The Impact of the Black Death on the Adoption of the Printing Press*

Noel D. Johnson † Andrew J. Thomas ‡ Alexander N. Taylor §

December 19, 2024

Abstract

We leverage plausibly exogenous variation in mortality from the Black Death (1347–52) across European cities to estimate the causal impact of market size on early print adoption. Using the universe of data from the Universal Short Title Catalogue we create a database linking early European printed material to historical city populations. We find that cities whose populations were more heavily impacted by the Black Death were less likely to be early adopters of the press and printed fewer unique book editions. We also provide evidence that beyond own-city mortality there were also spatial spillovers from the Black Death shock.

JEL Codes: N0, N9, O3, L1, Q5

Keywords: Pandemics; Black Death; Printing Press; Technical Change; Innovation; Media

*We've benefitted from conversations with Mark Koyama, Remi Jedwab, Stephan Heblich, Daniel Bogart, Cormac Ó Gráda, Neil Cummins, Noam Yuchtman, Hannes Malmberg, Walker Hanlon, Karen Clay, Randall Walsh, Andreas Ferrara, César Martinelli, Metin Coşgel, Melinda Miller, Stephen Ross, Remy Levin, and John Conlon. We received helpful comments from seminar audiences at the University of Pittsburgh, Virginia Tech, the University of Connecticut, the GMU Center for Micro-Economic Policy Research, the Free Market Institute at Texas Tech, the London School of Economics, the XIXth World Economic History Congress, the 2023 Meeting of the Economic History Association, and the University of Mississippi. The data on Black Death mortality rates used in this paper come from Jedwab et al. (2024) and Jedwab et al. (2022).

†Corresponding Author, Economics Department, George Mason University, njohnsoL@gmu.edu.

‡Economics Department, George Mason University, athoma5@gmu.edu.

§Economics Department, University of Evansville, at154@evansville.edu.

1 Introduction

The movable type printing press developed by Johannes Gutenberg around 1440 was one of the most important technologies of the last millennium. By drastically reducing the cost of book production information could be reproduced and disseminated faster than ever before (Eisenstein, 1980; Barbier, 2017). Recent research has established a causal link between a city being an early adopter of the printing press and subsequent growth (Dittmar, 2011). The press has also been shown to have facilitated the spread of the Protestant Reformation (Rubin, 2014). An important question, then, is why did some cities adopt the printing press earlier than others?

In this paper we exploit a natural experiment, the massive population shock caused by the Black Death to the urban system of Europe between 1347 and 1352, to identify a causal relationship between market size and early press adoption. Historians of the printing press note that large cities were early adopters yet, to our knowledge, there is no study establishing a causal link between city size and early press adoption. We create a new database linking data on early European printed material with historical city populations and rates of Black Death mortality and find that cities whose populations were more heavily impacted by the Black Death were less likely to adopt the press and printed fewer unique book editions during the period 1450–1600.

We begin by providing evidence that mortality at the city-level from the Black Death was plausibly random.[1] We then exploit this fact by using Black Death mortality as an instrument for city size in 1400. This allows us to estimate the causal impact of market size on early press adoption. Our identification strategy relies on our argument that city-level Black Death mortality rates were random and that, therefore, the predicted level of a city's population in 1400 using these rates will also be unrelated to any potential confounders that may also be correlated with press adoption.

We then link mortality-predicted city population in 1400 to printing press adoption using data from the Universal Short Title Catalogue (USTC), an online project collecting all known information on printed material during the first 200 years of the printing press in Europe. We scrape the USTC to create a data set of over 343,000 observations of individual editions published using the printing press between 1450 and 1600 across European cities. This gives us a city-level measure of year of press adoption, intensity of press usage and, since we know the subject matter of each edition, a measure of specialization in book topics.

In addition to exploiting plague mortality rates to predict city population, our identification strategy also includes controls for the presence of a university or a bishopric—two of the main sources of both demand and supply of books according to the historical literature. We also control for two other potential confounders: a city's trade potential (market access) prior to the Black Death

[1]The material in this paper on the exogeneity of Black Death mortality at the city-level is based on work in Jedwab et al. (2024), Jedwab et al. (2022), and Jedwab et al. (2019).

shock, and the agricultural fecundity of the area surrounding a city.[2]

It is well established in the literature on technical change that two important determinants of technological diffusion are the relative prices of inputs and market size (Acemoglu, 2002; Acemoglu and Linn, 2004). The Black Death could have affected press adoption through either of these channels. In the aftermath of the plague wages increased dramatically in Europe and did not return to pre-plague levels in many places until 1600 (Jedwab et al., 2022). It is possible that this increase in the relative price of labor incentivized the adoption of labor saving technologies, among them the press (Allen, 2011; Pamuk, 2007). Alternatively, the Black Death may have destroyed market potential in places where it was most severe and discouraged early entrepreneurs from bringing the press to those locations (Jedwab et al., 2024; Campbell et al., 2015).

We find strong evidence for the latter hypothesis—cities that lost more population because of the plague were also less likely to be early adopters of print technology. In our baseline IV estimates, a city with 10% lower population due to Black Death mortality in 1400 is predicted to be 3.5% less likely to have a press by 1500. We also investigate the intensive margin of print adoption using data on the number of unique editions published in each city in each year during the sample period. Our model suggests that a city with 10% lower population due to the plague published 17% fewer editions between 1450 and 1500. Using data on the subject matter of unique editions printed, we also find suggestive evidence that there was less specialization (less variety of subjects) published in cities impacted more by the Black Death.

We also investigate whether there were spatial spillovers from the Black Death that affected press adoption. Specifically, for each city we construct a measure of the change in its market access between 1300 and 1400.[3] This gives us a theory-based metric for the impact of the Black Death on the potential trading network for each city. When we extend the analysis to take into account the impact of the Black Death on the entire market network in which a city is embedded, we find a 10% decline in market access due to the plague led to between a 0.6% and 2.4% decrease in probability of press adoption by 1500. Measured on the intensive margin, a 10% decline in market potential due to the Black Death led to between 3% and 8% fewer editions being published. In other words, the demographic shock of the Black Death affected press adoption through both its direct impact on a city's population and the size of its potential trading network.

The rest of the paper proceeds as follows: in Section 2, we describe the construction of data on Black Death mortality and printed material in European cities. In Section 3 we lay out our estimation strategy. Section 4 describes our results, and Section 5 concludes.

[2] We will also show our results are unaffected by controlling for distance to Mainz since this is a popular instrument for press adoption (Dittmar, 2011; Rubin, 2014). We do not make any claims in this paper as to validity of the distance to Mainz instrument. We simply control for it in some regressions to alleviate potential concerns that it may be correlated with Black Death mortality.

[3] We exclude the own city when calculating market access so as to focus on inter-city trade potential.

2 Data

2.1 Black Death Mortality

The Black Death was a comparatively pure population shock: infrastructure was left intact, only people were killed, it did not explicitly target a sub-group of the population (e.g. intellectuals or a particular ethnic group), and there was no government or international organization that sponsored aid in the aftermath. As such, city-level mortality rates constitute a potentially powerful natural experiment to test the impact of a demographic shock on technology adoption.

The Black Death first appeared in Europe in the port of Messina, Sicily, in October 1347. Over the next three years it was responsible for the deaths of approximately 40% of the population.[4] One of the few robust predictors of the virulence of the Black Death was how early in the pandemic a city was infected. In general, epidemic diseases tend to be more virulent when they first enter into an immunologically naïve population. As people get infected and either die or recover with immunity, resistance to the disease emerges. As such, cities that were more connected to the point of origin of the Black Death in Europe, Messina, were more likely to experience high mortality rates. This did not necessarily mean geographic proximity to Messina drove mortality—infected fleas could be transported large distances by either their common black rat hosts or via the clothing of humans. As a result the spatial diffusion of the Black Death was characterized by large leaps that, in retrospect, are difficult to predict. For example, the disease arrived in Messina in October 1347, was in Valencia, Spain by May 1348, and reached Brighton, England by September 1348. In contrast, it didn't arrive in Köln Germany until December 1349 (see Appendix Figure 9). This suggests that there was considerable variation in mortality rates across European cities.

