlements that existed in year 500 CE within each country-cell.

2.4 Dependent variable III and IV: Nightlights and population density in 2010

As proxies of local economic activity today we rely on the intensity of lights at night and the level of population density, both measured in 2010.

The raw data for lights data come from satellites and sensors operated by the US Department of Defense's Version 4 Defense Meteorological Satellite Program Operational Linescan System (DMSP-OLS).[16] We use nighttime lights raw imagery at a resolution of 30 arc seconds and compute averages within each 1x1 country-cell within the contours of the Roman Empire in 117 CE. Figure 2 above shows the contemporary geographical distribution of nightlights in the area former covered by the Roman Empire.

For population density we use the UN-adjusted 2010 population count from the Gridded Population of the World version 4 database, that adjusts gridded population numbers to United Nations (UN) estimates of national population counts.[17] To construct population densities, we simply sum population numbers within each country-cell and divide the sum by the total country-cell area.[18]

3 The Roman Roads

3.1 Historical Introduction

The Roman road construction program during antiquity is generally considered to have been initiated in 312 BCE when censor Claudius Appius started the construction of a paved, all-weather road, subsequently named *Via Appia*, from Rome to Capua

[16] Data available at
http://www.ngdc.noaa.gov/eog/dmsp/downloadV4composites.html

[17] The alternative are unadjusted population levels, which are based on individual countries censuses and population registers, which we avoid with the aim of having more comparable data.

[18] We use the finest resolution data of 30 arc seconds to compute our variables. The raw data are available at http://sedac.ciesin.columbia.edu/data/collection/gpw-v4.

(Figure 4). The immediate reason for the construction of Via Appia was the ongoing Second Samnite War in which the Roman armies were trapped around Capua due to shortage of supplies from Rome. It is believed that the road was completed in 308 BCE. With the new and more efficient supply lines, the Romans defeated the Samnites in 304 BCE and Via Appia was eventually extended all the way to the southeastern port city of Brundisium a few decades later (Laurence, 1999).

Fig 4

Via Appia was neither the first road in the Mediterranean area (the Persians under Darius the Great had for instance constructed extensive royal roads in the 5th century BCE), nor in Roman territory (earlier, non-paved roads are mentioned by ancient Roman sources). Nonetheless, it would come to serve as a model for future road constructions, first on the Italian peninsula and later in the broader empire.[19] At the peak of the Roman empire at the death of Trajan in 117 CE, it is estimated that the empire hosted about 80,000 km of paved road (Gabriel, 2002). As Figure 2 shows, the road system connected regions in current Britain, Western Europe, Eastern Europe, North Africa and the Near East.

Because of their military purpose for achieving effective Roman control of a territory, the construction of these public highways (*viae publicae*) was carried out by Roman legions, and it was typically commissioned by a censor and administered by curatores in Rome.[20] Most of these roads were paved with stone and cement. Road building also included supporting public goods such as bridges, tunnels, guest houses, and drainage systems which required substantial engineering skills. Scholars have suggested that the construction of roads also fostered the use of ground surveys and maps (Davies, 1998). Ordinary citizens sometimes had to pay tolls at city gates and bridges, and the military always had priority. The *viae publicae* network was complemented by local roads, *viae vicinales*, which typically linked the major roads to a town or to other major roads. These roads were mainly the responsibility of local governments (Laurence, 1999).

[19]In Italy, for instance *Via Flaminia* (completed in 220 BCE) connected Rome with the Adriatic coast, whereas *Via Aemilia* (187 BCE) cut through the Po plain and made that imporant agricultural area available for Roman colonization.

[20]Censor and curator were public offices in Republican Rome. During Imperial times, road construction was mainly carried out by the emperors.

3.2 The Assignment of Roman Roads: Historical Priors

There are three main reasons why the Roman road construction program almost presents itself as a natural experiment, from the vantage point of the historical literature: i) The *military purpose* of the roads, ii) the preferred *straightness* of the construction, and iii) their construction in *newly conquered* and often undeveloped regions.

First, just as mentioned above with the early experience with the Appia during the war with the Samnites, the purpose of the roads was to increase the speed and the ease with which the legions could reach locations of military interest – including territories of ongoing campaigns, army bases and Roman colonies that provided the army with essential supplies. Viae publicae also played a key role for the consolidation of power and hegemony in newly conquered areas.[21] When the Roman *limes* stabilized along its northern and eastern frontiers, the road system was used to transport marching troops to the their legionary bases along the border. Very soon, the roads were also used by traders and for transportation of agricultural goods, but this was not the main intention.

Second, Roman roads were typically very straight over extensive distances. The ambition of the road engineers was typically to connect an existing point A in an area under Roman control with a specific point B in an area where power was less consolidated. The example of contemporary Rimini and Piacenza in Figure A1 in the Appendix illustrates this tendency. An obvious reason for choosing a straight road was the shorter distance and the lower costs in terms of building material and soldier efforts.[22] The straightness of the roads implied that they often passed right over hills and across difficult terrain. Via Appia, for instance, passed like a straight line right from Rome to the existing colony in Terracina through the Pontine marshes (Figure 4). Malaria was prevalent in this area and the Romans had to construct drainage systems in order to be able to get through. The marching armies were not necessarily much constrained by these difficult conditions, but it has been claimed that the steepness of the roads often made them unsuitable for commercial ox-drawn carts with agricultural goods (Mokyr,

[21]Laurence (1999) argues that a broader objective with the roads was also to demonstrate the general technological superiority and political commitment of the Romans to the peoples in neighboring areas. The construction of roads signalled an ability to even change the geography of landscapes, which presumably greatly impressed many of Rome's rivals.

[22]See Davies (1998) for an account of how the Romans managed to keep the roads straight between two points without modern surveying tools.

1990).

The fact that the major roads tended to be straight also suggests that in between the two connecting points A and B, the highway system often was not adjusted to take into account pre-existing local economic or other social characteristics. For instance, the Romans consciously avoided linking Via Appia to a number of existing Latin settlements in the vicinity (Laurence, 1999). In this sense, we argue that the straightness of the roads gives the road construction program the character of a random treatment on the pre-Roman countryside.

This relates to the third argument, namely that the Roman roads were often constructed in newly conquered areas without any extensive, or at least not comparable, existing network of cities and infrastructure. The Roman roads were laid out in territories in which they had limited prior knowledge and where they had the aim of quickly securing Roman hegemony.

As an illustration of this point, consider the case of Lugdunum (contemporary Lyon). Julius Caesar's conquest of Gaul north of the Mediterranean coast was completed relatively rapidly during a frantic campaign in the 50s BCE. There were many existing towns and cities in Gaul when the Romans arrived, but there was not a state in any sense comparable to the Roman polity, and most scholars refer to Gaul as proto-urban at the time (Woolf, 1998). In year 47 BCE, Caesar created a Roman colony in the important town of Vienne, 30 km south of contemporary Lyon in the Rhone valley and, at the time, the main settlement (referred to by Caesar as *oppida*) of the Gallic Allobroges tribe (see Figure A2 in the Appendix). In 43 BCE, the Romans were expelled from Vienne by the Allobroges. According to the Roman historian Dio Cassius, the Roman Senate then ordered the governor of Gallia Transalpina to found a new city for the refugees from Vienne to the north at the intersection of the Rhone and Saone rivers. This city became the Roman town of Lugdunum. According to Åhlfeldt (2017), this location was not an important existing oppida.

Shortly after the establishment of Lugdunum, Marcus Vipsanius Agrippa, the governor of Gallia Transalpina, initiated an extensive road building project in order to consolidate Roman rule in Gaul. Lugdunum was connected southwards along the Rhone to the important cities of Vienne, Avignon and Massilia. Agrippa also built extensive roads towards the Atlantic to the west, towards the North Sea, and towards the

Rhine to the east, thereby making Lugdunum a key hub in Roman Gaul (Gros, 1991). The city experienced a rapid growth as a result and soon eclipsed even the old Greek colonies to the south. It became the capital of and gave name to the Roman province of Gallia Lugdunensis, and served as the primary Roman city in Gaul for more than two centuries.

The example indicates that the Roman decision to make Lugdunum a hub of road construction was probably a combination of good geographical fundamentals (the Rhone and Saone intersection), the historical accident related to the hostility to Romans in the previously much more significant town of Vienne, and the need to quickly consolidate power in Gaul. We do not have strong reasons to believe that Roman road construction was based on an already existing network of prosperous towns in the area. Figure A2 shows the pre-Roman oppida in the Lugdunum area, as well as the subsequent Roman roads and settlements. At least around this key city, there are no indications that the Romans consciously tried to connect to older settlements.[23]

3.3 The Assignment of Roman Roads: Formal tests

In Table 1, we investigate determinants of road density. As described in Section 2.1, our units of analysis are country-cells, or grid cells of 1x1 latitude-longitude degrees within the borders of modern countries and territories covered by the extent of the Roman empire in 117 CE. The dependent variable is (log) Roman road density (or the percentage of a 5 km buffer around the Roman road system that lies wihin the total area of a country-cell). Accordingly, only cells featuring at least one road are in the sample.

The question we examine in this part is essentially the extent to which the received perception from the historical literature, suggesting a very limited influence of geography and pre-roman development on road investments, is accurate. Naturally, if geography does not play a significant role for road assignment, it lessens the need to control for it in the regressions to follow.

Table 1

[23]Michaels and Rauch (2016) find however that the location of existing oppida – pre-roman fortified towns – does seem to predict the location of Roman towns in Gaul.

In a number of instances the historical priors seem to be confirmed. Terrain ruggedness does not limit road density; on the contrary the correlation is positive and significant (column 2). Similarly, in sub-samples where we have proxies for pre-roman development (oppidas, in the case of Europe, and the timing of the Neolithic), we find very little evidence that such factors influence road density (columns 7 and 8). Even areas featuring mining activity during Roman times are not characterized by greater road density (column 9).

The expected militaristic motivation for road construction can also be confirmed. Areas further away from the borders of the empire feature less road density, and distance to Rome also matters in the expect way (column 6). We also observe from column 1 that road density was greater in the northern part of the empire, and to the east. This pattern is most likely found because these areas were more contested than the southern border areas. Finally, the fact that road density declines when moving away from navigable rivers may also be related to the needs of the military. Of course, roads ultimately linked up to army outpost which needed to be supplied with provisions. The transport of food and other necessities would be cheaper by sea transport, which may create the link between road density and distance to navigable rivers.

In other cases the historical priors seem to ring less true. For example, there appears to be a clear positive correlation between various measures of agricultural productivity and road density (column 4). Similarly, grid cells found at greater levels of elevation feature significantly lower levels of road density. Since these geographic features naturally may influence economic development, and the location of modern roads, in their own right they are essential controls in the remaining. Also significant are our various distance measures to waterways. While the significance of waterways may be consistent with a military motivation for road construction it is obvious that waterways may influence development in their own right.

Finally, in column 10 we study the collective explanatory power of geography on roman road density in our full sample. As can be seen geography does seem to matter: it accounts for about 40% of the variation in road density. Hence, while Table 1 does confirm important aspects of the historical priors it also clearly shows that geography needs to be controlled for when examining the persistence of roads and the link between ancient road density and economic activity today.

4 Ancient roads, Modern roads and Economic Activity

4.1 Empirical specification

We take to the data the following cross-sectional specification:

$$\log(y_{prc}) = \delta_c + \delta_r + \beta \cdot RRD_{prc} + \mathbf{X}'_{prc}\gamma + \epsilon_{prc}. \tag{1}$$

Our dependent variable, pertaining to pixel p in language region r and country c is denoted y_{prc}. The dependent variables of interest are, respectively, (log) modern day road density and (log) economic activity during antiquity and today. In the latter case we employ both nightlights (following Henderson et al, 2012) and population density (e.g., Rappaport and Sachs, 2003) in 2010. The independent variable of particular interest is log (1+) Roman road density, RRD_{prc}.[24]

In an effort to control for countrywide institutions we include a full set of country fixed effects, δ_c. In addition, since past research has documented important within-country variation in culture that affect economic activity (e.g., Tabellini, 2010; Michalopoulos and Papaiannou 2013), we rely on a full set of language fixed effects as a proxy, δ_r, following Andersen et al. (2016).

$\mathbf{X}_{prc}$ contains additional controls, which can broadly be partitioned into three categories. First, geographic variables that involve latitude, longitude, ruggedness, elevation and controls for soil quality. Second, proximity to waterbodies which involves distances to coast, major rivers and natural harbors. Third, a set of variables that control for distances to Rome, the border of the empire and the current capital. In addition, we also control for the location of historical mines. Finally, in *all* specifications we control for country-cell (*prc*-cell) area, as it varies with latitude and longitude due to the earth's curvature, modern country limits, and the borders of the former Roman Empire.

