

Supplemental Appendix

Long-Distance Trade and Long-Term Persistence - Sebastian Ellingsen - For online publication only

Contents

A	Data Details	2
A.1	Data Sources and Definitions	2
A.2	Estimating Transportation Times	5
A.2.1	Maritime transportation.	5
A.2.2	Land transportation.	7
A.3	Validation Exercises	9
B	Model Details	9
B.1	Model Ingredients	10
B.2	Existence and Uniqueness	13
B.3	Comparative Statics	14
C	Further Historical Background	15
D	Additional Results	20
E	Sensitivity Analysis	29
E.1	Synthetic Control Design	29
E.2	Robustness Value	31
E.3	Sensitivity Analysis	32

A Data Details

A.1 Data Sources and Definitions

Ports: The port locations and trading statuses are from Stangl (2019d) and Fisher (1997). New Orleans is omitted since it's not in the study region. Guyana is omitted since it is inland. The list of ports is displayed in Table 1 in the paper.

Cities: The sample of cities is based on a historical reconstruction of the urban structure during the early colonial period by Diaz-Cayeros (2022). He constructs a historical GIS of major cities during the early colonial period based on primary sources. The georeferencing of the cities mentioned in these sources is based on maps in López de Velasco (1570) and López de Velasco (1575). I obtain the longitude and latitude of the historical center as well as the legal status of each city from a gazetteer of colonial Spanish America (Stangl, 2019c). The longitude and latitude of each city are typically the main town square, church, or cathedral. I then keep cities designated *city* or *villa* to obtain places that were centers of Spanish settlement, governance, and economic activity before the 18th century.

Population: I use data on urban populations from Buringh (2015) made available by the Centre for Global Economic History at Utrecht University. The basic criterion for inclusion in the database was the availability of historical population data between 1500 and 1800. The authors apply a threshold rule and collect data on cities with a population exceeding 5,000 inhabitants in 1850 or 20,000 inhabitants in 2000. If one of these criteria was met the city was included in the dataset. Following Arroyo Abad and van Zanden (2016), I supplement and extend this dataset by consulting various regional and national sources. These sources are largely based on population and urbanization studies, colonial censuses, and regional economic studies. Data on Mexico is from Gerhard (1993a,b,c). Data on Buenos Aires is from Johnson and Seibert (1979). Information on Colombia is from Pinzón, Mora and Mora (1994). For urban population numbers for larger cities across the Americas, I use Morse (1974). For periods where population data are missing or the recorded year does not coincide with the year in the dataset, I follow Buringh (2015) and impute the value by assuming a uniform growth rate between the two years for that specific city.

Settlements: Data on settlements comes from a territorial gazetteer of places in Bourbon Spanish America (Stangl, 2019c). The data come from a variety of sources. "Sources

include archival material like census tables, mission reports, visitations of dioceses and provinces, but also more ephemeral documents like petitions of some city council which was mostly not written for giving geographic information but may touch one specific detail or incidentally exposes some relevant information. Non-archival contemporary sources include mostly highly systematic sources for information like so-called “Foreigner Guides” (printed calendar-manuals which included also lists of office holders of many parts of the Empire), maps, or geographical descriptions both printed and manuscripts.” (Stangl, 2018).

Logbook data: The logbook data is from Cliwoc 2.1 (García-Herrera et al., 2005) and contains daily information on ship location, reported wind speed, and wind direction. Wind speed is measured in tenths of a meter per second. Wind direction measures the direction from which the wind is blowing in whole degrees. I restrict the sample to entries of sailing ships, that are not anchored, and not in coastal waters. I further restrict the sample to contain entries between 1750 and 1855. The data is available at: <https://www.historicalclimatology.com/cliwoc.html>.

Wind data: Wind speed and wind direction at 0.5×0.5 degrees resolution are from The Global Forecasting System (GFS) atmospheric model provided by the National Oceanic and Atmospheric Administration (NOAA) and National Centers for Environmental Prediction (NCEP). The data was downloaded using the RWind library in R (Fernández-López and Schliep, 2019). Wind direction is stored as velocity components $\vec{v}$ (northward) and $\vec{u}$ (eastward). I define wind direction as the direction of the origin of the wind relative to true north measured in degrees. The wind speed is measured in meters per second. For each cell, I download data for every 14th day between 03/09/2011 and 09/09/2017. The data used in the analysis is averaged by grid cell over the whole period.

Transportation times: Author’s calculations. See section A.2 for details.

Trade flows: Data on the nominal value of imports and the number of ships arriving in Spain by origin in Spanish America for the period 1782 to 1796 is from Fisher (1985) (Appendix E). Data on the nominal value of exports from Spain to Spanish America by destination for the period 1797 to 1820 is from Fisher (1993) (Appendix E). Some origins/destinations list several ports of call. In these cases, I assign the origins/destinations to the first port mentioned. Data on total imports from Spanish America in constant prices is from Cuenca-Esteban (2008) (Appendix table 5).

Population flows: Data on the Argentinian population for 14 provinces distributed by province in 1869 is from the 1860 Argentinian census (Census, 1872). Available at: https://biblioteca.indec.gob.ar/bases/minde/1c1869_TU.pdf. Data on population by state in Mexico by state of birth is from the Mexican census of 1895 (Census, 1899). Available at: <https://www.inegi.org.mx/app/publicaciones/?p=1681>.

Wages: Annual wages in grams of silver for unskilled construction laborers or miners from Arroyo Abad, Davies and van Zanden (2012). The data is made available by the Global Price and Income History Group (UC Davis) and is available here: <https://gpih.ucdavis.edu/Datafilelist.htm#Latam> (under Living standards in Latin America).

Prices: Prices for wheat, wine, salt, and sugar for Spain (New Castile) between 1650 and 1800 are from Hamilton (1947). Prices for wheat, wine, salt, and sugar for Argentina, Bolivia, Chile, Colombia, Mexico, and Peru are from the Global Price and Income History Group (UC Davis) and are available here: <https://gpih.ucdavis.edu/Datafilelist.htm#Latam>. I convert the quantities into consistent units of measurement using the information provided in the Allen-Unger Global Commodity Prices Dataset (<http://www.gcpdb.info/data.html>). For Spain, I convert the monetary value into grams of silver using the conversion tables in Hamilton (1947). For each location i , the price index (P_i) is the average price across the available commodities. The price index for America is the average over all the locations: $P_{am} = \sum_{j \in am} P_j$. The measure depicted in the paper is P_{es}/P_{am} . Missing values are imputed using a multiple imputation procedure following Honaker and King (2010).

Province/state-level data: Province level population data is from Maloney and Valencia (2016). The data is primarily from Denevan (1992), Bruhn and Gallego (2011), and national censuses. See Maloney and Valencia (2016) for details.

Caloric yield: I use a measure of agricultural potential constructed by Galor and Özak (2015, 2016). The data measure the maximum attainable yield measured in calories that can be achieved for a variety of crops. Agricultural suitability measures are from the Global Agro-Ecological Zones (GAEZ) dataset (Food and Agriculture Organization of the United Nations, 2012) under rain-fed, low-input agriculture for six important staple and export crops. I consider wheat, maize, sugar cane, tobacco, and cocoa as the most important cash and staple crops.

Rivers, lakes, coastlines: Data on the location of rivers, lakes, and coastlines is from Natural Earth (2025) I remove Lake Gatun which is artificial and created in the 20th

century. I add shapefiles on the location of Lake Texcoco in the 18th century (Stangl, 2019a).

Mining centers: The location and type of the principal mining centers is from Figure 2 in Fisher (1997). The longitude and latitudes are found using Google Maps.

Post routes: Shapefiles of courier routes on land between 1745 and 1808 are from Stangl (2019b).

Potential vegetation: I use the Global Potential Vegetation Dataset dataset made available by the Center for Sustainability and the Global Environment at the University of Wisconsin-Madison to measure historical landcover. It measures 15 types of potential vegetation, that would exist in the absence of human intervention at the grid-cell level (Ramankutty and Foley, 1999). The data is available at <https://sage.nelson.wisc.edu/data-and-models/datasets/>.

Political borders: Viceroyalty and audiencia borders are from Stangl (2019e). Present-day province borders are from the Global Administrative Areas dataset Hijmans (2025). For some countries I used shapefiles from Ruggles et al. (2025). Present-day country borders are from Natural Earth (2025).

Elevation: Data on elevation is from the ETOPO Global Relief Model (Amante and Eakins, 2009).

Terrain ruggedness: Calculated using ETOPO Global Relief Model with the method following Elliot, DeGloria and Riley (1999). Ranges from 0 (level terrain surface) to 959 (extremely rugged surface).

The European economy: Data for Europe is given by Spain and its most important trading partners, France and the United Kingdom. The population data for the 18th century is taken from Bolt and Van Zanden (2024). The source for Spain is Álvarez Nogal and De La Escosura (2013), for France Ridolfi and Nuvolari (2021), and Broadberry et al. (2015) for the United Kingdom. The nominal wages are averaged over cities in these countries for which data is available (London, S. Eng. towns, Paris, Strasbourg, Madrid, and Valencia). Nominal wages are measured in grams of silver per day for building laborers between 1700 and 1749. This data is from Allen (2001).

A.2 Estimating Transportation Times

In this section, I provide further detail on how I reconstruct transportation times.

A.2.1 Maritime transportation.

In the first step, I calculate the average travel speed for the voyages in the CLIWOC 2.1 dataset (García-Herrera et al., 2005). Using the information on the date of the logbook entry as well the longitude and latitude, I infer ship i 's average speed between dates t and $t - 1$, which is denoted s_{it}^{sea} . From the entries at time t and $t - 1$, I calculate the direction of travel measured in degrees relative to true north. Using the information on the wind direction recorded in the logbook at time $t - 1$, I calculate the difference in the direction of movement from the direction of the origin of the wind in degrees. I denote this θ_{it} . I also keep the recorded wind speed in $t - 1$ measured in meters per second denoted w_{it} .

I then estimate the relationship between sailing speed s_{it}^{sea} , wind speed w_{it} , and the deviation of the wind direction to the direction of travel θ_{it} using the following regression model

$$s_{it}^{sea} = \alpha + \beta w_{it} + \omega \cos(\theta_{it}) + \lambda_{it}, \quad (\text{A.1})$$

where λ_{it} is an unobserved error term. To evaluate the model fit, I split the sample into a training set ($N_{\text{train}} = 179,255$) and a validation set ($N_{\text{test}} = 9,432$). The model is estimated with ordinary least squares on the training set which gives the estimates presented in Table A1. As expected, there is a positive relation between sailing speed and wind speed. The findings are robust to including voyage fixed effects. Moreover, the coefficients are found to be precisely estimated. Finally, I evaluate the fitted model on the training sample. I find that the mean squared error is 2.18. In sum, this exercise confirms that wind speed and wind direction predict sailing speeds in this context.

In the final step, I construct a grid of 0.16×0.16 degrees resolution. For each cell, the average wind direction and wind speed over the 2011-2017 period is calculated. Based on this grid, I construct a directed graph where each node is a grid cell and each edge is a link to an adjacent cell. For any adjacent pair of nodes i, j , the predicted time to sail from i to j is calculated using the fitted model version of Equation A.1, denoted $\hat{s}_{it}^{sea}$. This results in a cost surface of predicted sailing times between any two nodes spanning the western hemisphere. I denote the sparse matrix containing the attained speed between all nodes $\mathbf{S}$.

Table A1: Wind speed, direction, and sailing time.

Dependent variable:	Sailing speed (knots)	
	(1)	(2)
Wind speed	0.079 (0.003)	0.086 (0.002)
Wind angle	-0.226 (0.018)	-0.171 (0.015)
Constant	3.531 (0.029)	
Voyage FE		✓

Note: The table reports OLS estimates. The unit of observation is a daily logbook entry. There are 179,287 observations. *Dependent variable*: Sailing speed in knots. *Standard errors*: Clustered at the voyage level. 4,832 clusters.

A.2.2 Land transportation.

Mules were the most common means of bulk transportation for most of the colonial period and up to the second half of the 19th century. Furthermore, pack animals were typically accompanied by humans on foot (Verhagen, Joyce and Groenhuijzen, 2019). I therefore model the shipping time with pack animals using geographical features, drawing on least-cost analysis tools from archaeology. The walking speed attained when traveling between two locations i and j will depend on whether travel occurs on road, the slope of the terrain, the elevation, and the landcover. I consider the following model of walking speed from Weiss et al. (2018),

$$s_{ij}^{land} = 6\kappa_i e^{-3.5|\text{slope}_{ij}+0.05|-\gamma\text{elev}_i} \quad (\text{A.2})$$

where s_{ij}^{land} denotes the walking speed attained between locations i and j . slope_{ij} denotes the average slope between points i and j . elev_i denotes the elevation in meters above sea level. κ_i is a scalar that adjusts the walking speed to account for differences in landcover. The model closely resembles the Tobler hiking function (Tobler, 1993) but incorporates differences in elevation and landcover which could be important in pre-industrial contexts. It follows that the travel speed on flat terrain at sea level is approximately 5 kilometers per hour on road ($\kappa_i = 1$). I assume 8 hours of walking per

day. As a result, 40 kilometers can be traveled in a day in this ideal case.

I approximate the travel time by combining Equation A.2 with a different data sources. I use a digital elevation model to calculate the average slope and elevation. Following Weiss et al. (2018), I set $\gamma = -0.0001072$ to account for lower mobility at high altitudes. I approximate the landcover using data on potential vegetation from Ramankutty and Foley (1999). This dataset gives a measure of the type of vegetation of a cell in the absence of human intervention. There are 15 types of potential vegetation in the data. I construct a mapping from the potential vegetation to the terrain coefficient κ_i using information on the speed of walking in various terrain types from Weiss et al. (2018). Five terrain types have a natural mapping between the potential vegetation data and the terrain coefficients. These are tropical forests, temperate forests, deserts, savanna, and shrubland. The terrain factors are 0.324 in a tropical forest, 0.648 in a temperate forest, 0.97 in a savanna, 0.6 in shrubland, and 0.6 in deserts. Inland water such as rivers and lakes can be traversed at half the speed (Herzog, 2014). I assume that travel on road is affected by the slope and elevation, but not the landcover κ_i . Finally, the location of roads is approximated with the location of postal routes using data from Stangl (2019b).