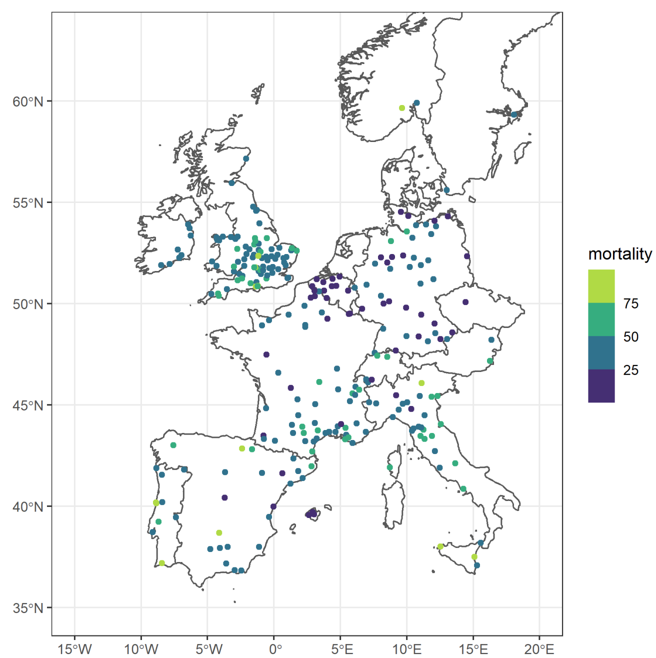
The was also nothing pre-ordained about the Black Death entering Europe through Messina and into the Mediterranean. Appendix Figure 10 shows counter-factual trading routes we reconstruct through various historical accounts through which the plague could have entered Europe. We know the trade center of Astrakhan was hit in 1345 and the Genoese colony of Kaffa was infected in 1346. It was from Kaffa that Genoese galleys traveled to Messina, the starting point of the disease in Europe. However, the disease could just as easily gone through the popular trade-route from Astrakhan to Moscow, then to Novgorod, and have entered through the Baltic rather than the Mediterranean. It could also have easily entered via Prague in Central Europe. This suggests there was nothing special about Messina as the entry point of the plague.

[4] As currently understood by scientists, the infection cycle was that fleas were infected by the bacterium *Yersinia pestis*, which would block their esophagus and make it impossible for them to feed. These unsatiated fleas would then bite humans and other animals (such as rats) in an attempt to feed, in the process regurgitating the bacterium into the host animal. In humans, within a week the bacteria would reach the lymph nodes and cause them to swell into painful buboes. Typically, about two-thirds of those infected would die within ten days of first contact with the disease.

We use data on city-level Black Death mortality compiled by Jedwab et al. (2019) and described in great detail in Jedwab et al. (2024). They build their dataset from Christakos et al. (2005, 117–122) who compile the mortality estimates from a wide range of sources including ecclesiastical and parish records, testaments, tax records, court rolls, chroniclers' reports, donations to the church, financial transactions, mortality of famous people, letters, edicts, guild records, hospital records, cemeteries and tombstones. Jedwab et al. (2019) check these mortality data against other sources including Ziegler (1969), Russell (1972), Gottfried (1983), and Benedictow (2005). This results in estimates of Black Death mortality for 274 towns and cities spread across 16 modern European countries.

These 274 observations are constructed from a variety of source types: 177 come from percentage estimates; 49 are derived from literary descriptions which Jedwab et al. (2019, 2024) rank according to their implied severity;[5] 19 are based on adjusted clergy mortality;[6] 29 cities have desertion rates.[7]

Figure 1



Notes: 274 cities with known Black Death mortality rates between 1347 and 1352. Source Jedwab et al. (2019).

Figure 1 shows the distribution of mortality rates across Europe. It is difficult to identify any major geographic patterns in the mortality data, suggesting that there was a considerable random

[5] 5% for "spared"/"escaped", 10% for "partially spared"/"minimal", 20% for "low", 25% for "moderate", 50% for "high", 66% for "highly depopulated", and 80% for "decimated".

[6] When cross-referenced against the subset of these 19 for which non-clergy mortality is known, clergy died at about an 8% higher rate than the general population. We, therefore, adjust the clergy mortality estimates downward accordingly in the main data set.

[7] Jedwab et al. (2019, 2024) follow Christakos et al. (2005) in lowering these rates for use in the main data as it is known that desertion rates were 1.2 times greater than mortality on average.

component to virulence. In section 2.3 below we confirm that city-level mortality was random with respect to a battery of measures of local physical geography, economic geography, and institutions.

2.2 City Populations

Since we are interested in the demographic impact of the Black Death, we intersect the 274 mortality cities with the Bairoch city population data set (Bairoch, 1988).[8] The Bairoch data contains estimates of the population of 1,801 European towns and cities in the years 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1750, 1800 and 1850. In order to be included in the data set in a given year city population has to be estimated at 1,000 people or greater. In 1300 there are 470 cities that satisfy this criterion. When we intersect the 274 mortality cities with the 470 Bairoch cities, we are left with 169 cities which form our baseline data set that we match with data on year of printing press adoption in Section 2.5.

With our 169 city data set, we can investigate the broad patterns of the impact of the Black Death on city populations in both the short- and long-run. We estimate a series of regressions of the form:

$$\Delta \text{Pop}_{i,t} = \alpha + \beta \text{Mort}_{i,1347-52} + \varepsilon_{i,t}, \tag{1}$$

where $\Delta\text{Pop}_{i,t}$ is the percentage change in city i's population over some time period.[9] $\text{Mort}_{i,1347-52}$ is Black Death mortality for city i or mean mortality for the modern country in which city i is located.

In Table 1 we report the β's from estimating Specification 1 on various samples. In Column 1 the dependent variable is city growth between 1300 and 1400. Assuming no city growth in the absence of the Black Death shock and zero recovery in the fifty years immediately following the shock, we would expect the coefficient to mechanically be equal to one. That is, a 10% mortality shock would be associated with a city shrinking in population by 10%. Our reported coefficient is -0.0085, which suggests that a city experiencing 10% mortality from the Black Death would still be 8.5% smaller in 1400. In other words, average demographic recovery was very slow in the immediate decades following the Black Death.[10] This is an important fact to establish as our study is premised on the assumption that the demographic impact of the plague persisted long enough to disrupt printing press adoption more than 100 years after the shock.

[8] We supplement the Bairoch data with those from Chandler and Fox (2013). This allows us to have estimates for cities on the "edges" of our main data, which will be important to ensure that the market access variables we create aren't biased downwards on the populated edges (e.g. Eastern Europe as opposed to the Atlantic edge).

[9] Since during our period of study there are both large positive and negative growth rates in city populations, we calculate percentage growth using the midpoint method.

[10] This claim would be even stronger if there was some positive city growth between 1300 and the onset of the Black Death.

Table 1: The Impact of the Black Death Shock

Dependent Variables:	Growth 1300-1400		Growth 1300-1600	Growth 1200-1300	Growth 1100-1200
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Mortality Rate	-0.0085***		-0.0024	0.0025	0.0010
	(0.0030)		(0.0032)	(0.0044)	(0.0029)
Mean Mortality Rate		-0.0145***			
		(0.0048)			
Fit statistics					
Observations	169	169	168	107	62

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Mortality measured from 0–100. Aggregate mortality is average modern-country mortality computed using all 274 morality observations. Population growth is calcuated using the midpoint method. Standardized beta coefficient for (1) is 0.26 and for (3) is -0.05.

In Column 2 we take a different approach to measuring the short-run impact of the plague and investigate whether there were spillover effects from the mortality shock. Population growth in medieval cities was driven mostly by in-migration from surrounding areas, so regions that had a higher average mortality may have recovered more slowly. We calculate the average mortality rate (using all 274 mortality cities) for the modern country in which city i is located and estimate Specification 1 using this as the measure of mortality for city i. When we do this, the reported coefficient is almost twice as big as when we measure mortality using just the own-city impact. This implies that a city in a region with 10% average mortality would be 14.5% smaller in 1400. This is even more evidence that the short-run impact of the plague persisted and recovery could possibly have been affected by how neighboring cities were impacted. We will exploit these potential spatial spillovers below.

In Column 3 we estimate the impact of own-city mortality on own-city growth between 1300 and 1600. The coefficient of -0.0024 is both statistically and economically insignificant. The standardized beta for the estimate in Column 3 is -0.05 (compared to -0.26 for Column 1). This is consistent with the accounts of historians that cities had, on average, recovered from the Black Death by 1600. However, as described in Jedwab et al. (2022) and Jedwab et al. (2024), there was a significant amount of heterogeneity in recovery. Some cities recovered quite quickly and flourished relative to their pre-pandemic position whereas other cities faded away and never returned to their position in the European urban network.

2.3 The Black Death as a Random Shock

We adopt several strategies to establish that city-level Black Death mortality was largely random.

First, in Table 1 in Columns 4 and 5 we report placebo regressions in which we investigate whether Black Death mortality can explain city growth in the two centuries *prior* to its arrival in Europe. Both estimated coefficients are small and statistically insignificant, as expected.

Second, we collect data on many variables which could potentially have been related to severity of the Black Death in different locations. We use these to test whether any city-level observables can explain Black Death mortality rates. The variables we construct fall into three categories (see Appendix Section 6.2 for detailed descriptions):

Local Physical Geography: Average temperature between 1500 and 1600, elevation, cereal suitability, potato suitability, pastoral suitability, within 10km of sea, within 10km of river, longitude, latitude.

