Silk Roads to Riches: Persistence Along an Ancient Trade Network

Ahmad, Zofia and Chicoine, Luke

University of British Columbia, Bates College and IZA

January 2021

Silk Roads to Riches:

Persistence Along an Ancient Trade Network*

Zofia Ahmad
University of British Columbia

Luke Chicoine†
Bates College and IZA

January 20, 2021

Latest Version Available Here

The Silk Roads were a decentralized network of trade routes that connected ancient cities across Eurasia. Goods, ideas, people, and technology moved along the roads for over 1,500 years. Using a detailed georeferenced map of the entire trade network, this paper finds that areas within 50 KM of the historic location of the Silk Roads have higher levels of economic activity today. The persistent effect of proximity to the ancient trade network is associated with increased access to modern transportation infrastructure and the historical diffusion of technology along the routes but cannot be explained by differences in contemporary or historical levels of population density. This analysis is complemented by individual-level data from 22 countries; we find that districts with populations closest to the Silk Roads have higher rates of inter-group marriage, suggesting a weakening of social boundaries between groups that might possess differential technological knowledge.

JEL classification: R11; R12; O18; O33; N75

Keywords: ancient trade network; nighttime light intensity; modern transportation infrastructure; technological diffusion; cultural persistence

*We would like to thank Antonella Bancalari, Kerstin Enflo, Alexander Moradi, Nathan Nunn, André Seidel, and Jeanet Sinding Bentzen for their extremely helpful feedback, as well as conference participants at the 24th ASREC Conference, the 2020 DSE Winter School, and the 2019 Liberal Arts Development Conference; and the faculty at Bates College and DePaul University for their insightful comments.

†Corresponding author: Department of Economics, Bates College and IZA - Institute of Labor Economics. Address: Pettengill Hall, 4 Andrews Rd, Lewiston, ME 04240. Email: lchicoin@bates.edu.

1 Introduction

We know that political institutions can greatly influence the wellbeing of people around the world long after the systems themselves have dissolved, and that infrastructure projects from past centuries can help explain local variation in contemporary economic development.[1] To this point, the literature has largely focused on the role of formal institutions as a central component of this type of persistence. Transcending the reach of any single political institution across both time and space, merchants moved goods along the overland trade network known as the Silk Roads for more than 1,500 years during the rise and fall of Rome, the great East Asian Dynasties, Persian Caliphates, and the Mongol Empire. Although overland trade began to fade from these paths approximately 500 years ago, today nearly one billion people live in areas immediately adjacent to the historic location of the Silk Roads. This setting provides a context in which we can examine the long-term consequences of economic activity along a set of trade routes that was not reliant on any single political institution.

In this paper, we examine the persistent impact of the overland Silk Roads on modern economic activity using georeferenced data of the entire trade network and satellite images of nighttime light intensity (NLI). We find that current economic activity is 13 percent higher in the area immediately around the Silk Roads, relative to the surrounding area which is shown to be similar across a number of geographic, historic, and climatic, and historic characteristics. The evidence suggests that the benefits of the ancient trade routes persist through two key mechanisms. First, we find that areas along the routes are significantly closer to modern major highways and railroads, reducing trade costs. Second, these areas are more likely to use technology that is known to have moved along the ancient trade network. Population density and the number of ethnicities are also higher in areas along the routes, yet these outcomes are unable to explain the increase in economic activity. There is no evidence of an effect on institutional quality along the routes, as measured by the provision of public goods.

The main analysis in the paper is conducted using a grid of half-degree by half-degree cells covering an area from China in the east to Istanbul in the west, and a vast majority of Kazakhstan in the north to Mumbai in the south. The area within 500 KM of the Silk Roads yields a sample of over 7,000 individual grid cells. Similar to Jedwab and Moradi (2016) and Jedwab, Kerby, and Moradi (2017), the central identification strategy focuses on a comparison of cells within 50 KM of the Silk Roads relative to neighboring cells 50

[1] The long-term importance of institutions is summarized by Acemoglu, Johnson, and Robinson (2005), with similar patterns found in a number historical settings (Acemoglu, Johnson, and Robinson, 2001; Dell, 2010; Michalopoulos and Papaioannou, 2013). Many papers have shown the returns to access to rail and other types of transportation infrastructure around the world, in Africa (Jedwab and Moradi, 2016), India (Donaldson, 2018), Sweden (Berger and Enflo, 2017), Germany (Hornung, 2015), and the United States (Donaldson and Hornbeck, 2016). Dalgaard, Kaarsen, Olsson, and Selaya (2018) found persistent and long-term benefits of Roman road construction, and Flückiger, Hornung, Larch, Ludwig, and Mees (2019) identified the role of the network in building economic integration across Europe.

to 100 KM from the trade network. We identify a sample of these two groups of bordering cells that are balanced on observable geographic and historic characteristics, and find significantly higher NLI in the cells immediately around the Silk Roads relative to the neighboring cells 50 to 100 KM from the routes.

All estimates employ Conley (1999, 2010) standard errors, and t-statistics for the main set of NLI estimates remain above three when allowing for spatial correlations between 500 and 3,000 KM using either uniform or Bartlett kernels. The results are robust to a number of specifications including flexible controls for the longitude and latitude of each cell, different sets of spatial fixed effects, and the findings cannot be explained by the location of pre-Silk Roads ancient cities. The effect on NLI is stronger in portions of the routes that have been in use for longer periods, and persist along a predicted path based on only the ease of movement across the terrain and 33 sites that existed along the Silk Roads during the first century CE. The magnitude and pattern of the Silk Roads results cannot be matched by constructed placebo routes between the largest ancient cities of the pre-Silk Roads period—even though the placebo routes run through areas with relatively advantageous geography and higher population density—by routes between 1,000 random combinations of these cities, or routes between 1,000 sets of randomly placed points.

Our geospatial analysis is supplemented by microdata data from 22 countries; we collapse more than 39 million individual observations into administrative districts. With these data, we again find that the location of the Silk Roads does not affect the provision of public goods; there is no evidence of improved access to public utilities or higher levels of educational attainment. In the labor market, we find that adults are less likely to be employed in agriculture in districts with populations closest to the Silk Roads, a result that is only partially offset by the expansion of the manufacturing industry. The net decline in employment along with the increased level of economic activity suggests an overall increase in productivity. Reductions in barriers between different demographic populations—as suggested by the increased rate of inter-group marriage in districts along the Silk Roads—may have led to an increased willingness to set aside traditional cultural practices and adopt new ideas and technologies (Giuliano and Nunn, 2021).

The conclusions of this paper are consistent with recent work by Blaydes and Paik (2020) who found that increased access to trade along the Silk Roads led to higher rates of population growth in cities along the routes between 1100 and 1800 CE, Michalopoulos, Naghavi, and Prarolo (2018) who documented the spread of Islam along a similar set of ancient trade routes, and Barjamovic, Chaney, Coşar, and Hortaçsu (2019) who provided previous evidence that the distribution of city size is persistent across millennia. Although overland trade along the Silk Roads was greatly reduced by competition from European seagoing vessels in the 15th century CE, Bleakley and Lin (2012) and Maloney and Valencia Caicedo (2016) showed that past fortunes, like those established along the Silk Roads more than 500 years ago, can remain vitally important in determining contemporary prosperity. One example of this persistence that links the Silk Roads to

modern NLI is the strong evidence that we find suggesting that the Silk Roads have greatly impacted the location of modern transportation infrastructure. The distribution of transportation networks has been shown to significantly impact economic development (Alder, 2019), while access to major highways and railroads connecting to international markets are known to greatly reduce trade costs (Coşar and Demir, 2016; Donaldson, 2018). In recent years, China has heavily invested in transportation infrastructure in the region under the Belt and Road initiative (BRI). This investment has been found to have increased market access (Reed and Trubetskoy, 2019), trade flows between participating countries (Baniya, Rocha, and Ruta, 2020), as well as income and GDP in the region (Bird, Lebrand, and Venables, 2020; de Soyres, Mulabdic, and Ruta, 2020). To ensure the effects of the BRI are not driving our results, we use variables from 2013—the year of the announcement of the initiative—and earlier whenever possible, this includes verifying our results with data for NLI as far back as 1992.

The evidence in this paper demonstrates that areas closer to the Silk Roads have higher contemporary levels of NLI, are more connected to international markets, and have higher levels of technology associated with the trading network. The magnitude of these relationships fade as distance from the routes increase. These findings are evidence that the support of a single political institution is not a necessary condition for promoting and sustaining persistent economic growth. The benefits of trade and the exchange of ideas along the Silk Roads are found to transcend both historic and contemporary political boundaries.

2 Background

The overland Silk Roads were a network of trade routes connecting cities, trading centers, and oasis towns. These centers managed investments needed to sustain their populations and supervised trade within their communities. The trade routes connected each destination to a vast network that stretched from the Mediterranean Sea in the West to China in the east, from the Steppes of Central Asia in the north and the coast of India in the south.[2]

The network known today as the Silk Roads is generally considered to have begun during the last century BCE. Although trade in these regions had of course existed prior to this period, the peace generated by the consolidation of power from the Han Dynasty in the east, the Roman Empire in the west, and the Parthian and Kushan Empires in between created the opportunity for significant increases in economic activity along the Silk Roads. With the rise and fall of empires throughout history, enforcement along the routes varied; however, trade never depended on any single political entity. Even during periods of consolidation of power

[2]For historical references mentioned throughout this section see Hill (2003, 2015); Dignas and Winter (2007); Hansen (2012); Renfrew (2014); Rossabi (2014); Williams (2014, 2015); Barisitz (2017).

along the routes—the Roman Empire to Han Dynasty; Tang Dynasty; Mongol Empire—no single entity controlled the entire extant of the trading network.

Trade along the routes declined in the centuries following the end of the Mongol Empire. The increased political instability during this period led to rising costs of the overland route, and combined with new competition from European seagoing vessels, long-distance overland trade between the Middle East and China was reduced to a trickle. Any persistent effect of the Silk Roads are the consequences of a network of trade routes that were in place for over 1,500 years, but fell from their peak more than five centuries ago.

The movement of goods along the Silk Roads often took the shape of piecemeal trade with goods being passed from one trader to the next prior to arriving at their final destination. At each stage of trade, the merchants—who carried new goods, ideas, and technologies along the routes—required support and amenities. Much like inns along modern day highway networks, numerous caravansaries have been discovered along the ancient routes. These stopovers needed a workforce and population, irrigation systems to help form a local agricultural base, and possibly a more formalized institutions to provide public goods. Persistence of these technologies and institutions, which may have supported the movement of nomadic traders, are prime examples of the channels through which the Silk Roads may continue to impact the world to this date.

3 Data

3.1 Unit of Observation

The area under study is defined by the rectangle bounded by the following coordinates: 19° N 27° E in the southwest to 53° N 123° E in the northeast. The region encompasses the coastline of modern day China in the east, Turkey in the west, Mumbai in the South, and a vast majority of Kazakhstan in the north; covering a distance of over 7,000 KM from Istanbul to Beijing. In order to analyze the effect of the Silk Roads network on a localized level, the region is divided into arbitrarily placed 0.5° × 0.5° degree grid cells (roughly 55 KM by 55 KM at the equator) that are used as the unit of observation.[3] It is also important to note that because cells are defined in degrees, the area of each cell is smaller as the sample moves further north (e.g. 360 degrees around the equator is much longer distance in kilometers than 360 degrees around the North Pole); therefore, the area of the cell will always be included as a control variable and population counts within each cell will be analyzed relative to the cell area (i.e. population density per square kilometer).

[3] A number of papers have used a similar approach to defining geospatial observations, these papers include Michalopoulos (2012); Jedwab and Moradi (2016); Jedwab et al. (2017); Berman, Couttenier, Rohner, and Thoenig (2017); König, Rohner, Thoenig, and Zilibotti (2017); Dalgaard et al. (2018); Dickens (2018).

3.2 Map of the Silk Roads

To geographically locate the Silk Roads we use a digitized and georeferenced map created by Williams (2014, 2015). The work maps and identifies over 40,000 kilometers of the continuous main routes that were once used along the trade network. It was commissioned by the International Council of Monuments and Sites in an effort to identify potential sites for consideration for the UNESCO World Heritage List, and is available at the World Map data repository developed by the Center for Geographic Analysis at Harvard University. Although there is not sufficient evidence to chronologically date each segment of the roads, Williams (2014) provided detailed archaeological data on which he based each segment of the trading network. Williams (2015) also noted that due to the "tremendous scale of the project" there likely exists some variability in the exact location of the route between known nodes, a buffer of 30 KM on either side of the route (60 KM wide) "worked most effectively in capturing most key elements" known to be related to activity along the Silk Roads.[4]

The map of the Silk Roads is shown in Figure 1; for reference the roads overlay a map of contemporary country boundaries. The land area is covered with the $0.5° \times 0.5°$ cells that correspond to the unit of observation discussed in the previous subsection. The extent of the Silk Roads is shown in green across the Eurasian continent. The routes are shown within an open space denoting cells with centroids within 50 KM of the routes, roughly approximating the buffer suggested by Williams (2015). The darkest shading is the second distance bin, representing cells with centroids between 50 and 100 KM, each successive 50 KM distance bin is shown in a lighter shade, out to 200 KM. Cells between 200 and 500 KM from the routes are hollow and outlined in black. This set of cells—within 500 KM of the Silk Roads—is used as the main sample for much of the paper's analysis. The remaining land area, beyond 500 KM from the route, is covered with empty cells outlined in gray.

This figure describes the identification strategy that will be more formally outlined in Section 4.1. After initially comparing cells that contain some segment of the Silk Roads to the rest of the sample, the output within cells in each of the 50 KM bandwidths will be compared in a strategy similar to Jedwab and Moradi (2016) and Jedwab et al. (2017). The main reference group will be the area from 200 KM to 500 KM, the empty cells outlined in black in Figure 1.

[4]In addition to the routes, Williams (2014, 2015) identified the location of 319 cities, most of which lie along established portions of the Silk Roads. From these cities we extract 83 cities known to have existed in the year 100 BCE. These data are supplemented using the location of ancient cities with at least 30,000 people that existed between the years 500 and 100 BCE (Reba, Reitsma, and Seto, 2016), adding another 39 cities to the final set of 122 ancient cities. More details on the selection of these cities can be found in Appendix Sections A.2 and A.3, respectively.

3.3 Economic Activity: Nighttime Light Intensity

To measure economic activity, our main analysis focuses on NLI in a cross-section of grid cells. Nordhaus and Chen (2015) demonstrated that the analysis of cross-sectional NLI contains significant amount of information, and Lessmann and Seidel (2017) showed that NLI can measure sub-national income inequalities. To avoid data challenges presented in time-series variation (Gibson, Olivia, and Boe-Gibson, 2020), we use NLI data from Zhang, Pandey, and Seto (2016) who generated a consistent time series for the years 1992 to 2012 using annual composites of satellite data, allowing us to link NLI to GDP within our study area. Additionally, lights resulting from gas flares (Elvidge, Ziskin, Baugh, Tuttle, Ghosh, Pack, Erwin, and Zhizhin, 2009) are removed from the data. The main analysis throughout the paper uses NLI data from the year 2010 to pre-date China's 2013 announcement of plans to reinvest along the traditional Silk Roads network under the concept of the "New Silk Road," results are shown to be robust to the use of any other year of NLI. Raw light intensity data range from 0 to 63, the censoring at 63 could mask additional variation. To ensure this limit is not driving our findings we reestimate our main results using the top-coding adjustment from Bluhm and Krause (2018).

3.3.1 Nighttime Light Intensity Along the Silk Roads

The NLI data are shown in Figure 2a, with the map of the Silk Roads overlaid on the output. This image illustrates the data used in the paper's main analysis, and provides a broad qualitative description of patterns of nighttime lights that can be seen along the Silk Roads. To provide a clearer idea of the shape of the light intensity along the route, the map is entirely blacked out, except for the corridor of cells within 50 KM of the route. The NLI in this 50 KM corridor along the Silk Roads can be seen in Figure 2b, this is the same corridor along the route shown in white in Figure 1. Not only is there visual evidence of economic activity across nearly the entire ancient trading route, but next to much of the route the lights are oriented along the Silk Roads corridor.[5]

We can also more formally visualize the unconditional pattern between NLI and distance from the Silk Roads using the scatter plots in Figure 3. The figure is constructed using the sample of all cells within 500 KM of the Silk Roads, panel (a), a sample of non-desert cells, panel (b), and a sample that also drops all cells that contain a river or are within 50 KM of an ancient city, panel (c). Each figure yields a similar pattern, NLI is highest in areas closest to the Silk Roads, and the level of light intensity declines through roughly the first 200 KM before increasing slightly and eventually leveling off.

Beyond the intensity of the effect fading as distance increases, we also see evidence of two additional

[5]Specific examples zoomed into more local sections of the map can found in Appendix Figures B.1 to B.4.

important characteristics. First, the level of economic activity is relatively consistent throughout what will be the reference area of the main analysis—200 to 500 KM. Second, a concern with finding heightened activity around the routes is that it could simply be due to a redistribution of activity that would have occurred elsewhere; however, the evidence in Figure 3 demonstrates that this is not occurring in a significant way. While there is a nadir of economic activity around the 200 KM mark, it is not significantly lower than the level of activity in the reference area. These patterns are exactly what one would expect if past activity along the Silk Roads was a significant determinant of the distribution of modern economic activity.

3.3.2 GDP and Nighttime Light Intensity

In Henderson, Storeygard, and Weil (2012), the elasticity between measured country-level GDP and NLI is found to be roughly 0.277. This estimate was found using GDP data measured in constant local currency units (LCU) from 188 countries around the world, for the years 1992 to 2008. Storeygard (2016) repeated the analysis and found a similar global elasticity of 0.284, and more directly related to our region of study, a sub-national relationship between prefecture GDP and NLI in China of 0.248. In India, Chanda and Kabiraj (2020) determined the elasticity between state-level GDP and NLI to be 0.241. Across each of these estimates the relationship between NLI and GDP is similar. However, to the best of our knowledge, this relationship has not been estimated for our area of study; therefore, before exploring the relationship between the Silk Roads and NLI we first confirm the relationship between country-level GDP and NLI in our region of study.

Replicating the methods of the previous studies, the following model is estimated:

$$Ln\left(GDP_{ct}\right) = \alpha + \beta \cdot Ln\left(NLI_{ct}\right) + \gamma_c + \theta_t + \varepsilon_{ct}. \tag{1}$$

$Ln\left(GDP_{ct}\right)$ is the natural log of constant GDP measured in LCU for country c in year t. For each country-year, the natural log of NLI is denoted as $Ln\left(NLI_{ct}\right)$, where NLI_{ct} is calculated as the average pixel value within each country's borders in year t. γ_c is a set of country-specific fixed effects and θ_t is a set of year-specific fixed effects; standard errors are clustered by country.

The estimates from equation (1) are shown in Table 1. The elasticity estimate that includes all available country-year observations is shown in column (1), the sample in column (2) is restricted to a balanced sample of countries with GDP data available for every year. Both elasticity estimates—roughly 0.24—are similar to the previous estimates in the region from Storeygard (2016) and Chanda and Kabiraj (2020). This evidence shows a strong association between NLI and country-level economic activity within the Eurasian study region.

3.4 Mechanisms: Potential Pathways from the Silk Roads to Economic Activity

Five different pathways through which the Silk Roads could influence contemporary economic activity are considered in Section 5.2. We examine the impact of the Silk Roads on each of these pathways, and investigate their ability to mediate the relationship between the Silk Roads and NLI. Additional detail for each variable discussed in this section can be found in Appendix Section A.

First, we consider how population densities in 2010 are impacted by the Silk Roads, and whether changes in NLI are explained by changes in population density. Second, we also explore the impact of the trade routes on the number of ethnicities within each cell. Increased diversity may have been be driven by connections associated with the Silk Roads, which might have led to an increased introduction and adoption of new ideas and technologies. Third, an increase in connectivity of an area could also take a more physical form, such as a major highway or railroad. Measures for each of these modes of transportation are examined separately using data from the year 2013.

The final mechanisms explored relate to characteristics necessary to support merchants and their caravans as they moved along the Silk Roads. Irrigation technologies were necessary to sustain the populations that built up around oases and caravanserais and supported the movement and people and goods along the trade routes (Williams, 2014; Li, Storozum, Wang, and Guo, 2017). As a measure of technology adoption we use an indicator for whether a cell includes irrigated land. Similar to technological investments, it is possible that the populations along the routes also required some type of improved institutional structure to support the rule of law and provision of public goods. To investigate this possible pathway we use the per capita measures of kilometers of local roads and the number of health care locations within each cell—similar to public goods measures used in Desmet, Gomes, and Ortuño-Ortín (2020) and Seidel (2020).[6]

3.5 Mechanisms: Additional Outcomes from Microdata

We supplement the geospatial analysis of potential mechanisms by using individual-level data from 22 countries within the study area. These data are collapsed into district level—the second sub-national administrative level—observations when possible; when this is not possible regional—first level—boundaries are used. The microdata allow for analysis of additional public goods such as access to public utilities and education, differences in labor market decisions and the structure of the economy, and the prevalence of inter-group marriage as a measure of willingness to set aside traditional cultural practices and adopt new ideas and

[6]Population data for 2010 are from the Gridded Population of the World data series (CIESIN, 2018). Ethnicity location data are from the publicly available Ethnic Power Relations (EPR) Dataset (Vogt, Bormann, Rüegger, Cederman, Hunziker, and Girardin, 2015). Data on irrigated areas are from the Global Map of Irrigation Areas - Version 5 dataset produced by the Food and Agriculture Organization of the United Nations (FAO)'s GeoNetwork. Data for local roads are from the Global Roads Inventory Project (GRIP) dataset (Meijer, Huijbregts, Schotten, and Schipper, 2018), and health care locations are from OpenStreetMaps, and curated by healthsites.io.

technologies.

The microdata are selected using the following set of rules, and when possible, the most recent pre-2013 data are used to minimize any impact of the BRI. First, census data are selected from the International Public Use Microdata Series (IPUMS; Minnesota Population Center, 2018). Census microdata from IPUMS are used from all countries with districts within the study area that include information on sub-national administrative location. IPUMS is the first-choice data source because the data service also provides the corresponding administrative boundaries. Second-level district boundaries are used when available, and first-level regional definitions are substituted when necessary. If any country or variable category is not available at the district level, then Demographic and Health Surveys (DHS) are used (ICF International, 2005-2015). Where available, survey cluster GPS locations are matched to district level boundaries provided by the Global Administrative Unit Layers (GAUL) maps (FAO GeoNetwork, 2015). At the regional level, IPUMS data are preferred due to the larger number of observations within each boundary, and regional level variables from the DHS are matched with boundaries provided by the DHS spatial data repository.

In total, the combination of these datasets yields information on more than 39 million adults living in the 22 countries. District-level data from IPUMS are available for 14 countries, missing variables from IPUMS are supplemented by district-level DHS data in nine countries, regional-level IPUMS data in two countries, and regional-level DHS data for four additional countries. For a number of countries the specific data source used differs from variable to variable, depending on availability. The number of observations and administrative districts in each country are outlined in Appendix Section G.1, along with the combination of datasets used for each variable.

4 Estimation and Summary Statistics

4.1 Econometric Specification

The initial analysis in the paper estimates the model described by the following regression:

$$Y_{gvc} = \alpha + \beta \cdot \mathbb{1}\left[SR_{gvc}\right] + \delta X_{gvc} + \rho_v + \gamma_c + \varepsilon_{gvc} \tag{2}$$

Where Y is the value of the dependent variable within each cell g, with potential vegetation v, located in country c. The variable $\mathbb{1}\left[SR_{gvc}\right]$ is an indicator variable that denotes the presence of the Silk Roads in grid cell g. X_{gvc} is a vector of cell specific control variables, ρ_v is a set of fixed effects for each type of potential vegetation, and γ_c is a set of fixed effects for each modern day country.[7] Standard errors allow

[7] Control variables included in all estimates are distance to the nearest coastline and river, the cell average and standard

for spatial correlations between cells with centroids within 1,500 KM of one another (Conley, 1999, 2010) using a uniform weighting matrix, and are calculated using the procedure from Colella, Lalive, Sakalli, and Thoenig (2019).[8]

In the main analysis of the paper, the dependent variable is the natural log of the average of NLI plus 0.01 within each half-degree grid cell. Following Michalopoulos and Papaioannou (2013, 2014), Hodler and Raschky (2014), and Dickens (2018), this definition of the NLI variable helps adjust for the non-normal distribution of the data while allowing for the inclusion of cells that have no visible NLI. In equation (2), β estimates the percent difference of NLI in cells containing the Silk Roads relative to all other cells in the sample. The relative difference in NLI can then be expressed in terms of economic activity using the elasticity estimates from Table 1.

Spatial Discontinuity *(Main Specification)* After initially examining whether there are higher levels of economic activity along the Silk Roads using equation (2), the main econometric strategy used throughout the paper compares cells within successive 50 KM bins around the Silk Roads. This specification is similar to that from the previous equation; however, the single Silk Roads indicator variable is now replaced by a set of four indicators, each representing one of the four distance levels shown in Figure 1,

$$\begin{aligned} Y_{gvc} \quad = \quad & \alpha + \theta_1 \mathbb{1}\left[SR\left(0-50\,KM\right)_{gvc}\right] + \theta_2 \mathbb{1}\left[SR\left(50-100\,KM\right)_{gvc}\right] \\ & + \theta_3 \mathbb{1}\left[SR\left(100-150\,KM\right)_{gvc}\right] + \theta_4 \mathbb{1}\left[SR\left(150-200\,KM\right)_{gvc}\right] \\ & + \delta X_{gvc} + \rho_v + \gamma_c + \varepsilon_{gvc}. \end{aligned} \tag{3}$$

The coefficient θ_1 represents the difference in the given outcome (Y) in cells with centroids within 50 KM of the Silk Roads, relative to the group of reference cells further than 200 KM from the routes. This set of cells is represented by the empty area along the Silk Roads in Figure 1. The same interpretation can be given to each remaining θ coefficient, for the respective set of cells moving progressively further from the routes. The set of control variables and fixed effects remain the same as equation (2), and standard errors for equation (3) are estimated using the same Conley (1999, 2010) standard error procedure as previously described.[9]

Similar to Jedwab and Moradi (2016) and Jedwab et al. (2017), the identifying assumption of this model

deviation of elevation, land area, cubics for longitude and latitude, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city. Details on variable construction and the source for each of these measures can be found in Appendix Section A.

[8]The spatial correlation adjustments in Colella et al. (2019) expand on previous from Hsiang (2010).