Finally, in terms of statistical inference, we follow Abadie et al. (2017), who argue that cluster adjustments for the standard errors should only be performed if there are strong theoretical priors to do so. In particular, the authors argue that clustering is only relevant to address an experimental design issue and/or a sampling design issue.

[24] Accordingly, since $\log(1+x) \approx x$ the coefficient β strictly speaking has the interpretation of a semi-elasticity.

Briefly, the former issue arises if the treatment focus occurs at a higher level of aggregation than the unit of observation, whereas the second one emerges if multi-level sampling is taking place (e.g., first in a sample of countries and then in a sample of regions within those countries). In the present case our sample consists of all the pixels within the Roman Empire that where treated by Roman roads, which means neither of the two issues arises. Accordingly, we rely on standard errors that are robust to heteroskedasticity throughout the empirical analysis in the paper.

4.2 Baseline results

In Table 2 we explore the link between road density during antiquity, and road density today, in our full sample.

Table 2

As column 1 of Table 2 shows, on average ancient roads can account for about 12 percent of the current differences in modern-day road density within our sample. At the level of raw partial correlation, an increase in Roman road density by one percent is associated with an increase in modern day road density of about 0.24 percent.

In column 2 we introduce country fixed effects and in column 3 we introduce simultanously country fixed effects as well as language fixed effects so as to partial out the influence from institutions and cultural value variation within nations. The economic significance of Roman road density declines, but only to a minor extent.

Adding the first set of geographic controls, involving e.g. measures of agricultural potential, makes more of a difference. Collectively, the controls adds about eight percent in explanatory power, and reduces the elasticity of Roman roads to about 0.15. As seen from the rest of the columns, the apparent persistent influence of ancient infrastructure on modern infrastructure remains when we add further controls, and all of the controls collectively. In total, historical road networks and geography account for about one fifth of the variation in contemporanous road network across grid cells.

Figure 5

Figure 5 demonstrates the partial correlation, corresponding to the model estimated in Table 2, column 7, using a binned partial residual plot in order better to assess the partial correlation in the present large sample. Consistent with our estimation approach, we use a linear fit to summarize the relationship between Roman roads and modern roads.[25] The positive link appears well determined.

In the Appendix, we further examine the robustness of the link between ancient roads and modern roads. In particular, we perturb the sample in various ways. We examine whether the results are affected by omitting Italy; all areas within 100 km of the ocean or the Roman border, respectively; if we add further climatic variables such as frost days or if we control for pre-Roman economic activity. Overall the results reported in Table 2 carry over.

Turning to the link between ancient roads and economic activity, Table 3 examines the link between Roman roads and economic activity around the collapse of the Western Roman Empire at the end of the fifth century. As a measure of economic activity we use the density of major settlements. The control strategy is similar to that invoked in Table 2. The general message from the table is that Roman road density is statistically strongly correlated with early economic activity, featuring elasticities between 0.5 and 1; roads are significant at the one percent level or better, regardless of which controls are added.

Table 3.

The controls themselves appear to enter in a meaningful way. Briefly, our results indicate that by the fifth century CE we find more major settlements at low levels of elevation and in areas with productive agriculture (column 4); close to the coast (column 5), and close to Rome (column 6). We also find that the density of major settlements declines as one moves from the southern parts of the empire and to the north, probably testifying to the importance of the Mediterranean basin during antiquity (Column 4, 7). In addition, the results indicate, more surprisingly, a weak tendency towards lower settlement density in the eastern part of the empire.

In the two subsequent tables, 4 and 5, we shift focus to contemporaneous economic activity, measured by nightlights (Table 4) and population density (Table 5).

Table 4, Table 5

[25] After partialing out controls we then divide the sample into 20 equal sized bins, average the residualized Roman road density and Modern road density within these bins, and plot the resulting reduced sample.

Once again, we find a statistically strong signal from ancient roads on economic activity. Regardless of controls, or exact choice of measure of economic activity, Roman road density is significant at the one percent level or better. Figure 6 and 7 depict the binned added variables plot for contemporaneous nightlights and population density, respectively, in the context of our full specification (Column 7 in the two tables). The strong positive (partial) correlations do not appear to be driven by outliers.

Figure 6 & 7

The point estimate for the controls are broadly consistent with priors. In both tables we find that economic activity tends to decline as one moves away from the ocean or navigable rivers. Also, economic activity tends to decline at higher levels of elevation, and with distance to the current capital. More unexpected is the positive correlation with ruggedness, which is found in both tables as well. These commonalities are consistent with the notion that both measures are reasonable proxies for economic activity.

At the same time, the result do not always line up. When we examine the determinants of nightlights we find a positive latitude gradient, but this is not the case when population density is used. On the other hand, the potential supply of calories seem to matter to population density, but does not help to explain the variation in nightlights. A potential explanation could be, that there need not be a perfect match between where people live and where they work. To see what this implies, suppose population density captures place of residency to a relatively greater extent than nightlights. Then the positive correlation with caloric suitablity could be due to the fact that cities historically usually were located near rich agrarian hinterlands (Henderson et al., 2016). The location of, say, a factory is potentially less path dependent than a city. Indeed, in recent times production has moved out of city centers to capitalize on lower land prices. Under a similar logic, latitude apparently influences productivity on-the-job more than the location of population centers.

Overall, Roman road density appears strongly associated with economic activity, both in the past and in the present. In every specification, statistical significance at the 1 percent level is attained, and the economic significance is quite substantial. In our full specification we find that economic activity during antiquity rises by about 0.6 percent for every percentage point increase in road density; in the modern day context we find elasticities in the range 0.5 - 1 depending on the indicator.

4.3 Exploring the channel: Persistence and Non-Persistence

A key question regarding the results above is whether they reflect a causal impact of ancient roads on modern roads, and ultimately economic activity today. Naturally, the Roman roads are strongly predetermined, so reverse causality is not a concern. But it seems hard to rule out that underlying structural charactaristics, perhaps notably of a geographic nature, could be driving both the intensity of Roman road treatment and the outcomes in focus. That is, despite our best efforts, the results may suffer from omitted variable bias.

In the present section we explore the likelihood that our results can be accounted for in this manner, by exploiting the remarkable abandonment of wheeled transport in the Middle East and North Africa (MENA) during the second half of the first millenium CE (Bulliet, 1990 [1975]). This event is an astonishing fact of world history. Perhaps especially since wheeled transport has had a very long history in the Middle East before its abandonment. The first instances of primitive two-wheeled carts, drawn by oxen or later by horses were found already in the earliest civilizations of ancient Mesopotamia, for example. Such transportation was clearly facilitated by roads. As mentioned above, notable roads were built in Persia during the Achaemenid period around 500 BCE. But during the Roman era the roads became more pervasive and better constructed. This frames the puzzle: why did wheeled transport decline and disappear under those circumstances?

4.3.1 Empirical strategy: The regional loss of wheeled transport

Bulliet (1990 [1975]) argues that the key proximate reason for the abandonment of the wheel was the emergence of the camel caravan as a more cost effective mode of transport of goods in the region. The cost advantage during antiquity can be supported by data from Diocletian's price edict in 301 CE, which suggests a roughly 20 percent cost advantage in transport of goods by way of camel, relative to oxen.[26] To an economist, this seems like a very reasonable explanation. But it immediately prompts the question of why the ox-carriage then continued to dominate land-based transport until the first

[26]The objective of the edict was to stabilize prices in the region, which makes it probable that the relative prices, stipulated by the edict, were based on relative cost differences. From the edict it can be calculated that the price of transporting a given amount of goods (in Roman pounds) over a given distance was 20 percent higher per oxen than per camel. See Bulliet (1990, Ch. 1) for further discussion.

half of the first millenium CE? After all, the domestication of the camel on the Arabian Peninsula pre-dates the Roman era by millenia (Almathen et al, 2016).

Bulliet's argument is that a series of developments had to take place before the camel could emerge as the dominant mode of transport in MENA. First, the emergence of a new type of camel saddle by 100 BCE made it possible for camel hearding tribesmen to utilize new types of effective weapons. This improved the military strength of ethnic groups that centuries earlier had perfected camel breeding, which allowed them to gradually gain control of the trade routes and, as a consequence, gain political power as well. Second, another important factor was the decline of Rome and the ultimate rise of Islam as a key power factor in the Mediterranean. As also forcefully argued by Henri Pirenne (2012 [1937]), the ensuring decline in long distance trade across the Mediterranean allowed for an increasing importance of inland trade routes within the former Roman empire, which, in the case of the MENA supported caravan transport. However, horse or ox-drawn carts remained the main mode of inland transport in Europe. Therefore it is not surprising that while the Roman roads continued to be maintained and in use in Europe (Glick, 2005 [1979}; Hitchener, 2012), where wheeled vehicles dominated land-based transport, the same does not seem to be have been the case in the remaining regions of the Roman empire where the caravan took over (Bulliet, 1990).

The implication of these developments is that since ancient roads fall into disrepair in the MENA region, to a much greater extent than in Europe, one should expect to see much less persistence in infrastructure density. The argument is simply that more than a millenium of disrepair most likely would erase the traces of the ancient infrastructure to a considerable extent, and when the importance of maintaining or building roads reappears in North Africa and the Middle East – with the advent of the automobile – the principles underlying road planning almost certainly *differed* from those that directed the planning of the Roman roads. In Europe, where the ancient roads persisted to a greater extent, modern roads are more likely to be built in place of the ancient roads. As a result, it would seem highly unlikely that modern road density would line up with ancient road density in the Middle East and North Africa whereas persistence would be more likely a priori in Europe.

The potentially differentiated degree of persistence in road density across regions of the Empire holds stark implications for the influence of Roman roads on comparative development: one should expect an influence from Roman road density on economic activity today only where persistence in infrastructure is found. Hence Roman roads should be of little importance to contemporary comparative development within the MENA region, while holding explanatory power within Europe. At the same time, one should expect a positive influence of Roman roads on economic activity in *all* regions during antiquity, *before* the abandonment of the wheel in North Africa and the Middle East.

These considerations lead to a straightforward test. We re-estimate equation (1) on subsamples: Europe and MENA, respectively. In this setting we expect to see persistence of an influence of Roman roads only within the European part of the Empire. This testing strategy allows us to assess the likelihood that our results above are driven by omitted variable bias. If indeed Roman roads do not predict modern road density in the MENA region, there is little reason why Roman roads should hold explanatory power vis-a-vis contemporary comparative development. If a significant link between past infrastructure investments and current economic activity arises in spite of this, the link is likely spurious or driven by unobserved geographic determinants of roads and economic development. Naturally, one might imagine that Roman roads could influence long run development through some type of cultural or institutional channel. But in light of our extensive controls for current institutions and cultural variation, through country fixed effects and language group fixed effects, such an account would seem to stretch the imagination. Accordingly, the abandonment of the wheel experiment in effect allows us to explore the channel through which our baseline results come about.

Before we turn to the results one further issue is worth raising. Today the MENA region is considerably poorer, on average, than Europe. Perhaps the factors that stifled economic development in this region would also serve to mollify the explanatory power of the past? That is, perhaps an absence of a "signal" from the past, in this region, would have little to do with the mechanism in focus: that the abandonment of the wheel diminished the persistence of infrastructure and therefore diminished the persistence in comparative economic development. The key aspect to notice, however, is that this concern involves comparative development *across* regions. The strategy employed below involves looking *within* regions; regions that ultimately followed seperate development trajectories overall. While the persistence of infrastructure in Europe *relative* to, say, the Middle East, may have many causes, the tests conducted below involve asking if infrastructure was persistent within the MENA region and, by extension, whether ancient infrastructure predicts comparative development within the MENA.

4.3.2 Empirical results

In Table 6 we begin by examining the correlation between Roman road density and our measure of economic activity by the end of antiquity. The control set is the one in our full specification (cf column 7 in Tables 2-5).

Table 6

As is evident, across country-cells within Europe and within MENA, repectively, there is more economic activity in places with greater density of Roman roads. A natural interpretation of these findings is that by the end of antiquity areas more connected to the Roman road network benefitted on net terms. Hence, prior to the abandonment of wheeled vehicles there is a positive influence from roads on comparative development, regardless of which region of the empire we focus on. If anything, the economic significance of the link appears stronger in MENA than within Europe.

If we then turn attention to contemporary outcomes, results change markedly. As seen from column 3-4 Roman road density holds statistically significant predictive power within Europe, with respect to modern-day road density, whereas the same is not true for MENA. The economic significance also declines, but the main result is that we can no longer reject the null that the observed positive link is a matter of chance. This is consistent with what one would expect in the aftermath of the abandonment of the wheel-experiment. As ancient roads are left to decay they ultimately become a less reliable predictor of modern road location in the MENA.