I again construct a measure of the travel time by setting up a grid of 0.16×0.16 degrees. For any adjacent pair of cells i, j , the predicted time to walk from i to j is calculated using Equation A.2 combined with the above data. This results in a matrix of approximated walking times between any two cells in North and South America. I denote the sparse matrix containing the attained speed between all nodes $\mathbf{L}$. The average speed of traversing a cell is 3.4 kilometers per hour.

A.3 Validation Exercises

Assessing the population data. I supplement and extend the city population data by consulting various regional and national sources. These sources are largely based on population and urbanization studies, colonial censuses, and regional economic studies and are discussed in Section A.1. In total, I adjust 70 year/city pairs. In this section, I use the national and regional sources to validate the data on cities in the Americas recorded in Buringh (2015). Reassuringly, there is a high correlation between the two datasets for the countries and period in question (ranging between 0.796 and 0.964). Furthermore, I re-estimate the models without the adjustments to the data as a robustness check. Unsurprisingly, I find very similar results as in the baseline case.

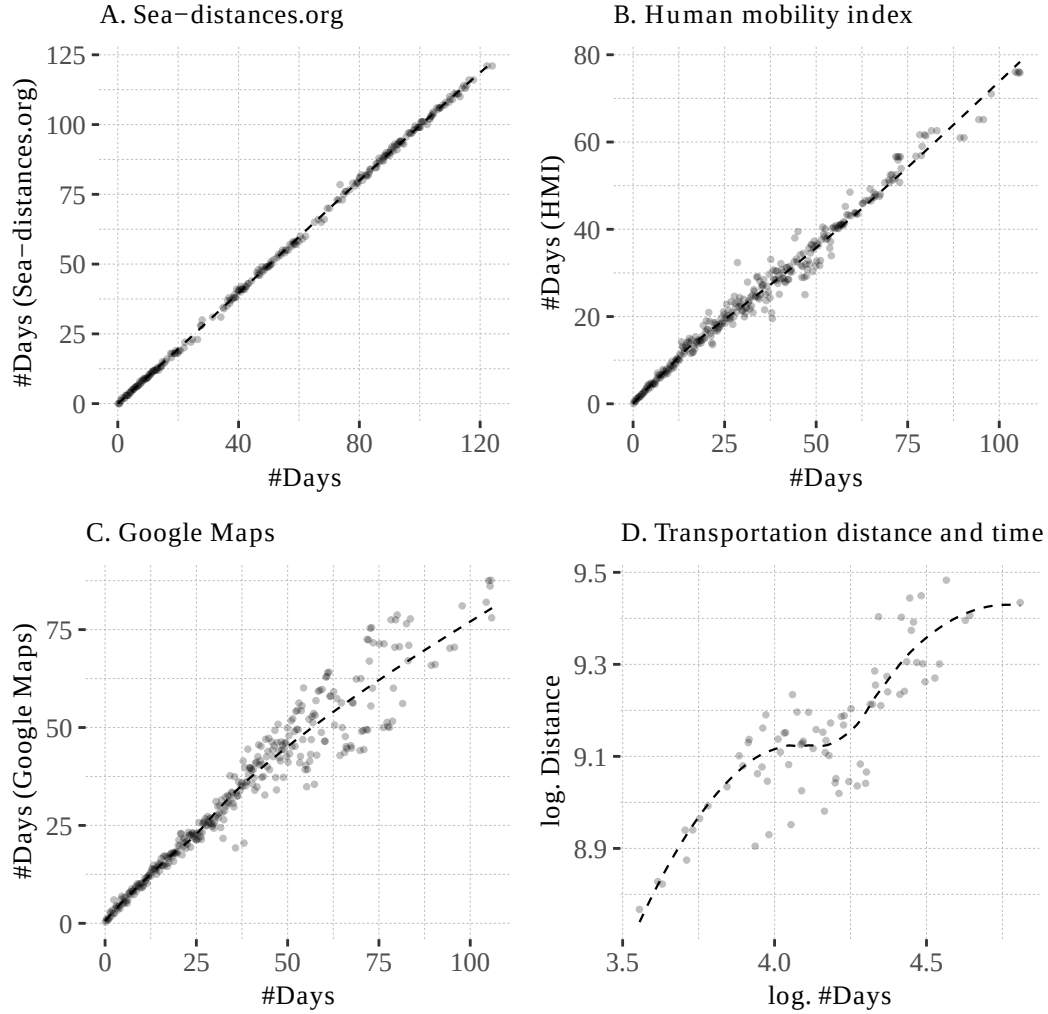


Figure A1: The figures depict the results from the main validation exercises. Panel A. depicts the relationship between estimated sailing times and sailing times from sea-distances.org for voyages between Cadiz and 21 ports in Spanish America. Panel B. depicts the relationship between the estimated travel times between large cities and travel times implied by the Human Mobility Index (Özak, 2010, 2018). Panel C. depicts the relationship between the estimated bilateral shipping times between major cities and google maps. Panel D. depicts the association between the number of days travel and distance between 1760 and 1810.

Assessing the transportation times. I validate the estimated transportation times for both land and maritime transportation. For maritime transportation, I compare the estimated shipping times to measures of sailing times from a database of bilateral sailing times.¹ For each port, I calculate the sailing time from Cadiz to all the ports in the dataset for which the website records information. The average speed of 4 knots is used, which is around the average speed of Spanish freight ships in 1750

¹A database of bilateral sailing times between major ports around the world. Data are available at seadistances.org.

(Kelly and Ó Gráda, 2019). For shipping on land, I compare the calculated shipping times with walking times using the Human Mobility Index (Özak, 2018) as well as Google Maps (downloaded in July 2021). Since there might be measurement error in the estimated transportation times, I consider an alternative approach where I calculate the least cost path on a cost surface without accounting for terrain or wind patterns. The association between the transportation time and resulting distance is provided in Panel D. Figure A1 shows that these alternative shipping time measures are correlated with the measure developed in this paper. Lastly, I assess the correlation between historical and contemporary wind patterns. There is a positive correlation between wind speed and direction in these two datasets (0.245 and 0.331).

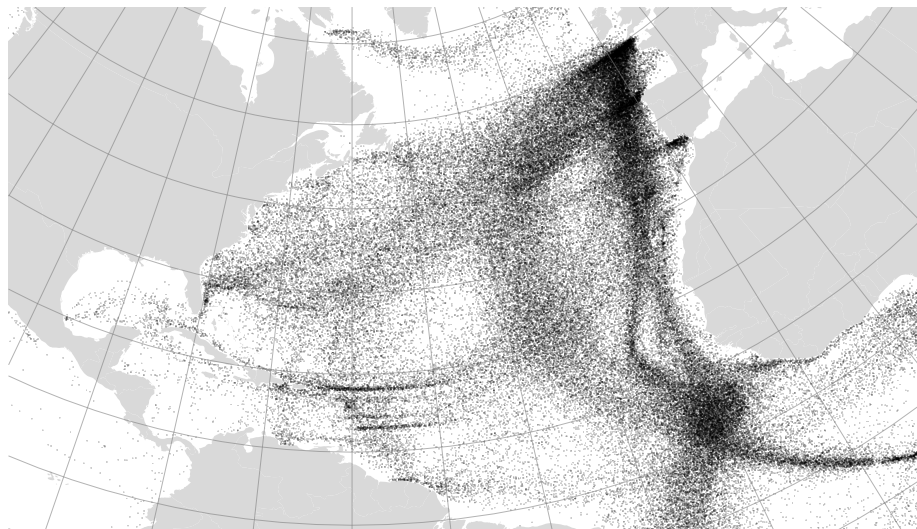


Figure A2: The figure depicts a snapshot of the logbook entries used to estimate the travel times in the main analysis. Each dot represents a logbook entry between 1750 and 1855. Source: CLIWOC 2.1.

B Model Details

This section provides a more complete description of the model, the equilibrium conditions, and the estimation strategy. The model follows Allen and Donaldson (2022) closely.

B.1 Model Ingredients

Geography and timing. The environment consists of R cities indexed by i or j . Each city differs in their location productivity $\mathcal{A} = \{\bar{A}_i\}_{i=1}^R$ and availability of land $\mathcal{H} = \{\bar{H}_i\}_{i=1}^R$. Locations also differ in their connections to other cities. Goods trade is subject to iceberg trade costs captured by the matrix $\mathcal{T}_t = \{\tau_{ijt}\}_{i,j \in R}$ where $\tau_{ijt} \geq 1$ for all i, j . Migration frictions are given by the matrix $\mathcal{M}_t = \{m_{ijt}\}_{i,j \in R}$. Both migration and trade costs are potentially time-varying. Finally, $\mathcal{L}_0 = \{L_{i0}\}_{i \in R}$ denotes the initial population across the R cities. Together, this defines the geography at time t , denoted $\mathcal{G}_t$. Following Allen and Donaldson (2022), each city is inhabited by young and old individuals. Old individuals supply labor and consume, give birth to one young individual, and then leave the model. Young individuals choose where to live when old to maximize their utility.

Preferences. The utility derived from consumption in city i at time t is defined over a composite of differentiated varieties C_{it} and a non-traded good f_{it} . The non-traded good can be thought of as food that is sourced from the city's immediate hinterland. The preferences are defined as follows,

$$u_{it} = \left(\frac{C_{it}}{\mu} \right)^\mu \left(\frac{f_{it}}{1 - \mu} \right)^{1-\mu}, \quad (\text{A.3})$$

where $C_{it} = \left(\sum_{j \in N} b_{ji}^\frac{1}{\sigma} c_{jit}^\frac{\sigma-1}{\sigma} \right)^\frac{\sigma}{\sigma-1}$. c_{jit} is the amount of the city j specific good consumed in city i and b_{ji} is an exogenous preference shifter. It follows that the expenditure shares on C_{it} and f_{it} are constant and given by μ and $1 - \mu$ respectively. The demand functions are given by $f_{it} = (1 - \mu)y_{it} / r_{it}$ and $C_{it} = \mu y_{it} / P_{it}$ where $P_{it} = \left(\sum_{j \in N} b_{ji} p_{jit}^{1-\sigma} \right)^\frac{1}{1-\sigma}$ is the price index for traded varieties, r_{it} the price of the non-traded good, and y_{it} income per capita. Moreover, the share of expenditure devoted to traded varieties spent on any particular variety is given by $\mu_{jit} = b_{ji}(p_{jit} / P_{it})^{1-\sigma}$. As a result, the individual utility derived from consumption in equilibrium for in city i is given by the indirect utility function, $V_{it} = w_{it} / \mu P_{it}^\mu r_{it}^{1-\mu}$. Finally, every individual supplies one unit of labor inelastically.

Technology and market structure. Traded goods q_{it} , are produced with a constant returns to scale technology $q_{it} = A_{it}L_{it}$ where L_{it} is the number of workers in city i and A_{it} the productivity level given by $A_{it} = \bar{A}_i L_{it}^{\alpha_1} L_{it-1}^{\alpha_2}$. There is assumed to be perfect competition in goods and factor markets. Since a worker in city i can produce A_{it}

units of a good and the nominal wage is w_{it} , the price of the good in city i is given by $p_{it} = w_{it} / A_{it}$. Let F_{it} denote the supply of the non-traded good. The aggregate demand is defined as $F_{it} = f_{it}L_{it}$. The non-traded good is produced using a linear technology where land is the only input. It follows that $F_i = \bar{H}_i$, where $\bar{H}_i$ denotes the availability of land in the city and the immediate hinterland. The price for the non-traded good in equilibrium is given by $r_{it} = (1 - \mu)Y_{it} / \bar{H}_i$ where Y_{it} denotes the total income of city i .

Trade. The traded varieties can be transported across cities subject to iceberg trade costs. The price faced by consumers in city i for good a good originating in j is therefore given by $p_{jit} = \tau_{jit}w_{jt} / A_{jt}$. The value of goods transported from city j to city i denoted X_{jit} , takes the gravity form given by

$$X_{jit} = b_{ji}\tau_{jit}^{1-\sigma} \left(\frac{w_{jt}}{A_{jt}} \right)^{1-\sigma} \mu Y_{it} P_{it}^{\sigma-1}. \quad (\text{A.4})$$

Using the expression for the local productivity A_{it} and that $\mu Y_{it} = w_{it}L_{it}$ gives the expression in the derivations that follow.

Migration. The utility of an individual moving from city i to city j is given by $V_{ijt} = \frac{V_{jt}}{m_{ij}}\epsilon_{jt}$ where ϵ_{jt} represents individual-level heterogeneity in preferences over cities, assumed to be an *i.i.d.* draw from a multivariate Fréchet-distribution with shape parameter θ . The CDF is thus given by $F(x) = e^{-x^{-\theta}}$. m_{ij} captures bilateral migration costs. An individual chooses to move from i to j if the realized utility of that city is higher, which happens with probability $Pr(V_{ijt}\epsilon_{jt} \geq V_{ikt}\epsilon_{kt} \forall k \neq j \in R)$. Since the idiosyncratic location preferences are *i.i.d.* and there is a continuum of agents in each location this probability corresponds to the share of individuals in city i moving to j , denoted π_{ijt} . It follows that the conditional probability of moving is,

$$\pi_{ijt}|\epsilon_{jt} = Pr\left(\epsilon_{kt} \leq \frac{V_{ijt}\epsilon_{jt}}{V_{ikt}} \epsilon_{kt} \forall k \neq j \in R\right) = \prod_{k \neq j} Pr\left(\epsilon_{kt} \leq \frac{V_{ijt}\epsilon_{jt}}{V_{ikt}}\right). \quad (\text{A.5})$$

As a result, the fraction of individuals moving from i to j is given by,

$$\pi_{ijt} = \frac{(V_{jt} / m_{ij})^\theta}{\sum_{k \in R} (V_{kt} / m_{ik})^\theta}. \quad (\text{A.6})$$

The number of people moving from i to j is the $L_{ijt} = \pi_{ijt}L_{it-1} = L_{it-1}V_{jt}^\theta \Pi_{it}^{-\theta} m_{ij}^{-\theta}$

where $\Pi_{it}^\theta = \sum_{j \in R} m_{ij}^{-\theta} V_{jt}^\theta$. This gives the expressions in the text.