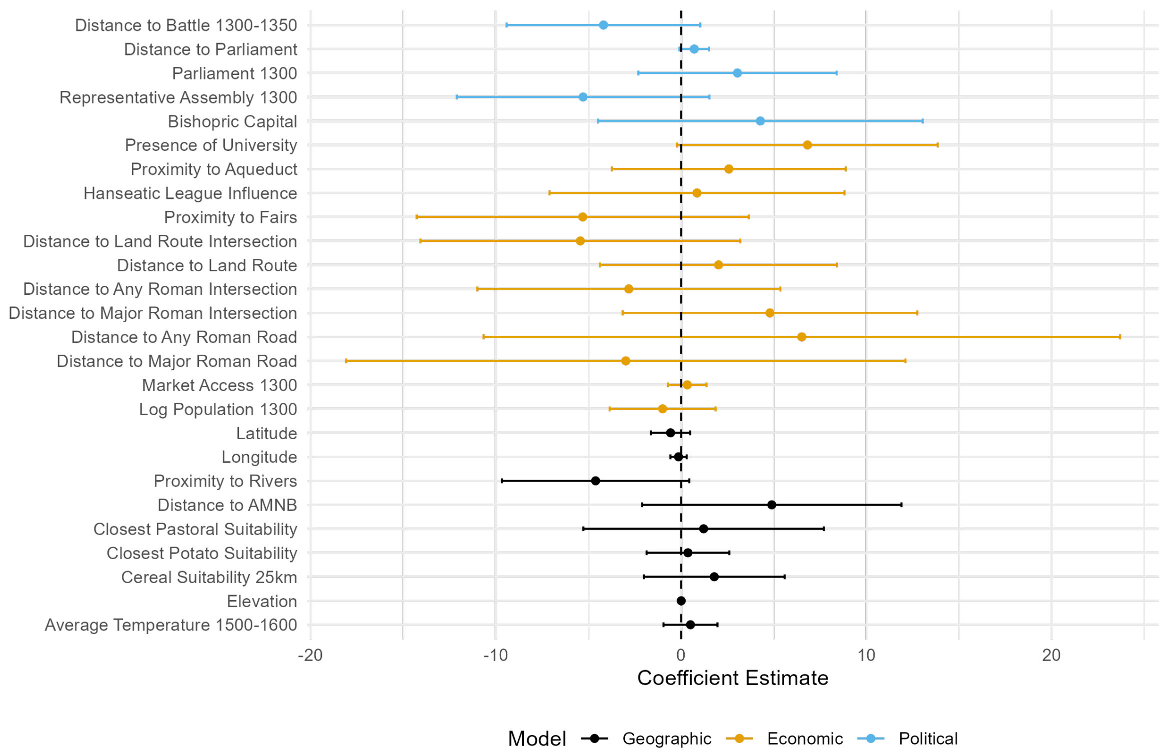
Local Economic Geography: City population in 1300, market access in 1300, within 10km of major Roman road, within 10km of any Roman road, within 10km of major Roman road intersection, within 10km of any Roman road intersection, within 10km of medieval road (from Shepherd Atlas), within 10km of medieval intersection (from Shepherd Atlas), medieval fair location, member of Hanseatic League, roman aqueduct, university.

Local Institutions: Capital city, representative institution in 1300, years parliament met in 14th century, distance to parliament, distance to battles between 1300 and 1350.

In Appendix Table 13 we regress the variables in each of these categories on mortality. We also regress them on mortality as a group. These coefficients with confidence intervals are given in Figure 2. In the regression on local physical geography, only proximity to rivers shows up as significant at the 10% level. However, it is no longer significant in Column 4 when we run the regression including all covariates. In Column 2 there are no variables associated with local economic geography that predict Black Death mortality rates. In Column 3 the only institutional variable that is significant is distance to parliament at the 10% level. This correlation also disappears in the regression using all covariates. Given that we are investigating twenty-six RHS variables, we would expect one or two should be significant by chance. Furthermore, in all balance regressions the adjusted R-squared is never greater than 10%, suggesting that observables do a very poor job explaining Black Death mortality.

Our third identification strategy is to construct measures of market access for each city in 1300 based on historical maps of transportation routes and the costs of using different travel technologies. Several sources suggest that the Black Death traveled along trade routes (see, e.g., Boerner and Severgnini (2014)). If the plague arrived earlier in places that were better connected to trade networks, then this could result in higher mortality rates for places better suited for trade. Trade networks were also likely correlated with press adoption. By accounting for a city's trade potential in 1300, fifty years before the Black Death, we are able to better focus on the variation in city and

Figure 2



Notes: Plot of coefficients from our full balance regression, as reported in Appendix Table 13. Full descriptions of control variables can be found in Appendix Section 6.2.

market size generated by Black Death mortality.

Under some standard assumptions in modern trade models, market access captures all the direct and indirect benefits of trade for a location (see, e.g., Donaldson and Hornbeck, 2016). The theoretical framework from which market access is derived accounts for the general equilibrium relationship between producers and consumers across the entire potential trade network for a city.

Formally, for city j market access is defined as:

$$MA_j = \sum_i \mathrm{N}_i \tau_{ji}^{-\sigma} \ , \tag{2}$$

Where N_i is the population of city i, τ_{ji} is the least cost cost of travel between city j and city i, and σ is a trade elasticity. We calculate market access in 1300 and, for MA_j we sum over the populations of all 470 - 1 = 469 cities for which populations are given Bairoch (1988) in 1300.

The most computationally challenging task in constructing the market access variable is to create a measure of the cost of travel τ between cities. To do this we begin by creating maps of Roman roads, medieval trade routes, major rivers, and seas.[11] Estimates from Bairoch (1988) allow us to assign the cost of transporting goods by each of these routes (portage is assumed to be used when there is no better alternative).[12] In Appendix Tables 20 and 21 we report estimates of our baseline regressions using three alternative parameterizations of travel costs from the literature.[13]

The trade elasticity, σ, measures the responsiveness of trade to changes in transportation costs between towns. Higher values reflect a more elastic response of trade to changes in transport costs. We follow Donaldson (2018) in using a value 3.8 for σ.[14] In Appendix Tables 22 and 23 we investigate alternative values for trade elasticity and find little effect on our baseline results.

To construct the market access measure we divide Europe into 5km x 5km grids and assign to each grid cell the lowest travel cost associated with the travel technologies located inside them. We then apply Djikstra's algorithm to determine the lowest cost of travel between all city pairs (van Etten, 2012). Appendix Figure 11 shows an example of the least cost travel path between Paris and Rome. Using the travel cost measures, we then create our variable measuring market access for each city using equation 2.[15]

2.4 The Printing Press

The movable type printing press was invented by Johannes Gutenberg in Germany around 1440, with the first print shop operating in Mainz by 1450. Though Gutenberg and his associates enjoyed an effective monopoly on movable type printing for several years, the technology spread rapidly in the following decades. By 1500 there were presses operating in 284 cities in Western Europe (*Incunabula Short Title Catalogue*, 1998).

Setting up and operating a press was expensive, requiring constant injections of funds and lucrative side contracts to support larger projects. While exact numbers are scarce, the primary fixed cost

[11] The data on river locations are from Nussli (2011). The data on Roman roads are from the *Digital Atlas of Roman and Medieval Civilizations*. It is available from: `https://darmc.harvard.edu`. We complement the Roman road data with data on medieval trade routes from digitized maps in Shepherd (1923). These data are especially helpful given that the Roman road coverage did not extend into the northeastern part of Europe.

[12] The normalized estimates for cost of travel we use from Bairoch (1988) are: porters = 1; roads = 0.81; rivers = 0.21; seas = 0.08

[13] Those three alternative parameterizations are: (1) Boerner and Severgnini (2014): porters = 1; roads = 0.50; rivers = 0.50; seas = 0.13; (2) Masschaele (2008): porters = 1; roads = 0.81; rivers = 0.51; seas = 0.10; (3) Galloway et al. (1996): porters = 1; roads = 0.81; rivers = 0.59; seas = 0.06.

[14] The appropriate σ depends on context. For modern and developed economies, researchers tend to estimate higher values. For example, Eaton and Kortum (2002) use 8.28 for OECD trade flows in 1995. Donaldson and Hornbeck (2016) estimate an average $\sigma = 8.22$ for trade flows in the U.S. in the second half of the 19th century. By contrast, Donaldson (2018) estimates $\sigma = 3.8$ for colonial India. Kopsidis and Wolf (2012) assume $\sigma = 1$ for their study of Prussian trade during the Industrial Revolution. This is also the value assumed by many earlier studies of 'market potential' or 'market access' (Harris, 1954).