[9]Kelly (2020) described the importance of considering spatial autocorrelation in the historical persistence literature. Although much of Kelly (2020) focused on country level analyses, to be prudent we follow each of the paper's main prescriptions. We include flexible latitude and longitude controls, spatial fixed effects, and Conley (1999, 2010) standard errors. Each of these prescriptions are adopted in our analysis; additionally, we demonstrate that the results from equation (3) are robust to the use of quadratic or quartic latitude and longitude controls, alternative spatial fixed effects (or no fixed effects), cutoff distances for the spatial kernel ranging from 500 to 3,000 KM, and the use of a Bartlett kernel over the same range.

is that the neighboring cells should have similar characteristics. Under this assumption, the crucial result from equation (3) is not the comparison of the area immediately adjacent to the Silk Roads to the reference area (θ_1), but the comparison of the area immediately adjacent to the Silk Roads relative to the neighboring area 50-100 KM from the route ($\theta_1 - \theta_2$). If the assumption that neighboring areas are similar holds, we can rule out pre-existing characteristics of the two adjacent areas as a cause of any difference between the coefficient of the first two distance bins.

4.2 Testing Balance in First Two Distance Bins and Further Identification

The differences between the first two distance bins—0 to 50 KM and 50 to 100 KM—are examined in Table 2. In Panel A, all non-desert cells are included in the sample. As a preview of the qualitative findings in the paper, at the top of the table we examine the unconditional mean of the logged value of NLI. Average NLI in 2010 is larger in the cells immediately surrounding the Silk Roads relative to the cells 50 to 100 KM away, a difference that is statistically significant at the 99 percent confidence level.[10] Following the visible patterns in Figure 2 and 3, this comparison adds additional evidence that is consistent with the conclusion that modern economic activity is higher in locations immediately along the Silk Roads.

The remainder of the comparisons test the underlying assumption of the model described in equation (3)—that the two bordering areas nearest to the Silk Roads have similar time invariant and pre-Silk Roads characteristics. However, the comparisons in Panel A yield consistent evidence that cells in the first two distance bins are not similar. Cells closer to the Silk Roads are lower in elevation, less rugged, more likely to contain a river, have higher levels of baseline population, are more likely to contain an ancient city, and are more suitable to agriculture and irrigation. With the consistent differences found in the bottom three sections of Panel A, for the full non-desert sample, it is not possible to conclude that the differences in NLI found at the top of the table are generated by proximity to the Silk Roads and not the differential in geographic and baseline population characteristics. However, following two simple rules likely to be associated with the location of the Silk Roads we identify a large and balanced sample.

In Panel B, on the right side of Table 2, we drop every cell containing a river and all cells with a centroid within 50 KM of the an ancient city. Once these two restrictions are in place, the remaining cells in the first two distance bins along the Silk Roads are similar along 11 of the 12 measured characteristics. In fact, remaining cells along the route are slightly higher in elevation and more rugged, slightly further from the coast and rivers, the population differences drop from roughly 0.25 standard deviations to 0.041 (greater than an 80% decline), and differences in agricultural suitability decline more than 50%. Importantly, the

[10] Desert cells are defined as cells in which the modal potential vegetation is tundra, desert, or polar desert. For each sample, counts of cells in each distance bin can be found in Appendix Table E.1, comparisons of additional modern outcomes can be found in Appendix Table E.2, and summary statistics for all cells within 500 KM of the Silk Roads in Appendix Table E.3.

differences in NLI remain large and statistically significant even with a balanced set of baseline and geographic characteristics. The main analysis used throughout this paper uses this balanced sample that satisfies the underlying identifying condition from equation (3). Additionally, the characteristics shown in Table 2 are included in the vector of control variables throughout the paper, and estimates of the impact of Silk Roads location on logged NLI are consistently similar, irrespective of the sample of cells selected.

4.2.1 Duration of Economic Activity - In and Away From Ancient City Centers

Estimates from equations (2) and (3) show evidence that NLI is higher in cells immediately adjacent to the Silk Roads. The hypothesis proposed in this paper is that the increased economic activity that exists today is a relic of the economic activity of merchants and their caravans carrying goods along the Silk Roads. If path dependence from areas of ancient economic activity is driving the higher levels of modern economic activity, we would expect to see larger returns in areas that were in use for a longer period of time. To test this hypothesis we use the location of 33 cities that are known to be part of the network during the first century CE, and define the ancient route from 2,000 years ago as the portion of Williams (2014, 2015)'s Silk Roads that connects these 33 ancient cities.[11] We test whether NLI along the ancient route is higher both relative to the reference area, and relative to the other paths along the Silk Roads. These results demonstrate that the increased levels of NLI are higher along the ancient portions of the Silk Roads—dating back to the first century CE—suggesting a cumulative effect of trade along the network of routes on today's economic activity.

4.2.2 Least Cost Path

One possible explanation for the positive association between the location of the Silk Roads and NLI is that the routes are known in areas that were more successful, and the routes that did not produce a similar effect have been lost to history.[12] To test whether the results are generated by selection, we construct a least

[11] The selected 33 ancient cities are largely based on writings of Han scholars of the period and translated by Hill (2003, 2015); the route to the west of Merv (37.66 N; 62.18 E) is defined cross referencing a number of additional sources (Dignas and Winter, 2007; Hansen, 2012; Williams, 2014; Barisitz, 2017; Whitfield, 2019). A map of the ancient cities and the route they define can be found in Appendix Figure E.1.

[12] Two characteristics of the scope of work by Williams (2014, 2015) make this argument less plausible. First, Williams (2014, 2015) identified over 150 sites that were part of the Silk Roads network that are no longer in use today, more than half of the dated sites in the dataset. Second, the main Silk Roads map used in this paper includes 40,068 KM of routes. Additionally, Williams (2014, 2015) identified 45,883 kilometers of secondary trade routes. For reference, the network of routes identified by Williams (2014, 2015) is longer than the entire interstate highway system in the United States. Any argument that these results are driven by selection would have to be an argument based on the importance of the 85,952nd most important kilometer of route; each kilometer that was more visible throughout history has been identified. It is unlikely that we should expect each and every kilometer of trade route to have a persistent effect throughout history. However, estimates using all 85,000 KM of routes from Williams (2014, 2015) yield the same patterns as estimates using the main routes. Although the estimates are smaller—which is unsurprising that secondary branches of the routes are not as important as the main routes—the results demonstrate that the extensive network of which we are aware, yields large and statistical significant results.

cost path (LCP) between the 33 ancient cities based only on the slope of the terrain and elevation.[13] This method yields an optimal walking path connecting the cities along the ancient Silk Roads, and yields a path that is similar to the routes selected by first century CE traders—as seen in the comparison of the ancient route and the LCP in Appendix Figure E.3.

Using this LCP and equation (3), we find evidence consistent with estimates using the actual location of the Silk Roads, suggesting that the results cannot be explained by a selection bias. However, with this method, we cannot differentiate between the benefits of geography that selects the location of the LCP and path dependence along the Silk Roads. To investigate the influence of geography, we also examine each 50 KM distance bin away from the LCP, to see if the effect of the main Silk Roads persists at different distances away from the LCP. We find that at every distance from the LCP, cells with the Silk Roads continuously have higher levels of NLI.

4.2.3 Placebo Routes

To gain some understanding of what a counterfactual route might look like, and to further ensure that the results are not being driven by geography we set up two falsification tests built around networks of placebo routes using the same LCP procedure described in Section 4.2.2. To give our placebo routes the largest possible chance of falling in areas with economic activity, we link two sets of known ancient cites with the LCP. The first route connects a set of 28 ancient cities from the 200 to 100 BCE period as defined by Reba et al. (2016). To construct the second placebo route we add an additional 20 cities from Reba et al. (2016) that existed during the years 500 to 200 BCE. In linking these ancient population centers, we are attempting to build routes that are the most likely alternatives to the Silk Roads for trade during this period.[14] In fact, the location of these placebo routes have several advantages relative to the Silk Roads: they are located at lower altitudes, in less rugged cells, with higher baseline populations, and higher levels of crop and irrigation suitability. The averages for cells containing each type of route can be found in Appendix Table E.4.

For both placebo routes we find no evidence of heightened economic activity along these geographically optimal paths. This demonstrates that optimal geography alone cannot explain the LCP results, although the cost layer that we use is able to generate results when linking ancient Silk Roads cities. Furthermore, these findings are initial evidence that our model is not identifying statistical noise, and is capable of producing a null result.

[13] The cost of overland movement between cities is calculated using information on slope and elevation, a function for the cost of walking across different slopes from Llobera and Sluckin (2007) and a penalty for elevations above 2000 meters suggested by Herzog (2014). The cost layer used can be seen in Appendix Figure E.2, and the route was constructed using ArcMap's cost connectivity tool.

[14] These 48 cities are a subset of the cities used to define the balanced sample in Table 2. The locations of the two sets of cities and the LCPs connecting them can be found in Appendix Figure E.4.

4.2.4 Random Routes

Finally, to ensure that our model is not simply capturing spatial noise in our data, we estimate again equation (3) using 2,000 routes with randomly selected nodes (i.e. cities) connected by the same LCP process as used in Section 4.2.2. We do this using two distinct sets of nodes, first we select 1,000 different combinations of 33 of the 48 cities from the 500 BCE to the 100 BCE period (Reba et al., 2016). We select 33 nodes to connect with the LCP because 33 cities are used to define the ancient route and we are able to show that NLI is significantly higher along an LCP between this number of Silk Roads cities. The second set of 1,000 routes each connect a different set of 33 randomly selected points within the boundary that outlines Figure 1. Compared to three different measures of the Silk Roads—the main Silk Roads; the ancient subset; or the LCP connections between ancient Silk Roads cities—none of the 2,000 random routes are able to match the magnitude of the increase in NLI within the first distance bin (θ_1) or the difference in NLI between the first two distance bins ($\theta_1 - \theta_2$).

5 Results

5.1 Silk Roads and Nighttime Light Intensity

The estimated impact of Silk Roads location on logged nighttime light intensity is shown in Table 3. The output from equation (2) is shown in columns (1) to (3), the coefficient of interest compares cells containing some portion of the Silk Roads to all other cells in the sample. The estimated coefficients for each of the four Silk Roads distance bins from equation (3) are shown in columns (4) to (6), along with the difference between the first two distance bins and the p-value from testing the null hypothesis that the coefficient values for these bins are equal. Each set of results are shown for the full sample of cells within 500 KM of the Silk Roads, columns (1) and (4), the sample of non-desert cells, columns (2) and (5), and the balanced sample in which cells are dropped if they contain a river or are within 50 KM of an ancient city, columns (3) and (6).

In each case, the model finds evidence of large increases in NLI for the cells closest to the historical location of the Silk Roads. In the first three columns, each sample of cells yields evidence that the NLI is between 87 and 96 percent higher along the Silk Roads. The type of cells included in the sample seems to have a minimal impact on the estimates; in fact, restricting the sample to the cells balanced on observables—column (3)—yields slightly a larger point estimate. This is evidence that time invariant and baseline characteristics have little impact in determining the relationship between the Silk Roads and modern economic activity.[15]

The results in the final three columns are similar. These estimates suggest that NLI intensity is between

[15]Estimates using NLI data with top coding corrections (Bluhm and Krause, 2018) yield similar results across all six columns, and can be found in Appendix Table E.5.

80 and 88 percent higher in cells with centroids within 50 KM of the Silk Roads, relative to the reference group of cells between 200 and 500 KM away from the routes. Importantly, we now see additional evidence of increased NLI in the cells closest to the routes relative to the neighboring cells, only 50 to 100 KM from the routes. The light intensity is found to be roughly 56% higher in the first distance bin relative to the second, a difference that is statistically significant at the 99 percent confidence level. Using the elasticity estimates from Table 1—relating NLI to GDP across the Eurasian region—the results in column (6) suggest that economic activity is 21% higher in areas near the Silk Roads, relative to the distant reference group.[16] A majority of this increase in economic activity occurs in the areas immediately adjacent to the routes, economic activity is 13% higher, relative to the cells in the second distance bin.[17]

In addition to the central finding of significant increases in economic production in areas immediately around the Silk Roads, the results in Table 3 match the patterns seen in Figure 3. The estimates fade towards zero in distance bins further away from the routes, with small coefficients in the furthest bin. Additionally, there is no evidence of the roads displacing economic activity. The evidence in Table 3 suggests an increase in economic activity in areas around the Silk Roads that was not generated by production moving away from the surrounding areas.

While the evidence in columns (3) and (6) of Table 3 suggest that persistent benefits of agglomeration within and nearby ancient cities is not driving the results, we further extend the exclusion radius around cities to ensure that this effect is not the cause of our results. From the 50 KM radius used in the balanced sample we extend the exclusion radius out to 200 KM, reestimating the model at each 50 KM interval. These three additional estimates—100 KM; 150 KM; 200 KM—yield no evidence of a weakening of the effect of the Silk Roads. The first distance bin remains statistically significant at the 99 percent confidence level with point estimates ranging from 0.808 to 0.834, and the difference between the first two distance bins also remain statistically significant at the 99 percent confidence level, ranging from 0.575 to 0.671.[18]

To this point the results have demonstrated that the findings along the Silk Roads are not a consequence of persistence around ancient city centers. However, it does raise the question of whether the benefits of the trade routes also exist in urban areas, or if the results we find are due to the creation of markets that

[16] The percentage change in economic activity, as measure by GDP, is approximately the percentage change in NLI $\times$ 0.243 = $\theta_1 \times 0.243 = 0.885 \times 0.243 = 0.215$.

[17] A number of robustness checks for these results are shown in the appendix, the above conclusions hold throughout the full set of alternative specifications. For each sample, equation (3) is reestimated using the year 0 CE population density, quadratic and quartic latitude and longitude polynomials, without country fixed effects, with sub-national regional fixed effects, and with fixed effects for the largest ethnicity. The output can found in Appendix Tables C.1 to C.3. Using the balanced sample and equation (3), alternative standard errors are calculated using spatial correlation ranges from 500 to 3,000 KM, t-statistics for the first distance bin are shown in Appendix Figure C.1. This is repeated using a Bartlett Kernel for the same set of ranges, t-statistics for the first distance bin are shown in Appendix Figure C.2. Finally, the model is reestimated for each year of available NLI data, the coefficient and 95 percent confidence interval of the first distance bin are shown in Appendix Figure C.3.

[18] The estimates can be found in Appendix Table E.6.

supported weary travelers and the resulting commerce that occurred in what would otherwise be desolate regions. The impact of the Silk Roads in and around ancient city centers—within 200 KM—is explored in detail in Appendix Section D. The 122 cities used throughout this study all existed in 100 BCE, or earlier. The evidence we have suggests ancient cities away from the Silk Roads were larger, on average, in both 1,000 BCE and 0 CE. This difference is consistently larger in the later—0 CE—period, suggesting that the cities along the Silk Roads were less developed during the initial period of the trade network. Similar to the evidence from Barjamovic et al. (2019), the effect of these ancient cities persists across millennia; cells with ancient cities away from the Silk Roads are found to have NLI that is 80% higher than cells without ancient cities that are also away from the trade routes. However, cells with both ancient cities and the Silk Roads have NLI levels that are 60% higher than the cells with ancients cities away from the Silk Roads; a result that is even larger when considering these cities were smaller during the early period of the Silk Roads. Furthermore, in the non-city cells within this radius around the ancient cities, the Silk Roads is again found to increase NLI by roughly 60%, relative to adjacent non-Silk Roads cells. The persistent benefits of the Silk Roads are found to be consistent and of a similar magnitude both away from and in the location of ancient cities, and are unaffected by the distance around these ancient urban centers that is dropped from the sample.

Economic Activity To test if increased economic activity along the Silk Roads is a plausible explanation for higher levels of NLI, we identify portions of the route in use during the earliest period of the Silk Roads trade network and examine whether the returns are larger in these areas. As outlined in Section 4.2.1, we do this by identifying 33 ancient cities that participated in trade along the routes in the first century CE and define the routes between these cities as the ancient portion of the Silk Roads. The location of these cities and the subsection of the route denoted as the Ancient Silk Roads can be found in Appendix Figure E.1. To test for differences between the ancient subset of routes and the complete network of paths along the Silk Roads, equation (3) is modified by including four additional distance bins for the distance from the ancient route.

The ancient route is defined as a strict subset of the main routes; therefore, the distance bins for the ancient routes can be constructed as an interaction between the distance bins for the main route and an indicator for whether the portion of the route that defines the distance bin is part of the ancient network. In this set up, the coefficients on the distance bins for the ancient route are directly interpreted as the difference in logged NLI for bins of a given distance along ancient routes relative to the same distance along the other portions of the main route. The estimated effect of the ancient portion of the routes at each distance is then the sum of the coefficients for the main route and the coefficient for the same distance from the ancient

route.

Using the balanced sample, the results for this modified version of equation (3) are shown in Table 4. All estimates in the table are from a single regression. The patterns in the results for distance from the main routes—column (1a)—are the same as those in Table 3. NLI is higher in areas immediately surrounding the Silk Roads, relative to both the second distance bin and the reference group; however, the magnitude of the effect is slightly muted with the removal of the ancient portions of the network. In column (1b), there is evidence that NLI is 38.8% higher in areas adjacent to the ancient routes, relative to similar areas near other portions of the trade network; this difference is statistically significant at the 99 percent confidence level. Finally, in the last column, the two sets of estimates are combined to yield the total effect along the ancient portion of the route. NLI is more than 100% higher along the ancient portion of the route relative to the reference area, and 63.7% higher relative to the second distance bin. With this increased duration of built up of economic activity, there is also evidence of spatial dispersion of the effect out to 150 KM away from the routes.

The evidence in this subsection links economic activity along the Silk Roads 2,000 years ago—during the first century CE—to economic activity today. It is consistent with the theory that places in which trade occurred along the Silk Roads over a longer duration have higher levels of economic activity today. These areas also see a more dispersed effect, out to 150 KM away from the Silk Roads. The set of findings in Tables 3 and 4, along with evidence of results both around, in, and away from ancient cities provide evidence that it was the actual trade that occurred along the Silk Roads network that has led to increased levels of economic activity. The gains are found to be similar along the trade routes, where trade promoted investments in new infrastructure such as caravansaries to support the movement of people and goods, around cities when the exclusion radii are shortened, and when directly comparing locations of ancient cities. In each scenario, the location with the Silk Roads have higher levels of nighttime light intensity, even centuries later.

Least Cost Path To rule out selection as a potential explanation of the relationship between the Silk Roads and NLI, the 33 ancient cities along Silk Roads are used to construct a slope and elevation based least cost path. This approximation for the ancient routes is explained in more detail in Section 4.2.2, and the paths can be seen in Appendix Figure E.3. Using the balanced sample and the optimal path connecting the ancient Silk Roads cities, the estimates produced by equation (3) can be found in column (1) of Table 5. These estimates yield a pattern similar to those from the ancient route they are approximating. The estimate in the first distance bin along the least cost path is large, with differences from both the reference group and the second distance bin that are both statistically significant. The results in column (1) are evidence that selection is not driving the Silk Roads estimates. These results are based on a path that is similar to the

ancient path, but predicted only using geographic features. Any type of selection effects cannot be driving these results.

There are two important characteristics of note from the results in column (1) of Table 5. First, while the LCP is able to rule out selection, it cannot differentiate between the effect of geography and the Silk Roads. Second, unlike previous results, the estimates do not fade to zero by the fourth distance bin. To address both of these points a new model is estimated in which each of distance bins around the LCP are interacted with an indicator that the cell is within 50 KM of the Silk Roads. The results are shown in Figure 4 for both non-desert and balanced samples.

As denoted by the black squares, NLI is consistently higher along the Silk Roads than the reference group, evidence that the Silk Roads do not rely on the ideal geography along the LCP.[19] In fact, NLI is higher in cells along the Silk Roads even relative to with the cells the same distance from the optimal path. Although this difference is not statistically significant in the cells immediately around the LCP, it is statistically significant at the 95 percent confidence level in seven of the eight other pairs, and at the 90 percent confidence level in the final grouping.

The results in Table 5 and Figure 4 provide evidence against two possible alternative explanations for the Silk Roads' association with NLI. The results are shown not to be driven by selection bias, as they remain similar when using an LCP to connect ancient cities that were part of trade along the routes during the first century CE. Additionally, they are likely not driven by geography, the effect of the Silk Roads on NLI persists at each distance from the LCP. It is also important to reiterate that when using the balanced sample, the first two distance bins from equation (3) are balanced across their geographic, climatic, and baseline characteristics. Any difference between those two bins cannot be attributed to geography.

Placebo Routes We construct a pair of falsification tests built around placebo routes between two sets of ancient cities using the same LCP procedure. The first set of ancient cities includes those with populations over 30,000 between the years 200 and 100 BCE (Reba et al., 2016), which includes 28 cities. The second set of cities expands the time frame back to 500 BCE, adding an additional 20 cities. This is our best estimate of an alternative trade route that would have existed during this early period. It is worth noting that these cities are larger than those along the Silk Roads, and that we know from the previous section that this LCP cost layer is capable of defining a route with high levels of NLI. Interestingly, the ancient cities from Reba et al. (2016) do have some overlap with the Silk Roads, 10 of the 48 cities are represented in the set of cities used to define the ancient Silk Roads; in total, 19 fall somewhere along the full extent of the Silk Roads.

[19] These findings also suggest why the results in column (1) of Table 5 do not decline towards zero. There are Silk Roads in each of the distance bins increasing the level of modern NLI. In the four distance bins closest to the LCR the sample size of the two groups is of a similar magnitude, while the reference group in the furthest distance bin contains between 7.5 and 10 times the number of cells compared to the area near the Silk Roads.

This intersection likely puts upward pressure on the placebo route estimates, and increases the probability of finding positive association with NLI. Alternatively, failing to find a statistically significant relationship between these placebo routes and NLI would further emphasize the precision of the model. The two sets of cities and placebo routes are mapped in Appendix Figure E.4, along with the Silk Roads.

Estimates using these placebo routes can be found in columns (2) and (3) of Table 5, for the 200 to 100 BCE and the 500 to 100 BCE set of cities, respectively. While the placebo routes yield positive estimates in the first distance bin, the largest estimate, in column (3), is roughly a quarter of the magnitude of the effect of the actual Silk Roads on NLI. Furthermore, none of the distance bin estimates are statistically significant, nor the differences between the first two distance bins. These estimates show that even if there is some overlap between the set of cities along the network, and therefore the location of routes themselves, the optimal network between major ancient cities does not yield a significant relationship with NLI. This is further evidence that the Silk Roads' results cannot be explained by ancient patterns of economic development and urbanicity, and these results are the initial evidence that the model does not generate large estimates for incorrectly placed routes.

Random Routes To ensure that our model is not capturing spatial noise, as described in Section 4.2.4, we build 2,000 additional LCP routes. These routes are constructed in two ways, the first 1,000 routes each connect a randomly selected 33 ancient cities from the set of 48 cities used above (Reba et al., 2016). The second set of routes use the LCP method to connect 33 random points placed throughout the study region, again this is done 1,000 times. We then reestimate equation (3) for each of these routes.

In Figures 5 and 6, for random cities and points, respectively, we study two characteristics of these paths. First, in panel (a) of each figure the density of the coefficient estimate of the first distance bin is shown, along with the corresponding estimates from our three Silk Roads models. For the cities, Figure 5a, this is a rather narrow distribution centered slightly above zero. For the random points, the distribution is much more dispersed and is centered around zero. In both cases, the coefficients from the actual Silk Roads lie significantly beyond the distributions. Outside of the paths used along the Silk Roads, there is no evidence that an alternative path definition—even when constructed using an LCP algorithm—lies along an area with a similar concentration of NLI. This magnitude of the Silk Roads estimates are unique to the actual trade routes, and the econometric model does not reproduce similar estimates using either type of randomly placed route.

Second, we also examine the magnitude of the difference between the first and second distance bins. Not only is there a concentration of NLI along the Silk Roads, but due to the fade out in the effect as distance from the routes grows, the estimate for the actual Silk Roads is significantly larger in the first distance bin.

In panel (b) of both figures we plot the density distribution of the random routes using both the magnitude of the coefficient on the first distance bin (x-axis) and the difference between the coefficients on the first two distance bins (y-axis). The values corresponding to the three Silk Roads estimates are also shown. In both cases, there is no example of a random route being able to produce the characteristics shown in the Silk Roads estimates. In fact, not only are the magnitudes not matched by the random LCPs, but of the 2,000 routes only 39 have (i) a positive coefficients for the first two distance bins, with (ii) a first distance bin coefficient larger than the second, and (iii) a difference that is statistically significant between the two distance bins, even at the 90 percent confidence level. These results show that paths connecting the random nodes are not able to reproduce the magnitude of the estimated effect of the Silk Roads on NLI; furthermore, the characteristics defining the differences between the first two distance bins in the Silk Roads estimates are exceedingly hard to produce with these random routes.

The Silk Roads estimates follow an expected pattern. Nighttime light intensity is higher in the areas where the Silk Roads trade network was once active, and the effect fades out over roughly 200 KM. We show that NLI is significantly higher in cells within 50 KM of the Silk Roads relative to those 50 to 100 KM away, even when the areas have the same observable characteristics. This pattern exists in a similar way both when the area around ancient cities are excluded, and in the area immediately surrounding the ancient cities; evidence that the results are not driven by persistence around ancient economic centers. In fact, routes connecting 1,000 different combinations of ancient cities fail to yield results similar to those found along the Silk Roads. The Silk Roads connected the old world in a unique way, and areas in which trade occurred along these routes during its earliest period—2,000 years ago—had even larger returns to modern day economic activity than other portions of the routes. Additional results have demonstrated that the results cannot be a product of geographical advantage, selection effects, or spatial noise.

5.2 Mechanisms: Silk Roads and Modern Outcomes

The above results link trade along the Silk Roads as early as the first century CE to today's economic activity. However, the mechanisms through which ancient economic activity influences the distribution of today's NLI remains unclear. To further investigate this we explore five potential mechanisms. For each of these mechanisms, we estimate a subset of the models used throughout Section 5.1 to examine whether the Silk Roads are able to explain variation in these modern mechanisms. We then test to see if these mechanisms are able to mediate the relationship between Silk Roads and NLI. These two pieces of evidence together lend insight into which mechanisms are the most likely paths through which the Silk Roads have led to increased levels of modern day economic activity.

First, we look at population density in 2010 CE, to explore the possibility that the areas along the Silk Roads are simply more populated, and therefore have higher levels of economic activity. Second, we examine whether the Silk Roads led to increased ethnic diversity, as measured by the number of ethnicities present. An increase in the number of peoples from different backgrounds could help spur innovation leading to increased levels of economic growth.

The next two potential explanations are motivated by the idea that as areas around the Silk Roads grew they would need increased levels of technology and institutional quality to be able to support the passing merchants, or possibly growing populations. We measure technology using an indicator for whether a cell contains irrigated agriculture. Second, institutions are proxied by the provision of public goods as measured by kilometers of local roads and the number of health care locations, each per 1,000 people.

Finally, we explore the possibility that the connections established by the routes have persisted over time, and that areas that were once part of the ancient trade network remain more connected to international markets today. We measure this outcome using two separate indicators for whether a cell contains a major highway and whether a cell contains a railroad. We also investigate whether these are two separate results, and not an artifact of the Silk Roads, railroads, and highways following the same path along the routes.