Equilibrium and steady-state. The equilibrium for period t , given the geography $\mathcal{G}_t$, is characterized by the following four equations.

1. $w_{it} L_{it} = \sum_{k \in R} X_{ikt}$ (goods market clearing).
2. $\sum_{k \in R} X_{ikt} = \sum_{k \in R} X_{kit}$ (balanced trade).
3. $L_{it} = \sum_{k \in R} L_{kit}$ (total population equals the number arriving in the city).
4. $L_{it-1} = \sum_{k \in R} L_{ikt}$ (total population equals the number exiting the city).

Since trade is balanced and trade costs are quasi-symmetric, it follows that the origin and destination terms in the gravity equation are proportional (Allen and Arkolakis, 2014). Therefore $w_{it}^{1-\sigma} A_{it}^{\sigma-1} \propto w_{it} L_{it} P_{it}^{\sigma-1}$. Using the indirect utility, it follows that $P_{it}^{\sigma-1} \propto w_{it}^{\sigma-1} V_{it}^{\frac{1-\sigma}{\mu}} \bar{H}_i^{\frac{(1-\mu)(\sigma-1)}{\mu}} L_{it}^{\frac{(\mu-1)(\sigma-1)}{\mu}}$. Inserting this gives the following expression for the nominal wage,

$$w_{it} \propto A_{it}^{\tilde{\sigma}} L_{it}^{\tilde{\sigma}(\frac{1}{1-\sigma} + \frac{1-\mu}{\mu})} V_{it}^{\frac{\tilde{\sigma}}{\mu}} \bar{H}_i^{\frac{(\mu-1)\tilde{\sigma}}{\mu}}, \quad (\text{A.7})$$

where $\tilde{\sigma} = 1 - \sigma / 1 - 2\sigma$. Using the functional form of the agglomeration spillovers, the goods market clearing condition (1.) and the equilibrium restrictions on labor mobility (3. and 4.) then results in the following equations for the equilibrium of the model,

$$\begin{aligned} L_{it}^{\tilde{\sigma}(1-\sigma\frac{\mu-1}{\mu}-\alpha_1(\sigma-1))} V_{it}^{\frac{\tilde{\sigma}\sigma}{\mu}} &= \phi \bar{A}_i^{\tilde{\sigma}(\sigma-1)} \bar{H}_i^{\frac{(1-\mu)\tilde{\sigma}\sigma}{\mu}} L_{it-1}^{\alpha_2\tilde{\sigma}(\sigma-1)} \times \\ \sum_{j \in R} b_{ij} \tau_{ijt}^{1-\sigma} \bar{A}_j^{\tilde{\sigma}\sigma} L_{jt}^{\tilde{\sigma}(1+\alpha_1\sigma+\frac{(1-\sigma)(1-\mu)}{\mu})} V_{jt}^{\frac{\tilde{\sigma}(1-\sigma)}{\mu}} \bar{H}_j^{\frac{\tilde{\sigma}(\mu-1)(1-\sigma)}{\mu}} L_{jt-1}^{\sigma\tilde{\sigma}\alpha_2}, \end{aligned} \quad (\text{A.8})$$

$$\Pi_{it}^\theta = \sum_{j \in R} m_{ij}^{-\theta} V_{jt}^\theta, \quad (\text{A.9})$$

$$L_{it} V_{it}^{-\theta} = \sum_{j \in R} m_{ij}^{-\theta} \Pi_{jt}^{-\theta} L_{jt-1}, \quad (\text{A.10})$$

where ϕ is a constant that is pinned down by the total population size $\sum_{k \in R} L_{kt} = \bar{L}_t$. These expressions are also used to calculate the location fundamentals. Next, consider the long-run steady state of the model. There will still be migration in the model in the steady-state, but bilateral flows will cancel out leaving the relative size of all cities fixed. The long-run steady state is characterized by $L_{it} = L_{it-1}$ for all i . Using this

condition gives the following system of equations for the steady state of the model.

$$L_{it}^{\tilde{\sigma}\left(1-\sigma\frac{\mu-1}{\mu}-(\alpha_1+\alpha_2)(\sigma-1)\right)} V_{it}^{\frac{\tilde{\sigma}\sigma}{\mu}} = \bar{A}_i^{\tilde{\sigma}(\sigma-1)} \bar{H}_i^{\frac{(1-\mu)\tilde{\sigma}\sigma}{\mu}} \times \sum_{j \in R} b_{ij} \tau_{ijt}^{1-\sigma} \bar{A}_j^{\tilde{\sigma}\sigma} L_{jt}^{\tilde{\sigma}(1+(\alpha_1+\alpha_2)\sigma+\frac{(1-\sigma)(1-\mu)}{\mu})} V_{jt}^{\frac{\tilde{\sigma}(1-\sigma)}{\mu}} \bar{H}_j^{\frac{\tilde{\sigma}(\mu-1)(1-\sigma)}{\mu}} \quad (\text{A.11})$$

$$\Pi_{it}^{\theta} = \sum_{j \in R} m_{ij}^{-\theta} V_{jt}^{\theta} \quad (\text{A.12})$$

$$L_{it} V_{it}^{-\theta} = \sum_{j \in R} m_{ij}^{-\theta} \Pi_{jt}^{-\theta} L_{jt}. \quad (\text{A.13})$$

B.2 Existence and Uniqueness

The existence and uniqueness of the equilibrium and steady state of the model can be characterized using the results in Allen and Donaldson (2022) and Allen, Arkolakis and Li (2024). The system in equations A.8, A.9, and A.10 contain $3 \times R$ endogenous variables for each period t . Ordering the endogenous variables as L , V , and Π gives the following matrices of coefficients,

$$\mathbf{B} = \begin{bmatrix} \tilde{\sigma}(1 - \sigma\frac{\mu-1}{\mu} - \alpha_1(\sigma-1)) & \frac{\tilde{\sigma}\sigma}{\mu} & 0 \\ 0 & 0 & \theta \\ 1 & -\theta & 0 \end{bmatrix}, \quad (\text{A.14})$$

$$\mathbf{\Gamma} = \begin{bmatrix} \tilde{\sigma}(1 + \alpha_1\sigma + \frac{(1-\sigma)(1-\mu)}{\mu}) & \frac{\tilde{\sigma}(1-\sigma)}{\mu} & 0 \\ 0 & \theta & 0 \\ 0 & 0 & -\theta \end{bmatrix}. \quad (\text{A.15})$$

Allen, Arkolakis and Li (2024) show that there exists a unique equilibrium if the largest eigenvalue of $\mathbf{A}$ is less than 1, where $\mathbf{A}_{hh'}$ is the absolute value of the corresponding element of $\mathbf{\Gamma B}^{-1}$. As the equations pinning down the steady state are the same as for the equilibrium except for the parameters, the existence and uniqueness follow directly from the above condition where α_1 is replaced by $\alpha_1 + \alpha_2$. For the baseline parameters of the model I find that both the equilibrium and steady state are unique.

B.3 Comparative Statics

The deterministic component of indirect utility is given by $V_{it} = w_{it} P_{it}^{-\mu} r_{it}^{\mu-1} / \mu$ where again $r_{it} = (1 - \mu) w_{it} L_{it} / \bar{H}_i$ and assuming quasi-symmetric trade costs $w_{it}^{1-\sigma} A_{it}^{\sigma-1} \propto w_{it} L_{it} P_{it}^{\sigma-1}$. As a result, $V_{it}^{-\sigma} = w_{it}^{-\mu\sigma} P_{it}^{\mu\sigma} L_{it}^{\sigma(1-\mu)} \bar{H}_i^{\sigma(\mu-1)}$ and $w_{it}^{-\sigma\mu} \propto A_{it}^{\mu(1-\sigma)} L_{it}^{\mu} P_{it}^{\mu(\sigma-1)}$. This gives,

$$L_{it}^{\mu+\sigma(1-\mu)+\alpha_1\mu(1-\sigma)} \propto V_{it}^{-\sigma} \bar{A}_i^{\mu(\sigma-1)} L_{it-1}^{\alpha_2\mu(\sigma-1)} P_{it}^{\mu(1-2\sigma)} \bar{H}_i^{\sigma(1-\mu)}, \quad (\text{A.16})$$

which in turn gives,

$$\begin{aligned} \nu \ln L_{it} = \kappa - \sigma \ln V_{it} + \alpha_2 \mu (\sigma - 1) \ln L_{it-1} - \mu (2\sigma - 1) \ln P_{it} + \\ \mu (\sigma - 1) \ln \bar{A}_i + \sigma (1 - \mu) \ln \bar{H}_i, \end{aligned} \quad (\text{A.17})$$

where $\nu = \mu + \sigma(1 - \mu) + \alpha_1\mu(1 - \sigma)$ and κ is a constant. Next, the expression is used to solve for the current population as a function of the full path of endogenous and exogenous variables,

$$\begin{aligned} \ln L_{it} = \psi + \left(\frac{\alpha_2 \mu (\sigma - 1)}{\nu} \right)^t \ln L_{i0} - \frac{\sigma}{\nu} \sum_{k=0}^{t-1} \left(\frac{\alpha_2 \mu (\sigma - 1)}{\nu} \right)^k \ln V_{it-k} - \\ \frac{\mu (2\sigma - 1)}{\nu} \sum_{k=0}^{t-1} \left(\frac{\alpha_2 \mu (\sigma - 1)}{\nu} \right)^k \ln P_{it-k} + \tilde{A}_i + \tilde{H}_i, \end{aligned} \quad (\text{A.18})$$

where $\tilde{A}_i = \ln \bar{A}_i \mu (\sigma - 1) \sum_{k=0}^{t-1} \left(\frac{\alpha_2 \mu (\sigma - 1)}{\nu} \right)^k$ and $\tilde{H}_i = \ln \bar{H}_i \sigma (1 - \mu) \sum_{k=0}^{t-1} \left(\frac{\alpha_2 \mu (\sigma - 1)}{\nu} \right)^k$.

To analyze the comparative statics underlying the reduced form exercise consider the above equation. I consider the direct effect of a reduction in the trade cost to Europe in period k for city i (that is a change in τ_{eik}), while ignoring general equilibrium effects. First note that the elasticity of the city's population size today with respect to the price index in period k is given by

$$\frac{\partial \ln L_{it}}{\partial \ln P_{ik}} = - \frac{\mu (2\sigma - 1)}{\nu} \left(\frac{\alpha_2 \mu (\sigma - 1)}{\nu} \right)^k < 0. \quad (\text{A.19})$$

Next, lower trade costs to Europe in period k lead to a lower price index for city i .

$$\frac{\partial P_{ik}}{\partial \tau_{eik}} = P_{ik}^{\sigma} \tau_{eik}^{-\sigma} b_{ei} \left(\frac{w_{ek}}{A_{ek}} \right)^{1-\sigma}. \quad (\text{A.20})$$

As a result, a lower trade cost to Europe has a persistent and positive effect on the population size in location i , that is $\frac{\partial \ln L_{it}}{\partial \tau_{eik}} < 0$. Moreover, when $\alpha_2 \mu (\sigma - 1) / \nu < 1$, it follows that $\lim_{l \rightarrow \infty} \partial \ln L_{it} / \partial \tau_{eil} = 0$. Finally, consider the effect of changes in trade for cities differing in initial market size (captured by A_{ik}). Since $\partial P_{ik} / \partial A_{ik} < 0$, it follows that the marginal impact of τ_{eik} declines as A_{ik} becomes larger.

C Further Historical Background

This section provides an extended discussion about the historical background for the analysis. As in the main body of the text, I discuss the background and the motivation for the trade reform as well as the historical relationship between trade and economic development within the Spanish Empire.

A central aim of commercial policy in the 18th century was to promote state wealth acquisition through trade surpluses (Findlay and O'Rourke, 2007). In the Spanish context, this was achieved through a range of policies restricting trade. First, trade was limited to four ports in the Americas (Cartagena de Indias, El Callao, Portobello/Nombre de Dios, and Veracruz) and only Seville/Cádiz in Europe. Further, the frequency of travel and the routes were restricted. Typically, only two fleets left Spain every year: the New Spain *flota* destined for Veracruz, and the *Tierra Firme galeones* destined for Cartagena and Portobello. In the Pacific, shipping was conducted by *Armada del Sur*, which carried goods from the trade fairs in Portobello to Pacific ports in South America (Walker, 1979). Moreover, the Manilla galleon would sail between Acapulco and Manilla. Official information was carried by *aviso* ships, which were light carriers operating separately from the commercial system and were not permitted or equipped to carry freight. Third, participation in Atlantic trade was restricted to Spanish merchants. Finally, there were high tax rates on imports and exports. The duties typically depended on the origin of the goods, with lower rates on goods originating from Spain. These measures effectively monopolized trade in the merchant guilds in Seville (later Cadiz), Mexico City, and Lima, and only the merchant guilds of these cities were allowed to buy and sell goods at the trade fairs at Veracruz and Portobello. These locations then in turn managed trade with other locations in their respective vicerealties, typically transported by third parties using mule trains or wagons depending on road conditions.

The system limited trade with Europe across large parts of the Spanish empire in America, however, there was still some maritime communication and trade occurring

in locations too remote relative to the large trade routes. In addition to dispatch ships (*avisos*), ships sailing under special permission of the crown (*registros*) occasionally supplied ports that were too remote relative to the large trade routes. However, this was never done at a sufficiently large scale (Walker, 1979) and increased the reliance on contraband trade which was sizeable (Christelow, 1942). While as a rule, there were no restrictions on inter-regional trade (Elliott, 2006, p. 111), there were cases where inter-regional was discouraged. For example, there were policies in place to limit trade between the Viceroyalties of Peru and New Spain to reduce the demand for the goods of the Manilla Galleon in Peru. Another example is the erection of a customs barrier in Córdoba (Argentina) in 1618 (Scobie, 1971, p. 53).