[15] Theory says market access should be logged (see, e.g., Donaldson and Hornbeck (2016)), which is why we use log market access in our regressions. We also follow Donaldson and Hornbeck (2016) in excluding own city population in calculating market access so as to mitigate potential endogeneity concerns.

faced by entering firms was the cost of movable type, which was 4 to 10 years of skilled wages in the mid-1500s (Dittmar, 2011). The Gutenberg Bible took two years to complete and required both a financier and side contracts for print runs of small but popular schoolbooks, pamphlets, and broadsheets to support the business. Some of the largest printing operations required capital contributions equivalent to over 1,000 years of unskilled wages and at the turn of the 15th century the estimated cost of running a large printing house was potentially as high as 200 ducats a month (Dittmar and Seabold, 2019; Pettegree, 2010).[16]

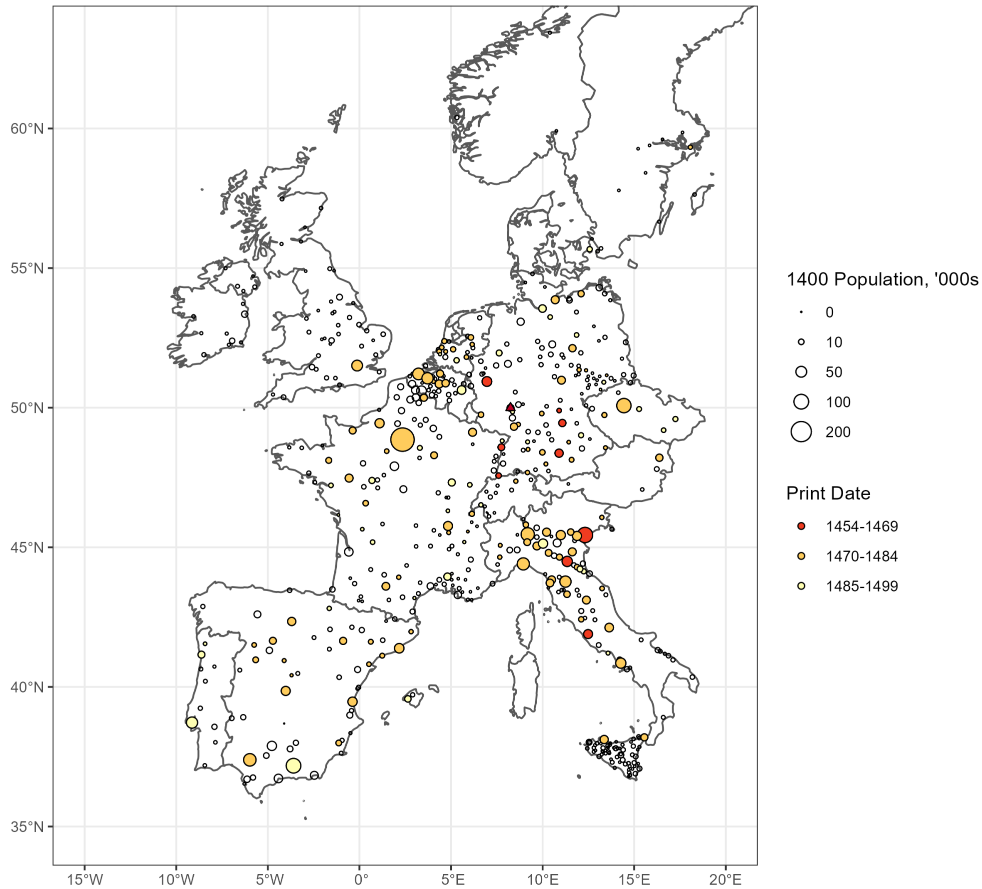
These high fixed costs of production suggest that a sufficient demand pool was necessary for printers to turn a profit. In theory printers could have produced to meet local demand or for trade with consumers in other cities. For example, Gutenberg's former partner and financier, Johann Fust, established a bookstore in Paris in 1466 hoping to find a new market for books printed back in Mainz. Yet within 4 years a local press was established (Pettegree, 2010, chap. 2). In general, such inter-city trade appears to have been limited by high transportation costs over land and the relative fragility of early books (Febvre and Martin, 1976, p. 222-3). Transportation costs gave local print shops a substantial advantage against sellers transporting books from other cities, particularly as competition in the printing market increased (Barbier, 2017, p. 186). Dittmar and Seabold (2019) also uses microdata on book prices to demonstrate the salience of the within-city book market compared to the inter-city market, finding a positive relationship between transport distance and book price and large effects on within-city book prices from the entry and exit of printing firms. Overall, the rapidly expanding market for books demanded a much higher quantity than inter-city booksellers could profitably provide.

Previous scholarship has relied on the geographically limited nature of the market for printed materials to demonstrate a positive impact of the printing press on a city's economic growth. Dittmar (2011) argues that cities that adopted a press in first 50 years received a significant growth advantage in the subsequent century. He suggests that this was due to the localized impact of the press and presents some evidence that the circulation of educational printed materials (e.g. commercial arithmetics, merchant manuals) gave these cities a human capital advantage. For example, double-entry bookkeeping was first described and spread through printed materials in the late 15th century.[17]

[16]For comparison, in 1499, according to the notebooks of Leonardo da Vinci, one ducat was enough to rent a room in Milan for 1.3 months.

[17]Both Dittmar (2011) and Rubin (2014) use distance to Mainz as an instrumental variable for early adoption of the press. In Appendix Table 14 we show that our baseline results are unaffected by including distance to Mainz as a control variable.

Figure 3



Notes: 193 cities with at least one print record before 1500 in the *Universal Short Title Catalogue* that are matched with populations from Bairoch (1988). The red triangle in western Germany is Mainz, where Gutenberg established the first printing press. Cities with no fill are those with population data in Bairoch that did not adopt a press before 1500.

2.5 Data from the Universal Short Title Catalogue

We link Black Death mortality and city population data to information on the spread of the printing press. To do this, we scrape all the data from the Universal Short Title Catalogue (USTC), a project collecting all known information on printed material during the first 200 years of the printing press in Europe.[18] Each book edition has a unique page on the site with specific information about the edition. To scrape these data, first the HTML code for each edition page was saved. Next, basic information about each edition was extracted from the saved HTML code, including print year, print location, imprint, subject, and author. This results in an initial dataset of 826,084 observations.

Some entries are incomplete, so we remove any observations lacking a place or year. This leaves

[18]See `https://www.ustc.ac.uk` and Pettegree and Kemp (2017)

us with 712,982 observations, of which 343,989 occur up to 1600. Each record comes with an 'imprint', a transcription from the title page of a book which also typically contains a year, but a small number of records have a discrepancy between this imprint year and the year encoded in the metadata. We eliminate any records with a discrepancy greater than 10 years, giving us a final working sample of 343,660 records from 1450 to 1600. From this, 26,012 observations are in the early print adoption period of 1450 to 1500.

The chief challenge in matching the USTC data to the Bairoch data is the lack of regularized city names across the two datasets. USTC records contain a 'place' mentioned in the object imprint, but no further details. We utilized a combination of direct and fuzzy matching, along with hand coding, to link just under half of USTC places with our base data. First, USTC data were joined directly by city name. There are a few instances of the same city name occurring in different countries in Bairoch; for example, Brest is both a city in France and in modern day Belarus. We verify the match from USTC by checking the language editions are printed in (in this case Russian), and remove the match to Brest, France. However, many cities have alternate names or spellings, particularly when anglicizing languages with diacritical marks. We institute another round of matching of remaining places by allowing for some string distance between names. We utilize a text matching algorithm via the *fuzzyjoin* package in R, using the Optimal String Alignment distance metric, a variant of the Levenshtein distance, that counts the number of insertions, deletions, and substitutions necessary to convert one string to another while allowing for transposition of adjacent characters. This returns some false matches, which we verify and remove. Finally, we hand match any remaining places that used different names across data sources such as Prague (Praha) and Aix-en-Provence (Aix).

Of 1,312 unique place-names present in the USTC data, we find matches to 631 cities in our base data. We find 193 cities from 1450–1500, and 475 additional cities from 1450–1600.

One concern is that printers may have been operating in places too small to be included in the Bairoch data. Looking at the remaining unmatched cities with entries before 1600, we find that most unmatched cities have only a single edition present in USTC, or otherwise comprise a small portion of the total records matched. According to Table 2 74% of the unique USTC cities between 1450 and 1500 are matched to Bairoch cities. This corresponds to 98% of the total editions published during that time. Appendix Figures 12 and 13 show these numbers in more detail.

Figure 3 shows the pre-1500 joined Bairoch-USTC data of 193 cities with both their estimated population in 1400 and timing of print adoption. The origin of the printing press, Mainz, is represented by the red triangle. There is a clear pattern of larger cities gaining presses in the initial 50-year period, and smaller cities not adopting. During the first 15 years of the press its

Table 2

	1450–1500	1450–1600
Matched USTC Cities	193	475
Missing USTC Cities	68	340
Share of Cities Matched	0.74	0.58
Matched USTC Editions	25,494	329,658
Missing USTC Editions	518	14,002
Share of Editions Matched	0.98	0.96

Notes: Counts and percentages of cities and editions matched to cities in Bairoch. The strongest overlap is in the 1450-1500 period, with 74% of cities printing 98% of all editions matched. The extended period from 1450-1600 still sees 58% of cities printing 96% of editions matched, indicating that unmatched cities were more likely to print at a much smaller scale in both periods.

adoption appears regionally limited to southern Germany and a handful of Italian cities (Venice, Bologna, and Rome), but by just 1485 there is evidence of a press in all corners of Western Europe.