In the remaining columns we turn attention to modern day economic activity. It is evident that whereas Roman roads hold strong predictive power over comparative development within Europe, both the economic and statistical significance are dramatically smaller within the MENA sample. In light of the absence of persistence in road density, these results are revealing, strongly suggesting that the explanatory power of Roman roads on current economic development is driven by the persistence of the road network.

Overall the results provide interesting perspectives on the roots of comparative development. While previous research has demonstrated how the observed persistence in economic development can arise due to variations in geographic initial conditions, either directly or indirectly via cultural or institutional change, the above results draw attention to an important role for shocks with persistent influence. From the point of view of any given geographical sub-region the emergence of the Roman Empire, and with it the Roman road network, is best viewed as external. The modest importance of geography in dictating the location of roads (cf Section 3) illustrates that *second nature* processes can, to some extent independently of geography (or *first nature* processes), have a substantial impact on long-run comparative development. In the present context the persistence of the shocks, and thereby in comparative development, arise via a remarkable degree of persistence in road density across several millennia, in regions where the roads were deemed economically useful. Evidently, persistence in infrastructure investments is a potential source of persistence in comparative development.

5 Concluding remarks

The existing literature on comparative development has drawn attention to a remarkable pattern of persistence in economic activity: places featuring comparatively high levels of economic development long before the industrial revolution often seem to feature high levels of comparative development today. In the present project we examine the persistence of an important proximate source of economic activity: Infrastructure investments.

Our analysis reveals that, within regions that used to be part of the Roman empire, infrastructure density is highly persistent. That is, Roman road density is generally a strong predictor of modern day road density. Moreover, Roman road density is generally a predictor of contemporary economic activity. These results are statistically strong and resillient to extensive controls including for contemporary institutions and cultural values. Taken at face value, these results suggest that infrastructure may be one important channel through which the persistence in comparative development comes about.

In examining whether our core results, linking early infrastructure to current-day infrastructure and economic activity, are likely to reflect causal relationships, we examine the remarkable historical case of the abandonment of the wheel that occurred in the Middle-East and North Africa (MENA) during the second half of the first millenium CE. We find that in the MENA region, Roman roads lose predictive power vis-a-vis modern day roads. Moreover, Roman road density does not predict current day economic activity within the MENA region. In contrast, in the European region, where the roads were maintained, our baseline results carry over. These results suggest quite strongly that our reduced form results, linking Roman road density to current comparative development, are importantly caused by the persistence of infrastructure over a remarkable period of 2000 years.

References

[1] Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). "When Should You Adjust Standard Errors for Clustering?" (No. w24003). National Bureau of Economic Research.

[2] Almathen, F., Charruau, P., Mohandesan, E., Mwacharo, J. M., Orozco-terWengel, P., Pitt, D., ... & Magee, P. (2016). "Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary" *Proceedings of the National Academy of Sciences*, 113(24), 6707-6712.

[3] Andersen, T. B., Dalgaard, C. J., & Selaya, P. (2016). "Climate and the emergence of global income differences" *Review of Economic Studies*, 83(4), 1334-1363.

[4] Angelucci, C., Meraglia, S., & Voigtländer, N. (2017). "The Medieval Roots of Inclusive Institutions: From the Norman Conquest of England to the Great Reform Act" (Working Paper No 23606). National Bureau of Economic Research.

[5] Ashraf, Q., & Galor, O. (2013). "The Out of Africa hypothesis, human genetic diversity, and comparative economic development" *American Economic Review*, 103(1), 1-46.

[6] Ashraf, Q. and O. Galor (forthcoming). The Macrogenoeconomics of Comparative Development, *Journal of Economic Literature*.

[7] Banerjee, A., E. Duflo and N. Qian (2012) "On the Road: Access to Transportation Infrastructure and Economic Growth in China" mimeo.

[8] Bird, J. and S. Straub (2014) "The Brasilia Experiment: Road Access and the Spatial Pattern of Long-term Local Development in Brazil" *TSE Working Paper*, n. 14-495, May 2014

[9] Bosker, M., & Buringh, E. (2017). "City seeds: Geography and the origins of the European city system" *Journal of Urban Economics*, 98, 139–157

[10] Bosker, M., E. Buringh, and J.L. van Zanden (2013) "From Baghdad to London: Unraveling Urban Development in Europe, The Middle East and North Africa 800-1800" *Review of Economics and Statistics* 95(4), 1418-1437.

[11] Bowman, A. and A. Wilson (2009) *Quantifying the Roman Economy: Methods and Problems*. Oxford University Press.

[12] Bulliet, R. W. (1990 [1975]). *The Camel and the wheel*. Columbia University Press. Morningside Edition.

[13] Chanda, A., Cook, C. J., & Putterman, L. (2014). "Persistence of fortune: Accounting for population movements, there was no post-Columbian reversal" *American Economic Journal: Macroeconomics*, 6(3), 1-28.

[14] Chen, T., J. Kung and C. Ma (2017) "Long Live Keju! The Persistent Effects of China's Imperial Examination System". Mimeo

[15] Chaves, I., S. Engerman, S., and J. Robinson (2014). "Reinventing the Wheel: The Economic Benefits of Wheeled Transportation in Early Colonial British West Africa" In E. Akyeampong, R. Bates, N. Nunn, and J. Robinson (eds.), *Africa's Development in Historical Perspective* (pp. 321-365). Cambridge: Cambridge University Press.

[16] Comin, D., Easterly, W., & Gong, E. (2010). Was the wealth of nations determined in 1000 BC?. *American Economic Journal: Macroeconomics*, 2(3), 65-97.

[17] Davies, H.E. (1998) "Designing Roman roads" Britannia, 29: 1-16.

[18] Donaldson, L. (2016) "Railroads of the Raj: Estimating the Impact of Transport Infrastructure" forthcoming *American Economic Review*.

[19] Fernald, J.G. (1999) "Roads to Prosperity? Assessing the Link Between Public Capital and Productivity" *American Economic Review* 89(3), 619-638.

[20] Finley, M. (1973) *The Ancient Economy*, Berkeley: University of California Press.

[21] Gabriel, R.A. (2002) *The Great Armies of Antiquity*, Westport: Greenwood.

[22] Galor, O. and Ö. Özak. "The agricultural origins of time preference." *American Economic Review* 106.10 (2016): 3064-3103.

[23] Glaeser, E. L., & Shleifer, A. (2002). Legal origins. *Quarterly Journal of Economics*, 117(4), 1193-1229.

[24] Glick (2005 [1979]). *Islam and Christian Spain in the early Middle Ages*. Brill.

[25] Gorodnichenko, Y., & Roland, G. (2017). "Culture, institutions, and the wealth of nations" *Review of Economics and Statistics*, 99(3), 402-416.

[26] Gros, Pierre (1991) *La France gallo-romaine*, Paris: Nathan.

[27] Jedwab, R. and A. Moradi (2016) "The Permanent Effects of Transport Revolutions in Poor Countries: Evidence from Africa" forthcoming *Review of Economics and Statistics*.

[28] Henderson, V., A. Storeygaard and D. Weil (2012) "Measuring economifc growth from outer space" *American Economic Review* 102(2), 994-1028.

[29] Hornung, E. (2015) "Railroads and Growth in Prussia" *Journal of the European Economic Association* 13(4), 699-736.

[30] Hitchener, R.B. (2012). 'Roads, integration, connectivity and economic performance in the Roman Empire',in (S.E. Alcock, J. Bodel and R.J.A. Talbert, eds.), Highways, Byways, and Road Systems in the Pre-ModernWorld, pp. 222–30, Malden, MA: Wiley-Blackwell.

[31] Jedwab, R., & Moradi, A. (2016). "The permanent effects of transportation revolutions in poor countries: evidence from Africa" *Review of Economics and Statistics*, 98(2), 268-284.

[32] Howard, P. (2006) *The Fall of the Roman Empire: A New History of Rome and the Barbarians*, Oxford University Press.

[33] Krugman, P. (1991) "History vs Expectations" *Quarterly Journal of Economics* 106(2), 651-667.

[34] Landes, D (1998). *The Wealth and Poverty of Nations*. Norton and Company: New York.

[35] Langas, S. (1995), "Completeness of the Digital Chart of the World (DCW) database". UNEP/GRID-Arendal, Project No. 205026.

[36] Laurence, R. (1999) *The Roads of Roman Italy: Mobility and Cultural Change*, Routledge: London.

[37] Maloney, W.F. and F. Valencia Caicedo (2016) "The Persistence of (Subnational) Fortune" *Economic Journal* 126(598), 2363-2401.

[38] McCormick, M., G. Huang, G. Zambotti, and J. Lavash (2013), "Roman Road Network (version 2008)", DARMC Scholarly Data Series, Data Contribution Series #2013-5. DARMC, Center for Geographic Analysis, Harvard University.

[39] Michaels, G. (2008) "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System" *Review of Economics and Statistics* 90(4), 683-701.

[40] Michaels, G. and F. Rauch (2016) "Resetting the Urban Network: 117-2012" forthcoming in *Economic Journal*.

[41] Michalopoulos, S., & Papaioannou, E. (2013). "Pre-Colonial ethnic institutions and contemporary african development" *Econometrica*, 81(1), 113-152.

[42] Mokyr, J. (1990) *The Lever of Riches: Technological Creativity and Economic Progress*. Oxford: Oxford University Press.

[43] New, M., D. Lister, M. Hulme and I. Makin (2002). "A high-resolution data set of surface climate over global land areas" *Climate research*, 21(1), 1-25.

[44] Nunn N. "Historical Development". In: Aghion P, Durlauf S Handbook of Economic Growth. Vol. 2. North-Holland ; 2014. pp. 347-402.

[45] Nunn, N., and D. Puga (2012). "Ruggedness: The blessing of bad geography in Africa" *Review of Economics and Statistics*, 94(1), 20-36.

[46] Olsson, O. and D.A Hibbs, Jr (2005) "Biogeography and Long-run Development" *European Economic Review* 49(4), 909-938.

[47] Pinhasi, R., J. Fort, and A. Ammerman (2005) "Tracing the Origins and Spread of Agriculture in Europe" *PLOS Biology* 3(12), 2220-2228.

[48] Pirenne, H. (2012 {1937]) *Economic and Social History of Medieval Europe*, Harcourt: Brace.

[49] 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.

[50] Rappaport J. and J. Sachs (2003) "The US as a coastal nation" *Journal of Economic Growth* 8, 5-46.

[51] Spolaore, E., & Wacziarg, R. (2013). How deep are the roots of economic development?. *Journal of Economic Literature*, 51(2), 325-369.

[52] Tabellini, G. (2010) "Culture and Institutions: Economic Development in the Regions of Europe" *Journal of the European Economic Association* 8(4), 677-716.

[53] Talbert, R.J. (2000) *Barrington Atlas of the Greek and Roman World*, Princeton University Press.

[54] Temin (2012) *The Roman Market Economy*, Princeton University Press.

[55] Wahl, F. (2015) "The Long Shadow of History: Roman Legacy and Economic Development - Evidence from the Roman Limes" *Hohenheim Discussion Papers in Business, Economics and the Social Sciences*, No 08-2015.

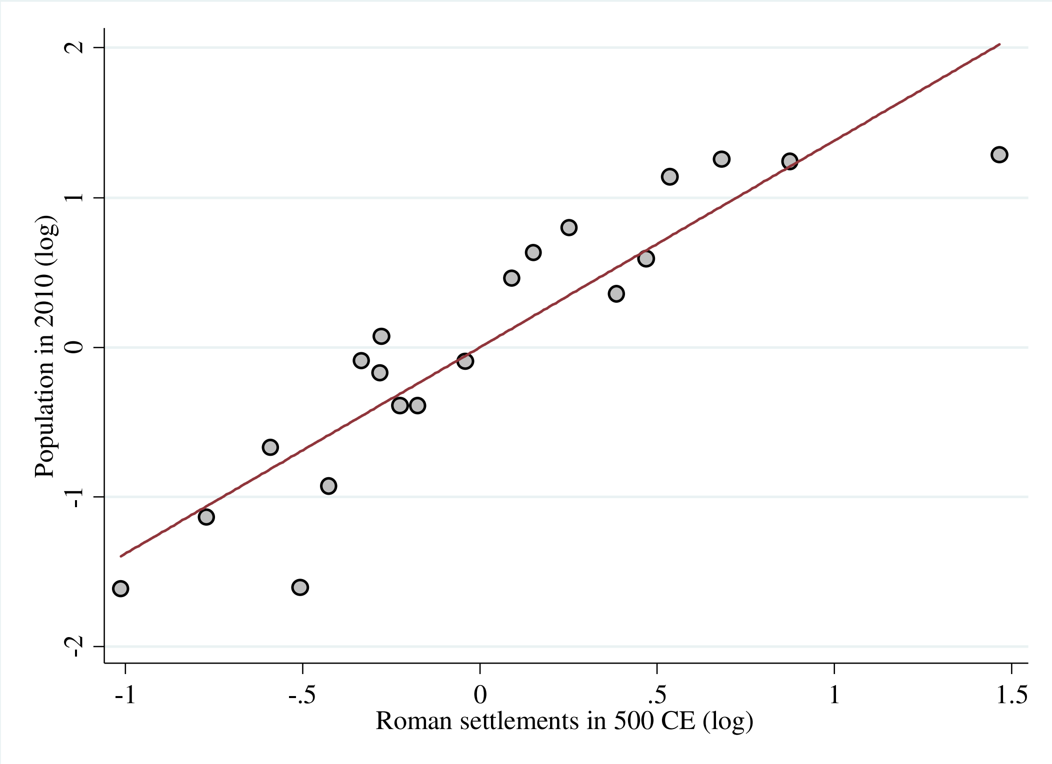
[56] Woolf, G. (1998) *Becoming Roman: The Origins of Provincial Civilization in Gaul.* Cambridge: Cambridge University Press.