5.2.1 The Silk Roads and Modern Outcomes

For any of these mechanisms to help explain the link between Silk Roads and NLI, they must also be outcomes that were influenced by the location of the Silk Roads. To examine the relationship between the Silk Roads location and seven different modern variables, representing the five mechanism categories, we estimate four previously used versions of equation (3) with these new dependent variables. The output is shown in Table 6.

For reference, the estimates using logged NLI as the dependent variable are shown in column (1), and the estimates for each of the alternative outcomes are shown in columns (2) through (8). In Panel A, the estimates for the first distance bin are shown using the non-desert sample of cells within 500 KM of the Silk Roads, along with the difference between the first and second distance bins. Estimates for the balanced sample are shown in Panel B. The difference between the first two panels highlights the consequence of removing ancient cities from the sample, and the role that persistence within urban centers plays in increasing the size of the coefficient estimates. The estimates in Panel C use the LCP connecting the ancient cities defined by Reba et al. (2016) for the 500 to 100 BCE period; these placebo routes are again used to search for evidence of a relationship between non-Silk Roads urban centers and modern outcomes. Finally, the value in Panel D represents the share of first distance bin estimates that are at least as large as those in Panel B when using

LCP routes between 1,000 sets of random points.[20]

In Panels A and B, five of the seven mechanism variables show evidence of increased levels along the Silk Roads. The two that show no evidence of a relationship with the Silk Roads are local roads and health care sites per 1,000 people, suggesting that increases in the provision of public goods and institutional quality is an unlikely explanation of why the Silk Roads continue to impact today's economic activity. The estimates for population density evidence that the result may be driven by persistence within urban centers. Not only is there significant decline in the magnitude of the estimates for the first distance bin between the first two panels, but also a more than 50% decline in the difference between the first two bins after ancient cities are dropped from the sample. Additionally, the coefficient along the placebo route is a similar magnitude to the result in Panel B, further suggests that persistence in areas surrounding ancient urban centers may have a role to play in explaining the effect on population density.

For each of the four remaining outcomes the difference between the first two distance bins is positive, large, and statistically significant. There is no significant evidence that these estimates can be explained by either the placebo or randomly placed routes.[21] Each of these four outcomes have higher levels directly along the Silk Roads, show no evidence of being driven by persistence in areas with ancient population centers, and cannot be explained by routes not associated with the Silk Roads. Technology and transportation infrastructure are the strongest initial candidates for mechanisms through which the Silk Roads impact today's economic activity. There is also evidence of increased ethnic diversity consistent with the conventional wisdom that people from different backgrounds often interacted along the Silk Roads, bringing with them new ideas, cultures, and religions (Hansen, 2012; Renfrew, 2014; Williams, 2014; Barisitz, 2017; Michalopoulos et al., 2018)—a plausible mechanism that will be investigated in more detail in Section 5.2.3.

5.2.2 Pathways from the Silk Roads to Economic Activity

The relationship between the Silk Roads and each of the modern outcomes is a necessary condition for the mechanism to be a link to NLI; however, this relationship alone is not sufficient. In Table 7, we again use equation (3) and the balanced sample, to explore each mechanism's ability to mediate the relationship between the Silk Roads and NLI. In all columns the dependent variable is logged NLI; for reference, the baseline estimate using no extra controls is shown in column (1). In columns (2) through (8), the mechanism

[20] The coefficient estimates for all four distance bins of the Panel A to C models can be found in Appendix Section F.1, along with figures of the distribution of the first bin's coefficient value for the 1,000 random point LCPs.

[21] To ensure that the overlap between major highways, railroads, and the Silk Roads is not generated by the three networks lying in exactly the sample place, possibly due to geographic constraints, the model is reestimated with the railroad indicator variable added as a control to the highway model, and vice versa. The estimates remain large and statistically significant for both major highways and railroads. These results can be found in Appendix Tables F.6 and F.7. Estimates for alternative measures of major highway and railroad—the logged distance to each type of transport—are also included in Appendix Tables F.8 and F.9, respectively.

denoted at the top of each column is added to the regression as a control variable; the full set of coefficients for all four distance bins are available in Appendix Table F.10.

The largest reductions in the magnitude of the coefficient for the first distance bin are generated by the inclusion of the irrigation, major highway and railroad controls. Each of the variables, measures of agricultural technology and transportation infrastructure, reduces the estimated relationship between the Silk Roads and NLI by at least 20%. Conversely, the two measures for local public goods provision have little impact on the coefficient of the first distance bin. Both population variables in the first two columns have some mediating effect, but these effects are much smaller than those of irrigation, major highway, and railroad.

These estimates reinforce the importance in both technology and transportation infrastructure as potential mechanisms through which the Silk Roads impact NLI and modern day economic activity. While evidence suggests that the number of ethnicities in an area is higher along the Silk Roads, this measure does not seem to be related to changes in economic activity. Importantly, the results in column (2) demonstrate that the increased levels of NLI along the Silk Roads are not generated by higher levels of population density, even with this control added to the regression, NLI is 51.7% higher in the first distance bin along the Silk Roads, relative to the second.[22]

5.2.3 Mechanisms: Additional Outcomes from Microdata

Further investigation of the mechanisms that link the Silk Roads and modern economic activity using geospatial variables is not possible due to the limited availability of geospatial data. In this subsection, we attempt to gain a deeper understanding of this connection by analyzing detailed microdata on the availability of public utilities, education, occupational choices, and the prevalence of inter-group marriage. To integrate these data with the Silk Roads analysis the individual data are collapsed into either first- or second-level sub-national administrative districts, as described in Section 3.5. These district observations are then used in a descriptive analysis to further investigate the characteristics of areas closest to the Silk Roads. Although we no longer have an arbitrary and consistent set of boundaries, this analysis provides insight into the characteristics that are most associated with proximity to the ancient trade route.

[22]To ensure that the estimated mediation effects seen in Table 7 are not specific to a single set of controls we select five variables—population density; number of ethnicities; irrigation; roads; railroads—one from each category. We then run 32 estimates, beginning with the baseline of no controls, and then adding every combination of the five control variables. We do this with one control at a time, two at a time, and so on. The estimates can be found in Appendix Table F.11. The mediation effect of adding each variable is then measured relative to the same model without that specific variable. For example, we compare the coefficient estimate on the first distance bin with population density and ethnicity controls added to the model (0.752***), to the model with population density, ethnicity, and railroad controls (0.482**). We then calculate the mediation effect of adding the railroad control to this model to be 35.9%. After the addition of each variable, we take the average across all 31 models with controls and calculate the average mediating effect for each of the five variables. For all five variables, the average mediation effect is within 1.2 percentage points of those shown in Table 7.

The econometric model used is similar to equation (3):

$$\begin{aligned} Y_{dvc} &= \alpha + \theta_1 \mathbb{1}\left[\bar{SR}\left(0-50\,KM\right)_{dvc}\right] + \theta_2 \mathbb{1}\left[\bar{SR}\left(50-100\,KM\right)_{dvc}\right] \\ &\quad + \theta_3 \mathbb{1}\left[\bar{SR}\left(100-200\,KM\right)_{dvc}\right] + \delta X_{dvc} + \rho_v + \gamma_c + \varepsilon_{dvc}. \end{aligned} \tag{4}$$

Distance from the Silk Roads $\left(\bar{SR}\right)$ is defined as the population weighted distance within each administrative district, d.[23] The first two distance bins are the same as equation (3), but the final distance bin is condensed into a single group to better balance the number of areas within each group, as seen in Appendix Table G.3. The set of control variables and fixed effects are also the same as in equation (3), and the analysis focuses on districts with a population weighted distance within 500 KM of the Silk Roads. When using microdata outcomes, districts are weighted by population, and all standard errors are again calculated using the spatial corrections suggested by Conley (1999, 2010).

To asses whether the analysis using administrative district observations is comparable to the previous sets of results we first estimate equation (4) using logged nightlights and the same set of GIS mechanisms previously analyzed in Sections 5.2.1 and 5.2.2 as the dependent variables. The only adjustment is for the irrigation variable. More than 99% of administrative districts in the sample have some irrigated land; therefore, the share of all land area irrigated is used to measure irrigation adoption. These estimates are shown in Table 8.

Encouragingly, the results are similar to those seen in the previous analysis. The estimates for NLI, major highway location, irrigation, and number of ethnicities all yield large and statistically significant estimates in the first distance bin, and the estimates fade towards zero as the distance from the Silk Roads increases. Estimates for both population and railroad location yield a similar pattern, but without generating statistically significant estimates. As with the previous set of results, there continues to be no evidence of increased access to local roads or healthcare in districts along the Silk Roads. These results demonstrate that this alternative set of district observations are able to capture the same patterns along the Silk Roads as the arbitrarily placed set of grid cells.

The next three tables—Tables 9 to 11—examine the variables derived from individual level data and collapsed into administrative district observations. The number of observations in each sample can vary widely depending on data availability; however, to ensure that the relationship between the Silk Roads

[23] The population weighted distance for each administrative district is calculated using the following procedure. First, using a raster of 30-arc-second population cells—about 1 sq. km.—from the Gridded Population of the World (CIESIN, 2018) the distance from each cell to the Silk Roads is calculated. Within each district the population weighted distance $\left(\bar{SR}_d\right)$ is then constructed using the following equation: $\bar{SR}_d = \left(1/\sum_g Pop_g\right) \sum_g Dist_g \times Pop_g$. Where Pop_g is the population within each 30-arc-second grid cell, and $Dist_g$ is the measure of kilometers from cell g to the Silk Roads.

and NLI persists within each of these samples the model is reestimated using five samples defined by the availability of different microdata outcomes. The coefficient estimate for the first distance bin is the largest and statistically significant at the 95 percent confidence level in four of the five samples, and the coefficients are smaller for each successive distance bin in all samples. Finally, all estimates shown in this section are redone after dropping districts that include the location of ancient cities to ensure that results are not driven by persistence around these ancient urban areas, and all microdata outcomes are also reestimated after controlling for logged NLI to ensure the estimates are not a product of differential levels of economic development. All findings discussed in this section hold in these alternative specifications, and the output can be found in Appendix Section G.2.

In Table 9, we utilize the microdata to further investigate measures of institutional quality and public goods provision. In the first three columns the outcome of interest is whether a household has access to different types of public utilities—electricity, piped water, and public sewers. In the next three columns, we look at different measures of education—literacy and the completion of primary and secondary school. Finally, we examine health outcomes in the final column, using the survival rate of children. Due to data availability, the survival rate is not conditional on age.[24] Across all seven measures, there is no consistent evidence of increased access to these public goods in districts closest to the Silk Roads.

The results throughout the paper have shown higher levels of economic activity along the Silk Roads, but we do not know how the labor market or the distribution of economic activity might differ in these more active regions. In Table 10, we further investigate the structure of these economies using data on individual employment outcomes. In column (1), there is evidence of a significant reduction in the likelihood that an adult is working in areas closest to the Silk Roads. This result suggests that productivity is significantly higher in the areas along the Silk Roads. In Table 7, it was shown that population density was unable to explain increases in NLI, and now in this section we find evidence that NLI is again significantly higher in districts around the Silk Roads while adults are less likely to be working.

At first this decline in employment may be puzzling; however, the results in columns (2) and (4) highlight that the decline is coming entirely from the agricultural sectors of the economy. It is important to note this does not mean that the districts are simply more urban; although, the number of individuals employed in agriculture is lower, the results from Table 8 demonstrate that a larger portion of these districts are (more productive) irrigated cropland. The results in column (3) show that there is no change in the share of adults working in the manufacturing jobs, while the estimate in column (5) suggests a significant increase in the

[24]The data on age of death for children necessary to calculated infant or under-5 mortality rates is only available in the DHS. However, the DHS does not have data from some of the most important countries along the trade routes, such as China, Iran, and Iraq. For this reason, we use the survival rate that is not conditioned on age. It is important to note that in the DHS surveys in which both the (under-1) infant survival rate and the unconditional survival rate can be calculated over two-thirds of all reported deaths of children occur by the age of one.

size of the industry. Taken together these results—(i) increased economic activity with a smaller workforce; (ii) expanded use of irrigation with significant declines in the agricultural workforce; (iii) expansion of the manufacturing industry without an increase in the number of manufacturers—produce evidence that productivity is higher in the areas around the Silk Roads.

One possible channel through which the Silk Roads may have led to increased access to new ideas and technologies could be through the increased interactions with other cultures and peoples. In Tables 6 and 8, we have seen evidence that there are an increased number of ethnicities in areas around the Silk Roads. While the presence of new perspectives and new ideas is valuable, it is not itself sufficient to promote the adoption of new technologies; there must the an openness and willingness on the part of the local community and industry to accept new ideas. For evidence of this, we explore rates of inter-group marriage in districts along the Silk Roads as a proxy for the weakening of boundaries between different populations.

In Table 11, we again see evidence that districts along the Silk Roads have more diverse ethnic populations—columns (3) and (4)—but not higher shares of foreign-born individuals—column (1). This suggests that the increases in ethnic diversity are not a phenomenon of the current generation. Most importantly, the estimates in columns (2) and (5) yield evidence that, in districts closest to the Silk Roads, natives are more likely to marry foreign-born individuals and the share of inter-ethnic marriages are higher. These increases in inter-group marriage suggests a decline in the social distance between groups, and possibly an increase in the willingness to adopt new productivity enhancing ideas. This openness could be a mechanism through which these districts have become more productive, through the adoption of new technologies and production methods from foreign merchants over the past two millennia.

6 Conclusion

This paper demonstrates that areas closer to the Silk Roads have higher contemporary levels of nighttime light intensity, are more proximate to major highways and railroads, and have higher levels of technology associated with the trading network. The magnitude of each of these relationships fades as the distance from the trade route increases. The results are stronger in areas in which the duration of trade was the longest, they exist both away from and within ancient city centers, are not explained by population densities, and cannot be matched by a set of over 2,000 placebo routes.

Future projects could focus on developing a theoretical framework to study the dispersion of technology along the routes, work to build a more explicit link between the adoption of culture and technology, and exploit ancient city locations and empire boundaries to investigate the importance of historical trade costs.

Our study finds evidence of significant progress in economic activity without the institutional investment

that is the focus of much of the previous literature. This suggests that there could be alternative innovation channels through which significant economic progress can be made, such as access to external trade networks. A deeper exploration of cases in which ancient institutions are not deterministic can have important implications for policy both in terms of understanding the nuances of the effects of the past upon the present and also in terms of implementing enduring policy actions that continue to be centered on access to trade and markets.

References

Acemoglu, Daron, Simon Johnson, and James A Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Acemoglu, Daron, Simon Johnson, and James A Robinson. 2005. "Institutions as a Fundamental Cause of Long-run Growth." *Handbook of Economic Growth* 1 385–472.

Alder, Simon. 2019. "Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development." Working Paper.

Baniya, Suprabha, Nadia Rocha, and Michele Ruta. 2020. "Trade Effects of the New Silk Road: A Gravity Analysis." *Journal of Development Economics* 146 102467.

Barisitz, Stephan. 2017. *Central Asia and the Silk Road. Economic Rise and Decline over Several Millenia*.

Barjamovic, Gojko, Thomas Chaney, Kerem Coşar, and Ali Hortaçsu. 2019. "Trade, Merchants, and the Lost Cities of the Bronze Age." *Quarterly Journal of Economics* 134 (3): 1455–1503.

Bentzen, Jeanet Sinding, Nicolai Kaarsen, and Asger Moll Wingender. 2017. "Irrigation and Autocracy." *Journal of the European Economic Association* 15 (1): 1–53.

Berger, Thor, and Kerstin Enflo. 2017. "Locomotives of Local Growth: The Short-and Long-term Impact of Railroads in Sweden." *Journal of Urban Economics* 98 124–138.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig. 2017. "This Mine is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review* 107 (6): 1564–1610.

Bird, Julia, Mathilde Lebrand, and Anthony J Venables. 2020. "The Belt and Road Initiative: Reshaping Economic Geography in Central Asia?." *Journal of Development Economics* 144 102441.

Blaydes, Lisa, and Christopher Paik. 2020. "Trade and Political Fragmentation on the Silk Roads: The Economic Effects of Historical Exchange between China and the Muslim East." *American Journal of Political Science* Forthcoming.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127 (2): 587–644.

Bluhm, Richard, and Melanie Krause. 2018. "Top Lights-bright Cities and Their Contribution to Economic Development." CESifo Working Paper No. 7411.

Center for International Earth Science Information Network (CIESIN) at Columbia University. 2018. "Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals, Revision 11."

Chanda, Areendam, and Sujana Kabiraj. 2020. "Shedding Light on Regional Growth and Convergence in India." *World Development* 133 104961.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig. 2019. "Inference with Arbitrary Clustering." IZA Discussion Paper No. 12584.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.

Conley, Timothy G. 2010. "Spatial Econometrics." In *Microeconometrics. The New Palgrave Economics Collection*, 303–313, Springer.

Coşar, A Kerem, and Banu Demir. 2016. "Domestic Road Infrastructure and International Trade: Evidence from Turkey." *Journal of Development Economics* 118 232–244.

Dalgaard, Carl-Johan, Nicolai Kaarsen, Ola Olsson, and Pablo Selaya. 2018. "Roman Roads to Prosperity: Persistence and Non-persistence of Public Goods Provision." CEPR Discussion Paper No. 12745.

Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (6): 1863–1903.

Desmet, Klaus, Joseph Gomes, and Ignacio Ortuño-Ortín. 2020. "The Geography of Linguistic Diversity and the Provision of Public Goods." *Journal of Development Economics* 143 102384.

Dickens, Andrew. 2018. "Ethnolinguistic Favoritism in African Politics." *American Economic Journal: Applied Economics* 10 (3): 370–402.

Dignas, Beate, and Engelbert Winter. 2007. *Rome and Persia in Late Antiquity: Neighbours and Rivals*. Cambridge University Press.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review* 108 (4-5): 899–934.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *Quarterly Journal of Economics* 131 (2): 799–858.

Elvidge, Christopher D, Daniel Ziskin, Kimberly E Baugh, Benjamin T Tuttle, Tilottama Ghosh, Dee W Pack, Edward H Erwin, and Mikhail Zhizhin. 2009. "A Fifteen Year Record of Global Natural Gas Flaring Derived from Satellite Data." *Energies* 2 (3): 595–622.

FAO GeoNetwork. 2015. "Global Administrative Unit Layers (GAUL)." Food and Agriculture Organization of the United Nations.

Fick, Stephen E, and Robert J Hijmans. 2017. "WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37 (12): 4302–4315.

Flückiger, Matthias, Erik Hornung, Mario Larch, Markus Ludwig, and Allard Mees. 2019. "Roman Transport Network Connectivity and Economic Integration." CESifo Working Papers No. 7740.

Gibson, John, Susan Olivia, and Geua Boe-Gibson. 2020. "Night lights in economics: Sources and uses." *Journal of Economic Surveys* 34 (5): 955–980.

Giuliano, Paola, and Nathan Nunn. 2021. "Understanding Cultural Persistence and Change." *Review of Economic Studies*, Forthcoming.

Hansen, Valerie. 2012. *The Silk Road: A New History*. Oxford University Press.

Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102 (2): 994–1028.

Herzog, Irmela. 2014. "Least-cost Paths–Some Methodological Issues." *Internet Archaeology* 36.

Hill, John E. 2003. *The Western Regions according to the Hou Hanshu*.

Hill, John E. 2015. *Through the Jade Gate: China to Rome: A Study of the Silk Routes During the Later Han Dynasty 1st to 2nd Centuries CE*. CreateSpace Independent Publishing.

Hodler, Roland, and Paul A Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics* 129 (2): 995–1033.

Hornung, Erik. 2015. "Railroads and Growth in Prussia." *Journal of the European Economic Association* 13 (4): 699–736.

Hsiang, Solomon M. 2010. "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America." *Proceedings of the National Academy of sciences* 107 (35): 15367–15372.

ICF International. 2005-2015. "Demographic and Health Surveys (various) [Datasets]." Funded by USAID. Rockville, Maryland: ICF [Distributor].

Jedwab, Remi, Edward Kerby, and Alexander Moradi. 2017. "History, Path Dependence and Development: Evidence from Colonial Railways, Settlers and Cities in Kenya." *Economic Journal* 127 (603): 1467–1494.

Jedwab, Remi, and Alexander Moradi. 2016. "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa." *Review of Economics and Statistics* 98 (2): 268–284.

Kelly, Morgan. 2020. "Understanding Persistence." CEPR Discussion Paper No. DP15246.

Klein Goldewijk, Kees, Arthur Beusen, and Peter Janssen. 2010. "Long-term Dynamic Modeling of Global Population and Built-up Area in a Spatially Explicit Way: HYDE 3.1." *The Holocene* 20 (4): 565–573.

König, Michael D, Dominic Rohner, Mathias Thoenig, and Fabrizio Zilibotti. 2017. "Networks in Conflict: Theory and Evidence from the Great War of Africa." *Econometrica* 85 (4): 1093–1132.

Lessmann, Christian, and André Seidel. 2017. "Regional Inequality, Convergence, and its Determinants–A View from Outer Space." *European Economic Review* 92 110–132.

Li, Yuqi, Michael J Storozum, Xin Wang, and Wu Guo. 2017. "Early Irrigation and Agropastoralism at Mohuchahangoukou (MGK), Xinjiang, China." *Archaeological Research in Asia* 12 23–32.

Llobera, Marcos, and Tim J Sluckin. 2007. "Zigzagging: Theoretical Insights on Climbing Strategies." *Journal of Theoretical Biology* 249 (2): 206–217.

Maloney, William F, and Felipe Valencia Caicedo. 2016. "The Persistence of (Subnational) Fortune." *Economic Journal* 126 (598): 2363–2401.

Meijer, Johan R, Mark AJ Huijbregts, Kees CGJ Schotten, and Aafke M Schipper. 2018. "Global Patterns of Current and Future Road Infrastructure." *Environmental Research Letters* 13 (6): 064006.

Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review* 102 (4): 1508–1539.

Michalopoulos, Stelios, Alireza Naghavi, and Giovanni Prarolo. 2018. "Trade and Geography in the Spread of Islam." *Economic Journal* 128 (616): 3210–3241.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. "Pre-colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81 (1): 113–152.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics* 129 (1): 151–213.

Minnesota Population Center. 2018. "Integrated Public Use Microdata Series, International: Version 7.1 [Dataset]." Minneapolis, MN: IPUMS.

Nordhaus, William, and Xi Chen. 2015. "A Sharper Image? Estimates of the Precision of Nighttime Lights as a Proxy for Economic Statistics." *Journal of Economic Geography* 15 (1): 217–246.

Ramankutty, Navin, and Jonathan A Foley. 1999. "Estimating Historical Changes in Global Land Cover: Croplands from 1700 to 1992." *Global Biogeochemical Cycles* 13 (4): 997–1027.

Reba, Meredith, Femke Reitsma, and Karen C Seto. 2016. "Spatializing 6,000 Years of Global Urbanization from 3700 BC to AD 2000." *Scientific Data* 3 (1): 1–16.

Reed, Tristan, and Alexandr Trubetskoy. 2019. "Assessing the Value of Market Access from Belt and Road Projects." World Bank Policy Research Working Paper 8815.

Renfrew, Colin. 2014. *Reconfiguring the Silk Road: New Research on East-West Exchange in Antiquity.* University of Pennsylvania Press.

Rossabi, Morris. 2014. "The "Decline" of the Central Asian Caravan Trade." In *From Yuan to Modern China and Mongolia*, 201–220, Brill.

Seidel, André. 2020. "A Global Map of Amenities: Public goods, Ethnic Divisions and Decentralization." SSRN Working Paper No. 3457141.

de Soyres, François, Alen Mulabdic, and Michele Ruta. 2020. "Common Transport Infrastructure: A Quantitative Model and Estimates from the Belt and Road Initiative." *Journal of Development Economics* 143 102415.

Storeygard, Adam. 2016. "Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa." *Review of Economic Studies* 83 (3): 1263–1295.

Vogt, Manuel, Nils-Christian Bormann, Seraina Rüegger, Lars-Erik Cederman, Philipp Hunziker, and Luc Girardin. 2015. "Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family." *Journal of Conflict Resolution* 59 (7): 1327–1342.

Whitfield, Susan. 2019. *Silk Roads: Peoples, Cultures, Landscapes*. University of California Press.

Williams, Tim. 2014. *The Silk Roads: An ICOMOS Thematic Study*. International Council of Monuments and Sites (ICOMOS).

Williams, Tim. 2015. "Mapping the Silk Roads." *The Silk Road: Interwoven History* 1 1–42.

Zhang, Qingling, Bhartendu Pandey, and Karen C Seto. 2016. "A Robust Method to Generate a Consistent Time Series from DMSP/OLS Nighttime Light Data." *IEEE Transactions on Geoscience and Remote Sensing* 54 (10): 5821–5831.

Figures and Tables

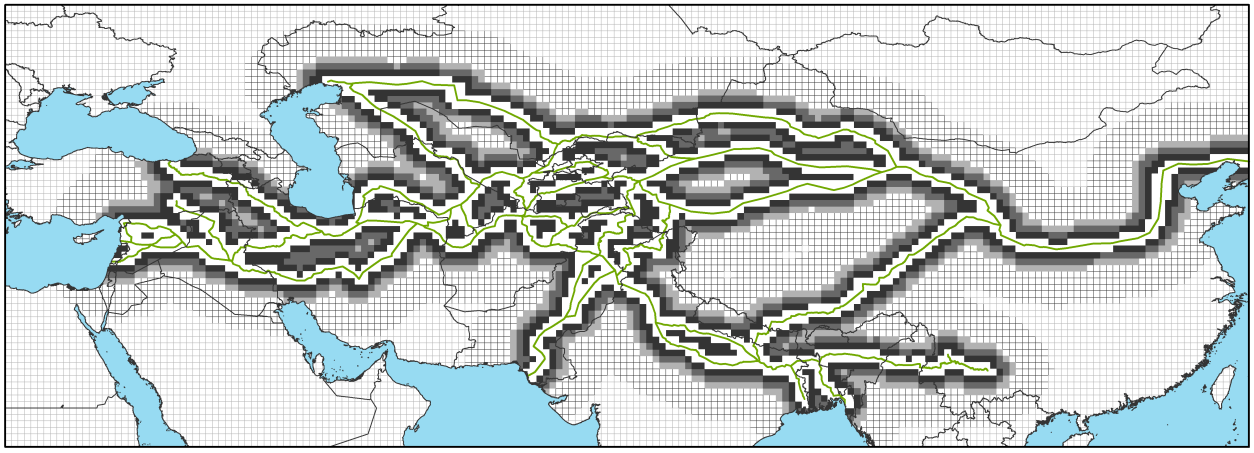
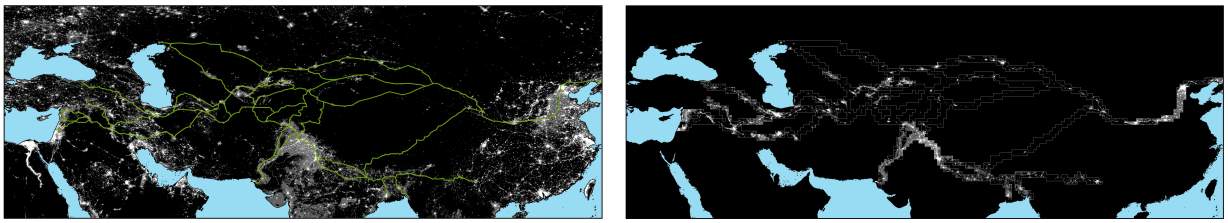


Figure 1: Map of Silk Roads; 50 KM Distance Bins; and the 500 KM Boundary

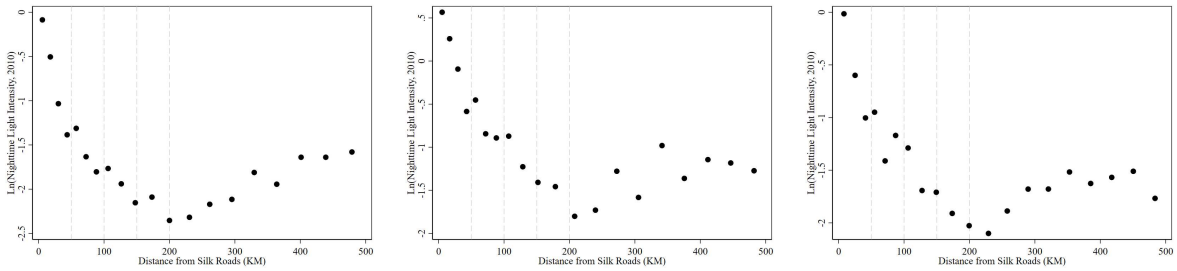
Note: The figure shows all 0.5° × 0.5° cells within the study area which is defined by the range from 53° N (northern border), 123° E (east), 19° N (south) and 27° E (west). The green line denotes the location of the Silk Roads from 200 BCE to the 15th century CE (Williams, 2014, 2015). The empty space around the Silk Roads denotes all cells with a centroid within 50 KM of the line. The darkest set of cells around the corridor represent cells with a centroid between 50 and 100 KM from the line. Each successive shade is another 50 KM bin. The empty cells outlined in black have a centroid within 500 KM of the line, while those outlined in gray are further than 500 KM. For reference, modern country boundaries are also shown.