Once we intersect our data sets containing city populations in 1400 (n=564) and Black Death mortality (n=247) we arrive at our base sample of 169 cities. This is the sample we use in the majority of the regressions below. When we intersect our base 169 cities with those that adopted the press by 1500 (n=193), then we have 73 cities who adopted the press and 96 that did not. When we extend our analysis to 1600, then we have 117 cities adopting the press and 52 that did not. See Appendix Table 6.1 for descriptive statistics for the main sample.

3 Estimation Strategy

The baseline model we estimate is:

$$\text{print}_i = \kappa + \mu \text{lpop1400} + \lambda X_i + \varepsilon_i \tag{3}$$

Where $print_i$ refers to some outcome of interest for city i, $lpop1400$ is the log of the population of city i in 1400, and X is a vector of controls. We always control for university or bishopric presence in a town. In other specifications we also control for market access in 1300 and potential cereal suitability.

The primary outcomes we explore are the extensive and intensive margins of print adoption. For the former we use a dummy variable for press adoption before 1500; for the latter we use the log number of editions published in a town between 1450 and 1500.[19] We investigate the longer run

[19]We argue that the number of editions published is a sufficient proxy for printing intensity, even though we do

impacts of the Black Death population shock by extending these measures to cover the period 1450–1600.

We also construct a proxy for the degree of specialization of a city's print market using a Herfindahl-Hirschman Index. Editions in the USTC are classified as falling into one of thirty-eight categories (see Appendix Figure 14). Using the sample of records with both year and subject present, we calculate the share of editions in each subject per city per decade. The HHI is calculated by summing the squares of each of these shares and ranges from zero to one. An index value of one means there is only one subject being printed in the local market.[20] Cities with a lower HHI have a wider variety of subjects being printed; alternatively, cities with a higher HHI are more dominated by fewer subjects. This may help alleviate some concern in using the raw number of editions by capturing another aspect of the local print market.

The estimates of μ in Specification 3 will likely be biased since there are many potential unobserved variables that are correlated with both print adoption and the size of a city in 1400. For example, cities with a more educated population may be richer (higher population density) and more likely to demand and supply books. Alternatively, the trade potential of a city may lead to both a higher population and a greater likelihood of receiving a press.

To address this concern, we will use mortality from the Black Death as an instrument for population in 1400. Figure 4 shows the bin scatter plot of mortality on population in 1400. It is clear from the Figure that the mortality instrument is highly relevant.[21] According to our discussion in Section 2.3 the mortality instrument should also be valid.

Finally, we will evaluate the impact of the Black Death on printing press adoption through its impact on the market potential of a city. Specifically, we calculate the log difference in market access for a city between 1300 and 1400 using Equation 2. To account for potential regional endogeneity in the Black Death population shock, we also adopt the strategy of recalculating our change in market access variable while systematically excluding any city i that falls within a 50,

not know the quantity of books printed of any edition nor of the number of pages of any given book. More editions being printed in a city could indicate either reprinting of existing books or runs of new books that had not yet been published locally. In the first case, new editions of existing books indicate high demand for those texts. Popular texts were frequently printed by multiple printers, resulting in more editions, and living authors would update their writings and have new editions printed. In the second case, a larger number of editions of different books indicates a diverse readership. While we do not explore the differences between these two cases here, neither poses problems for using the number of editions in a city as a proxy for printing intensity.

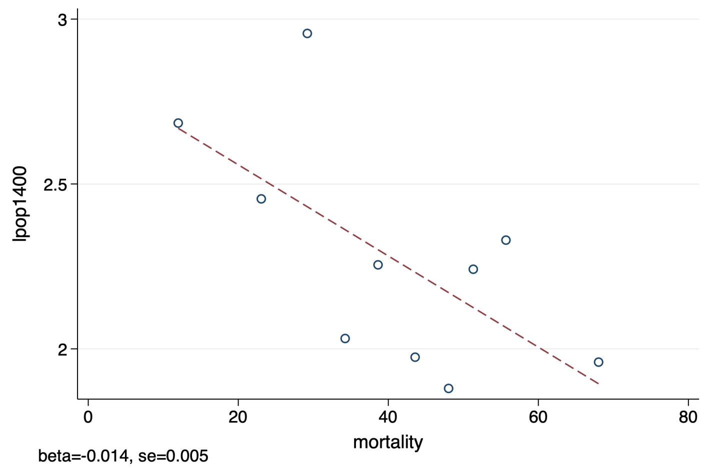
[20]More formally:

$$HHI_i = \sum_{j=1}^{J} s_j^2 \tag{4}$$

where s_j is the share of each subject of all categorized editions per city i for the time period.

[21]The specification estimated is $lpop1400_i = \alpha + \beta mortality_i + \gamma X_i + \varepsilon_i$, where X includes controls for university or bishopric, market access in 1300, and cereal suitability. The estimate of β is -0.014 with a standard error of 0.005.

Figure 4: Bin scatter of mortality regressed on log 1400 population



100, 150, 200, or 250 km radius of city j. This "doughnut" approach relies on Tobler's First Law of Economic Geography which states that "...everything is related to everything else, but near things are more related than distant things" (Johnson, 2018, 425). As the excluded radius of cities gets larger, potential unobserved confounders should become less correlated across cities.

Descriptive statistics for all variables used in the regressions can be found in Appendix Tables 11 and 12. Finally, as spatial correlation is a real concern, in all regressions we report Conley standard errors.[22]

4 Empirical Results

4.1 OLS

We begin by reporting basic OLS results for our measures of the extensive and intensive margins of print adoption in Table 3. In all specifications we see a positive effect of population size on both press adoption and the number of editions printed. Column 3, using our full specification and the press adoption dummy as the dependent variable, shows that a one standard deviation increase in log 1400 population is associated with a 25% greater chance that a city adopts the press by 1500.[23] On the intensive margin, the elasticity reported in Column 6 suggests that a 10% decrease in population in 1400 is associated with a 12.4% decrease in number of editions printed by 1500. Figures 5 and 6 show the bin scatter plots for Specifications 3 and 6 in Table 3.

[22]The distance cut-off we use is the same as that suggested by the function "vcov_conley" in the R package *fixest* and is designed to be robust to small sample issues and drawing the radius too small.

[23]A standard deviation in lpop1400 is about 1.06.

Table 3: OLS

Dependent Variables:	Printing Press in 1500			Log Unique Editions in 1500		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Log Population 1400	0.244***	0.247***	0.251***	1.216***	1.225***	1.241***
	(0.035)	(0.039)	(0.038)	(0.173)	(0.188)	(0.192)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	No	Yes	Yes	No	Yes	Yes
Cereal Suitability	No	No	Yes	No	No	Yes
Fit statistics						
R^2	0.3085	0.3300	0.3328	0.3864	0.3991	0.4007
F-test	24.54	20.20	16.26	34.64	27.23	21.79
Observations	169	169	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. See Appendix for further details.

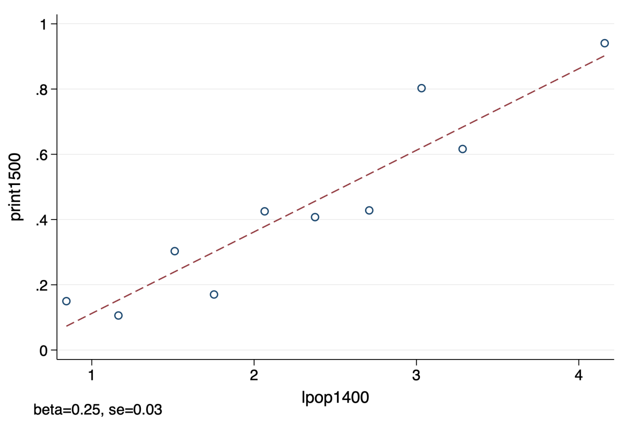


Figure 5: 1400 Population vs. Print Dummy

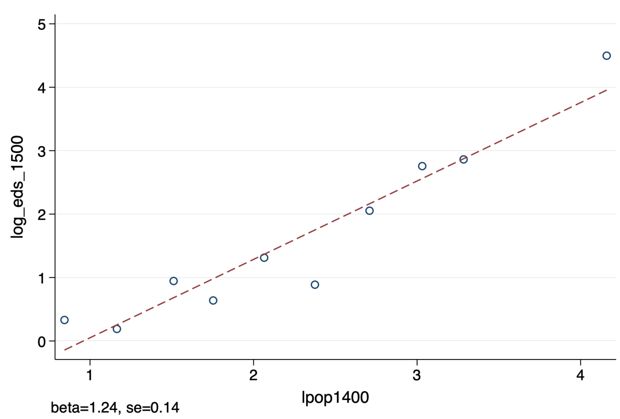


Figure 6: 1400 Population vs. No. Editions

Table 4: First Stage and Reduced Form

	First Stage			Reduced Form					
Dependent Variables:	Log Population 1400			Printing Press in 1500			Log Unique Editions in 1500		
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
Mortality Rate	-0.015***	-0.015***	-0.014***	-0.005***	-0.005***	-0.005***	-0.025***	-0.025***	-0.024***
	(0.005)	(0.005)	(0.004)	(0.002)	(0.002)	(0.002)	(0.009)	(0.008)	(0.008)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Cereal Suitability	No	No	Yes	No	No	Yes	No	No	Yes
Fit statistics									
R^2	0.1621	0.1638	0.1814	0.0965	0.1114	0.1115	0.1277	0.1351	0.1360
F-test	10.64	8.034	7.226	5.872	5.141	4.092	8.055	6.405	5.132
Observations	169	169	169	169	169	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

lost about 22% of a standard deviation in population in 1400.