[57] World GeoDatasets (2017), "Seamless Digital Chart of the World Base Map version 10.5". Available at http://www.worldgeodatasets.com/files/6214/9305/6335/WGDS_SDCW.pdf.

[58] Åhlfeldt, J. (2017) *Digital Atlas of the Roman Empire*. Web publication available at http://dare.ht.lu.se/. Lund University, Lund.

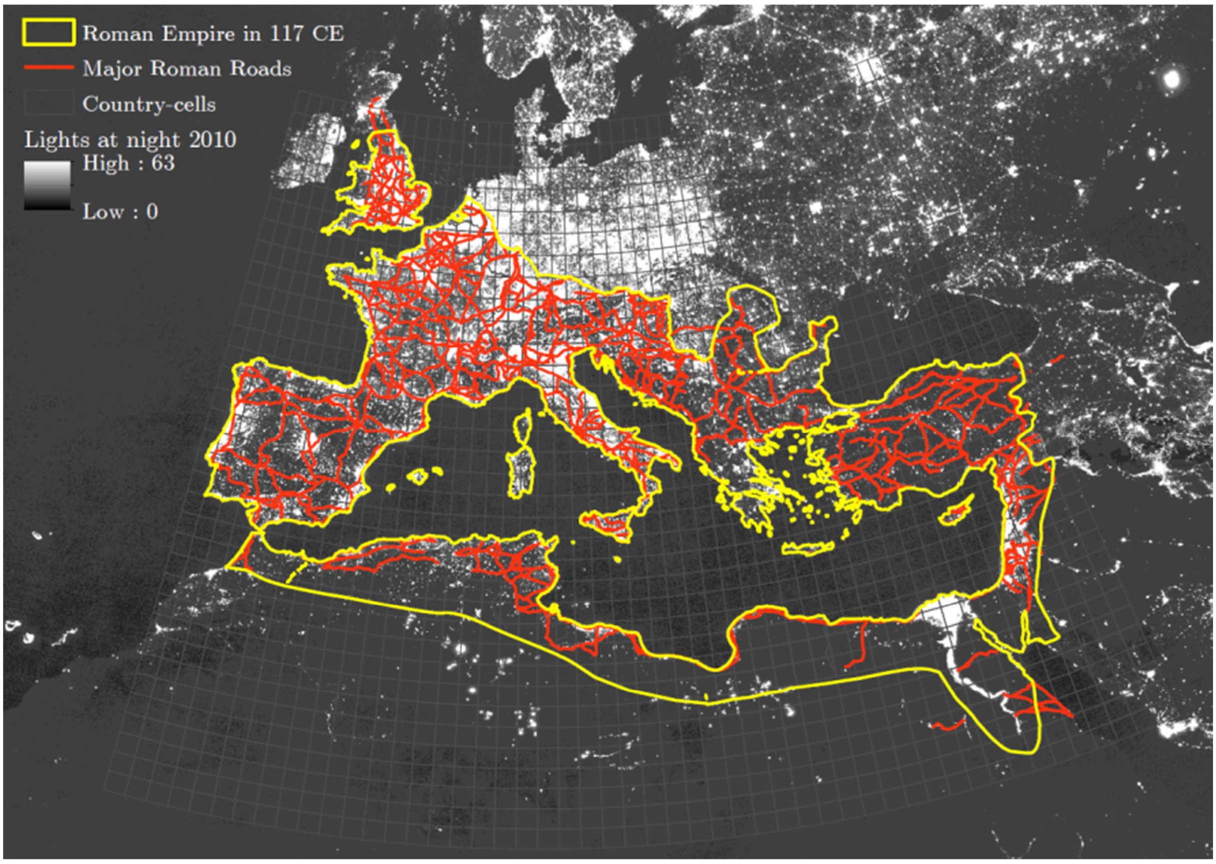
Figures

Figure 1: Conditional relationship between population in 2010 and the extent of Roman settlements in 500 CE within the former Roman Empire



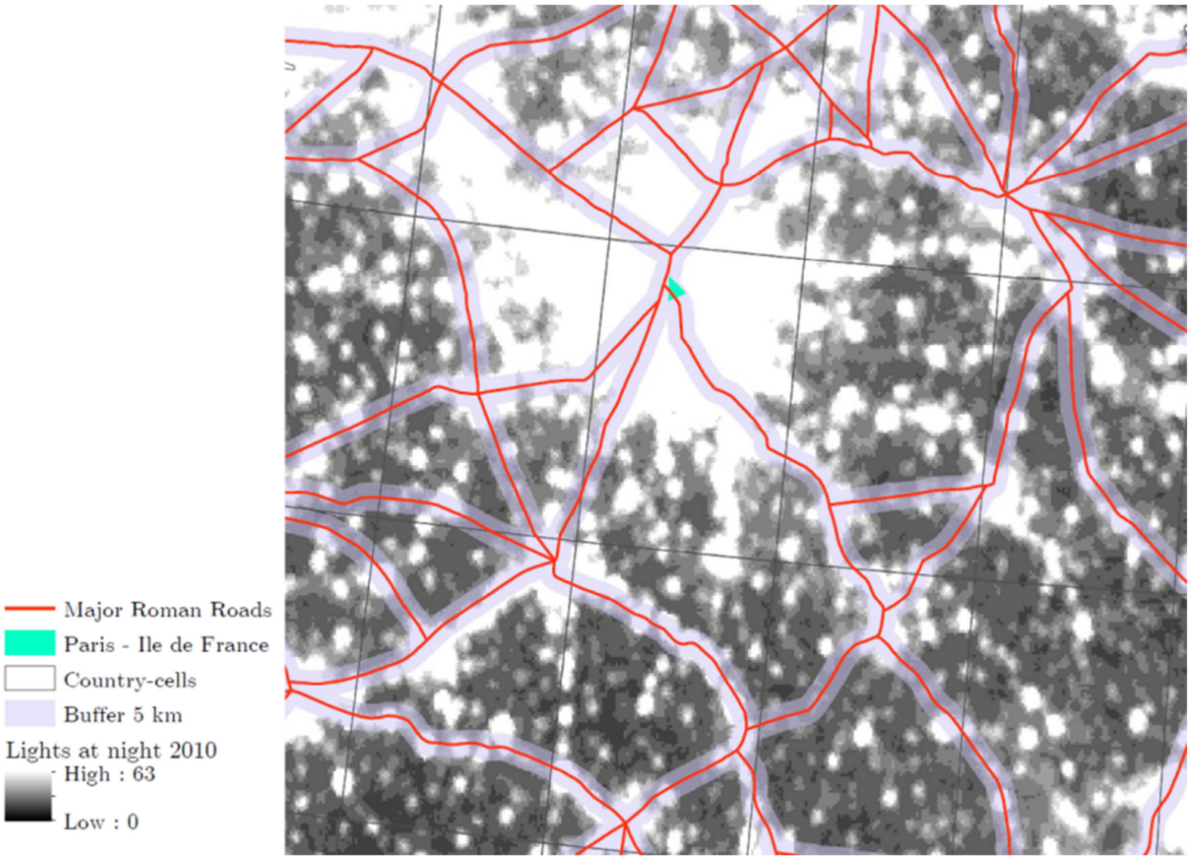
Note: The figure shows the conditional binned residual scatter plot of the relationship between population size (in logs) in 2010 and number of Roman settlements (in logs) in 500 CE for 693 country-cells within the former Roman empire. The binned scatter plot groups the x-axis variable into equal-sized bins, computes the mean of the x-axis and y-axis variables within each bin, then creates a scatterplot of these data points. The underlying regression controls for contemporary country fixed effects and hence estimates the within-country impact of historical Roman settlements on contemporaneous population levels.

Figure 2: Roman roads and contemporary night light intensity among 1000 country-cells within the Roman empire in 117 CE



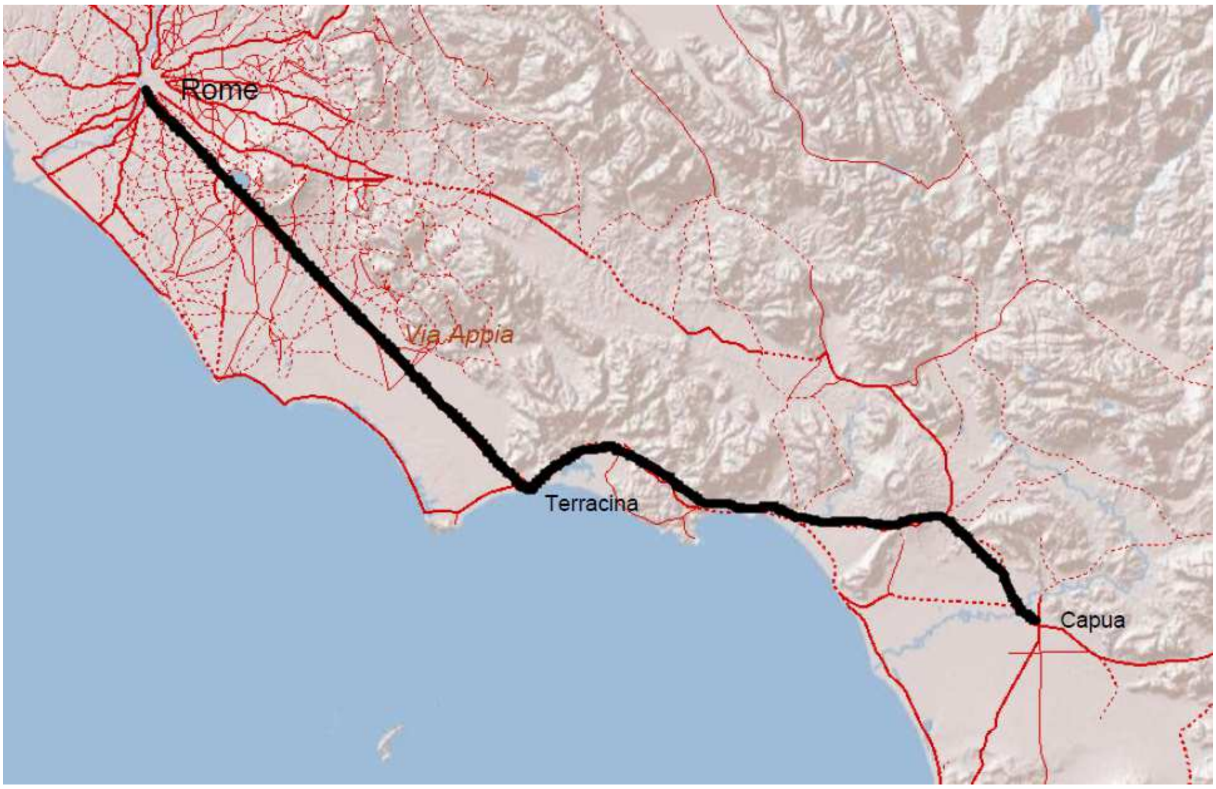
Note: The map shows major Roman roads (red lines) within the boundaries of the Roman Empire (green lines) in 117 CE with nightlights intensity in 2010 indicated by white color.