Restrictions on trade and high trade costs ensured that trade was limited to non-competing goods with a high value-to-weight ratio. Important exports during the period beyond precious metals were hides, tallow, sugar, indigo, and cochineal (Rahn Phillips, 1990). The slave trade was subject to different rules. Trade of slaves was allowed for British ships from the early to the mid-18th century as a result of the treaty of Utrecht, the *asiento* (Walker, 1979). These measures facilitated the naval defense of convoys and limited imports to the Americas, thus limiting the flow of bullion to other places than the Iberian Peninsula while keeping prices for Spanish exports artificially high. It also facilitated the managing of risk in a context where long shipping times and costly communication made it difficult to predict demand (Baskes, 2013). As a result, in addition to remittances directly controlled by the crown, private remittances to Spain were substantial (Cuenca-Esteban, 2008). However, a likely consequence of Spanish mercantilist policies before the liberalization in the late 18th century was the underdevelopment of peripheral areas in America (Fisher, 1997, p. 73). There were few changes to this system until the second half of the 18th century but there were some notable changes. In return for the support of France during the War of the Spanish Succession, French ships were allowed to trade along the Pacific coast for some time. Moreover, as part of the treaty of Utrecht, the English were granted the right to send a ship of 500 tons to the trade fairs.

Reforming transatlantic trade. Beginning in the 18th century, Spanish policymakers were induced by geopolitical considerations, originating mainly in Europe, to overhaul the external trading system (Elliott, 2006). In the immediate aftermath of Spain's defeat in the Seven Years' War, a special *junta* was appointed under Charles III to "review ways to address the backwardness of Spain's commerce with its colonies and foreign nations" Stein and Stein (2003). Drawing on ideas for reforming the system of govern-

ment in America that had been circulating for a long time, the *junta* proposed the abolition of the Cádiz monopoly as well as the fleet system. Further, it proposed opening 14 ports on the Iberian Peninsula as well as 35 ports in the Americas (Fisher, 1997). The ports that were opened on the Iberian peninsula in this period were Malaga, Almería, Cartagena, Alicante, Tortosa, Barcelona, Santander, Gijón, La Coruña, Palma de Mallorca, Santa Cruz de Tenerife. While the reform is believed to have a role in promoting the rise of the Barcelona textile industry, in the early 19th century, around 80 percent of Spanish trade with the Americas still went through the port of Cádiz (Fisher, 1997). Several ports in the Caribbean were opened already in 1765. Santo Domingo, Puerto Rico, Margarita, and Trinidad were opened for direct trade with Spain in 1765. Further, reform was slowed by the Esquilache riots in 1766 and the liberalization measures culminated in the decree of free trade in 1778, which opened several of the remaining ports. This was except Venezuela (Caracas), where it was believed the Caracas company's tobacco monopoly was worth protecting, and New Spain. Even so, especially Veracruz was affected by the changes before the late 1780s due to the abolition of the convoy system and the increased prevalence of register ships. In the 1780s, the remaining ports followed. Spanish communication with the Americas was disrupted during the Napoleonic wars (O'Rourke, 2006). Out of necessity, trade with neutral nations was therefore allowed. This marked the end of Spain's ability to enforce protected trade with the colonies. By the beginning of the 19th century, Spanish America enjoyed *de facto* although not *de jure* unrestricted trade with foreigners (Fisher, 1998). As a result, direct trade with Britain, not mediated through Spain, grew in importance (Prados de la Escosura and Casares, 1983). Independence was mostly followed by high tariffs, mainly driven by the revenue needs of post-independence governments (Coatsworth and Williamson, 2004).

The historical literature emphasizes the role of European interstate competition and the resulting increased need for a modernized imperial defense as motivating the reform. Thus, the drive to reform the Spanish commercial system can be understood as being motivated by the intense interstate competition between the European states of the 18th century (Kuethe and Andrien, 2014). Highlighted in the historical literature as an important impetus for the reform was the "humiliating" capture of Havana and Manila by the British during the Seven Years' War. This opened a window of opportunity for reform-minded policymakers in Spain who now could justify reforming the commercial system with concerns about the territorial integrity of the empire in what has been described in the historical literature as a "defensive modernization" (Stein and Stein, 2003). Furthermore, the commercial expansion of Havana during the British

occupation showcased the economic potential of the Spanish colonies.

The reform was therefore implemented rapidly after the Seven Years' War (Fisher, 1997). As a result, the timing of the reform is mainly driven by intensified interstate competition in Europe, rather than economic development in the Americas directly. Moreover, the reform was implemented from above, and no significant ports in which the policies were applied were excluded. This is also apparent from the fact that the policies were resisted by powerful interests in the Spanish Empire (Baskes, 2013). Finally, the selection of ports is unlikely to be driven by the perceived commercial potential of its hinterland. This is apparent when considering the case of New Spain. As the most important colony of the Spanish empire in America, it was believed New Spain would have diverted too much trade away from other regions (Fisher, 1997). Moreover, avoiding confrontation with merchants in New Spain whose resources were a key source of revenue for the crown. As a result, New Spain was not subject to the reform until the late 1780s.

It is generally agreed upon that the reform increased trade. This was recognized by contemporaries as well as in the historical literature. Floridablanca (minister under Charles III) wrote about a fortunate revolution (*feliz revolución*) when referring to Spanish export growth after 1778. When referring to Veracruz, went from "gloomy and ugly" to "elegant and growing" (Stein and Stein, 2003). The magnitudes in the economic history literature are contested (Cuenca-Esteban, 2008). Colonial imports to Spain increased tenfold and exports from Spain to the colonies fourfold according to Fisher (1985), while more modest estimates are found in Cuenca-Esteban (2008), also suggesting large effects. Fisher (1993) provides data on the composition of Spanish imports from Spanish America between 1782 and 1796 for the ports of Cadiz and Barcelona (which accounted for around 88 percent of imports from Spanish America). Precious metals still accounted for 56.4 percent of imports through this period. The other commodities were typically high-value agricultural commodities (tobacco 13.6, cacao 7.8, sugar 5.5, indigo 5.2, cochineal 4.2, hides 3.4 and cotton 0.4 percent) (Fisher, 1993). Cadiz remained the dominant port for trade with Spanish America between 1778 and 1796 (76.4 percent of total exports and 84.2 percent of imports). The remaining important ports were Barcelona (9.6 and 3.8 percent), Malaga (4.8 and 1.3 percent), Santander (3.3 and 2.6 percent), and La Coruña (3 and 6.8 percent) (Fisher (1993) p.20 and p.25).

Some accounts highlight that the lower trade costs induced by the reform promoted agricultural development. "...for the first time, the metropolis succeeded in unleashing the agricultural potential of its American possessions whilst also promoting the

continued expansion of mining production. The relationship between this economic growth and the liberalization of trade is abundantly clear”, (Fisher, 1997, p. 197). Moreover, lower trade costs induced by unrestricted sailing potentially allowed for specialization in a wider range of commodities, such as more perishable goods. However, bullion remained an important export commodity (Fisher (1997), p. 38). Moreover, it has been argued that the population and economies of previously stagnant peripheral colonies in Spanish America grew rapidly (Mahoney, 2010). In summary, the historical literature suggests the restrictions imposed on trade in goods with the Americas stunted economic development, and efforts induced by European interstate competition to relax these marked the beginning of a process that would have important effects on trade and economic development in the second half of the 18th century.

D Additional Results

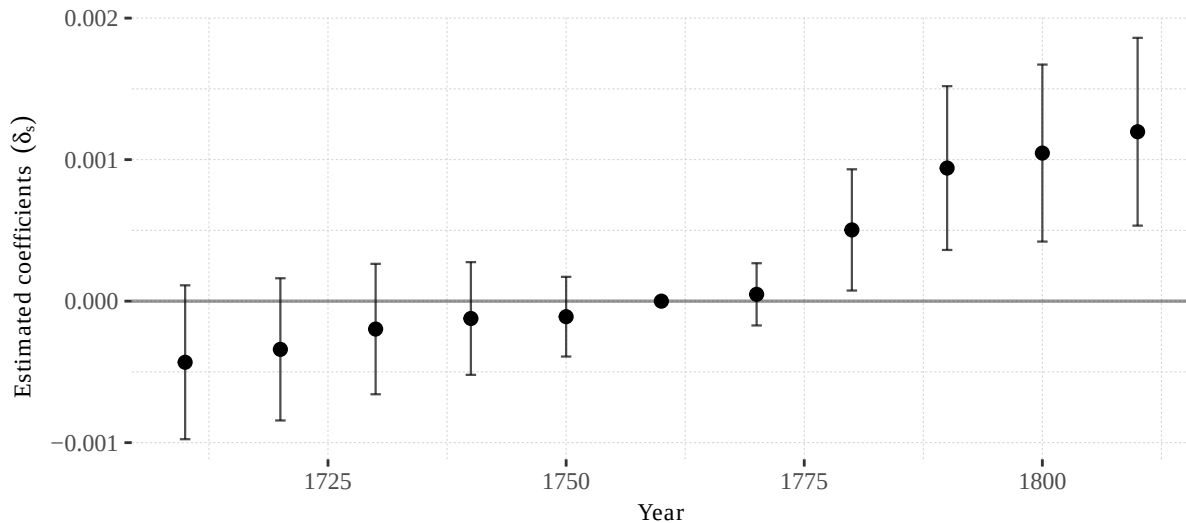


Figure A3: The figure depicts the estimated coefficients of the reduction in transportation times to Europe interacted with time indicators from Equation 6. *Dependent variable:* Indicator variable taking the value one if the grid cell contains a settlement. *Observations:* 5,3581 cells observed every decade between 1710 and 1810. $5,3581 \times 11 = 58,9391$ observations. *Standard errors:* Clustered at the grid-cell level (5,3581 clusters).

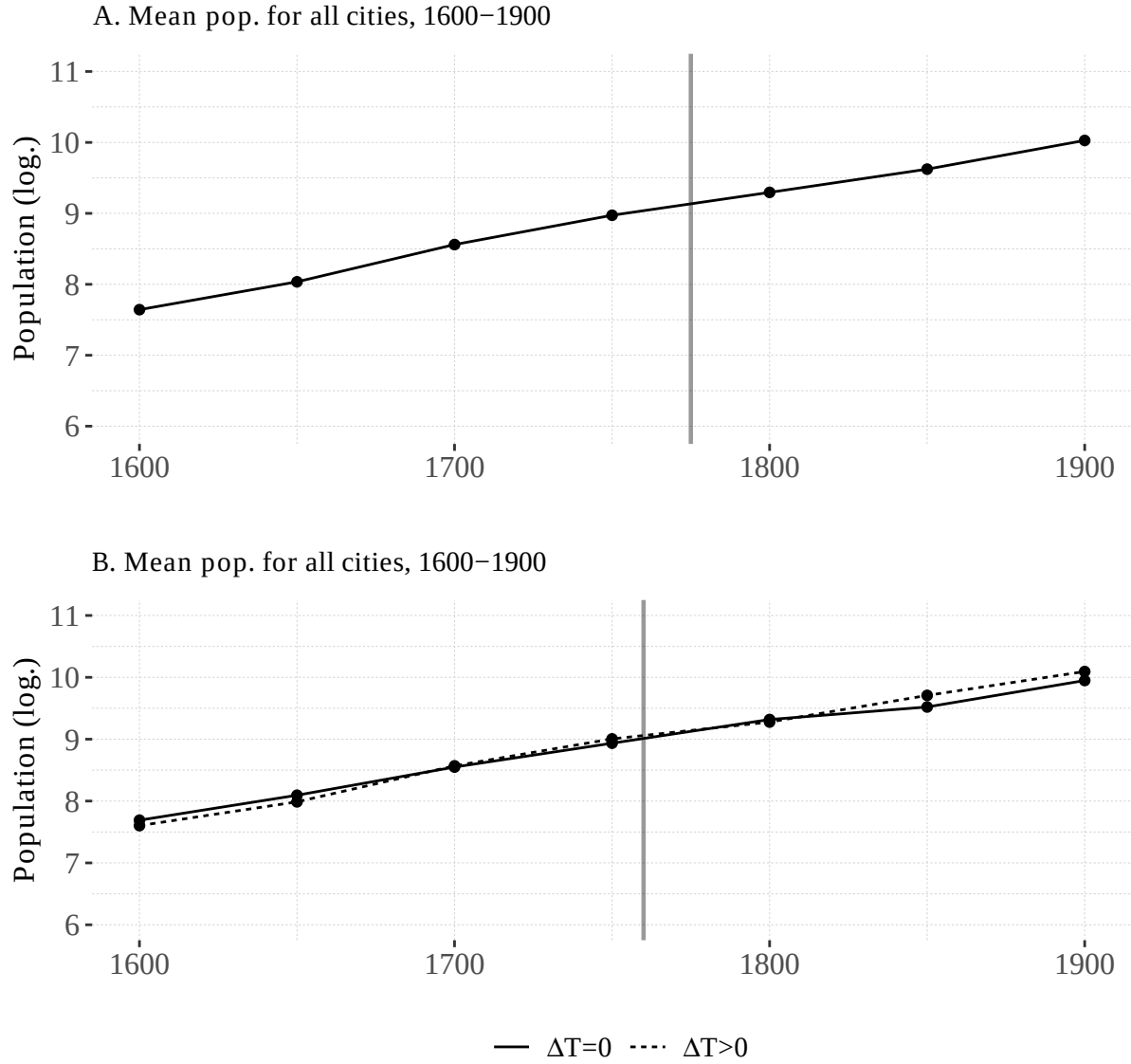


Figure A4: The figure depicts the average population across the cities in the full sample. Panel A. depicts the average population (log.) across all the cities. Panel B. depicts the average population (log.) by treatment status. $\Delta T_i > 0$ denotes cities that experienced changes in transportation times to Europe due to the reform. $\Delta T_i = 0$ denotes cities that did not experience changes in transportation times to Europe due to the reform.