Mortality also has a negative correlation with our measures of print adoption. In the reduced form regressions in Columns (6) and (9), a 10 percentage point higher Black Death mortality rate is associated with a city being 5% less likely to adopt the press by 1500, and printing about 24% fewer editions.

Table 5: Second Stage IV Models

Dependent Variables:	Printing Press in 1500			Log Unique Editions in 1500		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Log Population 1400	0.354**	0.340**	0.369**	1.706*	1.656**	1.770**
	(0.174)	(0.154)	(0.155)	(0.864)	(0.782)	(0.765)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	No	Yes	Yes	No	Yes	Yes
Cereal Suitability	No	No	Yes	No	No	Yes
Fit statistics						
R^2	0.25946	0.29478	0.27814	0.33903	0.36238	0.34706
F-test	5.8718	5.1413	4.0921	8.0551	6.4054	5.1317
Observations	169	169	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 5 reports the second stage results. The coefficients presented here are both economically and statistically significant. As reported in Column 3, a city with a one standard deviation smaller population due to Black Death mortality was about 37% percentage points less likely to adopt the printing press before 1500. On the intensive margin, the estimate in Column 6 suggests, again, that a 10% increase in Black Death mortality resulted in about 18% fewer editions being published by 1500. Encouragingly, both of these estimates are quite close to the OLS estimates in Table 3. The results in Table 5 support the idea that the Black Death predominantly affected press adoption through its impact on market size. We show the bin scatter of the regressions in Columns 3 and 6 of Table 5 in Figures 7 and 8.

Bin Scatters for 2SLS Impact of Black Death on Press Adoption and Editions Printed, 1450-1500

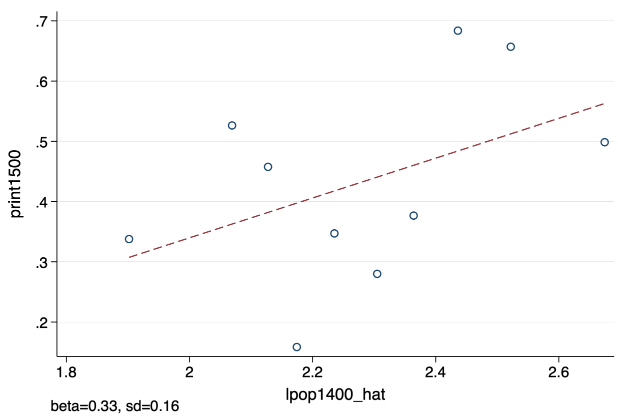


Figure 7: 1400 Population vs. Print Dummy

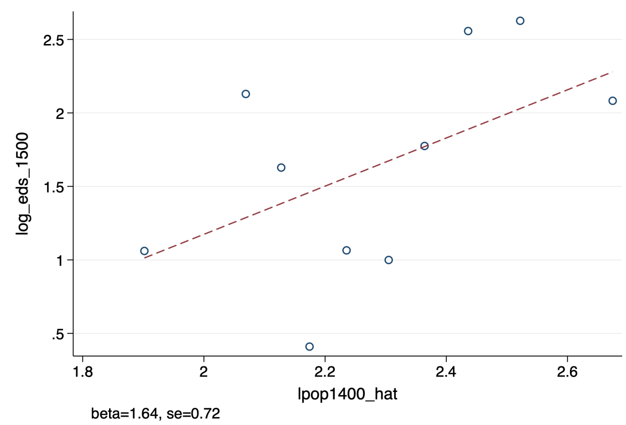


Figure 8: 1400 Population vs. No. Editions

Notes: Bin scatter plot of the impact of mortality predicted log population in 1400 on our extensive and intensive measures of print adoption.

There may still be concerns that the Black Death was such a significant event in European history that its impact on printing press adoption may have been felt through channels other than its direct impact on population. For example, we know the Black Death affected culture more broadly including art, religious belief, and antisemitic behavior (Herlihy, 1997; Jedwab et al., 2019). It is impossible to test this potential violation of the exclusion restriction directly, but we can investigate whether Black Death mortality has a direct impact on press adoption, conditional on population in 1400. In Appendix Table 26 we run these horse race regressions. They show that Black Death mortality has no direct impact on press adoption or number of editions printed by 1500 when population in 1400 is included in the regressions. This suggests that the primary channel through which the Black Death affected press adoption was through its impact on population.

4.3 Estimating the Population Threshold for Print Adoption

In order for market size to have mattered in the decisions of early print entrepreneurs, there must have also been a minimum efficient scale for printing. We attempt to estimate this using threshold regressions.[25] The procedure we implement uses conditional least squares to estimate if a threshold exists at some value of a user-inputted variable (in our case population in 1400). It implements rolling regressions over many possible values of the threshold and then chooses the threshold value(s) that minimizes the sum of squared residuals. Since we are interested in there being at least one printer in a city we use the dummy variable for print adoption by 1500 as the dependent variable. We allow for the existence of multiple thresholds (or none) and estimate

[25]We use the "threshold" command in STATA to make these estimates (`https://www.stata.com/manuals/tsthreshold.pdf`).

models in which we use actual population in 1400 or population in 1400 predicted using Black Death mortality (thereby approximating our 2SLS regressions above).[26] We report the regression output along with estimated population thresholds for print adoption in Appendix Table 15.

Using OLS we estimate thresholds of either 5,000 or 18,000 people depending on whether we use the minimal or maximal set of controls. This is a large range, corresponding to either the 25th or the 75th percentile in the city population distribution. When we use population predicted using Black Death mortality, however, we get a very stable estimate of 10,000 across specifications. This is our preferred estimate and corresponds to just over the 50th percentile in the city population distribution. In 1500 21% of cities with populations less than 10,000 in 1400 had a press. By contrast, 68% of cities above 10,000 population in 1400 had a press in 1500.[27]

Another way for us to think about this result is to ask which cities would have had a press in the absence of the Black Death shock. We do this in two ways. First, identify which cities had populations above 10,000 in 1300 but less than 10,000 in 1400. This approach assumes there would have been zero urban growth aside from the Black Death shock. Fourteen cities fall into this category, among them, Winchester, Norwich, Carcassonne, and Perpignan. We list all these cities along with their mortality rate and whether they had a press in 1500 in Appendix Table 16.

Our second approach is to assume cities would have continued to grow in the fourteenth century at the same rate as they had in the thirteenth, determine which cities would have had a population of more than 10,000 in this counter-factual world, and then pick out the ones that actually had less than 10,000 population (presumably due to the Black Death). This results in a much larger list of 54 cities that we reproduce in Appendix Table 17. Among the cities predicted to have missed out on press adoption due to the Black Death shock are Toulon, Kiel, Nottingham, Malmoe, and Wismar. For these cities median mortality was 48% and 28% had a press in 1500. By contrast, among the 39 cities predicted to have populations greater than 10,000 in 1400 that actually did have populations greater than 10,000 median mortality was 39% and 69% had a press in 1500.

4.4 Impact of Mortality on Alternative Measures of Press Adoption

We now investigate the robustness of these results across alternate measures of print adoption. First, we assess the persistent impact of the Black Death on press adoption beyond the early printing period. Table 6 shows the same specification as Table 5, but with the timeline for press adoption and printing extended through 1600. When including all controls (col. 3), cities with a one standard deviation smaller population due to the Black Death are 27% less likely to have adopted the printing press by 1600. On the intensive margin, a city with a 10% smaller population

[26] We estimate a first stage by regressing log population in 1400 on mortality and controls. Then we predict log population in 1400 and include this (as well as controls) in the threshold regressions.

[27] 43% of cities in total had a press by 1500.

due to the plague printed about 32% fewer editions.[28] While the results on the extensive margin are about the same for the pre-1500 period, the magnitude of the Black Death shock on the intensive margin is about double.