(a) Nighttime Light Intensity (b) Nighttime Light Intensity along the Silk Roads Corridor

Figure 2: Image of Nighttime Light Intensity and Silk Roads Corridor

Note: The nighttime light intensity data are from Zhang et al. (2016), derived from the NOAA's Earth Observation Group's DMSP/OLS annual composites. The data shown are for the year 2010. The line in panel (a) denotes the location of the Silk Roads, and the corridor in panel (b) corresponds to all 0.5° × 0.5° cells with a centroid within 50 KM of the Silk Road.



(a) All Cells (b) Non-Desert (c) Non-Desert + No Rivers; Ancient Cities > 50 KM

Figure 3: Average of Ln(Nighttime Light Intensity) within 500 KM of Silk Roads

Note: The sample in panel (a) includes all cells within 500 KM of the Silk Roads (N = 7,429), the sample in panel (b) includes all non-desert cells (N = 5,224), and the sample in panel (c) keeps in place the previous restriction while also dropping all cells that include a river or have a centroid within 50 KM of an ancient city (N = 3,634). For reference, the dashed gray lines are placed at successive 50 KM intervals.

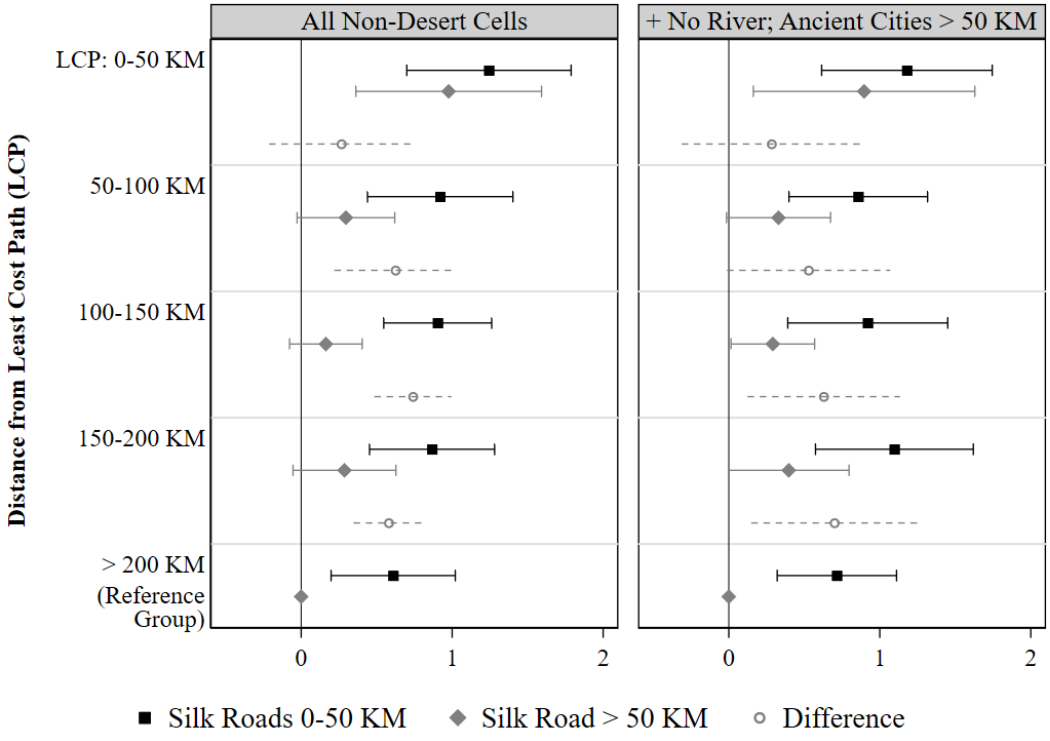
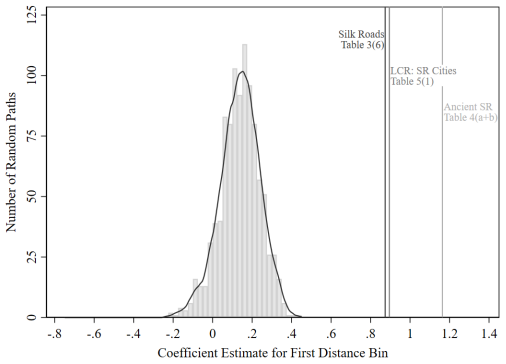
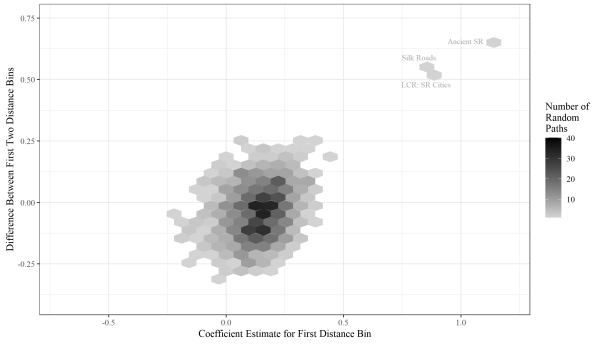


Figure 4: Nighttime Light Intensity (2010) – Distance from the LCP connecting Ancient Silk Roads' Cities and the Silk Roads

Note: The dependent variable is the log of nighttime light intensity plus 0.01; estimates are from two separate regressions. The sample in the panel on the left includes all non-desert cells within 500 KM of either the Silk Roads or the LCP between 33 ancient cities along the Silk Roads; the panel on the right further restricts the sample to cells not containing a river and with centroids at least 50 KM away from any ancient city. Each coefficient estimate shown in the figure is for a variable defined by the interaction between indicators for the noted distance from the LCP and Silk Road. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

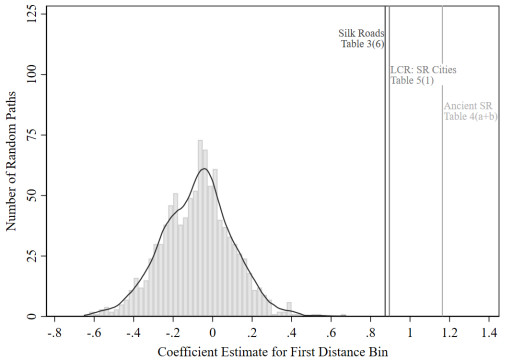


(a) Coefficient Estimates of $\mathbb{1}$ [0 - 50 KM]

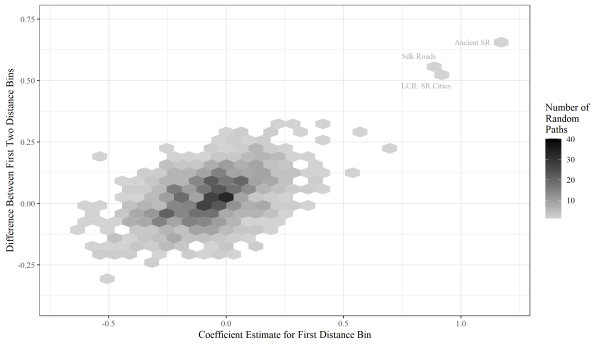


(b) Comparison of $\mathbb{1}$ [0 - 50 KM] and Difference Between First Two Distance Bins

Figure 5: Placebo Tests of LCP Connecting 33 Randomly Selected Ancient Cities (1,000 Reps) and Comparison to Silk Roads Estimates



(a) Coefficient Estimates of $\mathbb{1}$ [0 - 50 KM]



(b) Comparison of $\mathbb{1}$ [0 - 50 KM] and Difference Between First Two Distance Bins

Figure 6: Placebo Tests of LCP Connecting 33 Randomly Selected Points (1,000 Reps) and Comparison to Silk Roads Estimates

Note: Each of the above figures shows estimates from 1,000 repetitions of equation (3). In Figure 5, each path is defined by an LCP that connects 33 random cities from the set of 48 cities used in column (3) of Table 5. In Figure 6, each path is defined by an LCP that connects 33 points randomly placed within the study area defined in Figure 1. The density of estimates for the coefficient on the $\mathbb{1}$[Silk Road: 0 - 50 KM] variable is shown in panel (a) of each figure; in panel (b), the joint density of the coefficient on the first distance bin and the difference between the first two distance bins is shown. For all estimates, the dependent variable is the log of nighttime light intensity (2010) plus 0.01 and the sample includes all non-desert cells within 500 KM of either the Silk Roads or the defined path, not containing a river, and with centroids at least 50 KM away from any ancient city. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. For reference, three estimates for the Silk Roads are also shown in each figure.

Table 1: Nightlights and Economic Activity

	Ln(GDP)	Ln(GDP)
Ln(Nighttime Lights)	0.243***	0.236***
	(0.049)	(0.052)
N	919	861
Panel Type	Unbalanced	Balanced
Countries	45	41
Years	1992-2012	1992-2012
Country FE	Yes	Yes
Year FE	Yes	Yes

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable in both columns is the logged value of constant GDP. The main independent variable is the log of the average nighttime light intensity (after removing gas flares) across all pixels within each country's administrative borders. Each model includes country and year fixed effects. The sample in column (1) includes country-year observations of all countries with any in area in the study region, while the sample in column (2) includes only countries for which data are available in all years—the countries with missing GDP data are Afghanistan, Moldova, Qatar, and Syria. Robust standard errors clustered by country are shown in parentheses.

Table 2: Summary Statistics

	Centroid Distance to Silk Road					Centroid Distance to Silk Road				
	0-50 KM	50-100 KM	Difference	S.E.	N	0-50 KM	50-100 KM	Difference	S.E.	N
	A. All Non-Desert Cells					B. A + No River; Ancient Cities > 50 KM				
Nighttime Light Intensity										
Ln(NLI, 2010 CE)	0.039	-0.775	0.813***	0.104	1,877	-0.568	-1.216	0.649***	0.142	1,133
Geography										
Elevation (m)	903.3	1,045	-142.4***	45.710	1,877	1,055	1,049	6.850	59.39	1,133
Std. Dev. Elevation	227.7	261.5	-33.83***	11.940	1,877	271.7	260.9	10.80	15.98	1,133
Latitude	35.81	35.83	-0.020	0.293	1,877	37.04	36.99	0.055	0.364	1,133
Distance to Coast (km)	763.9	751.6	12.30	27.148	1,877	799.7	775.6	24.10	37.36	1,133
Distance to River (km)	69.41	75.16	-5.752*	3.465	1,877	104.4	101.5	2.918	4.551	1,133
$\mathbb{1}$ [Contains Coastline]	0.028	0.039	-0.011	0.008	1,877	0.023	0.044	-0.022**	0.011	1,133
$\mathbb{1}$ [Contains River]	0.380	0.301	0.079***	0.022	1,877	—	—	—	—	—
Population										
Population Density, 1,000 BCE (z)	0.696	0.444	0.252***	0.080	1,877	0.243	0.203	0.041	0.070	1,133
Population Density, 0 CE (z)	0.613	0.379	0.233***	0.075	1,877	0.182	0.145	0.037	0.066	1,133
$\mathbb{1}$ [Ancient City, 500 to 100 BCE]	0.049	0.009	0.041***	0.008	1,877	—	—	—	—	—
Climate										
Crop Suitability (z)	0.213	0.094	0.120**	0.047	1,877	0.002	-0.055	0.057	0.054	1,133
Increased Yield from Irrigation (z)	0.466	0.157	0.308***	0.059	1,877	0.128	0.049	0.078	0.067	1,133
Temperature (z)	0.288	0.187	0.101***	0.034	1,877	0.112	0.105	0.008	0.043	1,133
Precipitation (z)	0.123	0.240	-0.117**	0.053	1,877	0.029	0.035	-0.006	0.061	1,133

Note: *** p<0.01; ** p<0.05; * p<0.1. This table compares the averages of nighttime light intensity, geographic, historic, and climatic variables for cells with centroids within 50 KM of the Silk Roads to those with centroids between 50 and 100 KM from the route. All non-desert cells are included in Panel A; cells that contain a river or are within 50 KM of ancient cities are dropped from Panel B.

Table 3: Silk Roads and Nighttime Light Intensity (2010)

	Ln(Nighttime Light Intensity, 2010)					
	All Cells within 500 KM of Silk Roads	(1) + No Desert	(2) + No River; Ancient Cities > 50 KM	All Cells within 500 KM of Silk Roads	(4) + No Desert	(5) + No River; Ancient Cities > 50 KM
	(1)	(2)	(3)	(4)	(5)	(6)
𝟙 [Silk Roads]	0.877***	0.879***	0.962***			
	(0.136)	(0.196)	(0.211)			
𝟙 [Silk Roads: 0-50 KM]				0.798***	0.825***	0.885***
				(0.166)	(0.224)	(0.256)
𝟙 [Silk Roads: 50-100 KM]				0.238*	0.256	0.324
				(0.139)	(0.186)	(0.203)
𝟙 [Silk Roads: 100-150 KM]				0.151*	0.188	0.260
				(0.082)	(0.140)	(0.164)
𝟙 [Silk Roads: 150-200 KM]				0.015	-0.068	-0.044
				(0.100)	(0.171)	(0.172)
N	7,429	5,224	3,626	7,429	5,224	3,626
Difference Between First Two Distance Bins: Test of Equality [P-Value]				0.560	0.569	0.561
				[0.000]	[0.000]	[0.000]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. 𝟙[Silk Road] is an indicator variable equal to one if a given cell contains any portion of the Silk Roads. Each distance bin variable, 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the Silk Roads. All cells within 500 KM of the Silk Roads are included in the sample used in columns (1) and (4); all non-desert cells in columns (2) and (5); and cells containing a river or with centroids within 50 KM of an ancient city are dropped in columns (3) and (6). All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

Table 4: Silk Roads Location in the First Century CE
and Nighttime Light Intensity (2010): Balanced Sample

	Main Silk Roads	Difference between Ancient Silk Roads and Main Routes	Total Effect Along Ancient Silk Roads
	(1a)	(1b)	(1a) + (1b)
Distance from Route			
$\mathbb{1}$ [0 - 50 KM]	0.729***	0.388***	1.117***
	(0.231)	(0.136)	(0.261)
$\mathbb{1}$ [50 - 100 KM]	0.233	0.247	0.480**
	(0.172)	(0.180)	(0.241)
$\mathbb{1}$ [100 - 150 KM]	0.186	0.216	0.402**
	(0.197)	(0.274)	(0.198)
$\mathbb{1}$ [150 - 200 KM]	-0.118	0.221	0.103
	(0.217)	(0.275)	(0.169)
N	3,626		
Difference Between First	0.496		0.637
Two Distance Bins [P-Value]	[0.000]		[0.000]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; all output is from a single regression. Each distance bin variable, $\mathbb{1}$[0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the entirety of the Silk Roads in column (1a) and from the ancient portion of the route in column (1b). The sample includes all non-desert cells within 500 KM of the Silk Roads, not containing a river, and with centroids at least 50 KM away from any ancient city. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

Table 5: Geography, Selection, the Silk Roads and Nighttime Light Intensity (2010)

	Least Cost Paths Between Ancient Cities		
	Ancient Silk Roads Cities	Major Ancient Cities (Reba et al., 2016) (-200 to -100 BCE)	(-500 to -100 BCE)
	(1)	(2)	(3)
Distance from Route			
𝟙 [0 - 50 KM]	0.897***	0.126	0.223
	(0.259)	(0.258)	(0.189)
𝟙 [50 - 100 KM]	0.367**	0.098	0.092
	(0.155)	(0.134)	(0.136)
𝟙 [100 - 150 KM]	0.291**	0.021	-0.125
	(0.132)	(0.154)	(0.153)
𝟙 [150 - 200 KM]	0.394**	-0.056	-0.010
	(0.174)	(0.122)	(0.141)
Number of City Nodes	33	28	48
N	3,666	4,187	4,263
Difference Between First Two Distance Bins [P-Value]	0.530 [0.000]	0.028 [0.886]	0.131 [0.438]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of the route denoted at the top of each column. Each route is a least cost path between the set of cities defined at the top of each column. The sample includes all non-desert cells within 500 KM of either the Silk Roads or the denoted LCP, not containing a river, and with centroids at least 50 KM away from any ancient city. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

Table 6: Silk Roads and Other Modern Outcomes

	Ln(NLI, 2010)	z(Population Density, 2010)	Number of Ethnicities	𝟙 [Irrigation]	Local Roads (km) per 1,000	Health Care Sites per 1,000	𝟙 [Major Highway]	𝟙 [Railroad]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	A. All Non-Desert (SR < 500 KM)							
𝟙 [Silk Roads: 0 - 50 KM]	0.825***	0.186***	0.280***	0.098**	-0.487	-0.007	0.159***	0.244***
	(0.224)	(0.053)	(0.084)	(0.041)	(0.804)	(0.010)	(0.046)	(0.071)
Difference: First Two Distance Bins	0.569***	0.131***	0.123***	0.071**	0.186	0.013	0.111***	0.143***
	B. A + No River; Ancient City > 50 KM							
𝟙 [Silk Roads: 0 - 50 KM]	0.885***	0.118***	0.215***	0.114**	-0.365	-0.022**	0.186***	0.235***
	(0.256)	(0.042)	(0.082)	(0.049)	(1.227)	(0.009)	(0.053)	(0.064)
Difference: First Two Distance Bins	0.561***	0.064**	0.114***	0.078***	-0.233	0.003	0.152***	0.135***
	C. Placebo Route (Cities: 500 to 100 BCE)							
𝟙 [0 - 50 KM]	0.223	0.103**	-0.107*	0.007	5.969	-0.001	0.049	0.019
	(0.189)	(0.046)	(0.064)	(0.034)	(3.723)	(0.002)	(0.060)	(0.034)
	D. Random Points (1,000 repetitions)							
Share of First Distance Bin Greater Than Panel B	0.000	0.003	0.062	0.013	0.529	0.877	0.000	0.000

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted at the top of each column, and each column represents a separate regression. 𝟙[0 - 50 KM], is an indicator equal to one if the centroid of the cell is within 50 KM of the Silk Roads in Panels A and B, and the placebo route between cities from the 500 to 100 BCE period in Panel C. All non-desert cells within 500 KM of the Silk Roads are included in the sample used in Panel A, the remaining panels drop cells containing a river or with centroids within 50 KM of an ancient city. The sample in Panels C and D also include cells within 500 KM of the alternative path, if they fit all other criteria. All regressions include three additional indicators for 50 KM distance bins out to 200 KM, controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

Table 7: Silk Roads and Nighttime Light Intensity (2010): Balanced Sample With Modern Controls

Dependent Variable:	Ln(Nighttime Light Intensity, 2010)							
Mechanism Control Variable:	Ln(NLI, 2010)	z(Population Density, 2010)	Number of Ethnicities	𝟙 [Irrigation]	Local Roads (km) per 1,000	Health Care Sites per 1,000	𝟙 [Major Highway]	𝟙 [Railroad]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
𝟙 [Silk Road: 0 - 50 KM]	0.885***	0.804***	0.836***	0.703***	0.882***	0.899***	0.710***	0.560**
	(0.256)	(0.239)	(0.251)	(0.186)	(0.254)	(0.250)	(0.210)	(0.231)
Difference: First Two Distance Bins	0.561***	0.517***	0.535***	0.436***	0.560***	0.557***	0.418***	0.374***
N	3,626	3,626	3,626	3,626	3,619	3,619	3,626	3,626
	Average Reduction in First Distance Bin							
	—	9.15%	5.54%	20.56%	0.00%‡	-1.90%‡	19.77%	36.72%

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within 50 KM of the Silk Roads. The sample includes all non-desert cells within 500 KM of the Silk Roads, not containing a river, and with centroids at least 50 KM away from any ancient city. In addition to the control variable listed at the top of each column, all regressions include three additional indicators for 50 KM distance bins out to 200 KM, controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010). ‡Percent reductions are calculated relative to estimates restricting the column (1) model to exactly the sample available in these columns.

Table 8: Silk Roads and Modern Outcomes at the District Level (Geospatial Data)

	Ln(NLI, 2010)	z(Population Density, 2010)	Number of Ethnicities	Share of Area Irrigated	Local Roads (km) per 1,000	Health Care Sites per 1,000	𝟙 [Major Highway]	𝟙 [Railroad]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}\left[\bar{SR}: 0\text{-}50 \text{ KM}\right]$	0.545***	0.092	0.373*	0.111***	-0.992*	-0.623	0.161***	0.095
	(0.192)	(0.057)	(0.210)	(0.024)	(0.516)	(0.528)	(0.055)	(0.085)
$\mathbb{1}\left[\bar{SR}: 50\text{-}100 \text{ KM}\right]$	0.212	0.070	0.336	0.076***	-0.363	-0.602	0.098**	0.008
	(0.156)	(0.065)	(0.211)	(0.026)	(0.393)	(0.510)	(0.045)	(0.067)
$\mathbb{1}\left[\bar{SR}: 100\text{-}200 \text{ KM}\right]$	0.021	0.047	0.034	0.048**	-0.567	0.078	0.058	-0.003
	(0.163)	(0.039)	(0.190)	(0.024)	(0.463)	(0.191)	(0.036)	(0.048)
Difference Between First Two Distance Bins [P-Value]	0.332	0.022	0.037	0.035	-0.629	-0.021	0.063	0.086
	[0.001]	[0.277]	[0.756]	[0.000]	[0.058]	[0.775]	[0.093]	[0.010]
N	1,580	1,580	1,546	1,546	1,546	1,580	1,580	1,580
Mean of Dep. Var.	5.197	-0.113	3.42	0.194	3.122	0.212	0.299	0.506

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: 0\text{ - }50 \text{ KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table 9: Silk Roads and Public Goods at the District Level (Microdata)

	Utilities Connected to Home			Education			Health
	Electricity	Piped Water	Public Sewage	Literacy	Completed Primary	Completed Secondary	Child Survival Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.014	-0.011	0.037	0.003	0.010	0.022	0.001
	(0.049)	(0.027)	(0.024)	(0.014)	(0.017)	(0.016)	(0.004)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	-0.023	-0.020	0.020	-0.018	-0.006	-0.004	0.000
	(0.026)	(0.028)	(0.024)	(0.013)	(0.017)	(0.019)	(0.004)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	-0.007	-0.003	0.027	0.001	0.010	-0.010	-0.001
	(0.027)	(0.029)	(0.022)	(0.014)	(0.018)	(0.027)	(0.003)
Difference Between First Two Distance Bins [P-Value]	0.010	0.010	0.017	0.021	0.016	0.026	0.001
	[0.781]	[0.439]	[0.173]	[0.035]	[0.108]	[0.001]	[0.721]
N	1,531	1,233	1,356	1,380	1,430	1,430	1,724
Mean of Dep. Var.	0.872	0.308	0.138	0.730	0.644	0.302	0.929

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: \text{ 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table 10: Silk Roads and the Labor Market at the District Level (Microdata)

		Occupation		Industry	
	Currently Working	Agriculture; Fish; Forestry	Manufacturing	Agriculture; Fish; Forestry	Manufacturing
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.067***	-0.090***	0.002	-0.107***	0.021***
	(0.015)	(0.013)	(0.003)	(0.027)	(0.006)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	-0.063***	-0.065***	-0.002	-0.081***	0.016***
	(0.011)	(0.014)	(0.003)	(0.023)	(0.005)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	-0.056***	-0.046***	-0.002	-0.080***	0.009**
	(0.013)	(0.011)	(0.002)	(0.024)	(0.004)
Difference Between First Two Distance Bins [P-Value]	-0.004 [0.559]	-0.025 [0.012]	0.004 [0.017]	-0.027 [0.003]	0.005 [0.166]
N	1,299	1,200	1,200	1,121	1,121
Mean of Dep. Var.	0.572	0.234	0.039	0.283	0.048

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: 0 - 50 \text{ KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table 11: Silk Roads, Ethnicity, and Marriage at the District Level (Microdata)

	Native-Born		Ethnicity		
	Share of Population	Married to Foreign-Born	Number of Ethnicities	HHI of Ethnicity	Married to Diff. Ethnicity
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.005	0.019**	0.489**	-0.044**	0.027***
	(0.010)	(0.008)	(0.235)	(0.019)	(0.006)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	0.004	0.015**	0.358	-0.028	0.015***
	(0.011)	(0.007)	(0.234)	(0.037)	(0.006)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	0.004	0.010**	0.391*	-0.033*	0.019***
	(0.008)	(0.005)	(0.208)	(0.019)	(0.005)
Difference Between First Two Distance Bins [P-Value]	-0.010	0.004	0.131	-0.016	0.012
	[0.000]	[0.171]	[0.406]	[0.563]	[0.002]
N	694	694	255	255	255
Mean of Dep. Var.	0.979	0.016	4.306	0.588	0.038

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: \text{ 0-50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Silk Roads to Riches: Persistence Along an Ancient Trade Network

Zofia Ahmad Luke Chicoine

Appendix (For Online Publication)

Contents

A Detailed List of Data Sources

A.1 Nighttime Lights

Data are from Zhang et al. (2016) for the years 1992 to 2012. Zhang et al. (2016) used data from NOAA's Earth Observation Group's DMSP/OLS annual composites to generate consistent time series. The values for nighttime light intensity range from 0 to 63. After removing gas flares as defined by Elvidge et al. (2009), see below, the light intensity is then averaged within each 0.5 degree by 0.5 degree cell that defines the sample used throughout the paper.