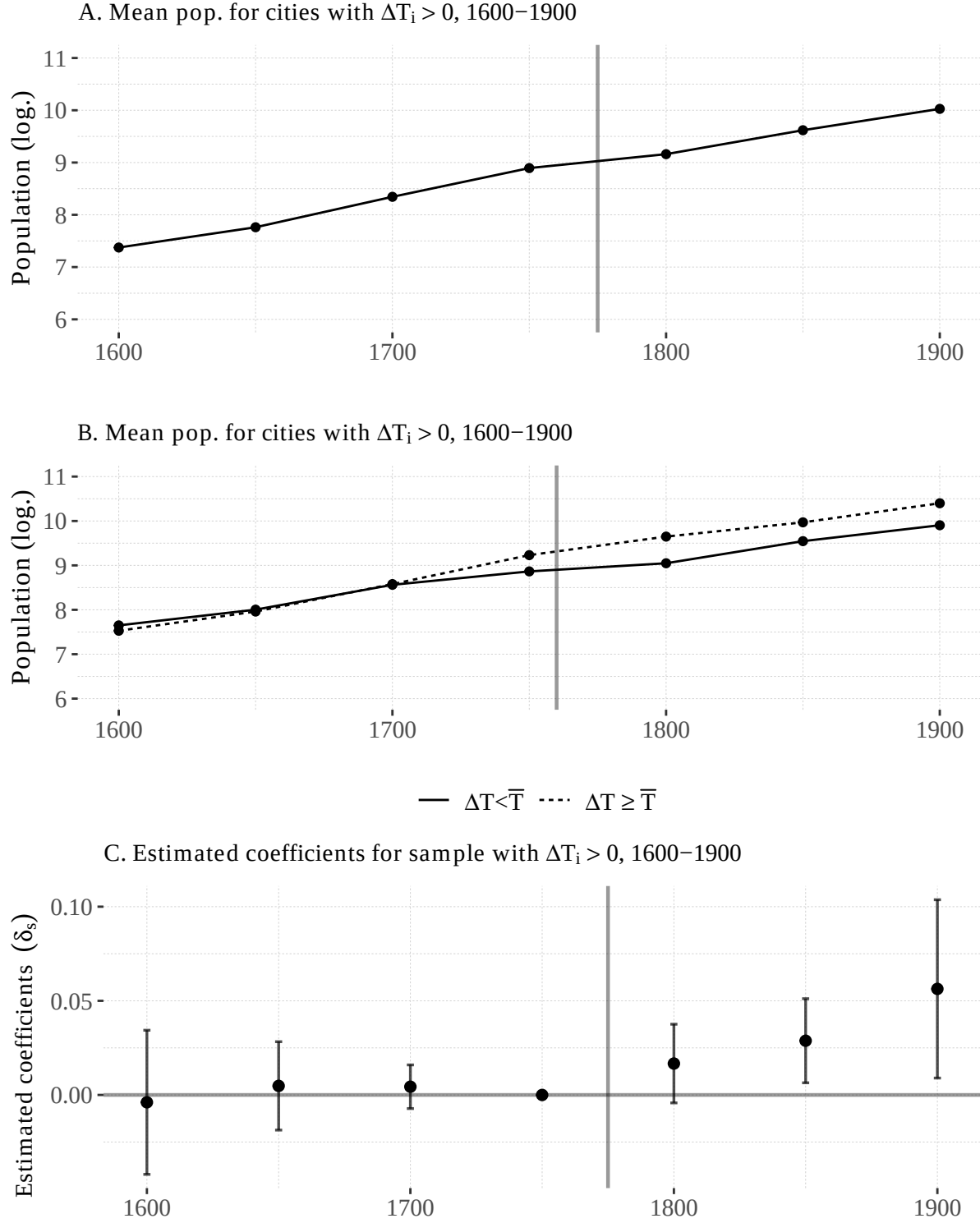


Figure A5: Panel A depicts the mean population for cities. Panel B depicts the average population by treatment intensity. $\bar{T}$ denotes the average change in transportation times. Panel C depicts the estimated coefficients of the reduction in transportation times to Europe interacted with time indicators from Equation 6. The sample is restricted to cities with $\Delta T_i > 0$.

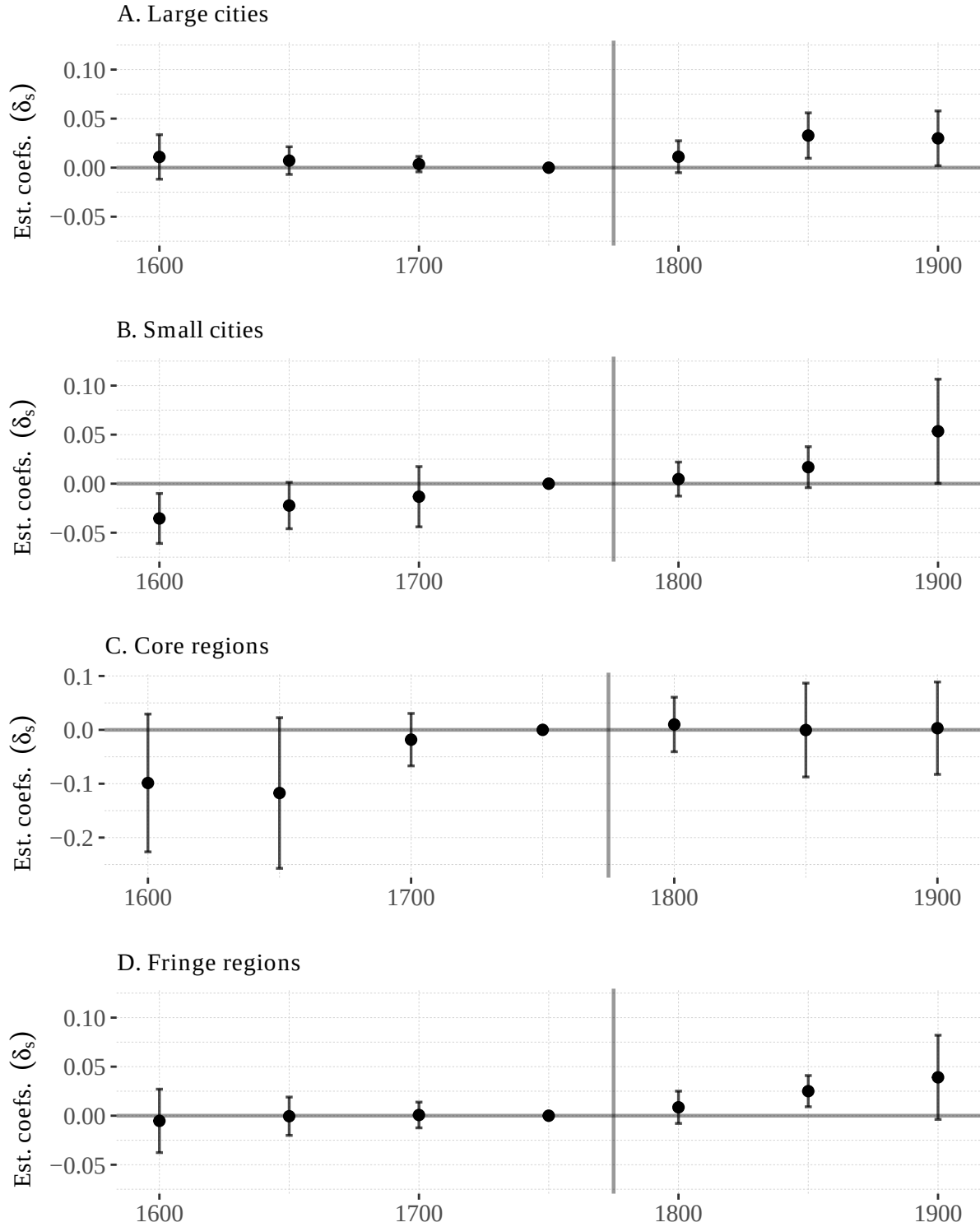


Figure A6: The figure depicts the estimated coefficients of the reduction in transportation times to Europe interacted with time indicators from Equation 6. Panel A/B is estimated on the sample of cities with a population above/below the median in 1750. Panel A/B is estimated on the sample of cities in the core/fringe regions.

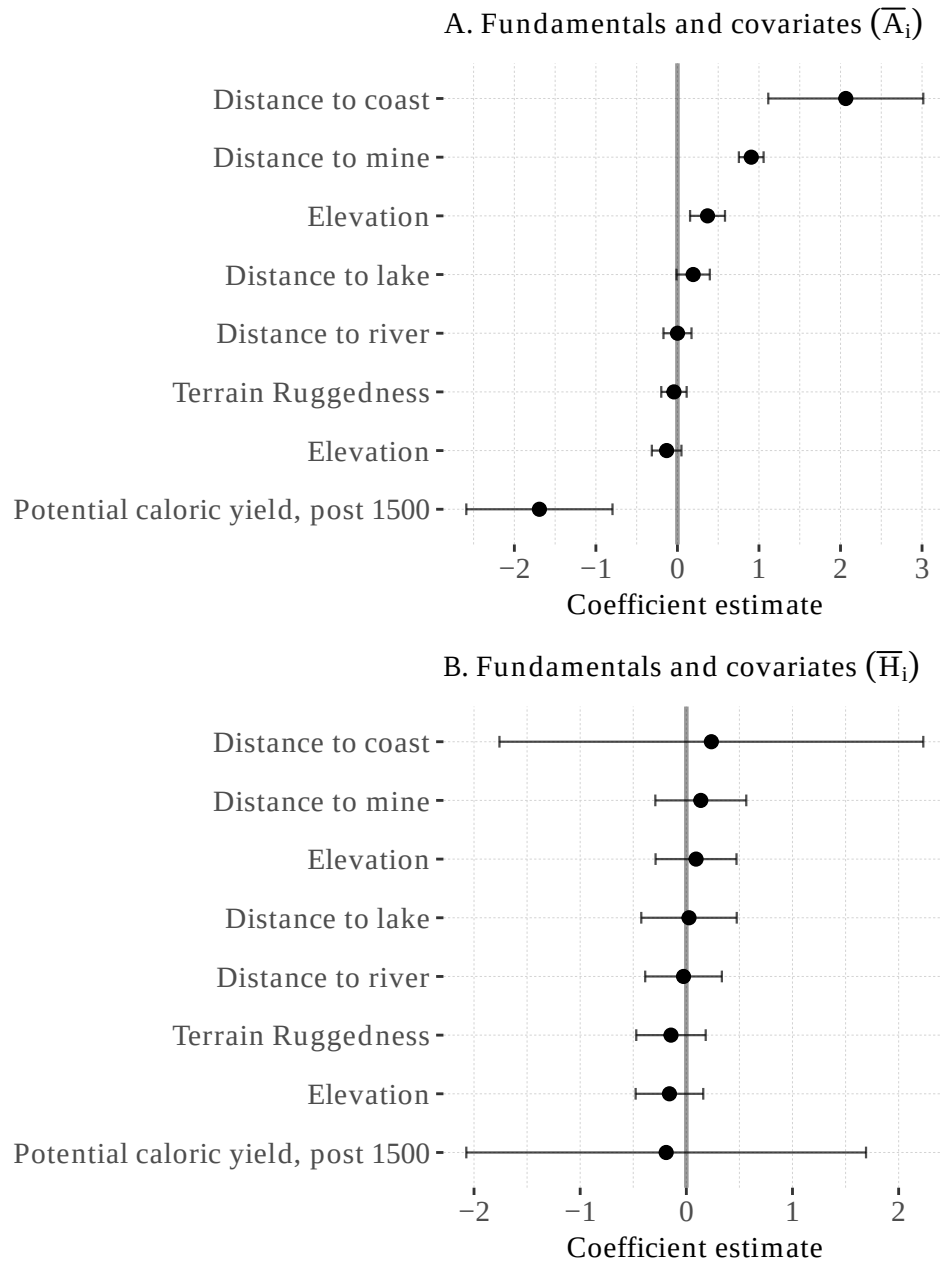


Figure A7: The figure depicts the correlates of the locational fundamentals implied by the model ($\bar{A}$ and $\bar{H}$) and 95 percent confidence intervals. The variables are standardized.

Table A2: Historical agglomeration spillovers

Dependent variable:	City population in 1800 (log.)				
	(1)	(2)	(3)	(4)	(5)
Pop. 1750 (log.)	0.144 (0.054)	0.045 (0.035)	0.050 (0.041)	0.046 (0.042)	0.043 (0.031)
Controls		✓		✓	✓
Country FE			✓	✓	
Viceroyalty FE					✓
Mean DV	9.296	9.296	9.296	9.296	9.296
N	62	62	62	62	62
R ²	0.973	0.993	0.994	0.996	0.993

Note: The table reports OLS estimates. The unit of observation is a city in 1800. Regressions are based on 62 cities. *Dependent variable:* The natural logarithm of the population size in 1800. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (62 clusters).

Table A3: Counterfactuals (fundamentals equalized)

Scenario:	Benchmark		Fixed migration frictions	
	Mean	25th/75th perc.	Mean	25th/75th perc.
<i>Panel A: All cities</i>				
$\Delta P(\%)$	-0.42	[-0.25, 0]	-0.36	[-0.21, 0.01]
$\Delta L(\%)$	0.95	[0.05, 0.64]	0.4	[-0.04, 0.23]
$\Delta V(\%)$	-0.13	[-0.1, -0.02]	0.11	[0.01, 0.06]
<i>Panel B: Core region</i>				
$\Delta P(\%)$	-0.11	[-0.01, 0]	-0.08	[0.01, 0.01]
$\Delta L(\%)$	0.27	[0.05, 0.06]	0.07	[-0.04, -0.03]
$\Delta V(\%)$	-0.05	[-0.03, -0.02]	0.03	[0.01, 0.01]
<i>Panel C: Fringe region</i>				
$\Delta P(\%)$	-0.61	[-0.53, 0]	-0.53	[-0.41, 0.01]
$\Delta L(\%)$	1.38	[0.06, 1.57]	0.61	[-0.03, 0.47]
$\Delta V(\%)$	-0.18	[-0.31, -0.03]	0.15	[0.01, 0.11]

Note: The table reports the results from the baseline counterfactual exercises. The percentage difference (measured from 0-100) across the various outcomes for the two scenarios is reported. The *Benchmark* assumes migration frictions change. *Fixed migration frictions* assume migration frictions remain fixed at 1760 values. The unit of observation is a city.