Table 6: Second Stage IV Models, 1600

Dependent Variables:	Printing Press in 1600			Log Unique Editions in 1600		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Log Population 1400	0.363***	0.342***	0.303***	3.418***	3.292***	3.222***
	(0.101)	(0.087)	(0.105)	(0.946)	(0.749)	(0.689)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	No	Yes	Yes	No	Yes	Yes
Cereal Suitability	No	No	Yes	No	No	Yes
Fit statistics						
Observations	169	169	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1600 takes on a value of 1 if a USTC record appears in that city up to the year 1600, or 0 otherwise. Log Unique Editions in 1600 is the natural log of the number of unique editions present in the USTC for a city up to the year 1600 plus 1. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Next we investigate two additional measures of press presence in a city. First we look at the year of press adoption as the dependent variable. To do this, we use the sample of cities that appear in the USTC up to 1600 (n=118). We also model our HHI measure of a city's subject specialization to investigate whether cities that shrank due to the Black Death published books on fewer topics than larger cities. This is motivated by the observation going back to Adam Smith that the degree of specialization in a market is limited by the extent of that market.

In Table 7 we report the second stage coefficients from our 2SLS estimates of the impact of mortality on these two alternate measures of the print market. We find that increasing a city's size by one standard deviation (1.10) would lead it to adopt the press about 25 years earlier. Though the estimates are imprecise, this suggests that higher Black Death mortality not only prevented some cities from adopting the press, but also delayed adoption in cities that eventually had a press by 1600.

[28] Bin scatters for regression 3 and 6 of Table 5 are in appendix figures 15 and 16.

Our results in Columns (4)–(6) suggest that population size is positively correlated with a greater diversity of subjects being printed in a city over the period 1450–1600. One interpretation is that, since individual publishers tended to specialize in subjects, cities with fewer printers also ended up publishing across fewer subjects. If cities more severely impacted by the Black Death could only support one or two printers, then this would explain our results.

Table 7: Impact on Year of Adoption and Market Specialization

Dependent Variables:	Year of First Printing Press			Topic HHI in 1600		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Log Population 1400	-26.075	-24.154	-25.601*	-0.164	-0.161*	-0.139**
	(21.808)	(18.505)	(14.922)	(0.100)	(0.092)	(0.064)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	No	Yes	Yes	No	Yes	Yes
Cereal Suitability	No	No	Yes	No	No	Yes
Fit statistics						
Observations	118	118	118	117	117	117

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Year of First Editions is the earliest year that a USTC edition appears in a city, up to the year 1600. Hirschman-Herfindahl Index is constructed at the decadal level for cities with at least 10 unique editions, then averaged over the period 1450-1600. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

4.5 Did Spatial Spillovers from The Black Death Discourage Press Adoption?

So far we have focused on the impact of the Black Death on own city population and how this discouraged adoption of the press. We now ask whether the decline in broader market potential (i.e. inter-city trade potential) for a city due to the Black Death may have also discouraged press adoption.

We create values for the change in market access of a city while systematically excluding neighbor cities in a 0, 50, 100, 150, 200, and 250 km radius to account for potential regional correlations with city mortality. As the size of the excluded radius increases we expect our measure of change in market access to become less biased but to also lose explanatory power.

We regress our measure of change in market access between 1300 and 1400 on the print dummy and the editions printed measure. Importantly, we exclude the own city in the calculation of market access. As such, we are focused in these regressions on the potential for trade in books beyond the city itself.

Table 8: Impact of Change in Market Access on Print Adoption up to 1500

Dependent Variable:	Printing Press in 1500					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Change in Market Access 1300-1400	0.078***					
	(0.026)					
Change in Market Access 1300-1400, 50km		0.271***				
		(0.071)				
Change in Market Access 1300-1400, 100km			0.254**			
			(0.108)			
Change in Market Access 1300-1400, 150km				0.176*		
				(0.091)		
Change in Market Access 1300-1400, 200km					0.084	
					(0.146)	
Change in Market Access 1300-1400, 250km						0.036
						(0.195)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	Yes	Yes	Yes	Yes	Yes	Yes
Cereal Suitability	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
R^2	0.12310	0.14345	0.12614	0.11507	0.11045	0.10966
F-test	13.028	15.541	13.395	12.066	11.522	11.430
Observations	470	470	470	470	470	470

Conley (240km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample consists of 470 cities for which a population estimate is recorded in 1300. Change in Market Access variables note the change in market access for a city from 1300 to 1400 based on full population data, excluding cities within a 50, 100, 150, 200, and 250 km radius of each observation. Printing Press 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details on market access construction. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 8 reports the effect of the decline in market potential between 1300 and 1400 on print adoption between 1450 and 1500. Due to endogeneity concerns, we expect the result in Column (1), where we only exclude the own city population change from the change in market access measure, to be biased. As we exclude cities within a 50, 100, and then 150 km radius, we expect to remove much of this bias without the estimates becoming too imprecise. From the estimate in Column 3 where we exclude cities within 100km, we calculate that a city with a one standard

Table 9: Impact of Change in Market Access on Print Intensity to 1500

Dependent Variable:	Log Unique Editions in 1500					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Change in Market Access 1300-1400	0.354**					
	(0.172)					
Change in Market Access 1300-1400, 50km		0.924***				
		(0.286)				
Change in Market Access 1300-1400, 100km			0.786*			
			(0.461)			
Change in Market Access 1300-1400, 150km				0.572		
				(0.405)		
Change in Market Access 1300-1400, 200km					0.369	
					(0.521)	
Change in Market Access 1300-1400, 250km						0.313
						(0.746)
Bishopric or University	Yes	Yes	Yes	Yes	Yes	Yes
Market Access in 1300	Yes	Yes	Yes	Yes	Yes	Yes
Cereal Suitability	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
R^2	0.14729	0.15455	0.13986	0.13356	0.13102	0.13047
F-test	16.030	16.965	15.090	14.305	13.992	13.924
Observations	470	470	470	470	470	470

Conley (240km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample consists of 470 cities for which a population estimate is recorded in 1300. Change in Market Access variables note the change in market access for a city from 1300 to 1400 based on full population data, excluding cities within a 50, 100, 150, 200, and 250 km radius of each observation. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details on market access construction. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

deviation market access shock is about 6% less likely to adopt the printing press by 1500.[29] This estimate is about triple the magnitude of the coefficient in Column 1 where no nearby cities are excluded.[30] This difference is possibly explained by potential confounders correlated with proximity to the own city and press adoption. For example, a simple explanation may be that there is a press "shadow effect" whereby potential press cities near a city that has already adopted are less likely to themselves adopt.

In Table 9 we investigate the intensive margin of printing by looking at the number of editions published between 1450 and 1500. The estimate reported in Column 3 where we exclude cities with 100 km of the own city suggest that a standard deviation decrease in market access between 1300 and 1400 is associated with 20% fewer editions being published.[31] Consistent with the coefficients in Table 8 and the potential for attenuation bias, as we exclude more cities from the market access change variables (e.g. 200 or 250 km radii) the coefficients gradually become less precise and approach zero.

In Appendix Tables 18 and 19 we estimate the impact of the change in market access between 1300 and 1400 on press adoption and number of editions published out to 1600. Consistent with the historical literature's claim that the early entrepreneurs of the press were less interested in the inter-city trade potential of a city, our estimated impact of decreased market potential on press diffusion are larger when we include the period after 1500. On the extensive margin, a city was 9% less likely to adopt the press between 1450 and 1600 if it suffered a one standard deviation reduction in market access (compared to 6% for 1450-1500). On the intensive margin, a one standard deviation decrease in market access led to 58% decrease in number of editions printed up to 1600 (compared to 20% for 1450-1500).

5 Conclusion

The Black Death and the invention of the printing press were two of the most important events of the late medieval period. The former devastated the human population but arguably broke the feudal equilibrium, making it a critical point in the long run economic growth of Western Europe. The latter enabled the rise of mass media and the diffusion of knowledge at an unprecedented level, revolutionizing how information and ideas were spread, manipulated, and shared. Here we have explored how these events are connected. While *a priori* the Black Death mortality shock could have impacted press adoption either positively (via higher relative wages) or negatively (via reduced market size), our results support the primacy of the latter.

[29] The bin scatter plot for this regression is in Appendix Figure 17.

[30] In Appendix Table 24 we show these results are robust to using different parameterizations for costs of travel in calculating market access. In Appendix Table 25 we show the results are robust to using alternative trade elasticity assumptions.

[31] The bin scatter plot for this regression is in Appendix Figure 18.

This has important implications for how we think about the relationship between the disease environment, population density, and economic development. While previous historical work has argued that the Black Death led to an increase in relative wages, incentivizing the adoption of less labor-intensive technologies, our results show that a sufficiently large market is an absolute necessity for technology adoption to occur. The direct demographic impact of the Black Death had a significant effect on technological diffusion. Finally, our results speak to the literature on technological innovation in general by providing causal evidence on the importance of market size.