Gas Flares

Gas flare polygons define the location of nighttime lights produced by the disposal of gas associated with oil production and processing. The region used in the study contains 22 countries that include gas flare polygons. The light produced in the area defined by these polygons is removed from the sample and set to zero.

A.2 Silk Roads Variables

Silk Roads - Main Routes

The UNESCO definition (Whitfield, 2019) of the Silk Roads is a set of geolocated routes documented by Williams (2014, 2015), and accessed for this project at the Harvard World Maps. Our main routes definition uses all 40,068 km of continuous portions of the main routes, as defined by Williams (2014, 2015).

Silk Roads - Cities

In the original work to define the location of the ancient trading activity along the Silk Roads, Williams (2014, 2015) documented the location of 594 sites along the routes. These sites include bazaars, bridges, burial sites, caravanserais, citadels, cities, fortresses, religious complexes, among a number of other categories. The largest category by far are the 319 cities identified. Using information from Williams (2014, 2015) we extract the location of 63 cities that were known to have existed in the year 100 BCE whether or not they were part of the Silk Roads during the routes' initial period. Cross-referencing with external sources (Hill, 2003, 2015; Dignas and Winter, 2007; Hansen, 2012; Barisitz, 2017; Whitfield, 2019) we identify 20 additional cities known to have been part of the Silk Roads during its initial period. These 20 cities were all identified by Williams (2014, 2015), but without specific time periods attached; this yields a total of 83 ancient cities from the Williams (2014, 2015) map. These city locations are one of two sources used to construct an indicator

for whether a cell contains the location of an ancient city—the second source of ancient city locations is discussed in Appendix Section A.3.

Ancient Routes

Hill (2003, 2015); Dignas and Winter (2007); Hansen (2012); Williams (2014); Barisitz (2017); Whitfield (2019) are used to define a subset of 33 of these cities that were actively part of the route during this early period, and are used to define the ancient network of routes for the first century CE. These 33 ancient cities are also used to construct a predicted least cost path as a proxy for the entire ancient network.

Silk Roads - Secondary Routes

Williams (2014, 2015) also mapped a set of secondary trading routes that were in use in the region during of study—100 BCE to about 1500 CE. When combined with the main routes, the total trade network documented by Williams (2014, 2015) defines the location of 85,951 km of the Silk Roads trade network across Eurasia.

A.3 Population and Peoples

Population Count

Population data are from NASA's Socioeconomic Data and Applications Center (SEDAC). From the Gridded Population of the World data series, we use version 4.11 of the 2010 UN WPP-Adjusted Population counts. Population counts are assigned to 30 arc-second grid cells, and adjusted to match the 2015 revision of the UN WPP country totals (CIESIN, CIESIN). Population counts are totaled within each 0.5 degree by 0.5 degree cell that defines the sample used throughout the paper.

Ethnicities

The georeferenced version of the Ethnic Power Relations (EPR) dataset—GeoEPR—provides geospatial information on every politically relevant ethnic group from around the world, and allows for the possibility of a number of ethnicities occupying the same area (Vogt et al., 2015). Groups that exist across modern national boundaries are linked using the EPR-TEK dataset.

Ancient City Locations

Locations of ancient cities from the BCE period are from Reba et al. (2016). We use these data as an additional source to Williams (2014, 2015) to define locations of major cities of the time not necessarily

associated with the Silk Roads. The data source allows us to locate 28 cities that existed within the area of study that had at least 30,000 inhabitants between the period of 200 BCE to 100 BCE. Five of these 28 cities were already included in the Williams (2014, 2015) set of cities, bringing the total number of ancient cities in the baseline period to 106. An additional 20 cities from Reba et al. (2016) for the period from 500 to 200 BCE are also added, four of which were already included in the Williams (2014, 2015) set of cities. This brings the total set of cities to 122; the balanced sample drops cells with a centroid within a 50 KM of any of these cities.

Historic Population Count

Rough estimates of population counts in 1,000 BCE—prior to the early period of the Silk Roads (Barisitz, 2017; Williams, 2014, 2015)—are from the History Database of the Global Environment (HYDE). The estimates are provided at the 5-arc minutes level, although estimates for this time period have uncertainty ranging between an assumed ±75% in 1 CE and ±100% in 10,000 BCE (Klein Goldewijk, Beusen, and Janssen, 2010). Estimates of ancient population counts are summed within each 0.5 degree by 0.5 degree unit of observation cell.

A.4 Modern Transportation Infrastructure, Irrigation, and Public Goods

Major Highways

The location of major highways are extracted from version 3.0.0 of the Roads shapefile from naturalearthdata.com. We use version 3.0.0 of the shapefile because it was posted on June 2, 2013. The initial announcement of the Belt and Road Initiative by China was announced in September of 2013; these data are not impacted by the explicit investment from this initiative. We then extract links of the polyline shapefile that are defined as "major highways" and construct two outcome variables (i) an indicator equal to one for cells that contain any portion of a major highway; (ii) the logged distance from a cell centroid to the nearest major highway.

Railroads

The location of railroads are from version 3.0.0 of the railroads shapefile from naturalearthdata.com. From these data we construct two outcome variables (i) an indicator equal to one for cells that contain any portion of the railroad; (ii) the logged distance from a cell centroid to the nearest railroad.

Irrigation - Area

Data are from the Global Map of of Irrigation Areas - Version 5 raster file which provides the number of hectares irrigated within each five arc-minute (0.0833 decimal degrees) pixel. The data show irrigated areas in 2005, and are available at the Food and Agriculture Organization of the United Nations (FAO)'s GeoNetwork. The area is summed within each observational grid cell for analysis, and is also used as an indicator to determine if any area is irrigated within the 0.5 degree by 0.5 degree observations.

Roads

Data on roads are from the Global Biodiversity Model for Policy Support's (GLOBIO) Global Roads Inventory Project (GRIP) dataset. We use data on secondary and tertiary roads to represent local investment in road infrastructure, these are the smallest levels for which data are found to be available in all parts of the world. As documented in Meijer et al. (2018), the extent of road coverage, especially the type of local roads used here, is much greater in the GRIP dataset than other available geospatial road data.

Health Sites

These data are geolocations of health care locations around the world as recorded on Open Street Maps. The set of health care locations are curated by healthsites.io to provide accurate information on health care location data around the world in the event of natural disaster or disease outbreak. As of October 30, 2020 there were 129,246 locations within the study area.

A.5 Geographic and Climatic Control Variables

Crop Suitability

Data are from the Suitability of Global Land Area for Rainfed Crops, using Maximizing Crop and Technology Mix raster file from the Food and Agriculture Organization of the United Nations (FAO)'s GeoNetwork database. The data are defined at the five arc-minute level, and contain a "suitability map using [the] maximizing crop and technology mix." The suitability values are then averaged within each 0.5 degree by 0.5 degree unit of observation cell.

Irrigation - Potential Returns

Data are from the Global Irrigation Impact Classes raster file from the Food and Agriculture Organization of the United Nations (FAO)'s Global Agro-Ecological Zones database. The file denotes 8 categories from

zero to seven; the areas labeled zero are undefined. The values from one to five denote the ranges for the increased yield that water beyond natural rainfall would generate (Bentzen, Kaarsen, and Wingender, 2017). For example, irrigation in category two areas would increase yields between 0 and 20 percent. We assign each area the midpoint of the range, the area with the greatest returns (>100%) is assigned a 100% return, undefined areas and water are set to missing, while areas with no returns or are not suitable to agriculture are set to zero. The values are then averaged within each 0.5 degree by 0.5 degree unit of observation cell.

Potential Vegetation

Data are from the Center for Sustainability and the Global Environment at the University of Wisconsin-Madison. The raster data categorizes each location into one of 15 different types of potential vegetation at the five arc-minute level (Ramankutty and Foley, 1999). The most common type of potential vegetation is assigned to each 0.5 degree by 0.5 degree unit of observation cell.

Elevation

Data are from Digital Elevation Data by Jonathan de Ferranti. The data are provided at the 15 arc-second level. Cell values used in the paper are the average elevation and the standard deviation of elevation within each cell. These elevation data are also used to construct the slope values that inform the least cost route analysis.

Temperature

Temperature data are monthly data from the World Clim 2.0 dataset. Average monthly temperature (Celsius) for the period from 1970 to 2000 is available at the 30 arc-second level (Fick and Hijmans, 2017). The average monthly temperatures are then averaged to calculate the average annual temperature within each pixel, and averaged within each 0.5 degree by 0.5 degree unit of observation cell.

Precipitation

Precipitation data are monthly data from the World Clim 2.0 dataset. Monthly temperature (mm) for the period from 1970 to 2000 is available at the 30 arc-second level (Fick and Hijmans, 2017). The average monthly precipitation are then summed to calculate the total annual precipitation within each pixel, and then averaged within each 0.5 degree by 0.5 degree unit of observation cell.

Coastline

The location of coastline around the world is from naturalearthdata.com. These data are used to denote whether an observation cell lies along a coastline, and the distance from each cell's centroid to the coast.

Rivers and Lakes

The centerlines of rivers and lakes from around the world are from naturalearthdata.com. These data are used to denote whether an observation cell contains a river, and the distance from each cell's centroid to the nearest river.

A.6 Administrative Boundaries

Country Boundaries

Country borders are from IPUMS-International's 2017 World Map.

Sub-national First-Level Boundaries

Administrative boundaries of the first-level of sub-national areas (i.e. province, region, state) are from the 2015 edition of the Food and Agriculture Organization of the United Nations (FAO)'s GeoNetwork's Global Administrative Unit Layers (GAUL) dataset.

B Silk Roads and NLI: Local Illustrative Examples

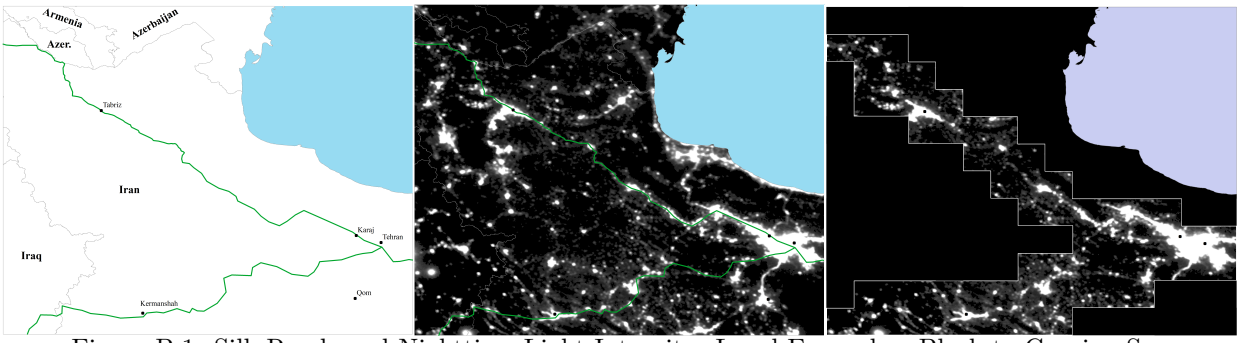


Figure B.1: Silk Roads and Nighttime Light Intensity: Local Examples: Black to Caspian Sea

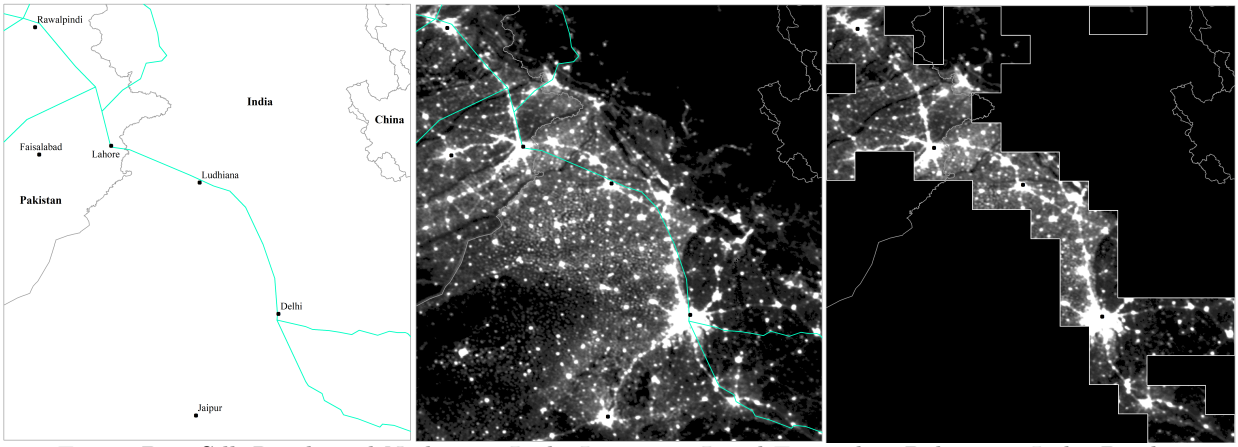


Figure B.2: Silk Roads and Nighttime Light Intensity: Local Examples: Pakistan - India Border

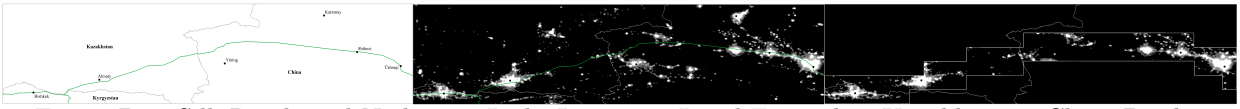


Figure B.3: Silk Roads and Nighttime Light Intensity: Local Examples: Kazakhstan - China Border

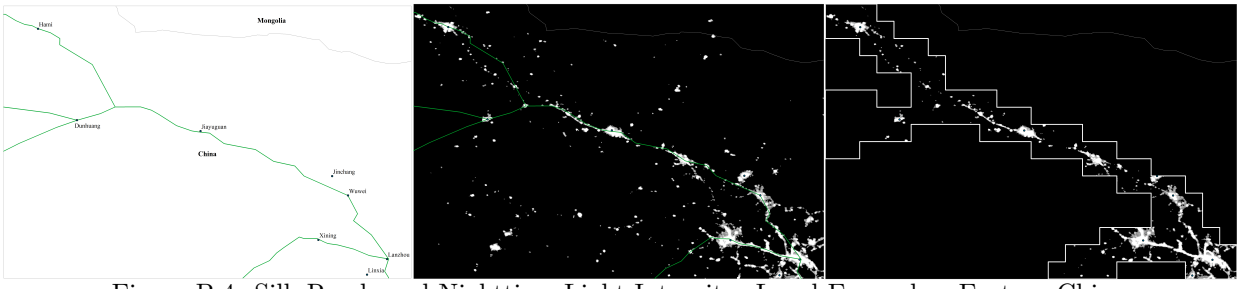


Figure B.4: Silk Roads and Nighttime Light Intensity: Local Examples: Eastern China

C Alternative Specifications of the Baseline Model

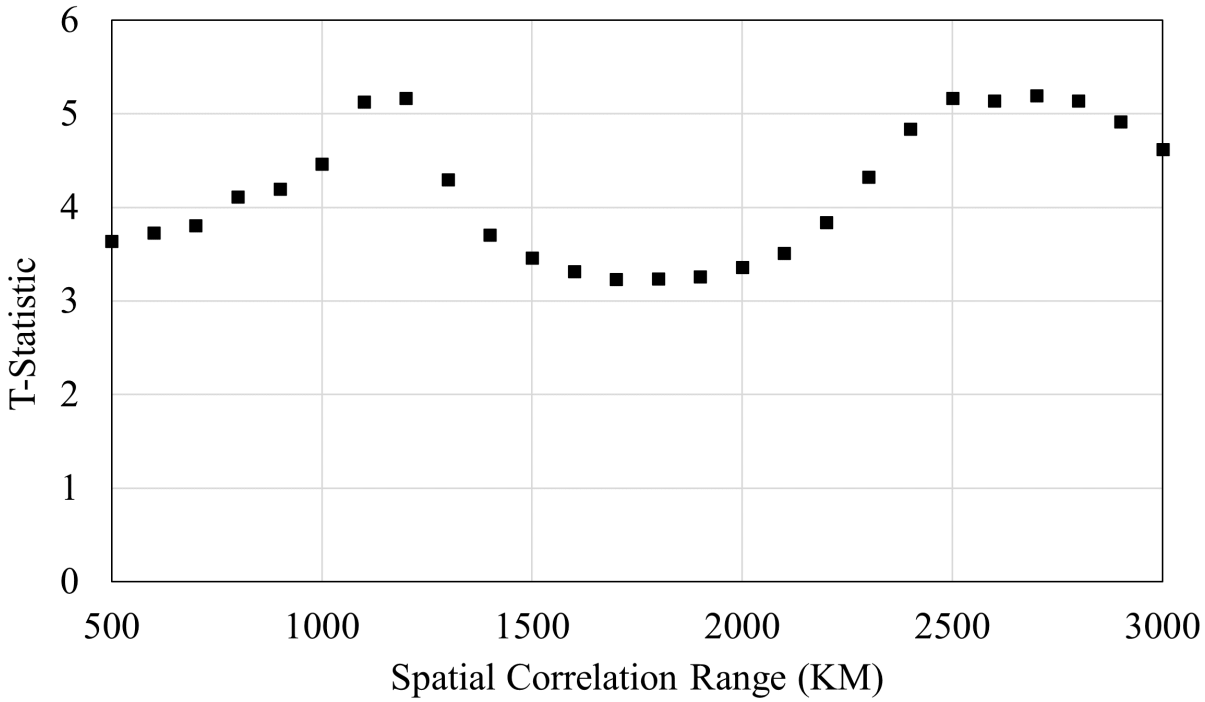


Figure C.1: T-Statistics for Coefficient on 𝟙 [Silk Road: 0-50 KM] in Equation (3): Uniform Kernel for Conley (1999, 2010) Standard Errors

Note: T-statistics for the coefficient on the 𝟙 [Silk Road: 0-50 KM] variable from equation (3) are shown. For each estimate, standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within the distance denoted in the figure (Conley, 1999, 2010). The dependent variable is the logged value of nighttime light intensity in 2010 plus 0.01. The sample includes non-desert cells within 500 KM of the Silk Roads, that do not contain a river and are at least 50 KM from any ancient city. All estimates include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country.

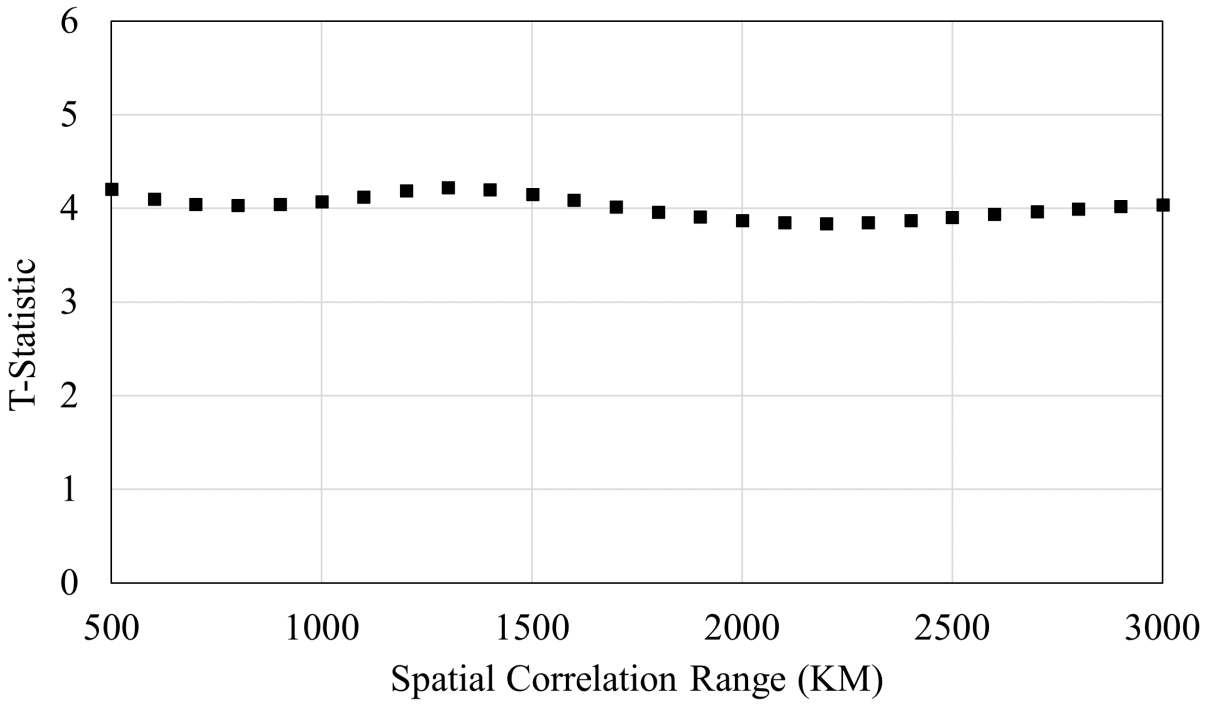


Figure C.2: T-Statistics for Coefficient on $\mathbb{1}$ [Silk Road: 0-50 KM] in Equation (3): Bartlett Kernel for Conley (1999, 2010) Standard Errors

Note: T-statistics for the coefficient on the $\mathbb{1}$ [Silk Road: 0-50 KM] variable from equation (3) are shown. For each estimate, standard errors are estimated using a using a linearly decaying Bartlett Kernel weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within the distance denoted in the figure (Conley, 1999, 2010). The dependent variable is the logged value of nighttime light intensity in 2010 plus 0.01. The sample includes non-desert cells within 500 KM of the Silk Roads, that do not contain a river and are at least 50 KM from any ancient city. All estimates include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country.

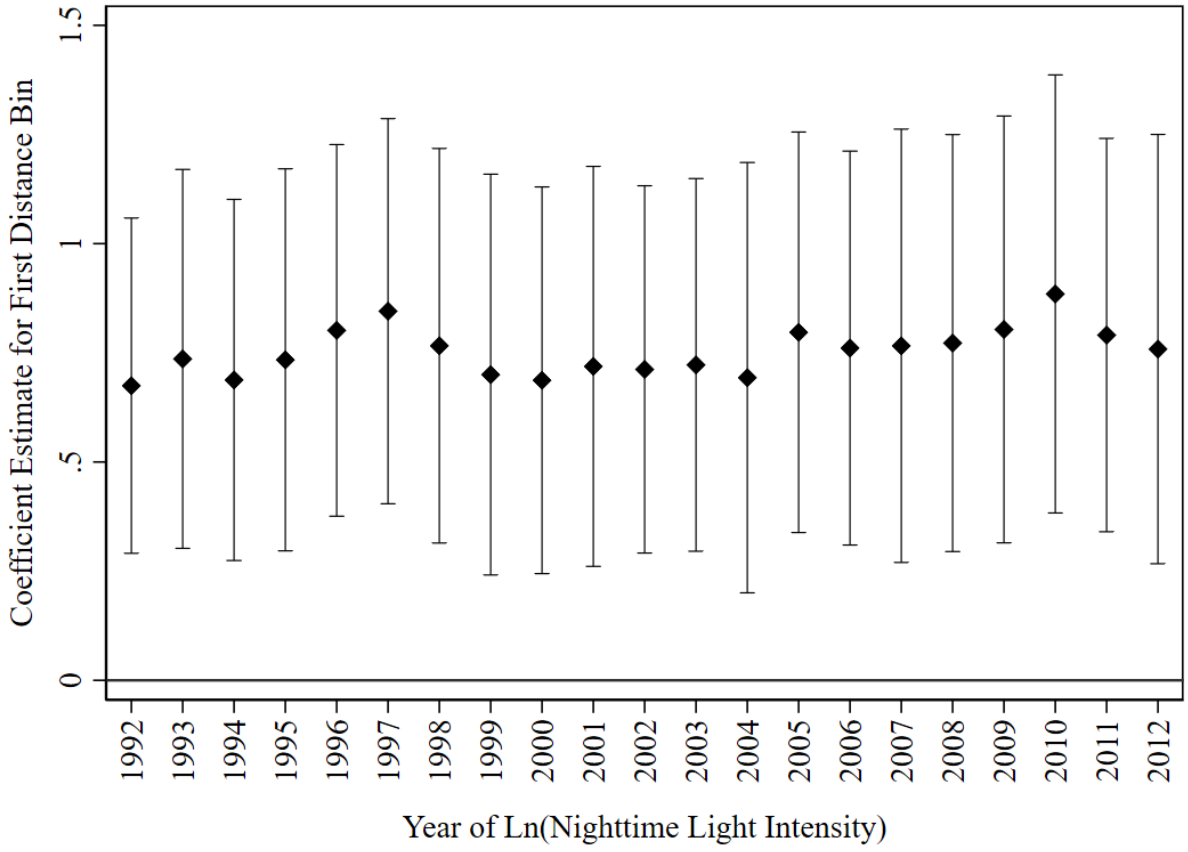


Figure C.3: Coefficient on $\mathbb{1}$ [Silk Road: 0-50 KM] in Equation (3):
95% Confidence Intervals Shown

Note: The dependent variable is the log of nighttime light intensity in plus 0.01 for the year denoted in the figure. Coefficient estimates and 95 percent confidence intervals for the $\mathbb{1}$ [Silk Road: 0-50 KM] variable from equation (3) are shown. The sample includes non-desert cells within 500 KM of the Silk Roads, that do not contain a river and are at least 50 KM from any ancient city. All estimates include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table C.1: Silk Roads and Nighttime Light Intensity (2010): All Cells within 500 KM of the Silk Roads

	Ln(Nighttime Light Intensity, 2010)						
					Spatial Fixed Effects		
	Baseline	Control: Pop. Dens. 0 CE	Lat/Lon Quadratic	Lat/Lon Quartic	None	Region	Major Ethnicity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
𝟙 [Silk Road 0-50 KM]	0.798***	0.801***	0.799***	0.797***	0.814***	0.942***	0.722***
	(0.166)	(0.166)	(0.159)	(0.151)	(0.239)	(0.151)	(0.170)
𝟙 [Silk Road 50-100 KM]	0.238*	0.241*	0.240*	0.245*	0.316*	0.352***	0.282**
	(0.139)	(0.134)	(0.138)	(0.145)	(0.191)	(0.096)	(0.135)
𝟙 [Silk Road 100-150 KM]	0.151*	0.151*	0.153*	0.166*	0.248**	0.214**	0.182***
	(0.082)	(0.077)	(0.093)	(0.087)	(0.099)	(0.100)	(0.070)
𝟙 [Silk Road 150-200 KM]	0.015	0.017	0.016	0.032	0.046	0.052	0.027
	(0.100)	(0.101)	(0.105)	(0.090)	(0.117)	(0.098)	(0.090)
N	7,429	7,429	7,429	7,429	7,429	7,429	7,429
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.560 [0.000]	0.560 [0.000]	0.559 [0.000]	0.551 [0.000]	0.498 [0.000]	0.590 [0.000]	0.441 [0.000]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. Each distance bin variable, 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the Silk Roads. The sample includes all cells within 500 KM of the Silk Roads. Unless otherwise noted at the top of the column, all regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table C.2: Silk Roads and Nighttime Light Intensity (2010):
All Non-Desert Cells within 500 KM of the Silk Roads

	Ln(Nighttime Light Intensity, 2010)						
					Spatial Fixed Effects		
	Baseline	Control: Pop. Dens. 0 CE	Lat/Lon Quadratic	Lat/Lon Quartic	None	Region	Major Ethnicity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
𝟙 [Silk Road 0-50 KM]	0.825***	0.821***	0.823***	0.842***	0.820***	0.968***	0.753***
	(0.224)	(0.219)	(0.207)	(0.219)	(0.282)	(0.192)	(0.196)
𝟙 [Silk Road 50-100 KM]	0.256	0.253	0.258	0.282	0.293	0.350*	0.328**
	(0.186)	(0.179)	(0.190)	(0.187)	(0.217)	(0.193)	(0.164)
𝟙 [Silk Road 100-150 KM]	0.188	0.185	0.191	0.229	0.247*	0.207	0.255**
	(0.140)	(0.137)	(0.153)	(0.139)	(0.146)	(0.201)	(0.115)
𝟙 [Silk Road 150-200 KM]	-0.068	-0.068	-0.067	-0.030	-0.078	-0.044	-0.010
	(0.171)	(0.171)	(0.169)	(0.161)	(0.166)	(0.166)	(0.135)
N	5,224	5,224	5,224	5,224	5,224	5,224	5,224
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.569	0.568	0.565	0.560	0.527	0.618	0.425
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. Each distance bin variable, 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the Silk Roads. The sample includes all non-desert cells within 500 KM of the Silk Roads. Unless otherwise noted at the top of the column, all regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table C.3: Silk Roads and Nighttime Light Intensity (2010): Balanced Sample

	Ln(Nighttime Light Intensity, 2010)						
					Spatial Fixed Effects		
	Baseline	Control: Pop. Dens. 0 CE	Lat/Lon Quadratic	Lat/Lon Quartic	None	Region	Major Ethnicity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
𝟙 [Silk Road 0-50 KM]	0.885***	0.886***	0.862***	0.889***	0.906***	0.989***	0.813***
	(0.256)	(0.251)	(0.250)	(0.245)	(0.347)	(0.248)	(0.258)
𝟙 [Silk Road 50-100 KM]	0.324	0.326*	0.308	0.331*	0.315	0.329	0.385**
	(0.203)	(0.198)	(0.197)	(0.195)	(0.249)	(0.209)	(0.178)
𝟙 [Silk Road 100-150 KM]	0.260	0.259	0.247	0.288*	0.340*	0.254	0.323**
	(0.164)	(0.161)	(0.169)	(0.151)	(0.175)	(0.237)	(0.133)
𝟙 [Silk Road 150-200 KM]	-0.044	-0.041	-0.049	-0.017	-0.027	0.011	0.013
	(0.172)	(0.175)	(0.168)	(0.158)	(0.141)	(0.188)	(0.118)
N	3,626	3,626	3,626	3,626	3,626	3,626	3,626
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.561	0.559	0.554	0.558	0.591	0.660	0.428
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. Each distance bin variable, 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the Silk Roads. The sample includes non-desert cells within 500 KM of the Silk Roads, that do not contain a river and are at least 50 KM from any ancient city. Unless otherwise noted at the top of the column, all regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

D The Silk Roads: In and Around Ancient Cities

The discussion throughout much of the main text focuses on the "balanced" sample that excludes a radius around 122 ancient cities, situated within 110 cells. However, the long term benefit of the Silk Roads is also found to exist within the cells that were home to these ancient cities.