Figure 3: Roman roads and contemporary nightlights intensity around Lutetia (Paris)



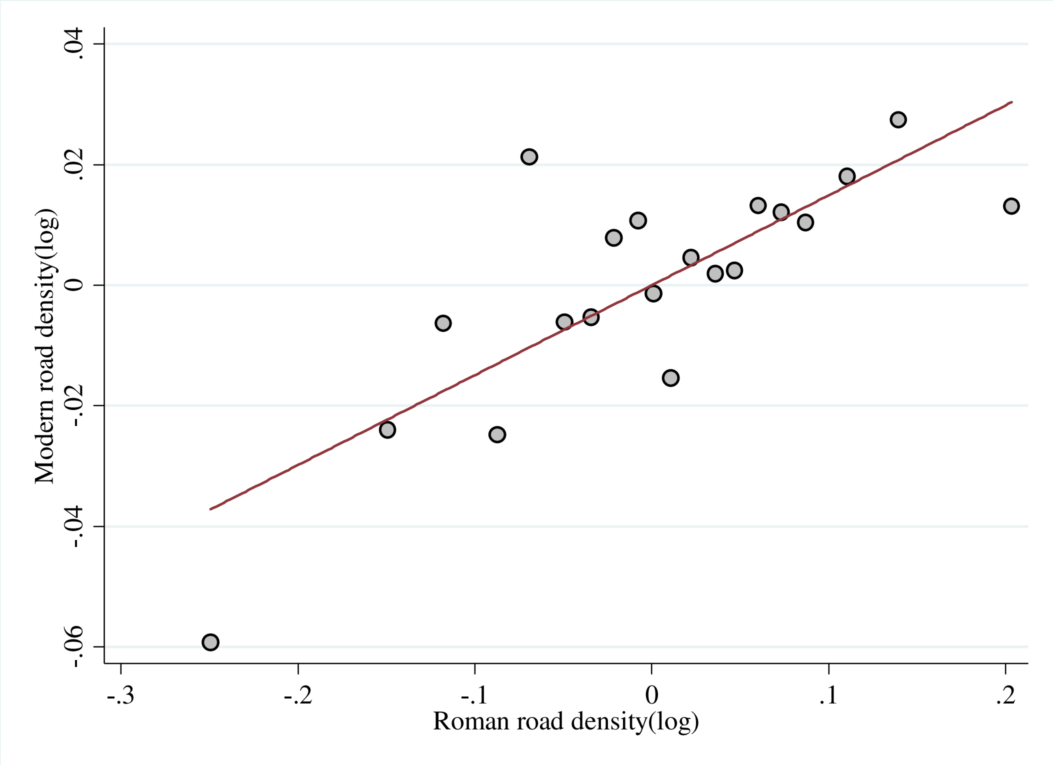
Note: The figure shows the rectangular 1x1 degree latitude/longitude country-cells including the hub of Roman roads around Lutetia (contemporary Paris – Ile de France), marked in light green. Buffer zones of 5 km on either side are shown around the roads. White color indicates the strength of modern nightlights intensity.

Figure 4: Via Appia from Rome to Capua in 312 BCE



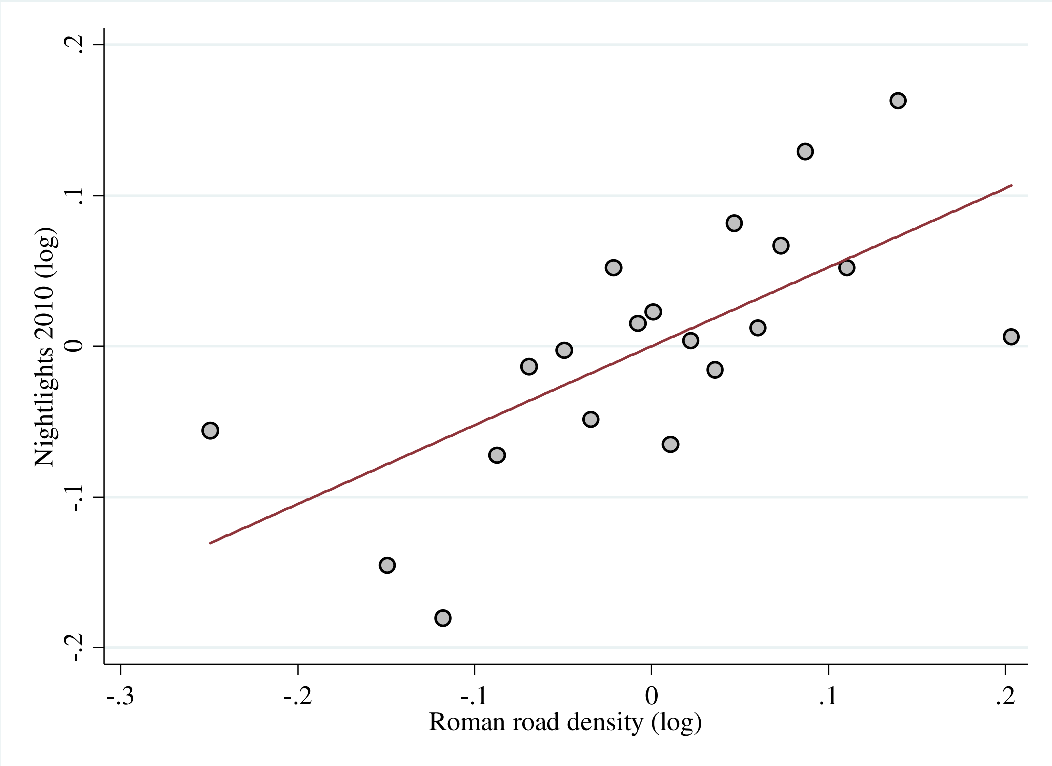
Note: The thick black line shows Via Appia whereas the red lines show other, later Roman roads.
Source: Created on the basis of data in Talbert (2000)

Figure 5: Conditional relationship between modern road density in 2000 and Roman road density in 117 CE within the former Roman Empire



Note: The figure shows the conditional binned residual scatter plot of the relationship between modern road density (in logs) in 2000 and Roman road density (in logs) in 117 CE for 675 country-cells within the former Roman empire. The binned scatter plot groups the x-axis variable into equal-sized bins, computes the mean of the x-axis and y-axis variables within each bin, then creates a scatterplot of these data points. The underlying regression in Table 2, column 7 controls for the full set of geographical controls, as well as for country and language fixed effects, and hence estimate the within-country impact of historical Roman settlements on contemporaneous population levels. See text and Appendix for exact variable definitions.

Figure 6: Conditional relationship between nightlight intensity in 2010 and Roman road density in 117 CE within the former Roman Empire



Note: The figure shows the conditional binned residual scatter plot of the relationship between nightlight intensity (in logs) in 2010 and Roman road density (in logs) in 117 CE for 675 country-cells within the former Roman empire. The binned scatter plot groups the x-axis variable into equal-sized bins, computes the mean of the x-axis and y-axis variables within each bin, then creates a scatterplot of these data points. The underlying regression in Table 4, column 7 controls for the full set of geographical controls, as well as for country and language fixed effects, and hence estimate the within-country impact of historical Roman settlements on contemporaneous population levels. See text and Appendix for exact variable definitions.

Tables

Table 1: Determinants of Roman Roads

	Dependent variable: Roman roads											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Latitude	0.165***									-0.122*	-0.406	0.224
	(0.040)									(0.069)	(0.374)	(0.142)
Longitude	0.031***									-0.009	-0.076	-0.009
	(0.007)									(0.013)	(0.112)	(0.025)
Ruggedness		0.025***								-0.000	0.024	0.006
		(0.010)								(0.010)	(0.021)	(0.017)
Elevation		-0.099***								-0.044***	-0.098***	-0.083***
		(0.010)								(0.013)	(0.034)	(0.026)
Africa dummy			-0.036*							0.039		
			(0.019)							(0.028)		
Middle East dummy			-0.034**							0.075**		0.092
			(0.015)							(0.030)		(0.056)
Pre-1500 caloric suitability				0.009***						0.014***	-0.026**	0.020**
				(0.003)						(0.003)	(0.010)	(0.008)
Agricultural suitability				0.104***						-0.016	0.026	-0.022
				(0.036)						(0.033)	(0.068)	(0.071)
Distance to major river					-0.031***					-0.017***	-0.014	-0.013
					(0.005)					(0.005)	(0.010)	(0.011)
Distance to coast					-0.052***					-0.037***	0.035**	0.013
					(0.005)					(0.007)	(0.016)	(0.012)
Distance to natural harbor					0.024***					0.019**	-0.058**	-0.027**
					(0.007)					(0.008)	(0.024)	(0.013)
Distance to Rome						-0.043***				-0.063***	-0.059	-0.047**
						(0.011)				(0.011)	(0.159)	(0.020)
Distance to Roman border						-0.022***				-0.033***	-0.048***	0.006
						(0.005)				(0.008)	(0.013)	(0.014)
Number of oppidas							0.591				0.743	
							(0.603)				(0.451)	
Years since Neolithic transition								0.017				0.242***
								(0.060)				(0.092)
Number of mines									-0.014	0.003	0.005	0.005
									(0.008)	(0.008)	(0.012)	(0.012)
Observations	693	693	693	675	693	693	183	182	693	675	183	176
R^2	0.034	0.193	0.011	0.056	0.186	0.095	0.005	0.001	0.002	0.363	0.365	0.287

Notes: This table documents that the density of Roman roads only to a limited extend was determined by geography and pre-existing development. Roman roads is the log of one plus the fraction of a 5 km buffer around the Roman road system that lies wihin the total area of a country-cell. The analysis is performed on country-cells within the Roman empire containing non-zero Roman roads. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table 2: Roman Roads and Modern Roads

	Dependent variable: Modern roads						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roman roads	0.240***	0.183***	0.197***	0.148***	0.175***	0.198***	0.149***
	(0.035)	(0.037)	(0.041)	(0.043)	(0.039)	(0.042)	(0.042)
Area	-0.003	-0.003	-0.006	-0.005	0.001	-0.006	-0.001
	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)
Latitude				-0.059			-0.004
				(0.116)			(0.135)
Longitude				-0.015			-0.010
				(0.021)			(0.022)
Ruggedness				0.013*			0.007
				(0.008)			(0.009)
Elevation				-0.025**			-0.009
				(0.011)			(0.013)
Post-1500 caloric suitability				0.009*			0.009*
				(0.005)			(0.005)
Agricultural suitability				0.097***			0.100***
				(0.033)			(0.032)
Distance to coast					-0.020***		-0.016***
					(0.004)		(0.006)
Distance to major river					0.001		0.002
					(0.005)		(0.005)
Distance to natural harbor					0.011*		0.011*
					(0.006)		(0.007)
Distance to Rome						0.006	0.015
						(0.012)	(0.013)
Distance to Roman border						0.003	-0.004
						(0.007)	(0.008)
Distance to capital						-0.009	-0.006
						(0.007)	(0.007)
Number of mines						-0.007	-0.007
						(0.006)	(0.006)
Country FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Country-language FE	No	No	Yes	Yes	Yes	Yes	Yes
Observations	693	693	693	675	693	693	675
R^2	0.123	0.465	0.548	0.581	0.570	0.551	0.592

Notes: This table documents the positive, statistically significant correlation between Roman roads and modern roads when accounting for geographical characteristics and country-language fixed effects. Roman roads and modern roads are defined as the log of one plus the fraction of a 5 km buffer around, respectively, the Roman and modern road system that lies wihin the total area of a country-cell. The analysis is performed on country-cells within the Roman empire containing non-zero Roman roads. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table 3: Roman Roads and Settlements

	Dependent variable: Settlements in 500 CE						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roman roads	0.998***	0.814***	0.831***	0.577***	0.733***	0.781***	0.536***
	(0.170)	(0.178)	(0.192)	(0.207)	(0.198)	(0.173)	(0.188)
Area	0.165***	0.171***	0.198***	0.209***	0.224***	0.200***	0.229***
	(0.019)	(0.021)	(0.017)	(0.020)	(0.020)	(0.022)	(0.023)
Latitude				-2.476***			-3.620***
				(0.820)			(0.826)
Longitude				-0.122			-0.275*
				(0.147)			(0.147)
Ruggedness				0.133***			0.108**
				(0.037)			(0.043)
Elevation				-0.180***			-0.160**
				(0.056)			(0.066)
Pre-1500 caloric suitability				0.034*			0.033*
				(0.019)			(0.019)
Agricultural suitability				0.612***			0.312*
				(0.190)			(0.178)
Distance to coast					-0.068***		-0.005
					(0.024)		(0.032)
Distance to major river					-0.026		-0.023
					(0.026)		(0.023)
Distance to natural harbor					-0.068		-0.023
					(0.044)		(0.046)
Distance to Rome						-0.833***	-0.863***
						(0.149)	(0.161)
Distance to Roman border						-0.026	-0.116***
						(0.030)	(0.032)
Distance to capital						0.006	0.018
						(0.037)	(0.037)
Number of mines						0.062	0.051
						(0.043)	(0.043)
Country FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Country-language FE	No	No	Yes	Yes	Yes	Yes	Yes
Observations	693	693	693	675	693	693	675
R^2	0.135	0.324	0.395	0.450	0.412	0.475	0.535