Table A4: Counterfactuals (pre-reform population equalized)

Scenario:	Benchmark		Fixed migration frictions	
	Mean	25th/75th perc.	Mean	25th/75th perc.
<i>Panel A: All cities</i>				
$\Delta P(\%)$	-0.63	[-0.74, 0.01]	-0.57	[-0.7, 0.01]
$\Delta L(\%)$	1.27	[-0.06, 1.34]	0.67	[-0.01, 0.83]
$\Delta V(\%)$	-0.11	[-0.07, 0.02]	0.16	[0, 0.18]
<i>Panel B: Core region</i>				
$\Delta P(\%)$	-0.13	[0, 0.01]	-0.12	[0, 0.01]
$\Delta L(\%)$	0.26	[-0.05, -0.02]	0.13	[-0.02, 0]
$\Delta V(\%)$	-0.02	[0.01, 0.02]	0.03	[0, 0.01]
<i>Panel C: Fringe region</i>				
$\Delta P(\%)$	-0.94	[-1.14, 0.01]	-0.85	[-0.9, 0.01]
$\Delta L(\%)$	1.92	[-0.06, 2.49]	1	[-0.01, 1]
$\Delta V(\%)$	-0.16	[-0.16, 0.02]	0.24	[0, 0.26]

Note: The table reports the results from the baseline counterfactual exercises. The percentage difference (measured from 0-100) across the various outcomes for the two scenarios is reported. The *Benchmark* assumes migration frictions change. *Fixed migration frictions* assume migration frictions remain fixed at 1760 values. The unit of observation is a city.

Table A5: Counterfactuals (size of shock equalized)

Scenario:	Benchmark		Fixed migration frictions	
	Mean	25th/75th perc.	Mean	25th/75th perc.
<i>Panel A: All cities</i>				
$\Delta P(\%)$	-0.36	[-0.41, -0.21]	-0.3	[-0.35, -0.16]
$\Delta L(\%)$	0.88	[0.62, 1.07]	0.33	[0.18, 0.4]
$\Delta V(\%)$	-0.16	[-0.18, -0.12]	0.08	[0.04, 0.1]
<i>Panel B: Core region</i>				
$\Delta P(\%)$	-0.3	[-0.4, -0.19]	-0.24	[-0.33, -0.14]
$\Delta L(\%)$	0.78	[0.58, 1.06]	0.27	[0.17, 0.38]
$\Delta V(\%)$	-0.16	[-0.17, -0.14]	0.07	[0.04, 0.09]
<i>Panel C: Fringe region</i>				
$\Delta P(\%)$	-0.4	[-0.46, -0.27]	-0.33	[-0.38, -0.2]
$\Delta L(\%)$	0.94	[0.72, 1.22]	0.38	[0.23, 0.44]
$\Delta V(\%)$	-0.16	[-0.19, -0.12]	0.09	[0.05, 0.11]

Note: The table reports the results from the baseline counterfactual exercises. The percentage difference (measured from 0-100) across the various outcomes for the two scenarios is reported. The *Benchmark* assumes migration frictions change. *Fixed migration frictions* assume migration frictions remain fixed at 1760 values. The unit of observation is a city.

Table A6: Counterfactuals (no trans-Atlantic migration)

Scenario:	Benchmark		Fixed migration frictions	
	Mean	25th/75th perc.	Mean	25th/75th perc.
<i>Panel A: All cities</i>				
$\Delta P(\%)$	-0.32	[-0.38, -0.15]	-0.32	[-0.38, -0.15]
$\Delta L(\%)$	0.19	[0, 0.26]	0.19	[0, 0.26]
$\Delta V(\%)$	0.19	[0.14, 0.2]	0.19	[0.14, 0.2]
<i>Panel B: Core region</i>				
$\Delta P(\%)$	-0.25	[-0.36, -0.14]	-0.25	[-0.36, -0.14]
$\Delta L(\%)$	0.11	[-0.02, 0.24]	0.11	[-0.02, 0.24]
$\Delta V(\%)$	0.17	[0.14, 0.19]	0.17	[0.14, 0.19]
<i>Panel C: Fringe region</i>				
$\Delta P(\%)$	-0.36	[-0.43, -0.22]	-0.36	[-0.43, -0.22]
$\Delta L(\%)$	0.24	[0.08, 0.32]	0.24	[0.08, 0.32]
$\Delta V(\%)$	0.2	[0.15, 0.21]	0.2	[0.15, 0.21]

Note: The table reports the results from the baseline counterfactual exercises. The percentage difference (measured from 0-100) across the various outcomes for the two scenarios is reported. The *Benchmark* assumes migration frictions change. *Fixed migration frictions* assume migration frictions remain fixed at 1760 values. The unit of observation is a city.

E Sensitivity Analysis

E.1 Synthetic Control Design

In this section, I explore an alternative approach by implementing a synthetic control design (Abadie, 2021). Specifically, I compare the growth in population size between cities that were directly affected by the reform ($\Delta T_i > 0$) and cities that were not ($\Delta T_i = 0$). I first average the population size across all treated cities and compare this to “never-treated” cities. The 22 never-treated cities constitute the donor pool. The synthetic group is constructed as a weighted average of cities in the donor pool, with weights chosen to minimize the mean squared error before 1750. The covariates chosen are the controls in the baseline model (caloric yield, distance to the coast, terrain ruggedness, distance to nearest lake, distance to nearest river, location of mining centers, maize suitability and wheat suitability) as well as the log. population size in 1600.

Panel A. Figure A8 displays the results. First, one can see that the pre-reform population is closely matched between the synthetic and control and treated cities. The gap between the two groups began diverging in 1800. The treatment effect is given by 0.338, suggesting a treatment effect of approximately 19 percent over the course of one century.

I assess the robustness of the estimates by constructing placebo SCMs by re-estimating the synthetic control after assigning each city in the donor pool to the treatment. I then calculate the ratio of the pre and post-RMSE for each placebo treatment group. Reassuringly, I find that the largest ratio of pre to post-RSME is obtained for the true treatment. This suggests that the reform promoted the growth of cities that experienced reductions in their transportation times to Europe as a result of the reform and independence from Spain in the 19th century. However, there are two important shortcomings of the analysis. First, there are relatively few pre and post-intervention periods for which data is available. Second, even locations that are not directly affected by lower transportation times might experience equilibrium spillovers as highlighted by the quantitative framework. Nevertheless, it is reassuring that the findings with this alternative approach are qualitatively similar to the findings in the main analysis.

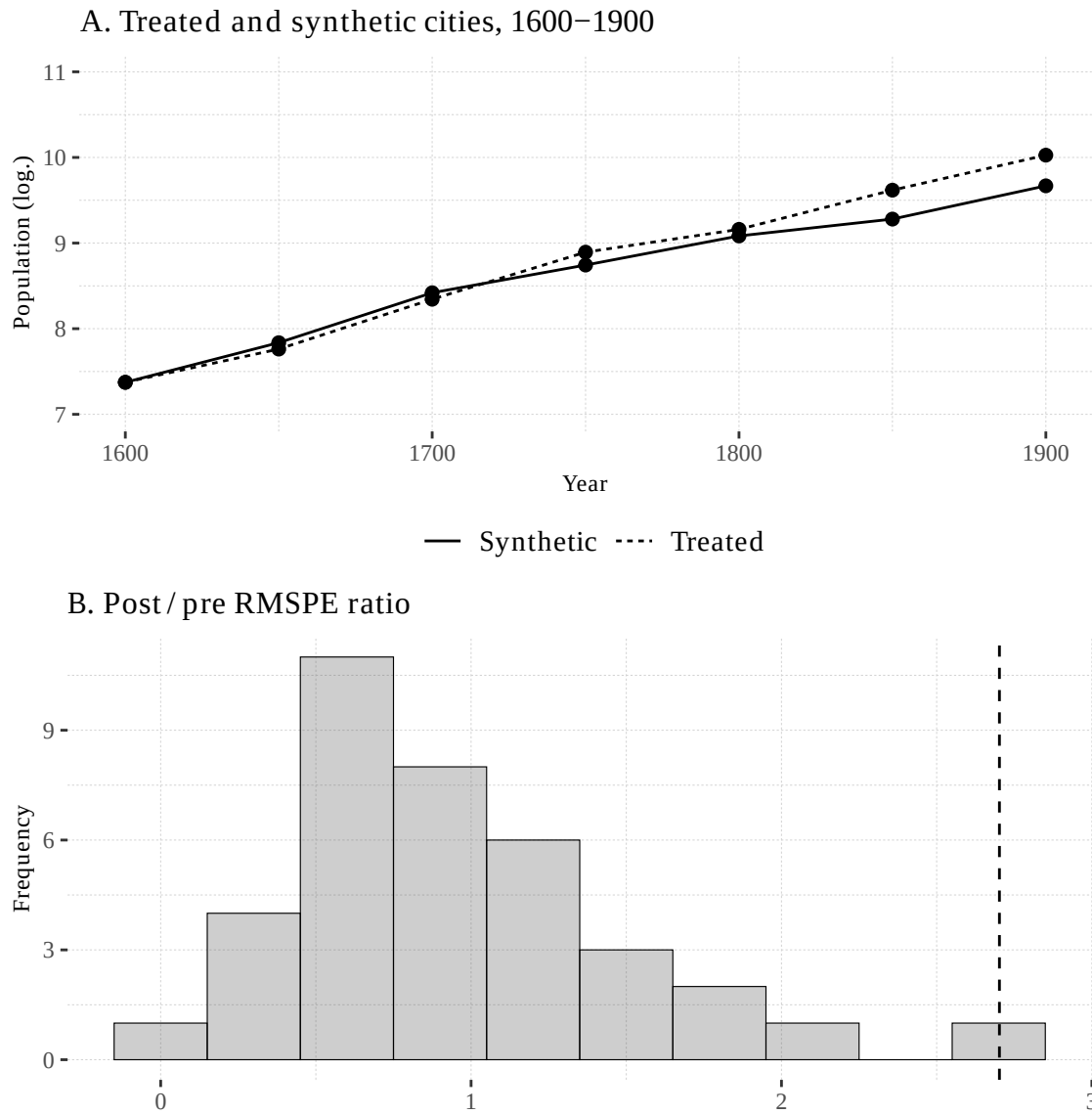


Figure A8: Panel A depicts the the results from the synthetic control design. The treatment group is the average population size for the cities in this group that experienced reductions in their transportation times to Europe. Panel B depicts a placebo exercise where each city in the donor pool is assigned the treatment.

E.2 Robustness Value

I consider the stability of the estimates in the baseline specification when including covariates. It is reassuring that the estimated coefficient is stable with the inclusion of these covariates. Moreover, the increase in the R^2 shows that the variables explain variation in the decline in shipping times and urban growth. A more systematic approach to sensitivity analysis is developed in Cinelli and Hazlett (2020). In this section, I estimate the sensitivity of the estimated impact of transportation times to Europe on population growth to unobserved confounders. Specifically, I follow the procedure developed in Cinelli and Hazlett (2020) and calculate the robustness value which quantifies the strength required by any unobserved confounder to account for the estimates in the baseline model. Alternatively, the strength needed from an unobserved confounder to render the coefficient statistically insignificant.

I find that a robustness value of 16.9 to account for a true effect of zero, and 7 to account for a true effect that is statistically insignificant. In other words, if there exists a confounder that explains 16.9 of the outcome and the treatment, then controlling for this confounder would make the effect of declining shipping time zero. Unfortunately, there are few time-varying covariates available that could serve as a benchmark for the magnitude of this effect. However, it should be noted that this is larger than the combined R^2 of the observable location fundamentals. This suggests that the estimated impact of the decline in shipping time is unlikely to be driven by an unobserved confounder.

E.3 Sensitivity Analysis

Table A7: Shipping time and urban population growth (interior cities)

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
Transportation time (T_{it})	−0.016 (0.009)	−0.024 (0.014)	−0.017 (0.018)	−0.001 (0.024)	−0.021 (0.016)
City & Year FE	✓	✓	✓	✓	✓
Controls × Year FE		✓			
Viceroyalty × Year FE			✓		
Country × Year FE				✓	
Viceroyalty × Time trend					✓
Mean DV	8.58	8.58	8.58	8.58	8.58
N	186	186	186	186	186
R^2	0.904	0.940	0.914	0.934	0.907

Note: The table reports OLS estimates. T_{it} denotes transportation time to Europe measured in days. The sample contains cities more than 228km from a port in 1750. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (31 clusters).

Table A8: Shipping time and urban population growth (removing outliers)

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
Transportation time (T_{it})	-0.018 (0.010)	-0.018 (0.011)	-0.014 (0.010)	-0.019 (0.013)	-0.017 (0.008)
City & Year FE	✓	✓	✓	✓	✓
Controls $\times$ Year FE		✓			
Viceroyalty $\times$ Year FE			✓		
Country $\times$ Year FE				✓	
Viceroyalty $\times$ Time trend					✓
Mean DV	8.502	8.502	8.502	8.502	8.502
N	306	306	306	306	306
R^2	0.900	0.920	0.907	0.934	0.900

Note: The table reports OLS estimates. T_{it} denotes transportation time to Europe measured in days. The sample contains 51 cities with population growth rates within the 5th and 95th percentiles. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (51 clusters).

Table A9: Shipping time and urban population growth, 1600-1800

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
Transportation time (T_{it})	-0.015 (0.013)	-0.011 (0.015)	-0.021 (0.017)	-0.014 (0.016)	-0.022 (0.014)
City & Year FE	✓	✓	✓	✓	✓
Controls $\times$ Year FE		✓			
Viceroyalty $\times$ Year FE			✓		
Country $\times$ Year FE				✓	
Viceroyalty $\times$ Time trend					✓
Mean DV	8.501	8.501	8.501	8.501	8.501
N	310	310	310	310	310
R^2	0.888	0.905	0.895	0.925	0.892

Note: The table reports OLS estimates. T_{it} denotes transportation time to Europe measured in days. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (62 clusters).