Summary Statistics The main analysis comparing the effect of the Silk Roads between ancient cities is employs a model comparing cells at different distances from the Silk Roads. Prior to this analysis we again want to compare cells on geographic and baseline characteristics to ensure that cells adjacent to the Silk Roads and those nearby are similar. Although the samples are small, the noticeable differences that persist across all samples are that the cities away from the Silk Roads tend to be larger, further from the coast, and situated in areas that are less suitable to agriculture and have less rain but have larger returns to irrigation. Finally, unlike the full sample, the unconditional mean of logged NLI in cells with ancient cities is smaller in those near the Silk Roads.

Empirical Methodology To focus on areas around cities, the sample is restricted to cells within 200 KM of each ancient city. To then estimate the differential levels of NLI depending on distance from the Silk Roads, we construct a model that yields estimates for both city and non-city sets of cells within this 200 KM radius. The model used to estimate these differences is the following:

$$\begin{aligned} Ln\left(NLI\right)_{gvc} \quad = \quad & \alpha + \pi_{11}\mathbb{1}\left[SR\left(0-50\,KM\right)_{gvc}\right] \times \mathbb{1}\left[Ancient\,City\right] \\ & + \pi_{21}\mathbb{1}\left[SR\left(50-200\,KM\right)_{gvc}\right] \times \mathbb{1}\left[Ancient\,City\right] \\ & + \pi_{31}\mathbb{1}\left[SR\left(>200\,KM\right)_{gvc}\right] \times \mathbb{1}\left[Ancient\,City\right] \\ & + \pi_{10}\mathbb{1}\left[SR\left(0-50\,KM\right)_{gvc}\right] + \pi_{20}\mathbb{1}\left[SR\left(50-200\,KM\right)_{gvc}\right] \\ & + \delta X_{gvc} + \rho_v + \gamma_c + \varepsilon_{gvc}. \end{aligned} \tag{D.1}$$

The first subscript for each π coefficient denotes the distance from the Silk Roads. The first distance bin is the same 0 to 50 KM used throughout the text. However, as seen in Appendix Table D.2, there are only 14 cities in cells between 50 and 200 KM from the Silk Roads, a number that is reduced as additional restrictions are placed on the data. For this reason, the second distance bin from the Silk Roads ranges from 50 to 200 KM, while the reference area is again set to cells more than 200 KM from the Silk Roads. The reference group for equation (D.1) is the set of cells that contain no city and are at least 200 KM from the Silk Roads. The second subscript value denotes whether a cell contains an ancient city. Therefore, comparisons of π_{11} with π_{21} and π_{31} measure the difference in logged NLI in cities with and without the Silk Roads, differentiating by the distance from the routes. While a comparison of π_{10} and π_{20} yields an analogous measure for nearby cells without Silk Roads. The set of controls and standard errors are the same

as described in the previous models.

Results The output from equation (D.1) is shown in Appendix Table (D.3). One of the more striking results is the magnitude of the city coefficient away from the Silk Roads. There is an extreme persistence in this set of cities that existed in the year 100 BCE, nightlights in these areas are at least 100% higher than in the reference cells which lack an ancient city and are at least 200 KM away from the Silk Roads. However, cities along the Silk Roads, which are smaller in the years 1,000 BCE and 0 CE, have nightlight intensity that is between 60 and 100% than the non-Silk Roads cities. This magnitude translates into roughly 15% higher levels of economic output in cities that lay along the Silk Roads, even though they were once smaller during the early period of the trade network. Additionally, the non-city areas see roughly the same benefit from being on the Silk Roads; a magnitude that is similar to the results seen throughout the main text of the paper.

Although much of the work in the paper focuses on ruling out persistence within ancient cities as a potential explanation, these results demonstrate that the impact of the Silk Roads trade was not isolated to the regions away from ancient cities. Even in comparing cells that contain the location of ancient cities, those that were part of the Silk Roads trade network have significantly higher levels of economic output today.

Table D.1: Summary Statistics - Cells with an Ancient City

	Centroid Distance to Silk Road					Centroid Distance to Silk Road					Centroid Distance to Silk Road				
	0-50 KM	> 50 KM	Difference	S.E.	N	0-50 KM	> 50 KM	Difference	S.E.	N	0-50 KM	> 50 KM	Difference	S.E.	N
	A. All Cells					B. All Non-Desert Cells					C. B + No River				
Nighttime Light Intensity															
Ln(NLI, 2010 CE)	1.132	1.826	-0.694**	0.329	110	1.806	1.959	-0.153	0.239	85	1.652	1.904	-0.252	0.338	46
Geography															
Elevation (m)	722.8	582.2	140.6	120.2	110	685.3	598.5	86.821	134.923	85	766.4	819.9	-53.47	192.1	46
Std. Dev. Elevation	161.8	185.6	-23.85	31.96	110	172.9	175.4	-2.542	34.578	85	242.7	223.8	18.84	49.51	46
Latitude	36.24	33.01	3.223***	0.998	110	35.48	33.33	2.155*	1.132	85	36.57	34.94	1.627	1.414	46
Distance to Coast (km)	1,023	403.5	619.7***	116.7	110	849.3	430.6	418.699***	118.488	85	816.0	361.1	454.9**	171.6	46
Distance to River (km)	77.84	91.48	-13.65	22.19	110	54.55	79.72	-25.170	19.986	85	92.56	131.7	-39.18	31.30	46
𝟙 [Contains Coastline]	0.043	0.175	-0.132**	0.056	110	0.038	0.212	-0.174**	0.067	85	0.071	0.222	-0.151	0.102	46
𝟙 [Contains River]	0.429	0.425	0.004	0.099	110	0.462	0.455	0.007	0.112	85	—	—	—	—	—
Population															
Population Density, 1,000 BCE (z)	0.983	1.426	-0.444	0.461	110	1.366	1.520	-0.154	0.553	85	1.100	1.636	-0.536	0.768	46
Population Density, 0 CE (z)	0.842	1.547	-0.705	0.430	110	1.226	1.688	-0.462	0.513	85	0.749	1.408	-0.659	0.576	46
𝟙 [Contains Ancient City, 100 BCE]	1.000	1.000	0.000	0.000	110	1.000	1.000	0.000	0.000	85	1.000	1.000	0.000	0.000	46
Climate															
Crop Suitability (z)	0.265	0.641	-0.376*	0.220	110	0.586	0.923	-0.337	0.240	85	0.478	0.560	-0.083	0.316	46
Increased Yield from Irrigation (z)	0.855	0.232	0.623***	0.230	110	1.049	0.298	0.751***	0.266	85	0.513	0.225	0.288	0.300	46
Temperature (z)	0.409	0.539	-0.130	0.097	110	0.486	0.473	0.014	0.106	85	0.417	0.333	0.084	0.127	46
Precipitation (z)	-0.242	0.161	-0.402***	0.148	110	-0.054	0.358	-0.412**	0.164	85	-0.057	0.184	-0.241	0.232	46

Note: *** p<0.01; ** p<0.05; * p<0.1. The sample includes cells that contain at least one of 122 ancient cities. This table compares the averages of both modern outcomes as well as geographic, historic, and climatic controls for cells with centroids within 50 KM of the Silk Roads to those with centroids between 50 and 100 KM from the route. All cells are included in Panel A; all non-desert cells are included in Panel B; cells that contain a river are dropped from Panel C.

Table D.2: Number of Cells in Each Category - Equation (D.1)

Distance from Silk Roads	Ancient City In Cell	Coefficient	All Cells < 200 KM from Ancient Cities	+ Non-Desert	+ No River
0 - 50 KM	Yes	π_{11}	70	52	28
50 - 200 KM	Yes	π_{21}	14	13	8
$>$ 200 KM	Yes	π_{31}	26	20	10
0 - 50 KM	No	π_{10}	765	545	350
50 - 200 KM	No	π_{20}	1,150	751	537
$>$ 200 KM	No	—	729	544	382

Table D.3: Silk Roads and Nighttime Light Intensity (2010): In and Around Ancient Cities

	Ln(Nighttime Light Intensity, 2010)		
	All Cells < 200 KM from Ancient Cities	(1) + Non-Desert	(2) + No River
	(1)	(2)	(3)
$\mathbb{1}$ [SR(0 - 50 KM)] × $\mathbb{1}$ [Ancient City](π_{11})	1.877***	1.565***	1.994***
	(0.438)	(0.362)	(0.400)
$\mathbb{1}$ [SR(50 - 200 KM)] × $\mathbb{1}$ [Ancient City](π_{21})	1.083***	0.734**	0.875***
	(0.276)	(0.346)	(0.291)
$\mathbb{1}$ [SR(> 200 KM)] × $\mathbb{1}$ [Ancient City](π_{31})	1.243***	0.976***	1.241***
	(0.269)	(0.234)	(0.412)
$\mathbb{1}$ [SR(0 - 50 KM)] (π_{10})	0.677**	0.431	0.743**
	(0.335)	(0.281)	(0.312)
$\mathbb{1}$ [SR(50 - 200 KM)] (π_{20})	0.029	-0.098	0.060
	(0.205)	(0.132)	(0.167)
N	2,754	1,925	1,315
Difference Between:			
$\pi_{11} - \pi_{21}$	0.794*	0.831	1.119*
$\pi_{11} - \pi_{31}$	0.634**	0.590***	0.753**
$\pi_{10} - \pi_{20}$	0.648***	0.529**	0.683

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. The distance bin variables, $\mathbb{1}$[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of the Silk Roads, and $\mathbb{1}$ [Ancient City] is an indicator equal to one if an ancient city was present within the cell. The sample includes all cells with a centroid within 200 KM on an ancient city, column (1), the sample is restricted to non-desert cells in column (2), and to cells without a river in column (3). All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline and river, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

E Supplementary Figures and Tables: Silk Roads and NLI

E.1 Maps of Ancient Silk Roads, Cost Layer, and LCPs

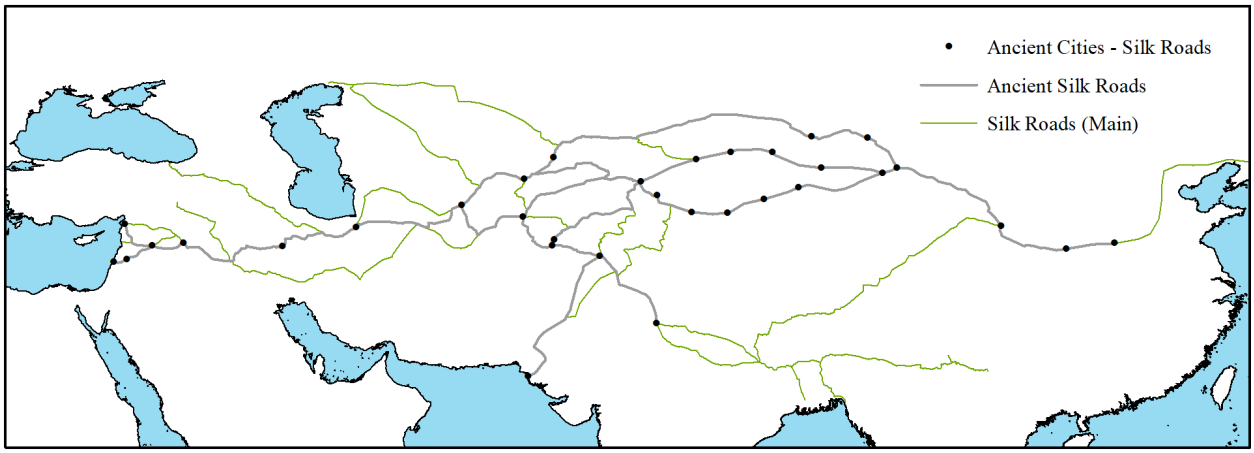


Figure E.1: Map of the Ancient Silk Roads and Cities – First Century CE

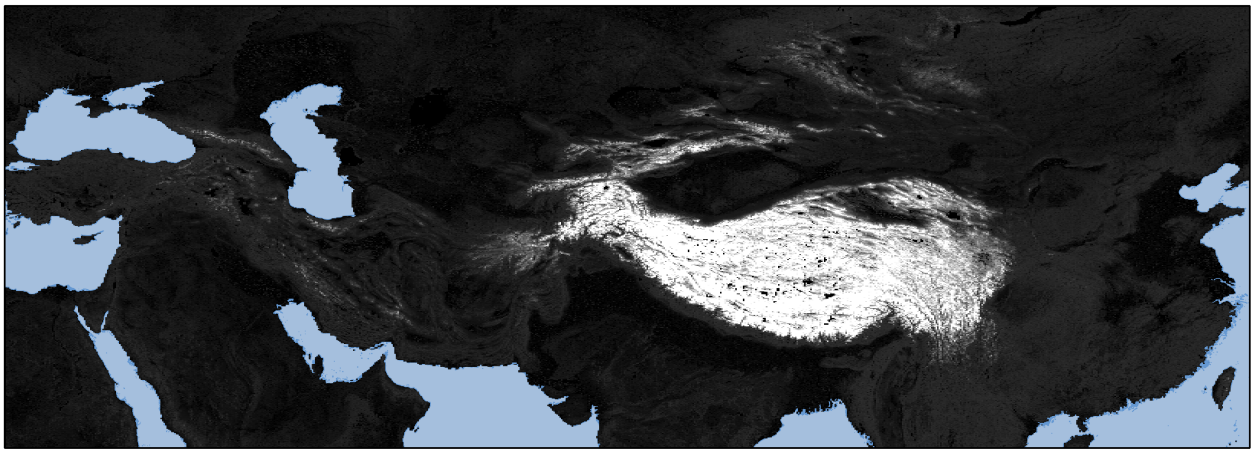


Figure E.2: Map of the Cost Layer Used to Construct Least Cost Paths

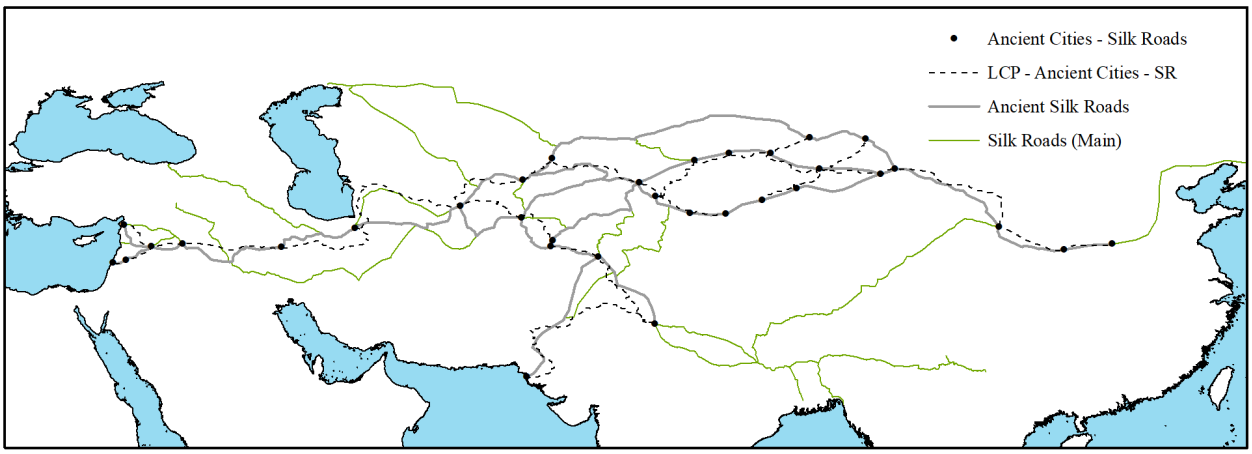


Figure E.3: Map of the Least Cost Path Between Ancient Silk Roads Cities

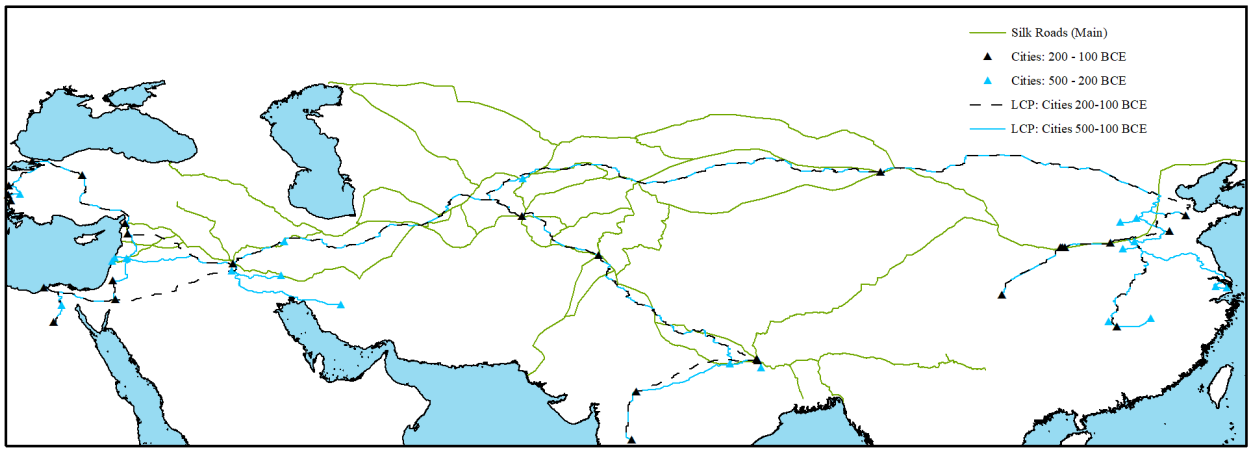


Figure E.4: Map of Ancient Cities from Reba et al. (2016) and the Placebo LCPs

E.2 Additional Summary Statistics

Table E.1: Number of Cells in Each Category - Equation (3)

Distance from Silk Roads	Coefficient	All Cells < 500 KM from Silk Roads	+ Non-Desert	+ No River; Ancient Cities > 50 KM
0 - 50 KM	θ_1	1,481	1,056	570
50 - 100 KM	θ_2	1,194	821	563
50 - 150 KM	θ_3	899	593	419
150 - 200 KM	θ_4	703	472	357
> 200 KM	—	3,152	2,282	1,717

Table E.2: Summary Statistics - Modern Outcomes

	Centroid Distance to Silk Road					Centroid Distance to Silk Road				
	0-50 KM	50-100 KM	Difference	S.E.	N	0-50 KM	50-100 KM	Difference	S.E.	N
	A. All Non-Desert Cells					B. A + No River; Ancient Cities > 50 KM				
Modern Outcomes										
Ln(Nighttime Lights, 2010)	0.039	-0.775	0.813***	0.104	1,877	-0.568	-1.216	0.649***	0.142	1,133
Population Density, 2010 CE (z)	0.436	0.164	0.271***	0.053	1,877	0.116	0.051	0.065	0.041	1,133
Number of Ethnicities	2.310	2.245	0.065	0.052	1,877	2.270	2.091	0.180***	0.066	1,133
𝟙 [Contains Major Highway, 2013]	0.331	0.192	0.139***	0.020	1,877	0.307	0.149	0.158***	0.025	1,133
𝟙 [Contains Railroad, 2013]	0.564	0.367	0.198***	0.023	1,877	0.470	0.329	0.142***	0.029	1,133
KM of Sec. / Tertiary Roads per 1,000, 2018	6.844	6.999	-0.154	0.863	1,877	7.765	8.143	-0.378	1.189	1,133
Health Care Location per 1,000, 2020	0.049	0.031	0.017**	0.009	1,877	0.034	0.034	-0.001	0.007	1,133
𝟙 [Contains Irrigated Cropland, 2013]	0.896	0.817	0.079***	0.016	1,877	0.835	0.748	0.087***	0.024	1,133
Pct. of Potential Crop Yield Produced (z)	0.048	0.012	0.036	0.036	1,737	0.099	0.033	0.065	0.051	994

Note: *** p<0.01; ** p<0.05; * p<0.1. This table compares the averages of modern outcomes for cells with centroids within 50 KM of the Silk Roads to those with centroids between 50 and 100 KM from the route. All non-desert cells are included in Panel A; cells that contain a river or are within 50 KM of ancient cities are dropped from Panel B.

Table E.3: Summary Statistics - All Cells

	Centroid Distance to Silk Road						Centroid Distance to Silk Road				
	0-50 KM	50-100 KM	Difference	S.E.	N		0-50 KM	50-100 KM	Difference	S.E.	N
Nighttime Light Intensity						Geography					
Ln(NLI, 2010 CE)	-0.746	-1.604	0.858***	0.097	2,675	Elevation (m)	1,279	1,442	-163.0***	53.99	2,675
Population Density, 2010 CE (z)	0.248	0.034	0.214***	0.039	2,675	Std. Dev. Elevation	248.3	263.2	-14.92	10.04	2,675
Number of Ethnicities	2.412	2.349	0.063	0.045	2,675	Latitude	36.18	36.25	-0.060	0.227	2,675
𝟙 [Contains Major Highway, 2013]	0.301	0.153	0.148***	0.016	2,675	Distance to Coast (km)	917.7	910.7	6.958	25.16	2,675
𝟙 [Contains Railroad, 2013]	0.462	0.272	0.190***	0.018	2,675	Distance to River (km)	86.76	94.91	-8.148**	3.886	2,675
KM of Sec. / Tertiary Roads per 1,000, 2018	10.86	13.26	-2.398	2.263	2,627	𝟙 [Contains Coastline]	0.021	0.032	-0.011*	0.006	2,675
Health Care Location per 1,000, 2020	0.044	0.032	0.012	0.007	2,628	𝟙 [Contains River]	0.342	0.256	0.085***	0.018	2,675
𝟙 [Contains Irrigated Cropland, 2013]	0.816	0.686	0.130***	0.016	2,675						
Pct. of Potential Crop Yield Produced (z)	0.155	0.196	-0.041	0.035	2,323						
Population						Climate					
Population Density, 1,000 BCE (z)	0.419	0.204	0.215***	0.059	2,675	Crop Suitability (z)	-0.050	-0.160	0.110***	0.036	2,675
Population Density, 0 CE (z)	0.341	0.145	0.197***	0.055	2,675	Increased Yield from Irrigation (z)	0.273	0.022	0.250***	0.045	2,675
𝟙 [Contains Ancient City, 100 BCE]	0.047	0.007	0.041***	0.006	2,675	Temperature (z)	0.079	-0.038	0.117***	0.034	2,675
						Precipitation (z)	-0.072	-0.002	-0.070*	0.040	2,675

Note: *** p<0.01; ** p<0.05; * p<0.1. This table compares the averages of both modern outcomes as well as geographic, historic, and climatic controls for cells with centroids within 50 KM of the Silk Roads to those with centroids between 50 and 100 KM from the route. Calculations include all cells within these distance ranges.