Notes: This table documents the positive, statistically significant correlation between Roman roads and Roman settlements when accounting for geographical characteristics and country-language fixed effects. Roman roads is log of one plus the fraction of a 5 km buffer around the Roman road system that lies wihin the total area of a country-cell. Roman settlements is log of one plus the number of major settlements within the country-cell in CE 500. The analysis is performed on country-cells within the Roman empire containing non-zero Roman roads. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table 4: Roman Roads and Nightlights

	Dependent variable: Nightlights in 2010						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roman roads	1.635***	1.349***	1.055***	0.580***	0.914***	0.963***	0.524***
	(0.186)	(0.161)	(0.153)	(0.152)	(0.153)	(0.152)	(0.153)
Area	-0.004	0.014	-0.018	0.019	0.014	-0.025	0.013
	(0.021)	(0.022)	(0.019)	(0.019)	(0.019)	(0.021)	(0.020)
Latitude				1.504***			1.182**
				(0.515)			(0.547)
Longitude				-0.089			-0.009
				(0.143)			(0.128)
Ruggedness				0.105***			0.117***
				(0.030)			(0.035)
Elevation				-0.356***			-0.344***
				(0.041)			(0.049)
Post-1500 caloric suitability				0.022			0.010
				(0.016)			(0.015)
Agricultural suitability				0.200			0.136
				(0.129)			(0.135)
Distance to coast					-0.089***		-0.013
					(0.021)		(0.024)
Distance to major river					-0.050***		-0.009
					(0.019)		(0.018)
Distance to natural harbor					-0.042		-0.054
					(0.033)		(0.034)
Distance to Rome						0.083	0.155
						(0.088)	(0.097)
Distance to Roman border						-0.033	-0.012
						(0.035)	(0.036)
Distance to capital						-0.169***	-0.169***
						(0.033)	(0.032)
Number of mines						-0.007	0.011
						(0.028)	(0.027)
Country FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Country-language FE	No	No	Yes	Yes	Yes	Yes	Yes
Observations	693	693	693	675	693	693	675
R^2	0.160	0.576	0.663	0.732	0.685	0.685	0.751

Notes: This table documents the positive, statistically significant correlation between Roman roads and nigthlights when accounting for geographical characteristics and country-language fixed effects. Roman roads is log of one plus the fraction of a 5 km buffer around the Roman road system that lies wihin the total area of a country-cell. Nightlights is log of the average light intensity measured at night by satelite in 2010. The analysis is performed on country-cells within the Roman empire containing non-zero Roman roads. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table 5: Roman Roads and population in 2010

	Dependent variable: Population in 2010						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Roman roads	3.231***	2.639***	2.559***	1.174***	2.202***	2.349***	1.078***
	(0.421)	(0.426)	(0.400)	(0.339)	(0.381)	(0.374)	(0.331)
Area	1.004***	1.029***	1.094***	1.137***	1.178***	1.061***	1.116***
	(0.061)	(0.069)	(0.040)	(0.035)	(0.038)	(0.041)	(0.038)
Latitude				1.382			1.018
				(1.255)			(1.379)
Longitude				-0.467			-0.250
				(0.302)			(0.269)
Ruggedness				0.335***			0.383***
				(0.069)			(0.078)
Elevation				-0.802***			-0.810***
				(0.109)			(0.128)
Post-1500 caloric suitability				0.242***			0.208***
				(0.061)			(0.059)
Agricultural suitability				0.281			0.069
				(0.351)			(0.343)
Distance to coast					-0.234***		-0.014
					(0.057)		(0.061)
Distance to major river					-0.112**		-0.026
					(0.048)		(0.045)
Distance to natural harbor					-0.092		-0.118
					(0.086)		(0.078)
Distance to Rome						0.098	0.248
						(0.168)	(0.174)
Distance to Roman border						0.031	0.054
						(0.075)	(0.079)
Distance to capital						-0.487***	-0.442***
						(0.083)	(0.077)
Number of mines						-0.046	-0.031
						(0.064)	(0.056)
Country FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Country-language FE	No	No	Yes	Yes	Yes	Yes	Yes
Observations	693	693	693	675	693	693	675
R^2	0.561	0.701	0.769	0.823	0.783	0.786	0.835

Notes: This table documents the positive, statistically significant correlation between Roman roads and population density when accounting for geographical characteristics and country-language fixed effects. Roman roads is log of one plus the fraction of a 5 km buffer around the Roman road system that lies wihin the total area of a country-cell. The analysis is performed on country-cells within the Roman empire containing non-zero Roman roads. All variables are in logs Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

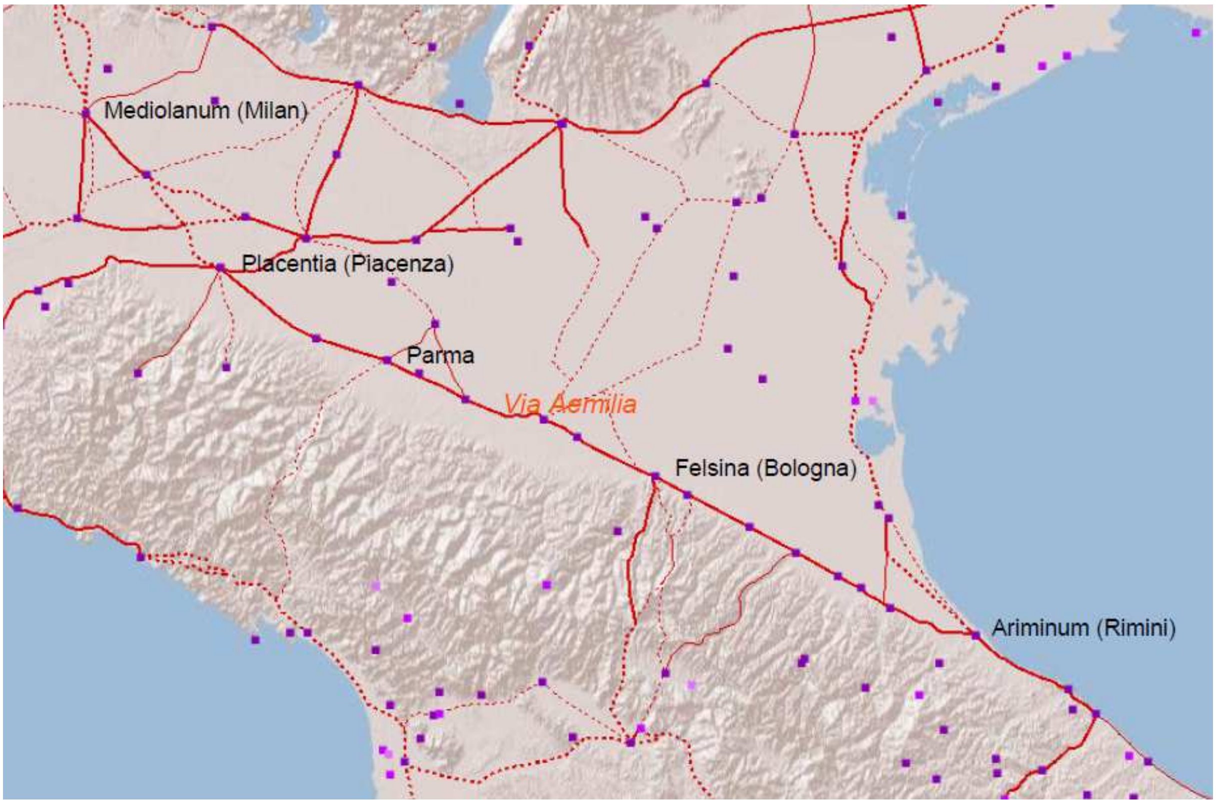
Table 6: Sample Split: Roman Roads and Development in Europe and MENA

Dependent variable:	Settlements in 500 CE		Modern roads		Nightlights in 2010		Population in 2010	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	Europe	MENA	Europe	MENA	Europe	MENA	Europe	MENA
Roman roads	0.399*	0.855**	0.160***	0.129	0.760***	0.109	1.394***	0.668
	(0.222)	(0.395)	(0.044)	(0.086)	(0.185)	(0.245)	(0.387)	(0.624)
Area	0.240***	0.210***	0.005	-0.025*	0.058***	-0.100***	1.177***	1.062***
	(0.026)	(0.052)	(0.005)	(0.013)	(0.022)	(0.032)	(0.038)	(0.098)
Latitude	-3.901***	-6.080***	0.068	-0.698**	1.505**	-0.163	0.353	0.977
	(0.995)	(1.807)	(0.108)	(0.297)	(0.729)	(1.053)	(1.425)	(3.427)
Longitude	-0.283*	0.005	-0.047**	0.186**	-0.082	0.581**	-0.334	0.537
	(0.159)	(0.389)	(0.019)	(0.073)	(0.152)	(0.233)	(0.298)	(0.602)
Ruggedness	0.015	0.105	0.009	-0.011	0.140***	-0.013	0.467***	-0.005
	(0.051)	(0.085)	(0.010)	(0.016)	(0.048)	(0.045)	(0.092)	(0.140)
Elevation	-0.138*	-0.100	-0.033**	0.046**	-0.380***	-0.271***	-0.908***	-0.666**
	(0.079)	(0.128)	(0.014)	(0.023)	(0.073)	(0.079)	(0.146)	(0.260)
Pre-1500 caloric suitability	-0.015	0.029						
	(0.020)	(0.023)						
Agricultural suitability	-0.151	1.256**	0.043	0.205**	0.094	0.827***	0.089	1.661*
	(0.183)	(0.584)	(0.033)	(0.100)	(0.167)	(0.293)	(0.378)	(0.924)
Distance to coast	0.021	-0.030	0.000	-0.031**	0.011	-0.006	0.029	-0.065
	(0.042)	(0.064)	(0.005)	(0.015)	(0.029)	(0.043)	(0.062)	(0.148)
Distance to major river	-0.002	-0.017	0.010***	0.003	0.025	-0.063*	0.052	-0.285***
	(0.021)	(0.068)	(0.004)	(0.017)	(0.020)	(0.037)	(0.044)	(0.107)
Distance to natural harbor	-0.034	0.063	0.001	0.003	-0.048	-0.140**	-0.076	-0.250
	(0.058)	(0.097)	(0.008)	(0.015)	(0.045)	(0.062)	(0.096)	(0.197)
Distance to Rome	-0.811***	-1.859***	-0.006	0.107	0.061	0.649**	0.119	1.001
	(0.155)	(0.551)	(0.011)	(0.093)	(0.095)	(0.326)	(0.177)	(1.147)
Distance to Roman border	-0.113***	0.161	-0.002	-0.110*	-0.034	-0.019	0.001	0.310
	(0.035)	(0.400)	(0.008)	(0.057)	(0.041)	(0.217)	(0.085)	(0.681)
Distance to capital	-0.016	0.006	-0.012*	-0.013	-0.142***	-0.280***	-0.356***	-0.763***
	(0.042)	(0.081)	(0.007)	(0.014)	(0.043)	(0.049)	(0.100)	(0.164)
Number of mines	0.010	0.184	-0.018***	0.037*	-0.029	0.094	-0.108*	0.151
	(0.041)	(0.152)	(0.006)	(0.020)	(0.028)	(0.070)	(0.056)	(0.178)
Post-1500 caloric suitability			0.009	0.012**	0.011	0.018	0.148***	0.205***
			(0.007)	(0.005)	(0.026)	(0.016)	(0.041)	(0.079)
Country-language FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	457	218	457	218	457	218	457	218
R^2	0.557	0.541	0.509	0.655	0.744	0.732	0.906	0.724

Notes: This table documents that the positive, statistically significant correlation between Roman roads and development persists in Europe but not in MENA (the Middle East and North Africa). Roman roads and modern roads are defined as the log of one plus the fraction of a 5 km buffer around, respectively, the Roman and modern road system that lies wihin the total area of a country-cell. Roman settlements is log of one plus the number of major settlements within the country-cell in 500 CE. Nightlights is log of the average light intensity measured at night by satelite in 2010 CE. The analysis is performed on country-cells within the Roman empire containing non-zero Roman roads. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