Table A10: Shipping time and urban population growth (restricted samples)

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Dropping Nueva Granada</i>					
Transportation time (T_{it})	-0.022 (0.010)	-0.003 (0.017)	-0.026 (0.016)	-0.021 (0.013)	-0.026 (0.014)
Mean DV	8.78	8.78	8.78	8.78	8.78
N	258	258	258	258	258
R^2	0.893	0.916	0.898	0.932	0.894
<i>Panel B: Dropping Rio de la Plata</i>					
Transportation time (T_{it})	-0.032 (0.020)	-0.052 (0.032)	-0.033 (0.021)	-0.031 (0.042)	-0.029 (0.020)
Mean DV	8.78	8.78	8.78	8.78	8.78
N	318	318	318	318	318
R^2	0.894	0.908	0.897	0.927	0.894
City & Year FE	✓	✓	✓	✓	✓
Controls $\times$ Year FE		✓			
Viceroyalty $\times$ Year FE			✓		
Country $\times$ Year FE				✓	
Viceroyalty $\times$ Time trend					✓

Note: The table reports OLS estimates. T_{it} denotes transportation time to Europe measured in days. The unit of observation is a city between 1550-1850. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level. Panel A: 42 clusters. Panel B: 53 clusters.

Table A11: Shipping time and urban population growth (weighted by population)

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
Transportation time (T_{it})	-0.024 (0.009)	-0.020 (0.014)	-0.029 (0.015)	-0.019 (0.011)	-0.030 (0.013)
City & Year FE	✓	✓	✓	✓	✓
Controls $\times$ Year FE		✓			
Viceroyalty $\times$ Year FE			✓		
Country $\times$ Year FE				✓	
Viceroyalty $\times$ Time trend					✓
Mean DV	8.688	8.688	8.688	8.688	8.688
N	372	372	372	372	372
R^2	0.877	0.896	0.884	0.919	0.879

Note: The table reports OLS estimates weighted by population size. T_{it} denotes transportation time to Europe measured in days. The unit of observation is a city between 1550-1850. Regressions are based on a balanced panel of 7 time periods (outcomes measured every 50 years) $\times$ 59 cities = 413 observations. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (62 clusters).

Table A12: Distance and urban population growth

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
Distance (D_{it})	-0.293 (0.164)	-0.189 (0.187)	-0.329 (0.176)	-0.274 (0.300)	-0.372 (0.163)
City & Year FE	✓	✓	✓	✓	✓
Controls $\times$ Year FE		✓			
Viceroyalty $\times$ Year FE			✓		
Country $\times$ Year FE				✓	
Viceroyalty $\times$ Time trend					✓
Mean DV	8.688	8.688	8.688	8.688	8.688
N	372	372	372	372	372
R^2	0.880	0.898	0.886	0.919	0.882

Note: The table reports OLS estimates. $Distance_{it}$ denotes the distance to Europe measured in kilometers (standardized). The unit of observation is a city between 1550-1850. Regressions are based on a balanced panel of 7 time periods (outcomes measured every 50 years) $\times$ 59 cities = 413 observations. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (62 clusters).

Table A13: Dynamic regression model (cities)

Dependent variable:	City population (ln)			
	(1)	(2)	(3)	(4)
$\Delta T \times \mathbb{1}(\text{year} = 1600)$	-0.015 (0.011)	-0.011 (0.025)	-0.017 (0.027)	-0.028 (0.048)
$\Delta T \times \mathbb{1}(\text{year} = 1650)$	-0.009 (0.010)	0.007 (0.019)	0.005 (0.018)	0.009 (0.040)
$\Delta T \times \mathbb{1}(\text{year} = 1700)$	-0.006 (0.011)	0.003 (0.012)	0.004 (0.012)	0.016 (0.029)
$\Delta T \times \mathbb{1}(\text{year} = 1800)$	0.008 (0.005)	0.010 (0.013)	0.009 (0.013)	-0.019 (0.026)
$\Delta T \times \mathbb{1}(\text{year} = 1850)$	0.024 (0.010)	0.027 (0.018)	0.031 (0.019)	0.013 (0.032)
City FE	✓	✓	✓	✓
Controls $\times$ Year FE		✓	✓	✓
Viceroyalty $\times$ Year FE			✓	
Country $\times$ Year FE				✓
N	372	372	372	372
R^2	0.871	0.899	0.905	0.936

Note: The table reports OLS estimates. The decline in shipping time is standardized. The unit is a city in a certain year. The omitted year is 1750 (the last period before the reform). *Dependent variable*: log of city population size. *Controls*: Elevation, crop suitability, terrain ruggedness, the location of active mines, and distance to the coastline. *Standard errors*: Clustered at the city-level (62 clusters).

Table A14: Dynamic regression model (settlements)

Dependent variable:	Indicator for cell containing a settlement			
	(1)	(2)	(3)	(4)
$\Delta T \times \mathbb{1}(\text{year} = 1710)$	−0.0004 (0.0003)	0.001 (0.0004)	0.0003 (0.0004)	0.002 (0.0004)
$\Delta T \times \mathbb{1}(\text{year} = 1720)$	−0.0003 (0.0003)	0.001 (0.0004)	0.0004 (0.0004)	0.002 (0.0004)
$\Delta T \times \mathbb{1}(\text{year} = 1730)$	−0.0002 (0.0002)	0.001 (0.0003)	0.0004 (0.0003)	0.001 (0.0004)
$\Delta T \times \mathbb{1}(\text{year} = 1740)$	−0.0001 (0.0002)	0.001 (0.0003)	0.0004 (0.0003)	0.001 (0.0004)
$\Delta T \times \mathbb{1}(\text{year} = 1750)$	−0.0001 (0.0001)	0.0003 (0.0002)	0.0001 (0.0002)	0.001 (0.0002)
$\Delta T \times \mathbb{1}(\text{year} = 1770)$	0.00005 (0.0001)	−0.00004 (0.0002)	0.00003 (0.0002)	−0.0002 (0.0002)
$\Delta T \times \mathbb{1}(\text{year} = 1780)$	0.001 (0.0002)	0.0003 (0.0003)	0.0002 (0.0003)	−0.0001 (0.0003)
$\Delta T \times \mathbb{1}(\text{year} = 1790)$	0.001 (0.0003)	0.001 (0.0003)	0.001 (0.0003)	−0.00003 (0.0004)
$\Delta T \times \mathbb{1}(\text{year} = 1800)$	0.001 (0.0003)	0.001 (0.0004)	0.001 (0.0004)	0.0001 (0.0004)
$\Delta T \times \mathbb{1}(\text{year} = 1810)$	0.001 (0.0003)	0.001 (0.0004)	0.001 (0.0004)	0.0002 (0.0004)
City FE	✓	✓	✓	✓
Controls $\times$ Year FE		✓	✓	✓
Viceroyalty $\times$ Year FE			✓	
Country $\times$ Year FE				✓
N	53,581	53,581	53,581	53,581

Notes: The table reports OLS estimates. The unit of analysis is at a $0.5^\circ \times 0.5^\circ$ grid-cell. The omitted year is the last decade prior to the treatment (1760). *Controls*: Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the distance to the nearest mine, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors*: Clustered at the grid-level (4,871 clusters).

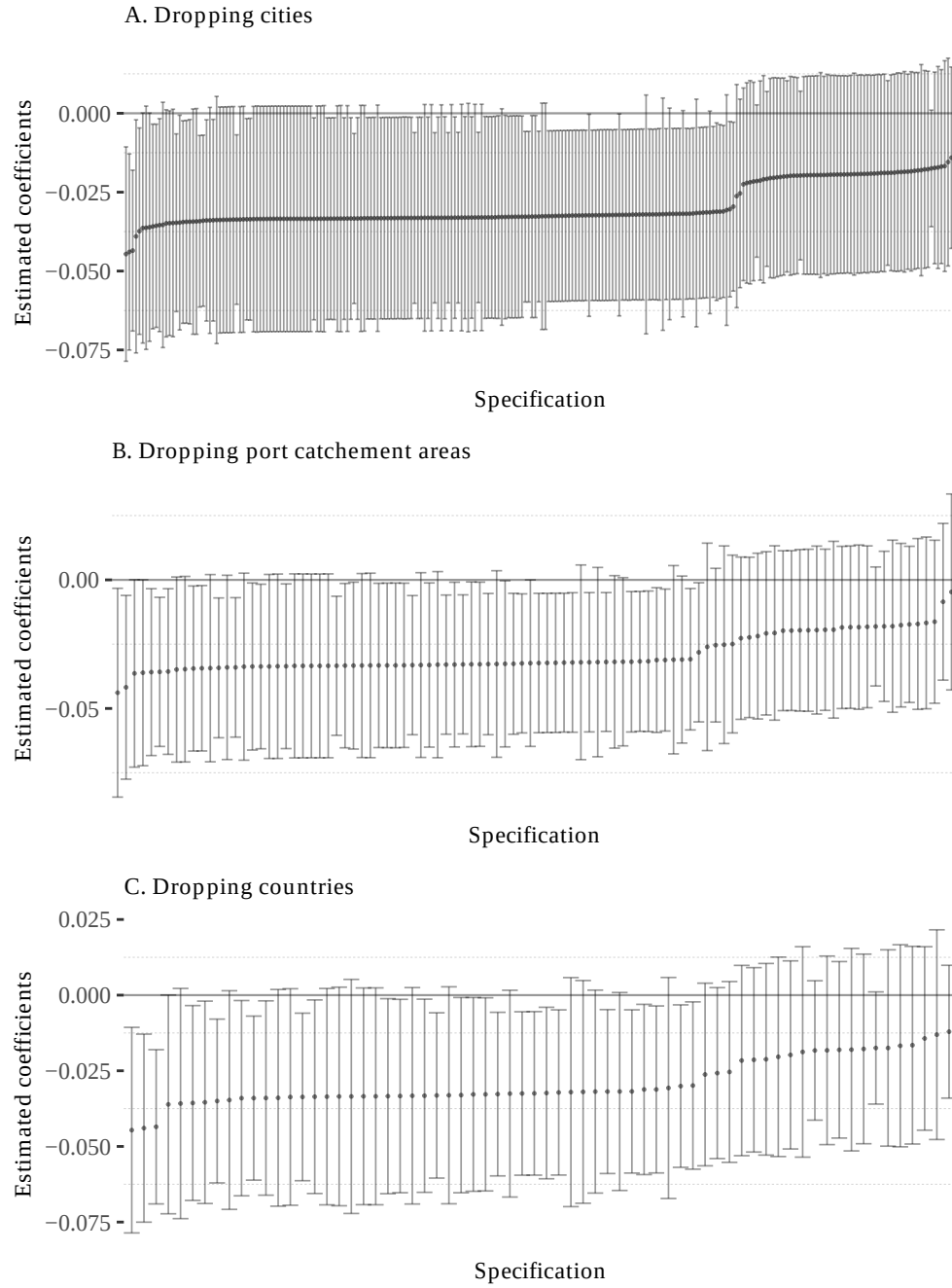


Figure A9: The figure shows the estimates in Table 3 for different subsamples. Panel A displays the model re-estimated after removing each port catchment area. Panel B displays the model re-estimated after removing each country. Panel C displays the model re-estimated after removing each vicerealty.

Table A15: Counterfactuals (parameter sensitivity)

	(1)	(2)	(3)
<i>Panel A: Elasticity of substitution</i>	$\sigma = 3$	$\sigma = 5$	$\sigma = 7$
$\Delta P(\%)$	-3.29	-0.63	-0.39
$\Delta L(\%)$	4.82	1.25	1.01
$\Delta V(\%)$	0.22	-0.1	-0.15
<i>Panel B: Shape parameter</i>	$\theta = 2$	$\theta = 3$	$\theta = 4$
$\Delta P(\%)$	-0.63	-0.63	-0.63
$\Delta L(\%)$	1.4	1.26	1.2
$\Delta V(\%)$	-0.17	-0.1	-0.07
<i>Panel C: Exp. share traded goods</i>	$\mu = 0.4$	$\mu = 0.5$	$\mu = 0.6$
$\Delta P(\%)$	-0.61	-0.63	-0.65
$\Delta L(\%)$	0.95	1.25	1.67
$\Delta V(\%)$	-0.15	-0.1	-0.03
<i>Panel D: Prod. spillover</i>	$\alpha_1 = 0.04$	$\alpha_1 = 0.055$	$\alpha_1 = 0.07$
$\Delta P(\%)$	-0.63	-0.63	-0.63
$\Delta L(\%)$	1.24	1.25	1.26
$\Delta V(\%)$	-0.1	-0.1	-0.09
<i>Panel E: Historical prod. spillover</i>	$\alpha_2 = 0.02$	$\alpha_2 = 0.045$	$\alpha_2 = 0.07$
$\Delta P(\%)$	-0.63	-0.63	-0.63
$\Delta L(\%)$	1.22	1.24	1.25
$\Delta V(\%)$	-0.1	-0.1	-0.1
<i>Panel F: Trade elasticity</i>	$\kappa = 0.28$	$\kappa = 0.56$	$\kappa = 1.12$
$\Delta P(\%)$	-1.18	-0.63	-0.59
$\Delta L(\%)$	2.19	1.25	1.2
$\Delta V(\%)$	-0.05	-0.1	-0.1

Note: The table reports the results from the benchmark counterfactual exercises under alternative parameter values. The percentage difference (mesured from 0-100) across the various outcomes for the two scenarios is reported. The counterfactuals assume migration frictions change. The unit of observation is a city.

Table A16: Shipping time and urban population growth (unadjusted population data)

Dependent variable:	City population (log.)				
	(1)	(2)	(3)	(4)	(5)
Transportation time (T_{it})	-0.022 (0.009)	-0.018 (0.013)	-0.026 (0.013)	-0.021 (0.009)	-0.026 (0.012)
City & Year FE	✓	✓	✓	✓	✓
Controls $\times$ Year FE		✓			
Viceroyalty $\times$ Year FE			✓		
Country $\times$ Year FE				✓	
Viceroyalty $\times$ Time trend					✓
Mean DV	8.646	8.646	8.646	8.646	8.646
N	372	372	372	372	372
R ²	0.893	0.910	0.900	0.930	0.895

Note: The table reports OLS estimates. T_{it} denotes transportation time to Europe measured in days. The unit of observation is a city between 1600-1850. Regressions are based on a balanced panel of 6 time periods (outcomes measured every 50 years) $\times$ 62 cities = 372 observations. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the location of mining centers, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (62 clusters).