Table E.4: Summary Statistics - LCP Routes: Non-Desert Cells with No River and at Least 50 KM from Ancient Cities

	Centroid Distance < 50 KM to:			
		Least Cost Path		
		Ancient Silk	Major Cities	
	Silk Roads	Roads Cities	(-200 to -100 BCE)	(-500 to -100 BCE)
Share of Cells with 50 KM of Silk Roads	1.00	0.58	0.36	0.31
Number of Cells	570	170	272	322
Geography				
Elevation (m)	1,055	858	757	677
Std. Dev. Elevation	271.7	200.5	164.9	155.4
Latitude	37.04	35.41	35.02	34.58
Distance to Coast (km)	799.7	754.6	610.0	531.2
Distance to River (km)	104.4	97.33	94.67	93.89
𝟙 [Contains Coastline]	0.023	0.006	0.026	0.059
𝟙 [Contains River]	—	—	—	—
Population				
Population Density, 1,000 BCE (z)	0.243	0.160	0.529	0.540
Population Density, 0 CE (z)	0.182	0.093	0.568	0.727
𝟙 [Contains Ancient City, 100 BCE]	—	—	—	—
Climate				
Crop Suitability (z)	0.002	-0.136	0.413	0.581
Increased Yield from Irrigation (z)	0.128	0.080	0.665	0.677
Temperature (z)	0.112	0.436	0.375	0.397
Precipitation (z)	0.029	-0.341	0.021	0.105

Note: This table compares the averages of geographic, historic, and climatic characteristics of non-desert cells within 50 KM of four different route types. All cells containing rivers or within 50 KM of ancient pre-Silk Roads cities are dropped. Each cell can be part of multiple samples, as shown with the share of cells in each category along the Silk Roads at the top of the table.

E.3 Supplementary Results

Table E.5: Silk Roads and Nighttime Light Intensity (2010) with Top Coding Adjustments

	Ln(Nighttime Light Intensity, 2010) with Top Coding Adjustments					
	All Cells within 500 KM of Silk Roads	(1) + No Desert	(2) + No River; Ancient Cities > 50 KM	All Cells within 500 KM of Silk Roads	(4) + No Desert	(5) + No River; Ancient Cities > 50 KM
	(1)	(2)	(3)	(4)	(5)	(6)
𝟙 [Silk Roads]	0.898***	0.886***	0.942***			
	(0.128)	(0.182)	(0.193)			
𝟙 [Silk Roads: 0-50 KM]				0.853***	0.904***	0.966***
				(0.190)	(0.234)	(0.236)
𝟙 [Silk Roads: 50-100 KM]				0.302**	0.364**	0.464***
				(0.141)	(0.174)	(0.175)
𝟙 [Silk Roads: 100-150 KM]				0.241***	0.345***	0.455***
				(0.078)	(0.129)	(0.157)
𝟙 [Silk Roads: 150-200 KM]				0.121	0.051	0.104
				(0.104)	(0.181)	(0.194)
N	7,429	5,224	3,626	7,429	5,224	3,626
Difference Between First Two Distance Bins: Test of Equality [P-Value]				0.551	0.539	0.502
				[0.000]	[0.000]	[0.000]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01 with top coding corrections from Bluhm and Krause (2018); each column represents a separate regression. 𝟙[Silk Road] is an indicator variable equal to one if a given cell contains any portion of the Silk Roads. Each distance bin variable, 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the Silk Roads. All cells within 500 KM of the Silk Roads are included in the sample used in columns (1) and (4); all non-desert cells in columns (2) and (5); and cells containing a river or with centroids within 50 KM of an ancient city are dropped in columns (3) and (6). All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

Table E.6: Silk Roads and Nighttime Light Intensity (2010): Balanced Sample

	Ln(Nighttime Light Intensity, 2010)			
	Baseline	Distance from Ancient Cities		
	> 50 KM	> 100 KM	> 150 KM	> 200 KM
	(1)	(2)	(3)	(4)
𝟙 [Silk Road 0-50 KM]	0.885***	0.820***	0.808**	0.834**
	(0.256)	(0.295)	(0.317)	(0.334)
𝟙 [Silk Road 50-100 KM]	0.324	0.246	0.207	0.163
	(0.203)	(0.190)	(0.146)	(0.108)
𝟙 [Silk Road 100-150 KM]	0.260	0.278*	0.250**	0.145
	(0.164)	(0.155)	(0.124)	(0.119)
𝟙 [Silk Road 150-200 KM]	-0.044	-0.053	-0.087	-0.057
	(0.172)	(0.160)	(0.148)	(0.092)
N	3,626	3,322	2,981	2,638
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.561	0.575	0.601	0.671
	[0.000]	[0.000]	[0.001]	[0.002]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the logged value of nighttime light intensity plus 0.01; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of the Silk Roads. The sample includes non-desert cells within 500 KM of the Silk Roads, that do not contain a river and are at least the denoted distance away from any ancient city. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table E.7: Silk Roads and Nighttime Light Intensity (2010) - Secondary Routes Included

	Ln(Nighttime Light Intensity, 2010)					
	All Cells within 500 KM of Silk Roads	(1) + No Desert	(2) + No River; Ancient Cities > 50 KM	All Cells within 500 KM of Silk Roads	(4) + No Desert	(5) + No River; Ancient Cities > 50 KM
	(1)	(2)	(3)	(4)	(5)	(6)
𝟙 [Silk Roads + Secondary Routes]	0.720***	0.644***	0.613***			
	(0.210)	(0.185)	(0.175)			
𝟙 [Silk Roads + Secondary: 0-50 KM]				0.499**	0.460***	0.401***
				(0.198)	(0.103)	(0.113)
𝟙 [Silk Roads + Secondary: 50-100 KM]				-0.066	-0.080	-0.087
				(0.089)	(0.122)	(0.108)
𝟙 [Silk Roads + Secondary: 100-150 KM]				-0.072	0.015	0.007
				(0.110)	(0.138)	(0.127)
𝟙 [Silk Roads + Secondary: 150-200 KM]				-0.167*	-0.126	-0.124
				(0.095)	(0.122)	(0.085)
N	9,003	6,524	4,658	9,003	6,524	4,658
Difference Between First Two Distance Bins: Test of Equality [P-Value]				0.565	0.540	0.488
				[0.003]	[0.001]	[0.001]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. 𝟙[Silk Road + Secondary] is an indicator variable equal to one if a given cell contains any portion of the main Silk Roads or secondary routes. Each distance bin variable, 𝟙[Silk Road: 0 - 50 KM], is an indicator equal to one if the centroid of the cell is within the stated distance of the main Silk Roads or secondary routes. All cells within 500 KM of either set of routes are included in the sample used in columns (1) and (4); all non-desert cells in columns (2) and (5); and cells containing a river or with centroids within 50 KM of an ancient city are dropped in columns (3) and (6). All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following Conley (1999, 2010).

F Additional Geospatial Mechanism Results

F.1 Mechanisms as the Dependent Variable

Table F.1: Silk Roads and z(Population Density, 2010)

	Silk Roads		
	All Non-Desert; Silk Roads < 500 KM	A + No River; Ancient City > 50 KM	Placebo Routes Cities: 500 to 100 BCE
	(A)	(B)	(C)
𝟙 [0 - 50 KM]	0.186***	0.118***	0.103**
	(0.053)	(0.042)	(0.046)
𝟙 [50 - 100 KM]	0.054**	0.054**	0.096***
	(0.024)	(0.023)	(0.028)
𝟙 [100 - 150 KM]	0.063*	0.064**	-0.007
	(0.037)	(0.026)	(0.012)
𝟙 [150 - 200 KM]	-0.005	-0.007	0.068*
	(0.024)	(0.021)	(0.040)
N	5,224	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.131 [0.009]	0.064 [0.030]	0.006 [0.769]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the z-score of population density in 2010; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.2: Silk Roads and Number of Ethnicities

	Silk Roads		
	All Non-Desert;	A + No River;	Placebo Routes
	Silk Roads < 500 KM	Ancient City > 50 KM	Cities: 500 to 100 BCE
	(A)	(B)	(C)
$\mathbb{1}$ [0 - 50 KM]	0.280***	0.215***	-0.107*
	(0.084)	(0.082)	(0.064)
$\mathbb{1}$ [50 - 100 KM]	0.157	0.102	-0.151**
	(0.102)	(0.107)	(0.067)
$\mathbb{1}$ [100 - 150 KM]	-0.004	-0.023	-0.238***
	(0.074)	(0.082)	(0.068)
$\mathbb{1}$ [150 - 200 KM]	-0.003	-0.039	-0.247***
	(0.062)	(0.056)	(0.065)
N	5,224	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.123 [0.000]	0.114 [0.000]	0.044 [0.433]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the number of ethnicities; each column represents a separate regression. The distance bin variables, $\mathbb{1}$[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.3: Silk Roads and 𝟙[Irrigation]

	Silk Roads		
	All Non-Desert;	A + No River;	Placebo Routes
	Silk Roads < 500 KM	Ancient City > 50 KM	Cities: 500 to 100 BCE
	(A)	(B)	(C)
𝟙[0 - 50 KM]	0.098**	0.114**	0.007
	(0.041)	(0.049)	(0.034)
𝟙[50 - 100 KM]	0.027	0.036	-0.041***
	(0.031)	(0.040)	(0.007)
𝟙[100 - 150 KM]	0.019	0.028	-0.060***
	(0.020)	(0.028)	(0.019)
𝟙[150 - 200 KM]	0.001	0.004	-0.038
	(0.025)	(0.034)	(0.024)
N	5,224	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.071 [0.017]	0.078 [0.000]	0.048 [0.209]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is an indicator equal to one if the cell contains irrigated cropland; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.4: Silk Roads and Local Roads (km) per 1,000 people

	Silk Roads		
	All Non-Desert;	A + No River;	Placebo Routes
	Silk Roads < 500 KM	Ancient City > 50 KM	Cities: 500 to 100 BCE
	(A)	(B)	(C)
$\mathbb{1}$ [0 - 50 KM]	-0.487	-0.365	5.969
	(0.804)	(1.227)	(3.723)
$\mathbb{1}$ [50 - 100 KM]	-0.673	-0.132	2.407
	(0.484)	(0.726)	(1.499)
$\mathbb{1}$ [100 - 150 KM]	1.852	3.405	8.326
	(1.989)	(2.778)	(5.801)
$\mathbb{1}$ [150 - 200 KM]	1.410	2.631	8.181
	(3.854)	(5.370)	(5.230)
N	5,216	3,618	4,258
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.186 [0.834]	-0.233 [0.871]	3.562 [0.267]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the kilometers of local roads per 1,000 people; each column represents a separate regression. The distance bin variables, $\mathbb{1}$[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.5: Silk Roads and Health Care Sites per 1,000 people

	Silk Roads		
	All Non-Desert; Silk Roads < 500 KM	A + No River; Ancient City > 50 KM	Placebo Routes Cities: 500 to 100 BCE
	(A)	(B)	(C)
𝟙 [0 - 50 KM]	-0.007	-0.022*	-0.001
	(0.010)	(0.013)	(0.009)
𝟙 [50 - 100 KM]	-0.020*	-0.025*	0.006
	(0.011)	(0.014)	(0.011)
𝟙 [100 - 150 KM]	-0.022	-0.034*	-0.006
	(0.016)	(0.020)	(0.006)
𝟙 [150 - 200 KM]	-0.023*	-0.034*	-0.014*
	(0.013)	(0.018)	(0.009)
N	5,217	3,619	4,256
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.013 [0.199]	0.003 [0.693]	-0.007 [0.459]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the number of health care locations per 1,000 people; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.6: Silk Roads and 𝟙 [Major Highway]

	Silk Roads			
	All Non-Desert; Silk Roads < 500 KM	A + No River; Ancient City > 50 KM	+ Control for Railroad	Placebo Routes: Cities: 500 to 100 BCE
	(A)	(B)	(B.2)	(C)
𝟙 [0 - 50 KM]	0.159***	0.186***	0.150***	0.049
	(0.046)	(0.053)	(0.044)	(0.060)
𝟙 [50 - 100 KM]	0.048	0.034	0.019	0.026
	(0.030)	(0.021)	(0.021)	(0.053)
𝟙 [100 - 150 KM]	0.090***	0.075***	0.068***	-0.016
	(0.032)	(0.023)	(0.021)	(0.034)
𝟙 [150 - 200 KM]	0.019	0.025	0.024	0.037
	(0.018)	(0.016)	(0.022)	(0.034)
𝟙 [Railroad]			0.152***	
			(0.041)	
N	5,224	3,626	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.111 [0.000]	0.152 [0.001]	0.131 [0.001]	0.024 [0.655]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is an indicator equal to one if the cell contains a major highway; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. The estimates in column (B.2) include an additional control variable, an indicator equal to one if the cell contains a major highway. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.7: Silk Roads and 𝟙 [Railroad]

	Silk Roads			
	All Non-Desert; Silk Roads < 500 KM	A + No River; Ancient City > 50 KM	+ Control for Major Highway	Placebo Routes: Cities: 500 to 100 BCE
	(A)	(B)	(B.2)	(C)
𝟙 [0 - 50 KM]	0.244***	0.235***	0.188***	0.019
	(0.071)	(0.064)	(0.060)	(0.034)
𝟙 [50 - 100 KM]	0.101**	0.100**	0.091**	0.046
	(0.050)	(0.043)	(0.043)	(0.031)
𝟙 [100 - 150 KM]	0.065	0.049	0.030	0.016
	(0.047)	(0.046)	(0.043)	(0.020)
𝟙 [150 - 200 KM]	0.020	0.005	-0.002	0.021
	(0.037)	(0.041)	(0.044)	(0.030)
𝟙 [Major Highway]			0.254***	
			(0.031)	
N	5,224	3,626	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	0.143	0.135	0.097	-0.027
	[0.003]	[0.004]	[0.025]	[0.584]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is an indicator equal to one if the cell contains a railroad; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. The estimates in column (B.2) include an additional control variable, an indicator equal to one if the cell contains a major highway. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

Table F.8: Silk Roads and Ln(Distance to Major Highway)

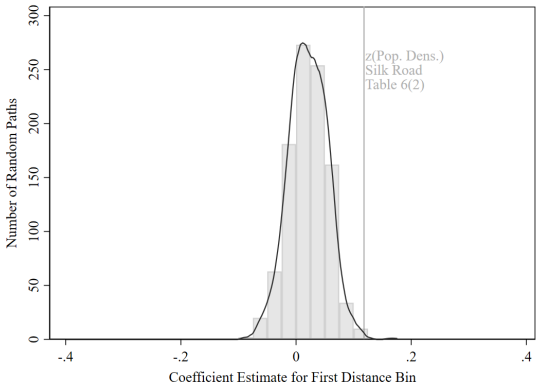
	Silk Roads			
	All Non-Desert; Silk Roads < 500 KM	A + No River; Ancient City > 50 KM	+ Control for Railroad	Placebo Routes: Cities: 500 to 100 BCE
	(A)	(B)	(B.2)	(C)
𝟙 [0 - 50 KM]	-0.513***	-0.533**	-0.350*	-0.182
	(0.192)	(0.217)	(0.199)	(0.243)
𝟙 [50 - 100 KM]	-0.217	-0.172	-0.064	-0.196
	(0.164)	(0.156)	(0.146)	(0.179)
𝟙 [100 - 150 KM]	-0.257*	-0.210	-0.146	-0.029
	(0.142)	(0.132)	(0.116)	(0.073)
𝟙 [150 - 200 KM]	-0.061	-0.032	-0.022	-0.152***
	(0.090)	(0.061)	(0.072)	(0.046)
Ln(Distance to Railroad)			0.217***	
			(0.033)	
N	5,224	3,626	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	-0.296	-0.361	-0.286	0.014
	[0.000]	[0.001]	[0.001]	[0.937]

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the logged distance to a major highway; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. The estimates in column (B.2) include an additional control variable, an indicator equal to one if the cell contains a major highway. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).

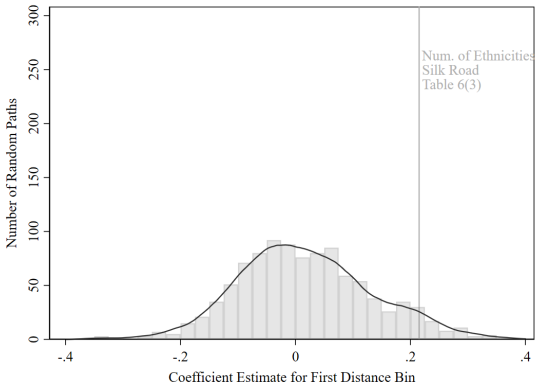
Table F.9: Silk Roads and Ln(Distance to Railroad)

	Silk Roads			Placebo Routes
	All Non-Desert; Silk Roads < 500 KM	A + No River; Ancient City > 50 KM	+ Control for Major Highway	Cities: 500 to 100 BCE
	(A)	(B)	(B.2)	(C)
𝟙 [0 - 50 KM]	-0.816***	-0.841***	-0.667***	-0.108
	(0.221)	(0.198)	(0.180)	(0.087)
𝟙 [50 - 100 KM]	-0.437***	-0.497***	-0.441***	-0.077
	(0.146)	(0.125)	(0.118)	(0.096)
𝟙 [100 - 150 KM]	-0.314*	-0.295*	-0.227	-0.032
	(0.163)	(0.173)	(0.154)	(0.064)
𝟙 [150 - 200 KM]	-0.075	-0.048	-0.037	-0.055
	(0.118)	(0.120)	(0.126)	(0.061)
Ln(Distance to Major Highway)			0.328***	
			(0.034)	
N	5,224	3,626	3,626	4,263
Difference Between First Two Distance Bins: Test of Equality [P-Value]	-0.379	-0.344	-0.226	-0.031
	[0.000]	[0.001]	[0.017]	[0.768]

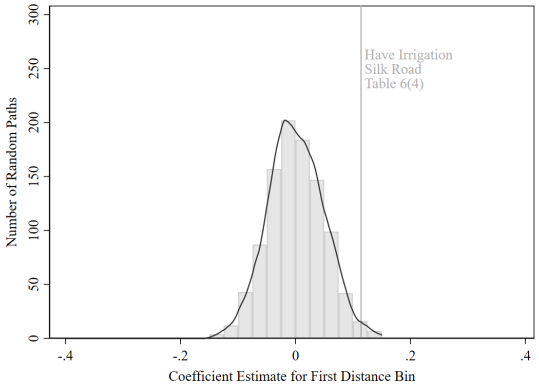
Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is logged distance to a railroad; each column represents a separate regression. The distance bin variables, 𝟙[Silk Road: 0 - 50 KM], are indicators equal to one if the centroid of the cell is within the stated distance of routes denoted at the top of the column. In column (A), the sample includes non-desert cells within 500 KM of the Silk Roads, in column (B) the sample is also restricted to cells that do not contain a river and are at least 50 KM from any ancient city. The sample in column (C) includes additional cells within 500 KM of the placebo route. All regressions include controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. The estimates in column (B.2) include an additional control variable, an indicator equal to one if the cell contains a major highway. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM (Conley, 1999, 2010).



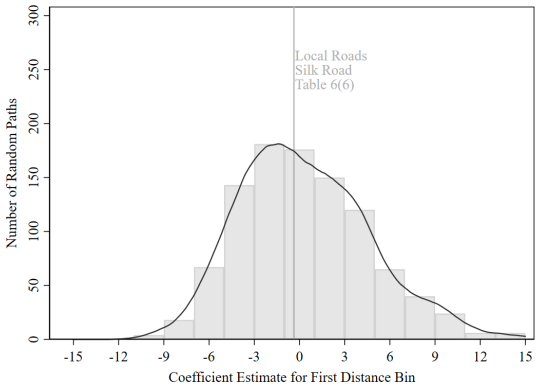
(a) z(Population Density)



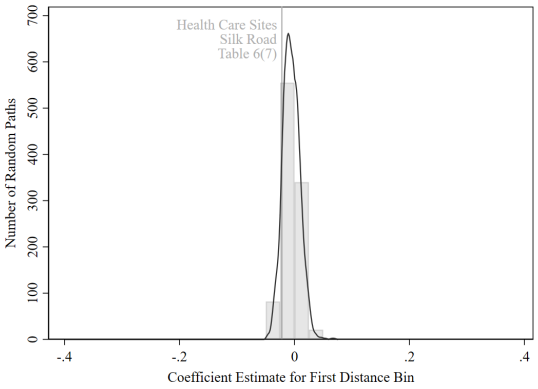
(b) Number of Ethnicities



(c) Have Irrigation



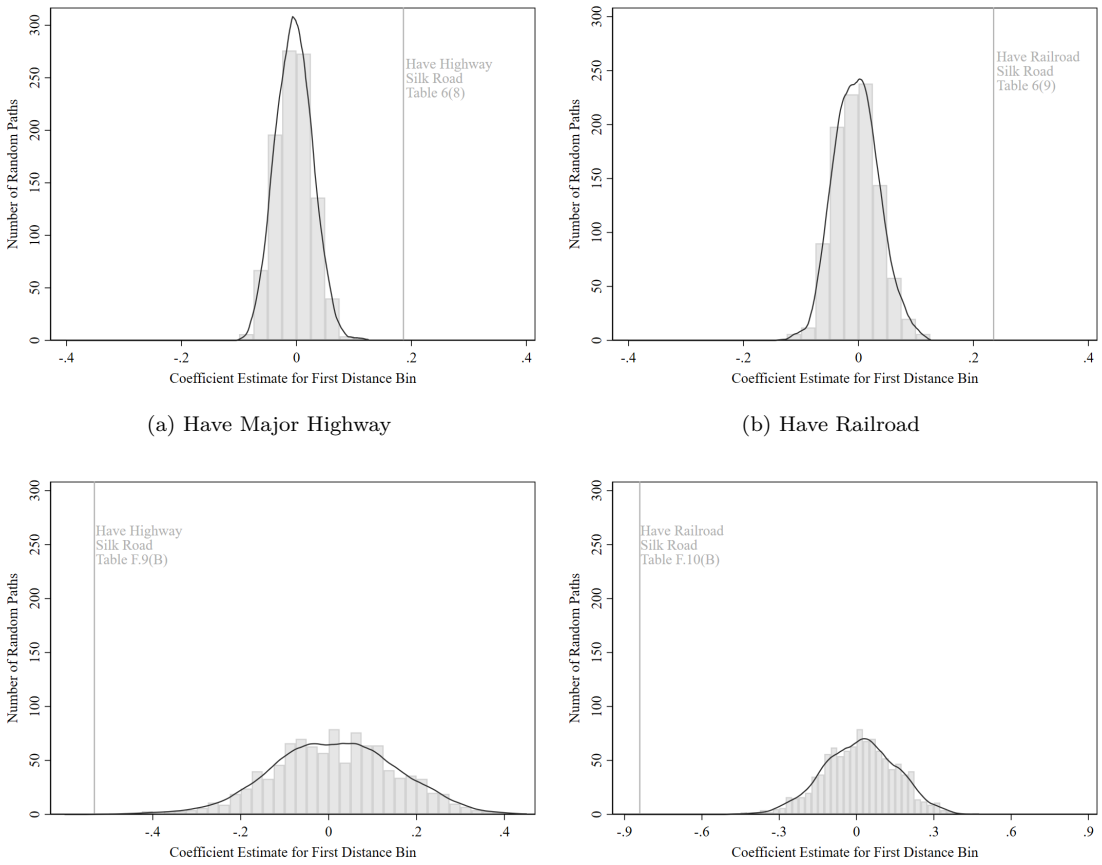
(d) Local Roads (km) per 1,000



(e) Health Care Locations per 1,000

Figure F.1: Placebo Tests of LCP Connecting 33 Randomly Selected Points (1,000 Reps): Mechanism Variables

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is defined in each panel; the figures show estimates from 1,000 paths defined by an LCP connecting 33 points randomly placed within the study area defined in Figure 1. The density of estimates for the coefficient on the first distance bin is shown, relative to the estimate along the Silk Roads using the balanced sample. Additional detail of the regression specification can be seen in Table 6.



(a) Have Major Highway

(b) Have Railroad

(c) Ln(Distance to Major Highway)

(d) Ln(Distance to Railroad)

Figure F.2: Placebo Tests of LCP Connecting 33 Randomly Selected Points (1,000 Reps): (Mechanism) Transportation Variables

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is defined in each panel; the figures show estimates from 1,000 paths defined by an LCP connecting 33 points randomly placed within the study area defined in Figure 1. The density of estimates for the coefficient on the first distance bin is shown, relative to the estimate along the Silk Roads using the balanced sample. Additional detail of the regression specification can be seen in Table 6.

F.2 Mediation Analysis

Table F.10: Silk Roads and Nighttime Light Intensity (2010): Balanced Sample With Modern Controls

Dependent Variable:	Ln(Nighttime Light Intensity, 2010)									
Mechanism Control Variable:	Ln(NLI, 2010)	z(Population Density, 2010)	Number of Ethnicities	𝟙 [Irrigation]	Local Roads (km) per 1,000	Health Care Sites per 1,000	𝟙 [Major Highway]	𝟙 [Railroad]	Ln(Distance to Major Highway)	Ln(Distance to Railroad)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
𝟙 [Silk Road: 0 - 50 KM]	0.885***	0.804***	0.836***	0.703***	0.882***	0.899***	0.710***	0.560**	0.659***	0.464**
	(0.256)	(0.239)	(0.251)	(0.186)	(0.254)	(0.250)	(0.210)	(0.231)	(0.194)	(0.217)
𝟙 [Silk Road: 50 - 100 KM]	0.324	0.287	0.301	0.266*	0.323	0.342*	0.292	0.186	0.251	0.076
	(0.203)	(0.194)	(0.190)	(0.147)	(0.202)	(0.198)	(0.201)	(0.191)	(0.171)	(0.186)
𝟙 [Silk Road: 100 - 150 KM]	0.260	0.216	0.265*	0.214*	0.256	0.282*	0.189	0.192	0.171	0.112
	(0.164)	(0.152)	(0.158)	(0.128)	(0.161)	(0.152)	(0.139)	(0.136)	(0.125)	(0.102)
𝟙 [Silk Road: 150 - 200 KM]	-0.044	-0.039	-0.035	-0.050	-0.040	-0.013	-0.067	-0.050	-0.058	-0.068
	(0.172)	(0.168)	(0.167)	(0.145)	(0.174)	(0.168)	(0.169)	(0.141)	(0.162)	(0.131)
Mechanism Control Variable	—	0.688***	0.225***	1.601***	0.000	0.757***	0.939***	1.381***	-0.424***	-0.500***
		(0.189)	(0.068)	(0.148)	(0.000)	(0.115)	(0.101)	(0.129)	(0.053)	(0.060)
N	3,626	3,626	3,626	3,626	3,619	3,619	3,626	3,626	3,626	3,626
Difference: First Two Distance Bins	0.561	0.517	0.535	0.436	0.560	0.557	0.418	0.374	0.408	0.389
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	Average Reduction in First Distance Bin									
	—	9.15%	5.54%	20.56%	0.00%‡	-1.90%‡	19.77%	36.72%	25.54%	47.57%

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01; each column represents a separate regression. 𝟙[Silk Road: 0 - 50 KM] is an indicator equal to one if the centroid of the cell is within 50 KM of the Silk Roads. The sample includes all non-desert cells within 500 KM of the Silk Roads, not containing a river, and with centroids at least 50 KM away from any ancient city. In addition to the control variable listed at the top of each column, all regressions include three additional indicators for 50 KM distance bins out to 200 KM, controls for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following Conley (1999, 2010). ‡Percent reductions are calculated relative to estimates restricting the column (1) model to exactly the sample available in these columns.