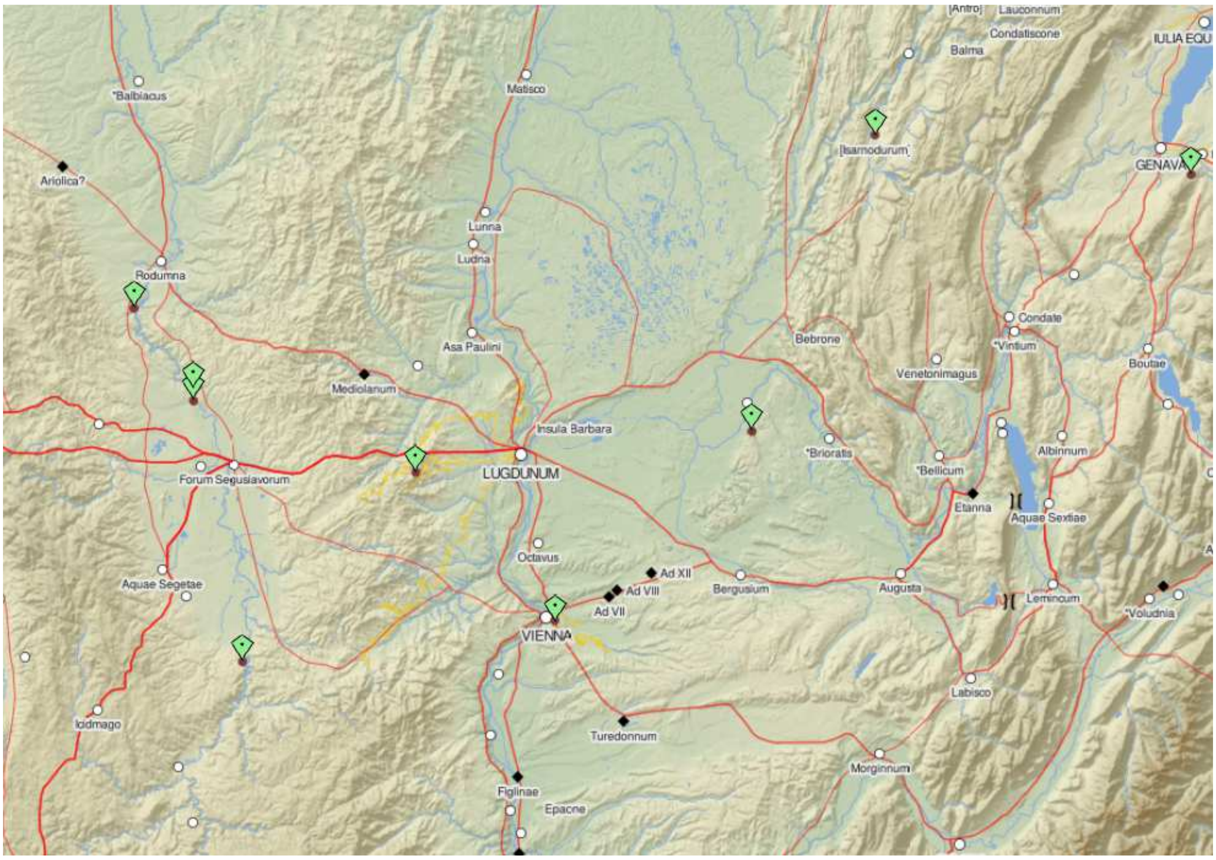
A Appendix: Additional figures

Figure A.1: Via Aemilia from Ariminum to Placentia in 187 BCE



Note: The figure shows Via Aemilia from Ariminum to Placentia, completed in 187 BCE. It also shows other later confirmed roads as thick red lines and unconfirmed roads as dotted lines. Quadratic symbols denote later Roman towns and settlements during antiquity.
Source: Created on the basis of data in Talbert (2000).

Figure A.2: Roman towns, pre-Roman towns and roads emanating from Lugdunum (Lyon) around 20 BCE



Note: The figure shows Roman towns and settlements after Caesar's conquest in 58-50 BCE as dark and white circles and pre-Roman Celtic settlements (oppida) as green arrows. Red lines show Roman roads, including Via Agrippa from hub city Lugdunum (Lyon) south along the Rhône.

Source: Created on the basis of data in Åhlfeldt (2015)

B Appendix: Tables

Table B.1: Summary Statistics

	Mean	Std	Min	Max	Obs
Main variables:					
Roman roads	27.8	21.6	0.0	100.0	675
Modern roads	73.5	18.6	0.2	100.0	675
Number of major settlements in 500 AD	1.4	2.8	0.0	28.0	675
Nightlights	15.0	11.0	3.0	63.0	675
Population in 2010 (1000s)	706.0	1085.6	0.0	10513.0	675
Geography:					
Area	5312.3	3530.3	0.0	11226.4	675
Latitude	41.6	6.1	24.5	55.1	675
Longitude	14.3	13.2	-9.2	39.3	675
Ruggedness (1000s)	169.8	159.9	1.0	1086.8	675
Elevation	488.3	446.4	0.0	2736.6	675
Post-1500 caloric suitability	7870.0	3426.8	0.0	15080.3	675
Pre-1500 caloric suitability	7153.1	2802.0	0.0	10549.4	675
Agricultural suitability (pct.)	55.9	30.6	0.0	99.8	675
Temperature	12.6	4.3	-1.0	24.8	675
Precipitation	58.9	28.1	0.2	189.5	675
Number of frost days	6.0	3.8	0.0	21.3	675
Waterways:					
Distance to coast	121.9	111.9	0.1	508.5	675
Distance to major river	174.6	248.9	0.1	1213.2	675
Distance to natural harbor	276.9	168.1	0.8	850.0	675
Location:					
Distance to Rome	1177.8	583.6	2.7	2731.8	675
Distance to Roman border	759.0	608.2	0.1	2246.2	675
Distance to capital	282.4	186.5	0.0	1126.1	675
Number of mines	0.7	1.8	0.0	17.0	675
Historical:					
Number of oppidas	0.9	1.7	0.0	11.2	183
Years since Neolithic transition	6844.8	1290.8	4790.0	10220.0	176

Notes: This table shows the summary statistics of the variables included in the analysis based on the observations included in the full specification of Tables 1 through 5. Variables are not in logs.

Table B.2: Roman Roads and Modern Roads - Robustness

	Dependent variable: Modern roads				
	(1) No Italy	(2) No coastal	(3) Roman road dummy	(4) No border	(5) More controls for climate
Roman roads	0.163*** (0.045)	0.155** (0.069)	0.131*** (0.040)	0.130*** (0.046)	0.150*** (0.042)
Area	-0.001 (0.006)	0.012 (0.009)	-0.001 (0.004)	0.002 (0.006)	-0.001 (0.005)
Latitude	-0.031 (0.164)	-0.194 (0.281)	-0.171 (0.152)	0.054 (0.156)	-0.003 (0.168)
Longitude	-0.012 (0.025)	-0.113** (0.045)	-0.031 (0.025)	-0.000 (0.023)	-0.006 (0.023)
Ruggedness	0.005 (0.010)	0.003 (0.018)	0.018* (0.010)	0.003 (0.009)	0.005 (0.010)
Elevation	-0.004 (0.014)	-0.025 (0.027)	-0.002 (0.012)	-0.003 (0.015)	-0.008 (0.019)
Post-1500 caloric suitability	0.008 (0.005)	0.006 (0.007)	0.017*** (0.004)	0.009* (0.005)	0.007 (0.006)
Agricultural suitability	0.105*** (0.034)	0.037 (0.046)	0.106*** (0.034)	0.110*** (0.035)	0.085** (0.039)
Distance to coast	-0.020*** (0.006)	0.009 (0.022)	-0.024*** (0.006)	-0.020*** (0.007)	-0.016*** (0.006)
Distance to major river	0.002 (0.005)	0.008 (0.006)	0.002 (0.005)	0.000 (0.006)	0.002 (0.005)
Distance to natural harbor	0.014 (0.008)	0.021 (0.021)	0.015* (0.009)	0.009 (0.007)	0.012* (0.007)
Distance to Rome	0.007 (0.041)	0.013 (0.069)	0.043** (0.018)	0.017 (0.013)	0.017 (0.013)
Distance to Roman border	-0.004 (0.008)	-0.017 (0.013)	-0.005 (0.009)	0.013 (0.019)	-0.003 (0.008)
Distance to capital	-0.006 (0.007)	0.001 (0.010)	-0.011 (0.007)	-0.003 (0.008)	-0.007 (0.007)
Number of mines	-0.008 (0.007)	-0.016* (0.008)	0.006 (0.007)	-0.005 (0.007)	-0.007 (0.006)
Roman road dummy			0.026* (0.014)		
Number of frost days					0.005 (0.028)
Temperature					0.033 (0.050)
Precipitation					0.017 (0.021)
Country-language FE	Yes	Yes	Yes	Yes	Yes
Observations	621	318	964	583	675
R^2	0.582	0.642	0.679	0.616	0.593

Notes: This table documents that the conclusion from Table 2 that Roman roads are postively and significantly correlated to modern roads is robust to changing the sample and including additional control variables. (1): Excludes Italy. (2): Excludes country-grid cells that lie within 100 km. of the coast. (3): Includes country-cells with zero Roman roads and adds a Roman road dummy that equals one if Roman roads is positive. (4): Excludes country-grid cells that lie within 100 km. of the border of the Roman empire. (5): Adds further controls for climate. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table B.3: Roman Roads and Ancient Settlements - Robustness

	Dependent variable: Settlements in 500 CE				
	(1) No Italy	(2) No coastal	(3) Roman road dummy	(4) No border	(5) More controls for climate
Roman roads	0.650***	0.684***	0.193	0.577**	0.531***
	(0.183)	(0.247)	(0.172)	(0.229)	(0.190)
Area	0.199***	0.146***	0.086***	0.246***	0.233***
	(0.020)	(0.026)	(0.015)	(0.029)	(0.022)
Latitude	-4.155***	-2.892*	-2.343***	-3.194***	-4.575***
	(0.969)	(1.549)	(0.604)	(1.008)	(1.030)
Longitude	-0.231	-1.027***	-0.243**	-0.265	-0.217
	(0.159)	(0.297)	(0.110)	(0.163)	(0.156)
Ruggedness	0.079**	0.073	0.050*	0.147***	0.085*
	(0.040)	(0.053)	(0.026)	(0.053)	(0.046)
Elevation	-0.142**	-0.182**	-0.058	-0.194**	-0.209**
	(0.066)	(0.086)	(0.038)	(0.077)	(0.089)
Pre-1500 caloric suitability	0.033*	0.010	0.017	0.030	0.014
	(0.017)	(0.022)	(0.011)	(0.019)	(0.021)
Agricultural suitability	0.326*	0.394*	0.322**	0.430**	0.271
	(0.183)	(0.222)	(0.144)	(0.217)	(0.192)
Distance to coast	-0.035	0.018	0.009	0.009	0.004
	(0.032)	(0.133)	(0.025)	(0.033)	(0.036)
Distance to major river	-0.012	0.002	-0.011	-0.017	-0.023
	(0.024)	(0.033)	(0.022)	(0.026)	(0.023)
Distance to natural harbor	0.001	-0.300**	-0.053	-0.011	-0.017
	(0.048)	(0.131)	(0.041)	(0.048)	(0.046)
Distance to Rome	-0.938***	-0.891**	-0.655***	-0.922***	-0.819***
	(0.212)	(0.345)	(0.160)	(0.175)	(0.162)
Distance to Roman border	-0.100***	-0.072*	-0.044	-0.009	-0.123***
	(0.031)	(0.043)	(0.027)	(0.113)	(0.033)
Distance to capital	-0.001	0.024	-0.069*	0.033	-0.000
	(0.036)	(0.059)	(0.036)	(0.042)	(0.038)
Number of mines	0.077*	0.086	0.118***	0.041	0.051
	(0.045)	(0.054)	(0.044)	(0.050)	(0.043)
Roman road dummy			0.086		
			(0.056)		
Number of frost days					0.070
					(0.167)
Temperature					0.028
					(0.211)
Precipitation					0.264**
					(0.125)
Country-language FE	Yes	Yes	Yes	Yes	Yes
Observations	621	318	964	583	675
R^2	0.488	0.541	0.480	0.543	0.541

Notes: This table documents that the conclusion from Table 3 that Roman roads are postively and significantly correlated to Roman settlements is robust to changing the sample and including additional control variables. (1): Excludes Italy. (2): Excludes country-grid cells that lie within 100 km. of the coast. (3): Includes country-cells with zero Roman roads and adds a Roman road dummy that equals one if Roman roads is positive. (4): Excludes country-grid cells that lie within 100 km. of the border of the Roman empire. (5): Adds further controls for climate. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table B.4: Roman Roads and Nightlights - Robustness