Table A17: Shipping time and urban population growth (viceroyalty fixed effects)

Sample:	Full sample		$\Delta T_i < \text{Median}$		$\Delta T_i \geq \text{Median}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Pop. 1750 (log.)	0.544 (0.142)	0.556 (0.183)	0.534 (0.213)	0.777 (0.281)	0.562 (0.187)	0.585 (0.322)
Mean DV	12.89	12.89	12.83	12.83	12.94	12.94
N	62	62	28	28	34	34
R ²	0.251	0.376	0.282	0.638	0.274	0.313
Viceroyalty FE	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓

Note: The table reports OLS estimates. T_{it} denotes transportation time to Europe measured in days. The unit of analysis is a city. *Dependent variable:* The natural logarithm of the population size. *Controls:* Elevation, caloric yield pre-1500, crop suitability (wheat, maize), terrain ruggedness, the location of mining centers, distance to the nearest coastline, distance to the nearest major river, and distance to the nearest lake. *Standard errors:* Clustered at the city level (62 clusters).

References for the Online Appendix

- Abadie, Alberto. 2021. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature* 59(2):391–425.
- Allen, Robert C. 2001. "The Great Divergence in European Wages and Prices from the Middle Ages to the First World War." *Explorations in Economic History* 38(4):411–447.
- Allen, Treb and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics* 129(3):1085–1140.
- Allen, Treb, Costas Arkolakis and Xiangliang Li. 2024. "On the Equilibrium Properties of Spatial Models." *American Economic Review: Insights* (forthcoming).
- Allen, Treb and Dave Donaldson. 2022. "Persistence and Path Dependence in the Spatial Economy." *Manuscript*.
- Amante, C and B.W. Eakins. 2009. "ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis." *NOAA Technical Memorandum NESDIS NGDC-24*. (National Geophysical Data Center, NOAA).
URL: <https://www.ncei.noaa.gov/products/etopo-global-relief-model>
- Arroyo Abad, Leticia, Elwyn Davies and Jan Luiten van Zanden. 2012. "Between Conquest and Independence: Real Wages and Demographic Change in Spanish America, 1530–1820." *Explorations in Economic History* 49(2):149–166.
- Arroyo Abad, Leticia and Jan Luiten van Zanden. 2016. "Growth under Extractive Institutions? Latin American Per Capita GDP in Colonial Times." *The Journal of Economic History* 76(4):1182–1215.
- Baskes, Jeremy. 2013. *Staying Afloat: Risk and Uncertainty in Spanish Atlantic World Trade, 1760–1820*. Stanford: Stanford University Press.
- Bolt, Jutta and Jan Luiten Van Zanden. 2024. "Maddison Project Database 2023.".
URL: <https://doi.org/10.34894/INZBF2>
- Broadberry, S. N., B. M. S. Campbell, Alexander Klein, Mark Overton and Bas van Leeuwen. 2015. *British Economic Growth, 1270–1870*. Cambridge: Cambridge University Press.
- Bruhn, Miriam and Francisco A. Gallego. 2011. "Good, Bad, and Ugly Colonial Activities: Do They Matter for Economic Development?" *The Review of Economics and Statistics* 94(2):433–461.

- Buringh, Eltjo. 2015. "The Clio-Infra Database on Urban Settlement Sizes: 1500-2000." URL: <https://web.archive.org/web/20200207093254/http://www.cgeh.nl/urbanisation-hub-clio-infra-database-urban-settlement-sizes-1500-2000>
- Census, Argentina. 1872. *Primer Censo de la República Argentina (1869). Resultados*. Buenos Aires: Imprenta del Porvenir.
URL: https://biblioteca.indec.gob.ar/bases/minde/1c1869_TU.pdf
- Census, Mexico. 1899. *Censo General de la República Mexicana (1895). Resumen General*. Mexico: Oficina Tip. de la Secretaria de Fomento.
URL: <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825410100>
- Christelow, Allan. 1942. "Contraband Trade between Jamaica and the Spanish Main, and the Three Port Act of 1766." *The Hispanic American Historical Review* 22(2):309.
- Cinelli, Carlos and Chad Hazlett. 2020. "Making Sense of Sensitivity: Extending Omitted Variable Bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82(1):39–67.
- Coatsworth, John H and Jeffrey G Williamson. 2004. "Always Protectionist? Latin American Tariffs from Independence to Great Depression." *Journal of Latin American Studies* 36(2):205–232.
- Cuenca-Esteban, Javier. 2008. "Statistics of Spain's Colonial Trade, 1747–1820: New Estimates and Comparisons with Great Britain." *Revista de Historia Económica / Journal of Iberian and Latin American Economic History* 26(3):323–354.
- Denevan, William M. 1992. *The Native Population of the Americas in 1492*. Madison, Wisconsin: University of Wisconsin Press.
- Diaz-Cayeros, Alberto. 2022. The Future of Latin American and Caribbean Cities: Urban Bias and Political Fragments in Place. In *The Routledge Handbook of Urban Studies in Latin America and the Caribbean*. Routledge.
- Elliot, Robert, Stephen DeGloria and Shawn Riley. 1999. "A terrain ruggedness index that quantifies topographic heterogeneity." *Intermountain Journal of Sciences* 5(1-4):23–27.
- Elliott, J. H. 2006. *Empires of the Atlantic World: Britain and Spain in America, 1492-1830*. New Haven: Yale University Press.
- Fernández-López, Javier and Klaus Schliep. 2019. "rWind: download, edit and include wind data in ecological and evolutionary analysis." *Ecography* 42(4):804–810.

- Findlay, Ronald and Kevin H. O'Rourke. 2007. *Power and Plenty: Trade, War, and the World Economy in the Second Millennium*. Princeton: Princeton University Press.
- Fisher, John R. 1993. *El Comercio entre España e Hispanoamérica: 1797-1820*. Number 27 in "Estudios de historia económica" Madrid: Banco de España.
- Fisher, John R. 1997. *The Economic Aspects of Spanish Imperialism in America, 1492-1810*. Liverpool: Liverpool University Press.
- Fisher, John R. 1998. "Commerce and Imperial Decline: Spanish Trade with Spanish America, 1797-1820." *Journal of Latin American Studies* 30(3):459-479.
- Fisher, J.R. 1985. *Commercial Relations Between Spain and Spanish America in the Era of Free Trade, 1778-1796*. Liverpool: Centre for Latin-American Studies, University of Liverpool.
- Food and Agriculture Organization of the United Nations. 2012. "Global Agro-Ecological Zones (GAEZ v3.0) Database."
 URL: <https://gaez.fao.org/>
- Galor, Oded and Ömer Özak. 2015. "Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability." *Manuscript* .
- Galor, Oded and Ömer Özak. 2016. "The Agricultural Origins of Time Preference." *American Economic Review* 106(10):3064-3103.
- García-Herrera, R., G. P. Können, D. A. Wheeler, M. R. Prieto, P. D. Jones and F. B. Koek. 2005. "CLIWOC: A Climatological Database for the World's Oceans 1750-1854." *Climatic Change* 73(1):1-12. Dataset.
 URL: <https://www.historicalclimatology.com/cliwoc.html>
- Gerhard, Peter. 1993a. *A Guide to the Historical Geography of New Spain*. Oklahoma: University of Oklahoma Press.
- Gerhard, Peter. 1993b. *The North Frontier of New Spain*. Oklahoma: University of Oklahoma Press.
- Gerhard, Peter. 1993c. *The Southeast Frontier of New Spain*. Oklahoma: University of Oklahoma Press.
- Hamilton, Earl Jefferson. 1947. *War and Prices in Spain, 1651-1800*. Cambridge: Harvard University Press.
- Herzog, Irmela. 2014. "Least Cost Paths – Some Methodological Issues." *Internet Archaeol* 36.
 URL: <https://intarch.ac.uk/journal/issue36/5/5-5.html>

- Hijmans, Robert J. 2025. "Database of Global Administrative Areas (GADM)." Database of Global Administrative Areas, Museum of Vertebrate Zoology, University of California, Berkeley. Accessed April 2024.
URL: <https://gadm.org/data.html>
- Honaker, James and Gary King. 2010. "What to Do about Missing Values in Time-Series Cross-Section Data." *American Journal of Political Science* 54(2):561–581.
- Johnson, Lyman L. and Sibila Seibert. 1979. "Estimaciones de la población de Buenos Aires en 1744, 1778 y 1810." *Desarrollo Económico* 19(73):107–119.
- Kelly, Morgan and Cormac Ó Gráda. 2019. "Speed Under Sail During the Early Industrial Revolution (c. 1750–1830)." *The Economic History Review* 72(2):459–480.
- Kuethe, Allan J. and Kenneth J. Andrien. 2014. *The Spanish Atlantic World in the Eighteenth Century: War and the Bourbon Reforms, 1713–1796*. Cambridge: Cambridge University Press.
- López de Velasco, Juan. 1570. "Demarcación y división de las Indias [Manuscrito].".
URL: <https://bdh.bne.es/bnearch/detalle/3427025>
- López de Velasco, Juan. 1575. "Demarcacion y nauegaciones de Yndias.".
URL: https://jcb.lunaimaging.com/luna/servlet/detail/JCBMAPS_1_1_1100_102700001:-Demarcacion-y-nauegaciones-de-Yndi
- Mahoney, James. 2010. *Colonialism and Postcolonial Development: Spanish America in a Comparative Perspective*. Cambridge: Cambridge University Press.
- Maloney, William F. and Felipe Valencia. 2016. "The Persistence of (Subnational) Fortune." *The Economic Journal* 126(598):2363–2401.
- Morse, Richard M. 1974. "Trends and Patterns of Latin American Urbanization, 1750–1920." *Comparative Studies in Society and History* 16(4):416–447.
- Natural Earth. 2025. "Natural Earth Data: Free Vector and Raster Map Data.". Public domain dataset; accessed January 2020.
URL: <https://www.naturalearthdata.com/downloads/>
- O'Rourke, Kevin H. 2006. "The Worldwide Economic Impact of the French Revolutionary and Napoleonic Wars, 1793–1815." *Journal of Global History* 1(1):123–149.
- Pinzón, H.T., J.A.T. Mora and C.E.T. Mora. 1994. *Convocatoria al poder del número: censos y estadísticas de la Nueva Granada, 1750-1830*. Serie Historia Bogotá, Colombia: Archivo General de la Nación.

- Prados de la Escosura, Leandro and Gabriel Tortella Casares. 1983. "Tendencias a Largo Plazo del Comercio Exterior Español, 1714–1913." *Revista de Historia Económica / Journal of Iberian and Latin American Economic History* 1(2):353–367.
- Rahn Phillips, Carla. 1990. The Growth and Composition of Trade in the Iberian Empires, 1450–1750. In *The Rise of Merchant Empires: Long-Distance Trade in the Early Modern World, 1350–1750*, ed. James D. Tracy. Cambridge: Cambridge University Press.
- Ramankutty, Navin and Jonathan A. Foley. 1999. "Global Potential Vegetation Dataset.". Center for Sustainability and the Global Environment, University of Wisconsin–Madison.
URL: <https://sage.nelson.wisc.edu/data-and-models/datasets/#vegetation>
- Ridolfi, Leonardo and Alessandro Nuvolari. 2021. "L'histoire immobile? A Reappraisal of French Economic Growth Using the Demand-Side Approach, 1280–1850." *European Review of Economic History* 25(3):405–428.
- Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Sarkar, Matthew Sobek, Derek Burk, Dan Ehrlich, Jane Lee and Nate Merrill. 2025. "Integrated Public Use Microdata Series, International: Version 7.6 [dataset].".
URL: <https://doi.org/10.18128/Do20.V7.7>
- Scobie, James R. 1971. *Argentina: A City and a Nation*. Oxford: Oxford University Press.
- Stangl, Werner. 2018. "'The Empire Strikes Back'?: HGIS de las Indias and the Postcolonial Death Star." *International Journal of Humanities and Arts Computing* 12(2):138–162.
- Stangl, Werner. 2019a. "Data: Lago Texcoco, 18th century.".
URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DOT5EY>
- Stangl, Werner. 2019b. "Data: Mail land routes of Spanish America, 1745–1808.".
URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/W4C9H7>
- Stangl, Werner. 2019c. "Data: Places Gazetteer of Spanish America, 1701–1808.".
URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/FUSJD3>
- Stangl, Werner. 2019d. "Data: Ports of Spanish America, 1701–1808.".
URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/UXDJLQ>
- Stangl, Werner. 2019e. "Data: Territorial Gazetteer for Spanish America, 1701–1808.".
URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/YPEU5E>
- Stein, Stanley J. and Barbara H. Stein. 2003. *Apogee of Empire: Spain and New Spain in the Age of Charles III, 1759–1789*. Baltimore: Johns Hopkins University Press.

- Tobler, Waldo. 1993. "Three Presentations on Geographical Analysis and Modeling." *Technical Report* .
- Verhagen, Philip, Jamie Joyce and Mark R. Groenhuijzen, eds. 2019. *Finding the Limits of the Limes: Modelling Demography, Economy and Transport on the Edge of the Roman Empire*. Computational Social Sciences Cham: Springer International Publishing.
- Walker, Geoffrey J. 1979. *Spanish Politics and Imperial Trade, 1700-1789*. London: Palgrave Macmillan.
- Weiss, D. J., A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, E. Poyart, S. Belchior, N. Fullman, B. Mappin, U. Dalrymple, J. Rozier, T. C. D. Lucas, R. E. Howes, L. S. Tusting, S. Y. Kang, E. Cameron, D. Bisanzio, K. E. Battle, S. Bhatt and P. W. Gething. 2018. "A global map of travel time to cities to assess inequalities in accessibility in 2015." *Nature* 553(7688):333–336.
- Álvarez Nogal, Carlos and Leandro Prados De La Escosura. 2013. "The rise and fall of Spain (1270–1850)." *The Economic History Review* 66(1):1–37.
- Özak, Ömer. 2010. "The Voyage of Homo-œconomicus: Some Economic Measures of Distance." *Manuscript* .
- Özak, Ömer. 2018. "Distance to the Pre-Industrial Technological Frontier and Economic Development." *Journal of Economic Growth* 23(2):175–221.