Table F.11: Silk Roads and Nighttime Light Intensity (2010): Balanced Sample With Combinations of 5 Modern Controls

	Coef. and S.E. of 0 – 50 KM Bin	P-Value of F-Test of First vs. Second Dist. Bins	Difference of 0 – 50 KM Bin Coefficient Relative to Same Regression Excluding the Noted Variable: z(Population Density, 2010) (1)	Number of Ethnicities (2)	𝟙 [Irrigation] (3)	Local Roads (km) per 1,000 (4)	𝟙 [Railroad] (5)
R.1	0.885*** (0.256)	[0.000]					
R.2	0.804*** (0.239)	[0.000]	9.15% (R.1)				
R.3	0.836*** (0.251)	[0.000]		5.54% (R.1)			
R.4	0.703*** (0.186)	[0.000]			20.56% (R.1)		
R.5	0.882*** (0.254)	[0.000]				0.34% (R.1)	
R.6	0.560** (0.231)	[0.000]					36.72% (R.1)
R.7	0.752*** (0.235)	[0.000]	10.05% (R.3)	6.47% (R.2)			
R.8	0.622*** (0.168)	[0.000]	11.52% (R.4)		22.64% (R.2)		
R.9	0.801*** (0.237)	[0.000]	9.18% (R.5)			0.37% (R.2)	
R.10	0.520** (0.214)	[0.000]	7.14% (R.6)				35.32% (R.2)
R.11	0.672*** (0.185)	[0.000]		4.41% (R.4)	19.62% (R.3)		
R.12	0.833*** (0.248)	[0.000]		5.56% (R.5)		0.36% (R.3)	
R.13	0.525** (0.225)	[0.000]		6.25% (R.6)			37.20% (R.3)
R.14	0.700*** (0.185)	[0.000]			20.63% (R.5)	0.43% (R.4)	
R.15	0.429** (0.175)	[0.000]			23.39% (R.6)		38.98% (R.4)
R.16	0.557** (0.229)	[0.000]				0.54% (R.6)	36.85% (R.5)
R.17	0.588*** (0.170)	[0.000]	12.50% (R.11)	5.47% (R.8)	21.81% (R.7)		
R.18	0.748*** (0.233)	[0.000]	10.20% (R.12)	6.62% (R.9)		0.53% (R.7)	
R.19	0.482** (0.208)	[0.000]	8.19% (R.13)	7.31% (R.10)			35.90% (R.7)
R.20	0.619*** (0.166)	[0.000]	11.57% (R.14)		22.72% (R.9)	0.48% (R.8)	
R.21	0.387** (0.155)	[0.000]	9.79% (R.15)		25.58% (R.10)		37.78% (R.8)
R.22	0.517** (0.212)	[0.000]	7.18% (R.16)			0.58% (R.10)	35.46% (R.9)
R.23	0.407** (0.174)	[0.000]		5.13% (R.15)	22.48% (R.13)		39.43% (R.11)
R.24	0.668*** (0.183)	[0.000]		4.57% (R.14)	19.81% (R.12)	0.60% (R.11)	
R.25	0.523** (0.224)	[0.000]		6.10% (R.16)		0.38% (R.13)	37.21% (R.12)
R.26	0.426** (0.174)	[0.000]			23.52% (R.16)	0.70% (R.15)	39.14% (R.14)
R.27	0.584*** (0.167)	[0.000]	12.57% (R.24)	5.65% (R.20)	21.93% (R.18)	0.68% (R.17)	
R.28	0.362** (0.154)	[0.000]	11.06% (R.23)	6.46% (R.21)	24.90% (R.19)		38.44% (R.17)
R.29	0.480** (0.207)	[0.000]	8.22% (R.25)	7.16% (R.22)		0.41% (R.19)	35.83% (R.18)
R.30	0.384** (0.153)	[0.000]	9.86% (R.26)		25.73% (R.22)	0.78% (R.21)	37.96% (R.20)
R.31	0.404** (0.172)	[0.000]		5.16% (R.26)	22.75% (R.25)	0.74% (R.23)	39.52% (R.24)
R.32	0.359** (0.153)	[0.000]	11.14% (R.31)	6.51% (R.30)	25.21% (R.29)	0.83% (R.28)	38.53% (R.27)
			Average Reduction of Coefficient Value in First Distance Bin *(% of R.1)*				
			9.79%	5.93%	22.76%	0.55%	37.52%

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is logged value of nighttime light intensity, and each row is from a separate regression. The coefficient shown in each row (R.#) is for the 𝟙[Silk Road: 0 - 50 KM] indicator which is equal to one if the centroid of the cell is within the stated distance of the Silk Roads. Although only the first distance bin is shown, all regressions include four distance bin indicators. All non-desert cells within 500 KM of the Silk Roads that do not contain a river, and are at 50 KM from any ancient city are included in each regression; for the distance to the nearest coastline and river, the cell average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for coastline, river, and ancient city, as well as fixed effects for potential vegetation and country. In each row, a combination of modern mechanism variables are added as controls. The value in each row and column denotes the percent difference between the coefficient estimate on the first distance bin indicator for the regression in the corresponding row and the coefficient value in the same regression if the mechanism variable in the corresponding column were to be removed. The coefficient from which the difference is calculated is noted below the difference value in each row and column. The average percent difference from including each mechanism is shown at the end of the table. Standard errors are estimated using a using a uniform weighting matrix that allows for contemporaneous spatial correlations between cells with centroids within 1,500 KM of one another following (Conley, 1999, 2010).

G Description of Microdata and Supplementary District-Level Results

G.1 Description of Microdata

Table G.1: Summary of Data Used to Construct District Dataset

	Data	Administrative	Number of:		Data	Administrative	Number of:	
Country	Source	Level	Districts	Individuals	Source	Level	Districts	Individuals
Afghanistan	.	.	.	.	DHS	1	34	40,221
Armenia	IPUMS	1	11	180,389	DHS	2	35	7,506
Azerbaijan	.	.	.	.	DHS	1	9	11,002
Bangladesh	IPUMS	2	64	3,676,693	DHS	2	61	14,767
China	IPUMS	2	180	6,897,222	.	.	.	.
Egypt	IPUMS	2	234	3,762,319	DHS	2	260	16,527
India	IPUMS	2	322	248,854	DHS	2	505	811,808
Iran	IPUMS	2	330	701,992	.	.	.	.
Iraq	IPUMS	2	77	779,506	.	.	.	.
Israel	IPUMS	2	14	267,859	.	.	.	.
Jordan	IPUMS	2	42	239,420	DHS	2	58	21,302
Kazakhstan	.	.	.	.	DHS	1	6	6,240
Kyrgyzstan	IPUMS	2	56	296,728	DHS	2	44	10,621
Mongolia	IPUMS	2	63	118,391	.	.	.	.
Myanmar	.	.	.	.	DHS	2	38	17,622
Nepal	IPUMS	2	72	1,520,672	.	.	.	.
Pakistan	IPUMS	2	12	5,464,305	DHS	2	104	10,023
Russia	IPUMS	1	25	4,368,016	.	.	.	.
Tajikistan	.	.	.	.	DHS	2	54	9,656
Turkey	IPUMS	2	112	1,812,393	.	.	.	.
Uzbekistan	.	.	.	.	DHS	1	5	4,415
Vietnam	IPUMS	2	300	7,802,984	.	.	.	.

Note: Data are either from national census surveys accessed through IPUMS or the DHS. The administrative level refers to either the first sub-national administrative level (i.e. region, province, or state), or the second sub-national administrative level (i.e. district or country). The number of observations of individuals are restricted to include data from adults aged between 20 to 60 in the census data, or those recorded in the Women or Men datasets of the DHS. The number of individual observations shown in the final column are collapsed into the stated number of districts; these districts are used in the final analysis.

Table G.2: Summary of Microdata Collected and Collapsed to Administrative Level

	Variable Category (Data Source; Map of Administrative Boundaries; Administrative Level)																							
Country	Geographic Data			Electricity			Piped Water			Public Sewer			Literacy			Education Level			Country					
Afghanistan	.	DHS	1	D.15	DHS	1	D.15	DHS	1	D.15	DHS	1	D.15	DHS	1	D.15	DHS	1	Afghanistan					
Armenia	.	GAUL	2	D.10	GAUL	2	D.10	GAUL	2	D.10	GAUL	2	I.11	IPUMS	1	D.10	GAUL	2	Armenia					
Azerbaijan	.	DHS	1	D.06	DHS	1	D.06	DHS	1	D.06	DHS	1	.	.	.	D.06	DHS	1	Azerbaijan					
Bangladesh	.	IPUMS	2	D.07	GAUL	2	D.07	GAUL	2	D.07	GAUL	2	I.11	IPUMS	2	I.11	IPUMS	2	Bangladesh					
China	.	IPUMS	2	.	.	.	.	.	.	.	.	.	I.00	IPUMS	2	I.00	IPUMS	2	China					
Egypt	.	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	Egypt					
India	.	IPUMS	2	D.15	GAUL	2	D.15	GAUL	2	D.15	GAUL	2	I.09	IPUMS	2	I.09	IPUMS	2	India					
Iran	.	IPUMS	2	I.06	IPUMS	2	.	.	.	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	Iran					
Iraq	.	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	Iraq					
Israel	.	IPUMS	2	.	.	.	.	.	.	.	.	.	.	.	.	I.95	IPUMS	2	Israel					
Jordan	.	IPUMS	2	I.95	IPUMS	2	.	.	.	.	.	.	I.95	IPUMS	2	I.95	IPUMS	2	Jordan					
Kazakhstan	.	DHS	1	D.99	DHS	1	D.99	DHS	1	.	.	.	.	.	.	D.99	DHS	1	Kazakhstan					
Kyrgyzstan	.	IPUMS	2	D.12	GAUL	2	D.12	GAUL	2	D.12	GAUL	2	I.09	IPUMS	2	I.09	IPUMS	2	Kyrgyzstan					
Mongolia	.	IPUMS	2	I.00	IPUMS	2	.	.	.	.	.	.	I.00	IPUMS	2	I.00	IPUMS	2	Mongolia					
Myanmar	.	GAUL	2	D.15	GAUL	2	D.15	GAUL	2	D.15	GAUL	2	D.15	GAUL	2	D.15	GAUL	2	Myanmar					
Nepal	.	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	Nepal					
Pakistan	.	GAUL	2	D.06	GAUL	2	D.06	GAUL	2	D.06	GAUL	2	I.98	IPUMS	2	I.98	IPUMS	2	Pakistan					
Russia	.	IPUMS	1	.	.	.	.	.	.	.	.	.	I.10	IPUMS	1	I.10	IPUMS	1	Russia					
Tajikistan	.	GAUL	2	D.12	GAUL	2	D.12	GAUL	2	D.12	GAUL	2	.	.	.	.	.	.	Tajikistan					
Turkey	.	IPUMS	2	.	.	.	I.00	IPUMS	2	.	.	.	I.00	IPUMS	2	I.00	IPUMS	2	Turkey					
Uzbekistan	.	DHS	1	D.96	DHS	1	D.96	DHS	1	D.96	DHS	1	.	.	.	.	.	.	Uzbekistan					
Vietnam	.	IPUMS	2	I.09	IPUMS	2	I.09	IPUMS	2	.	.	.	I.09	IPUMS	2	I.09	IPUMS	2	Vietnam					

Table G.2: ... (continued) Summary of Microdata Collected and Collapsed to Administrative Level

	Variable Category (Data Source; Map of Administrative Boundaries; Administrative Level)																							
Country	Child Survival Rate			Work Status			Occupation			Industry			Nativity			Ethnicity			Country					
Afghanistan	D.15	DHS	1	D.15	DHS	1	D.15	DHS	1	.	.	.	.	.	.	D.15	DHS	1	Afghanistan					
Armenia	D.10	GAUL	2	D.10	GAUL	2	D.10	GAUL	2	I.11	IPUMS	1	I.11	IPUMS	1	I.11	IPUMS	1	Armenia					
Azerbaijan	D.06	DHS	1	D.06	DHS	1	D.06	DHS	1	.	.	.	.	.	.	D.06	DHS	1	Azerbaijan					
Bangladesh	D.07	GAUL	2	I.11	IPUMS	2	.	.	.	.	.	.	.	.	.	.	.	.	Bangladesh					
China	I.00	IPUMS	2	.	.	.	.	.	.	.	.	.	.	.	.	I.00	IPUMS	2	China					
Egypt	D.08	GAUL	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	.	.	.	Egypt					
India	D.15	GAUL	2	I.09	IPUMS	2	I.09	IPUMS	2	I.09	IPUMS	2	.	.	.	.	.	.	India					
Iran	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	I.06	IPUMS	2	.	.	.	Iran					
Iraq	I.97	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	I.97	IPUMS	2	.	.	.	Iraq					
Israel	.	.	.	I.95	IPUMS	2	.	.	.	I.95	IPUMS	2	I.95	IPUMS	2	.	.	.	Israel					
Jordan	D.07	GAUL	2	I.95	IPUMS	2	I.95	IPUMS	2	I.95	IPUMS	2	.	.	.	.	.	.	Jordan					
Kazakhstan	D.99	DHS	1	D.99	DHS	1	D.99	DHS	1	.	.	.	.	.	.	D.99	DHS	1	Kazakhstan					
Kyrgyzstan	I.09	IPUMS	2	I.09	IPUMS	2	D.12	GAUL	2	I.09	IPUMS	2	I.09	IPUMS	2				Kyrgyzstan					
Mongolia	.	.	.	I.00	IPUMS	2	I.00	IPUMS	2	I.00	IPUMS	2	I.00	IPUMS	2	I.00	IPUMS	2	Mongolia					
Myanmar	D.15	GAUL	2	D.15	GAUL	2	D.15	GAUL	2	.	.	.	.	.	.	.	.	.	Myanmar					
Nepal	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	I.11	IPUMS	2	Nepal					
Pakistan	D.06	GAUL	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Pakistan					
Russia	.	.	.	I.10	IPUMS	1	.	.	.	.	.	.	I.10	IPUMS	1				Russia					
Tajikistan	D.12	GAUL	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Tajikistan					
Turkey	I.00	IPUMS	2	I.00	IPUMS	2	I.00	IPUMS	2	I.00	IPUMS	2	I.00	IPUMS	2	.	.	.	Turkey					
Uzbekistan	D.96	DHS	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Uzbekistan					
Vietnam	I.09	IPUMS	2	I.09	IPUMS	2	I.09	IPUMS	2	I.09	IPUMS	2	.	.	.	.	.	.	Vietnam					

Note: Each variable category has three pieces of information. In the first column, information on the data source is shown, either IPUMS (I) or DHS (D), and the year of the sample. In the second column, the map into which the microdata is collapsed is noted. In the third column, the administrative level that is used to is recorded; this is either the first sub-national administrative level (i.e. region, province, or state), or the second sub-national administrative level (i.e. district or country).

Table G.3: Number of Districts in Each Distance Bin

Population Weighted Distance from Silk Roads $(\bar{SR})$	Number of Districts
0 - 50 KM	447
50 - 100 KM	348
100 - 200 KM	327
$>$ 200 KM	458

G.2 Supplementary District-level Results

Table G.4: Silk Roads and NLI at the District Level: Alternative Samples (Geospatial Data)

	Dependent Variable: Ln(NLI, 2010)				
	(1)	(2)	(3)	(4)	(5)
Variable Defining Sample Selection	Public Sewage	Education	Industry	Nativity	Ethnicity
$\mathbb{1}\left[\bar{\text{SR}}\text{: 0-50 KM}\right]$	0.506	0.497***	0.429**	0.894***	1.161***
	(0.367)	(0.168)	(0.210)	(0.212)	(0.384)
$\mathbb{1}\left[\bar{\text{SR}}\text{: 50-100 KM}\right]$	0.308	0.197	0.184	0.625***	0.844**
	(0.305)	(0.130)	(0.167)	(0.200)	(0.336)
$\mathbb{1}\left[\bar{\text{SR}}\text{: 100-200 KM}\right]$	0.156	0.007	-0.045	-0.088	-0.337*
	(0.296)	(0.164)	(0.175)	(0.258)	(0.200)
Difference Between First Two Distance Bins [P-Value]	0.198	0.300	0.246	0.268	0.317
	[0.100]	[0.001]	[0.002]	[0.000]	[0.138]
N	1,356	1,430	1,121	694	255
Mean of Dep. Var.	5.189	5.235	5.42	5.575	3.415

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is the log of nighttime light intensity plus 0.01. The distance bin variables, $\mathbb{1}\left[\bar{\text{SR}}: \text{ 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), restricted to included the only districts with information on the "variable defining sample selection," and the samples are restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between district with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.5: Silk Roads and Modern Outcomes at the District Level:
No Districts with Ancient City Locations (Geospatial Data)

	Ln(NLI, 2010)	z(Population Density, 2010)	Number of Ethnicities	Share of Area Irrigated	Local Roads (km) per 1,000	Health Care Sites per 1,000	𝟙 [Major Highway]	𝟙 [Railroad]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
𝟙 [SR̄: 0-50 KM]	0.600***	0.086	0.166***	0.111	524.740***	0.111***	-0.889*	-0.571
	(0.202)	(0.055)	(0.060)	(0.089)	(176.220)	(0.025)	(0.534)	(0.542)
𝟙 [SR̄: 50-100 KM]	0.267	0.058	0.111**	0.011	379.224***	0.076***	-0.267	-0.515
	(0.164)	(0.065)	(0.048)	(0.067)	(136.663)	(0.026)	(0.445)	(0.500)
𝟙 [SR̄: 100-200 KM]	0.060	0.042	0.062*	0.000	195.373	0.048**	-0.481	0.126
	(0.172)	(0.038)	(0.037)	(0.048)	(143.024)	(0.024)	(0.521)	(0.195)
Difference Between First Two Distance Bins [P-Value]	0.333	0.029	0.056	0.100	145.516	0.035	-0.622	-0.056
	[0.000]	[0.096]	[0.149]	[0.009]	[0.083]	[0.000]	[0.068]	[0.509]
N	1,518	1,518	1,518	1,518	1,484	1,484	1,484	1,518
Mean of Dep. Var.	5.176	-0.114	0.286	0.500	790.350	0.194	3.173	0.200

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, 𝟙 [SR̄ : 0 - 50 KM], are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads that do not contain the location of an ancient city. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline and river, as well as fixed effects for potential vegetation and country. Standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.6: Silk Roads and Public Goods at the District Level: Additional Control for NLI (Microdata)

	Utilities Connected to Home			Education			Health
	Electricity	Piped Water	Public Sewage	Literacy	Completed Primary	Completed Secondary	Child Survival Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.049	-0.018	0.012	-0.012	-0.006	0.009	-0.001
	(0.039)	(0.025)	(0.027)	(0.011)	(0.014)	(0.016)	(0.004)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	-0.039*	-0.019	0.009	-0.019	-0.006	-0.004	0.000
	(0.023)	(0.027)	(0.022)	(0.012)	(0.015)	(0.018)	(0.004)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	-0.017	-0.003	0.019	0.006	0.016	-0.005	-0.000
	(0.022)	(0.030)	(0.023)	(0.012)	(0.016)	(0.024)	(0.004)
Difference Between First Two Distance Bins [P-Value]	-0.010	0.000	0.003	0.007	0.000	0.013	-0.001
	[0.697]	[0.990]	[0.852]	[0.413]	[0.993]	[0.061]	[0.623]
Control: Ln(NLI, 2010)	X	X	X	X	X	X	X
N	1,531	1,233	1,356	1,380	1,430	1,430	1,724
Mean of Dep. Var.	0.872	0.308	0.138	0.730	0.644	0.302	0.929

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: \text{ 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the log of nighttime light intensity plus 0.01, distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.7: Silk Roads and Public Goods at the District Level:
No Districts with Ancient City Locations (Microdata)

	Utilities Connected to Home			Education			Health
	Electricity	Piped Water	Public Sewage	Literacy	Completed Primary	Completed Secondary	Child Survival Rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.013	-0.031	0.001	0.006	0.012	0.019	-0.001
	(0.048)	(0.027)	(0.025)	(0.021)	(0.026)	(0.022)	(0.004)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	-0.023	-0.037	-0.013	-0.007	0.004	0.004	-0.002
	(0.027)	(0.027)	(0.027)	(0.023)	(0.029)	(0.023)	(0.004)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	-0.007	-0.013	0.000	0.012	0.023	0.003	-0.002
	(0.028)	(0.028)	(0.022)	(0.019)	(0.023)	(0.027)	(0.004)
Difference Between First Two Distance Bins [P-Value]	0.010	0.006	0.014	0.013	0.008	0.016	0.001
	[0.778]	[0.663]	[0.258]	[0.370]	[0.569]	[0.068]	[0.609]
N	1,488	1,202	1,315	1,327	1,375	1,375	1,663
Mean of Dep. Var.	0.870	0.307	0.138	0.731	0.643	0.303	0.928

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: \text{ 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads that do not contain the location of an ancient city. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline and river, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.8: Silk Roads and the Labor Market at the District Level: Additional Control for NLI (Microdata)

		Occupation		Industry	
	Currently Working	Agriculture; Fish; Forestry	Manufacturing	Agriculture; Fish; Forestry	Manufacturing
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\left[\bar{\text{SR}}\text{: 0-50 KM}\right]$	-0.067***	-0.078***	0.000	-0.093***	0.018***
	(0.014)	(0.014)	(0.003)	(0.026)	(0.005)
$\mathbb{1}\left[\bar{\text{SR}}\text{: 50-100 KM}\right]$	-0.063***	-0.058***	-0.003	-0.072***	0.014***
	(0.011)	(0.014)	(0.003)	(0.020)	(0.005)
$\mathbb{1}\left[\bar{\text{SR}}\text{: 100-200 KM}\right]$	-0.056***	-0.039***	-0.003	-0.070***	0.007
	(0.013)	(0.010)	(0.002)	(0.021)	(0.005)
Difference Between First Two Distance Bins [P-Value]	-0.004 [0.506]	-0.020 [0.054]	0.003 [0.014]	-0.021 [0.039]	0.004 [0.247]
Control: Ln(NLI, 2010)	X	X	X	X	X
N	1,299	1,200	1,200	1,121	1,121
Mean of Dep. Var.	0.572	0.234	0.039	0.283	0.048

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{\text{SR}}\text{: 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the log of nighttime light intensity plus 0.01, distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.9: Silk Roads and the Labor Market at the District Level:
No Districts with Ancient City Locations (Microdata)

		Occupation		Industry	
	Currently Working	Agriculture; Fish; Forestry	Manufacturing	Agriculture; Fish; Forestry	Manufacturing
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.076***	-0.090***	0.000	-0.111***	0.018***
	(0.015)	(0.014)	(0.003)	(0.025)	(0.006)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	-0.071***	-0.065***	-0.004	-0.085***	0.012***
	(0.012)	(0.016)	(0.003)	(0.021)	(0.005)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	-0.062***	-0.045***	-0.003	-0.085***	0.007*
	(0.012)	(0.012)	(0.002)	(0.023)	(0.004)
Difference Between First Two Distance Bins [P-Value]	-0.005	-0.025	0.004	-0.027	0.006
	[0.490]	[0.014]	[0.029]	[0.004]	[0.137]
N	1,150	1,054	1,054	976	976
Mean of Dep. Var.	0.580	0.249	0.037	0.301	0.047

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: \text{ 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads that do not contain the location of an ancient city. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline and river, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.10: Silk Roads, Ethnicity, and Marriage at the District Level: Additional Control for NLI (Microdata)

	Native-Born		Ethnicity		
	Share of Population	Married to Foreign-Born	Number of Ethnicities	HHI of Ethnicity	Married to Diff. Ethnicity
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	0.002	0.015*	0.462**	-0.042**	0.028***
	(0.012)	(0.008)	(0.228)	(0.021)	(0.006)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	0.008	0.013*	0.351	-0.027	0.016***
	(0.011)	(0.007)	(0.230)	(0.037)	(0.006)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	0.003	0.010**	0.425*	-0.037**	0.019***
	(0.007)	(0.004)	(0.217)	(0.017)	(0.004)
Difference Between First Two Distance Bins [P-Value]	-0.006 [0.011]	0.002 [0.499]	0.111 [0.458]	-0.014 [0.588]	0.012 [0.002]
Control: Ln(NLI, 2010)	X	X	X	X	X
N	694	694	255	255	255
Mean of Dep. Var.	0.979	0.016	4.306	0.588	0.038

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: 0 - 50 \text{ KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads. All regressions include controls for the log of nighttime light intensity plus 0.01, distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline, river, and an ancient city, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).

Table G.11: Silk Roads, Ethnicity, and Marriage at the District Level:
No Districts with Ancient City Locations (Microdata)

	Native-Born		Ethnicity		
	Share of Population	Married to Foreign-Born	Number of Ethnicities	HHI of Ethnicity	Married to Diff. Ethnicity
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}\left[\bar{SR}\text{: 0-50 KM}\right]$	-0.009	0.019**	0.613**	-0.062*	0.026***
	(0.010)	(0.008)	(0.311)	(0.032)	(0.007)
$\mathbb{1}\left[\bar{SR}\text{: 50-100 KM}\right]$	0.000	0.015**	0.371	-0.028	0.014***
	(0.010)	(0.007)	(0.333)	(0.048)	(0.005)
$\mathbb{1}\left[\bar{SR}\text{: 100-200 KM}\right]$	0.000	0.009*	0.385*	-0.036*	0.020***
	(0.007)	(0.005)	(0.215)	(0.020)	(0.005)
Difference Between First	-0.010	0.004	0.243	-0.034	0.012
Two Distance Bins [P-Value]	[0.000]	[0.158]	[0.171]	[0.265]	[0.005]
Control: Ln(NLI, 2010)	674	674	232	232	232
N	0.980	0.016	4.496	0.575	0.038

Note: *** p<0.01; ** p<0.05; * p<0.1. The dependent variable is denoted above each column. The distance bin variables, $\mathbb{1}\left[\bar{SR}: \text{ 0 - 50 KM}\right]$, are indicators equal to one if the population weighted average distance to the Silk Road is within the stated distance. Each observation is an administrative area (district or region), and the sample is restricted to districts within 500 KM of the Silk Roads that do not contain the location of an ancient city. All regressions include controls for the distance to the nearest coastline and river, average and standard deviation of elevation, cubics for longitude and latitude, land area, the z-scores of population density in 1,000 BCE, average temperature, precipitation, agricultural suitability, and returns to irrigation, indicators for whether the district contains coastline and river, as well as fixed effects for potential vegetation and country. Districts are weighted by population, and standard errors are estimated using a Bartlett weighting matrix that allows for contemporaneous spatial correlations between districts with centroids within 1,500 KM of one another following Conley (1999, 2010).