	Dependent variable: Nightligts in 2010				
	(1)	(2)	(3)	(4)	(5)
	No Italy	No coastal	Roman road dummy	No border	More controls for climate
Roman roads	0.558***	0.645***	0.614***	0.406**	0.523***
	(0.160)	(0.201)	(0.141)	(0.171)	(0.151)
Area	-0.000	0.052**	0.014	0.006	0.009
	(0.020)	(0.023)	(0.010)	(0.024)	(0.019)
Latitude	0.750	1.174	0.886*	0.380	2.097***
	(0.630)	(1.091)	(0.458)	(0.676)	(0.718)
Longitude	-0.011	-0.133	-0.003	0.016	0.046
	(0.134)	(0.200)	(0.100)	(0.130)	(0.132)
Ruggedness	0.136***	0.110**	0.048*	0.109***	0.107***
	(0.036)	(0.052)	(0.029)	(0.039)	(0.038)
Elevation	-0.349***	-0.259***	-0.214***	-0.364***	-0.257***
	(0.050)	(0.080)	(0.038)	(0.053)	(0.073)
Post-1500 caloric suitability	0.008 (0.015)	0.010 (0.018)	0.016 (0.011)	0.015 (0.015)	0.006 (0.018)
Agricultural suitability	0.069 (0.143)	0.097 (0.157)	0.171 (0.114)	0.030 (0.150)	0.048 (0.148)
Distance to coast	-0.025	-0.018	-0.043**	-0.016	-0.005
	(0.026)	(0.084)	(0.021)	(0.027)	(0.026)
Distance to major river	-0.009 (0.018)	-0.030 (0.022)	-0.018 (0.017)	-0.012 (0.019)	-0.015 (0.017)
Distance to natural harbor	-0.028 (0.040)	0.157* (0.087)	-0.037 (0.031)	-0.072** (0.036)	-0.053 (0.034)
Distance to Rome	-0.004	-0.152	0.169**	0.201*	0.144
	(0.159)	(0.232)	(0.086)	(0.108)	(0.096)
Distance to Roman border	-0.012 (0.038)	-0.041 (0.040)	-0.023 (0.030)	-0.153 (0.101)	-0.008 (0.036)
Distance to capital	-0.187***	-0.180***	-0.172***	-0.196***	-0.173***
	(0.031)	(0.048)	(0.028)	(0.036)	(0.033)
Number of mines	0.012	-0.010	0.032	-0.004	0.010
	(0.027)	(0.030)	(0.025)	(0.030)	(0.027)
Roman road dummy			0.026		
			(0.046)		
Number of frost days					-0.181
					(0.123)
Temperature					0.092
					(0.175)
Precipitation					-0.001
					(0.096)
Country-language FE	Yes	Yes	Yes	Yes	Yes
Observations	621	318	964	583	675
R^2	0.752	0.826	0.755	0.751	0.753

Notes: This table documents that the conclusion from Table 4 that Roman roads are postively and significantly correlated to nightlights is robust to changing the sample and including additional control variables. (1): Excludes Italy. (2): Excludes country-grid cells that lie within 100 km. of the coast. (3): Includes country-cells with zero Roman roads and adds a Roman road dummy that equals one if Roman roads is positive. (4): Excludes country-grid cells that lie within 100 km. of the border of the Roman empire. (5): Adds further controls for climate. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

Table B.5: Roman Roads and Population - Robustness

	Dependent variable: Population in 2010				
	(1)	(2)	(3)	(4)	(5)
	No Italy	No coastal	Roman road dummy	No border	More controls for climate
Roman roads	1.061***	1.011**	0.528	1.045***	1.064***
	(0.356)	(0.392)	(0.434)	(0.374)	(0.328)
Area	1.105***	1.153***	0.907***	1.142***	1.102***
	(0.043)	(0.041)	(0.078)	(0.047)	(0.037)
Latitude	0.515	0.680	3.451	0.009	4.327**
	(1.693)	(2.978)	(2.594)	(1.674)	(1.864)
Longitude	-0.322	-0.685	0.078	-0.111	-0.042
	(0.295)	(0.463)	(0.259)	(0.277)	(0.272)
Ruggedness	0.410***	0.334***	0.362***	0.398***	0.345***
	(0.085)	(0.105)	(0.104)	(0.091)	(0.084)
Elevation	-0.846***	-0.741***	-0.463**	-0.892***	-0.486***
	(0.136)	(0.157)	(0.182)	(0.144)	(0.177)
Post-1500 caloric suitability	0.201***	0.189***	0.132***	0.208***	0.205***
	(0.060)	(0.036)	(0.040)	(0.060)	(0.072)
Agricultural suitability	-0.074	-0.159	0.266	-0.167	-0.137
	(0.353)	(0.377)	(0.323)	(0.376)	(0.349)
Distance to coast	-0.028	-0.167	-0.020	-0.028	0.022
	(0.070)	(0.237)	(0.092)	(0.069)	(0.060)
Distance to major river	-0.035	-0.055	-0.092	-0.043	-0.049
	(0.046)	(0.054)	(0.071)	(0.050)	(0.044)
Distance to natural harbor	-0.049	0.402	-0.271***	-0.176**	-0.125*
	(0.100)	(0.251)	(0.086)	(0.079)	(0.075)
Distance to Rome	-0.349	-0.755	0.602***	0.373**	0.194
	(0.416)	(0.621)	(0.212)	(0.188)	(0.176)
Distance to Roman border	0.042	0.002	0.111	-0.093	0.060
	(0.085)	(0.078)	(0.103)	(0.223)	(0.077)
Distance to capital	-0.466***	-0.444***	-0.603***	-0.515***	-0.453***
	(0.077)	(0.128)	(0.083)	(0.090)	(0.080)
Number of mines	-0.026	-0.047	0.146*	-0.079	-0.031
	(0.058)	(0.069)	(0.083)	(0.063)	(0.055)
Roman road dummy			0.590***		
			(0.226)		
Number of frost days					-0.797***
					(0.273)
Temperature					-0.021
					(0.492)
Precipitation					-0.072
					(0.273)
Country-language FE	Yes	Yes	Yes	Yes	Yes
Observations	621	318	964	583	675
R^2	0.821	0.917	0.803	0.807	0.838

Notes: This table documents that the conclusion from Table 5 that Roman roads are postively and significantly correlated to population in 2010 is robust to changing the sample and including additional control variables. (1): Excludes Italy. (2): Excludes country-grid cells that lie within 100 km. of the coast. (3): Includes country-cells with zero Roman roads and adds a Roman road dummy that equals one if Roman roads is positive. (4): Excludes country-grid cells that lie within 100 km. of the border of the Roman empire. (5): Adds further controls for climate. All variables are in logs. Heteroskedasticity robust standard errors are reported in parentheses. *** denotes statistical significance at the 1 pct. level, ** at the 5 pct. level, and * at the 10 pct. level.

C Appendix: Additional data definitions and sources

Oppida were Celtic settlement sites that functioned as economic and political centres during the last two centuries BCE continuing to the first century CE before in Britain and Europe. The settlement area of oppidas was often hundreds of hectares, and they typically accommodated thousands of regular inhabitants. They were also characterized by being heavily fortified, and having a timber-laced stone-faced murus gallicus or a 'Gallic wall'.

We take data for the location of oppida during the La Tène CD period preceding the beginning of Roman expansion, from the Archaeology Data Service at the University of York.[1] Given that oppida were present only in Britain and continental Europe, we create an envelope of the buffer of 100 km around all oppida, as an estimate of the influece area of the La Tène culture. Within this area, similar to our procedure to estimate the area of influence of roads, we construct a buffer of 5 km around each oppidum, and compute the percentage of that area within each country-cell.[2]

Language areas are a complete set of language fixed effects, or indicator variables of ethnic languages recorded in the World Language Mapping System Database version 3.01 (WLMS, which is a dataset containing polygons for the linguistic homelands of more than 7,000 ethnic languages around the globe). We follow Andersen et al (2016) in the construction of these variables and, basically, we consider the predominant ethnic language to be the one with the largest area in cases where a country-cell has more than one ethnic language; and we assign a separate dummy variable that represents the excluded language category in each country, in country-cells where there are no specific ethnic languages recorded in WLMS.

Ruggedness is drawn from Nunn and Puga (2012) and reflects small-scale variations in elevation. They compute the index as the sum of the squared differences in elevation between the cell in question and eight adjacent cells. The ruggedness index is available at 30 arc-second cells. We aggregate the index by averaging across smaller cells within each 1 by 1 degree cell.

Elevation is computed using the Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). This data set contains the average elevation in meters at the level of 30 arc-second cells. We aggregate elevation by averaging across smaller cells within each 1 by 1 degree cell.

Caloric suitability is computed by Galor and Özak (2016) as the maximum potential crop yield in calories per hectare per year. It is constructed based only on geographical characteristics and not actual historical yields which makes suitable as an exogenous control variable. We compute the average of caloric suitability across smaller grid cells within each 1 by 1 degree grid cell. Since the types of crops available in different regions across the globe changed markedly after the Columbian Exchange, Galor and Özak (2016) provide two types of variables, one for the pre-1500 era and one for the post-1500 era. We use the first measure for regressions pertaining to the period prior to 1500 AD and the second measure for the regressions pertaining to the modern period.

Agricultural suitability is an index reflecting the suitability of the climate and geography for agriculture computed by Ramankutty et. al. (2002). The index takes on values between 0 and 1 indicating the probability

[1] The complete documentation can be found at http://archaeologydataservice.ac.uk/archiveDS/archiveDownload?t=arch-408-1/dissemination/zip/Atlas/GIS_data/documentation/ATLAS_METADATA.pdf.

[2] A buffer of 5 km for an oppidum is broadly consistent an area of about 7000 ha of agricultural land, outside a wall of 6 km lenght securing an area of 300 ha.

of cultivation. It is computed as follows: First, the relationship between actual cultivation and exogenous data on soil and climate is estimated in a statistical model. Then the probability of cultivation is predicted for each grid cell based on climatic and geographical characteristics. The index is available across 0.5 by 0.5 decimal degree grid cells. We average this to the level of 1 by 1 degree grid cells.

Distance to nearest river is computed using the location of rivers provided by CIA World Databank II. We compute the distance in km. from the centroid of each 1 by 1 degree grid cell to the nearest major river.

Distance to coast is based on the location of shorelines from Natural Earth (2017). We compute the distance in km. from the centroid of each 1 by 1 degree grid cell to the nearest coast.

Time since Neolithic transition is the number of years elapsed since the earliest evidence of agriculture in the grid cell. We compute it based the list of archaeological Neolithic sites compiled by Pinhasi et al (2005) who use it to trace the spread of the transition from hunting and gathering to agriculture throughout Europe and the Middle East. Each site is provided with a radio-carbon date measured in years before present. We define the time since the Neolithic transition as the years since the earliest Neolithic site within each 1 by 1 degree grid cell.

Climatic variables (temperature, precipitation, frost days) are computed using data of New et al (2002). They use data from weather stations across the globe to construct climate variables at the level of 10 by 10 arc-minute grid cells. Each variable is computed as monthly means of the period from 1961-1990. We first compute the yearly average of temperature in degrees Celsius, precipitation in mm pr. month and the number of days with ground frost per month. We then average across 10 by 10 arc-minute cells within each 1 by 1 degree cell unit.

Ancient mines. We use the Digital Atlas of the Roman Empire (DARE) database constructed by Åhlfeldt (2017) to compute the number of ancient mines in each cell. This variable is used as a control to ensure the estimated effect of roads on economic activity is not confounded by mineral deposits that were constructed before or simultaneously with the roads. For each country-pixel, we count the number of mines that existed prior to year 500 AD.

References

[1] Andersen, T. B., Dalgaard, C. J., & Selaya, P. (2016). Climate and the emergence of global income differences. *Review of Economic Studies*, 83(4), 1334-1363.

[2] Galor, O. and Ö. Özak. (2016) "The agricultural origins of time preference." *The American Economic Review* 106(10): 3064-3103.

[3] Natural Earth (2017) http://www.naturalearthdata.com/

[4] New, M., D. Lister, M. Hulme and I. Makin (2002). A high-resolution data set of surface climate over global land areas. *Climate research*, 21(1), 1-25.

[5] Nunn, N., and D. Puga (2012). Ruggedness: The blessing of bad geography in Africa. Review of Economics and Statistics, 94(1), 20-36.

[6] Pinhasi, R., J. Fort, and A. Ammerman (2005) "Tracing the Origins and Spread of Agriculture in Europe" *PLOS Biology* 3(12), 2220-2228.

[7] Åhlfeldt, J. (2017) *Digital Atlas of the Roman Empire*. Web publication available at http://dare.ht.lu.se/. Lund University, Lund.