References

Acemoglu, Daron (2002), 'Directed Technical Change', *Review of Economic Studies* **69**(4), 781–809.

Acemoglu, Daron and Joshua Linn (2004), 'Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry', *The Quarterly Journal of Economics* **119**(3), 1049–1090.

Allen, Robert C (2011), 'Why the industrial revolution was british: commerce, induced invention, and the scientific revolution', *The Economic History Review* **64**(2), 357–384.

Bairoch, Paul (1988), *Cites and Economic Development, from the dawn of history to the present*, University of Chicago Press, Chicago. translated by Christopher Braider.

Barbier, Frédéric (2017), *Gutenberg's Europe: The book and the invention of western modernity*, John Wiley & Sons.

Benedictow, Ole J. (2005), *The Black Death 1346–1353: The Complete History*, The Boydell Press, Woodbridge.

Boerner, Lars and Battista Severgnini (2014), Epidemic trade. LSE Economic History Working Paper No. 212/2014.

Campbell, Bruce MS, Alexander Klein, Mark Overton and Bas van Leeuwen (2015), *British economic growth, 1270–1870*, Cambridge University Press.

Chandler, Tertius and Gerald Fox (2013), *3000 years of urban growth*, Elsevier.

Christakos, George, Ricardo A Olea, Marc L Serre, Lin-Lin Wang and Hwa-Lung Yu (2005), *Interdisciplinary public health reasoning and epidemic modelling: the case of black death*, Springer.

Dittmar, Jeremiah E. (2011), 'Information Technology and Economic Change: The Impact of The Printing Press', *The Quarterly Journal of Economics* **126**(3), 1133–1172.

Dittmar, Jeremiah E. and Skipper Seabold (2019), New media and competition: printing and Europe's transformation after Gutenberg., *in* 'CEP Discussion Papers (1600)', London, UK.

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', *The Quarterly Journal of Economics* **131**(2), 799–858.

Eaton, Jonathan and Samuel Kortum (2002), 'Technology, geography, and trade', *Econometrica* **70**(5), 1741–1779.

Eisenstein, Elizabeth L (1980), *The printing press as an agent of change*, Vol. 1, Cambridge University Press.

Febvre, L and HJ Martin (1976), *The Coming of the Book: The Impact of Printing 1450-1800*, Verso.

Galloway, James A, Derek Keene and Margaret Murphy (1996), 'Fuelling the city: production and distribution of firewood and fuel in london's region, 1290-1400 1', *The Economic History Review* **49**(3), 447–472.

Gottfried, Robert S. (1983), *The Black Death: Natural and Human Disaster in Medieval Europe*, The Free Press, New York.

Harris, Chauncy D (1954), 'The market as a factor in the localization of industry in the United States', *Annals of the Association of American Geographers* **44**(4), 315–348.

Herlihy, David (1997), *The Black Death and the transformation of the West*, Harvard University Press.

Incunabula Short Title Catalogue (1998), London: British Library.

Jedwab, Remi, Noel D Johnson and Mark Koyama (2019), 'Negative shocks and mass persecutions: evidence from the black death', *Journal of Economic Growth* **24**(4), 345–395.

Jedwab, Remi, Noel D. Johnson and Mark Koyama (2022), 'The Economic Impact of the Black Death', *Journal of Economic Literature* **60**(1), 132–178.

Jedwab, Remi, Noel D Johnson and Mark Koyama (2024), 'Pandemics and cities: Evidence from the black death and the long-run', *Journal of Urban Economics* **139**, 103628.

Johnson, Noel D (2018), Geospatial information systems, *in* 'An Economist's Guide to Economic History', Springer, pp. 425–432.

Kopsidis, Michael and Nikolaus Wolf (2012), 'Agricultural productivity across Prussia during the Industrial Revolution: a Thünen perspective', *The Journal of Economic History* **72**(03), 634–670.

Masschaele, James (2008), 'Transport costs in medieval england', *The Economic History Review* **46**, 266 – 279.

Nussli, Christos (2011). URL: http://www.euratlas.com/about.html.

Pamuk, Şevket (2007), 'The black death and the origins of the 'great divergence'across europe, 1300–1600', *European Review of Economic History* **11**(3), 289–317.

Pettegree, Andrew (2010), *The Book in the Renaissance*, Yale University Press.

Pettegree, Andrew and Graeme Kemp (2017), 'The universal short title catalogue. creating and exploiting a major bibliographical resource', *La Bibliofilía* **119**(1), 159–172.

Rubin, Jared (2014), 'Printing and Protestants: An Empirical Test of the Role of Printing in the Reformation', *The Review of Economics and Statistics* **96**(2), 270–286.

Russell, Josiah Cox (1972), *Medieval Regions and their Cities*, David & Charles, Newton Abbot, Devon.

Shepherd, William R. (1923), *Historical Atlas*, Henry Holt and Company, New York.

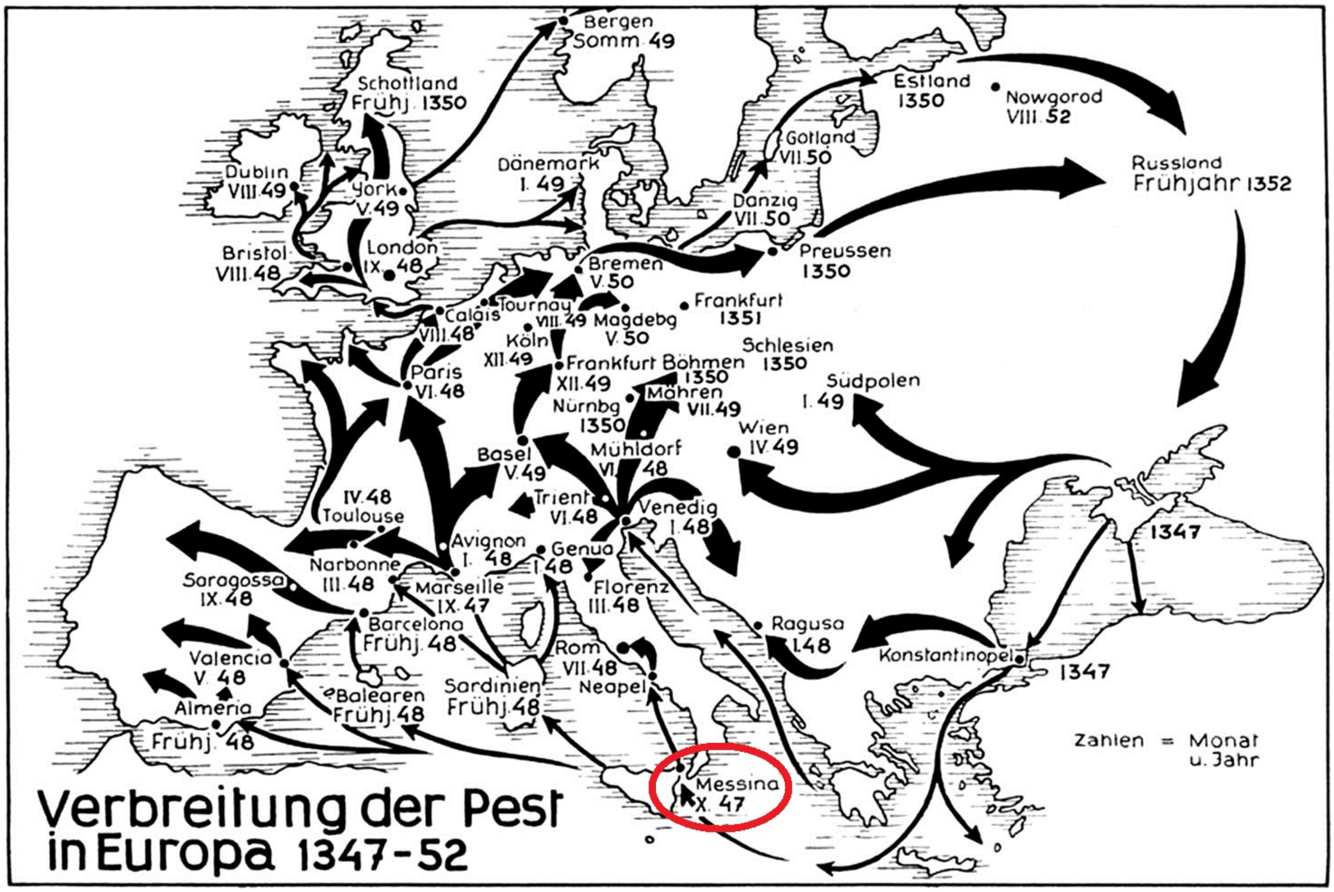
van Etten, Jacob (2012), ‘gdistance: distances and routes on geographical grids. r package version 1.1-4’.

Ziegler, Phillip (1969), *The Black Death*, Penguin Books.

6 Appendix: For Online Publication

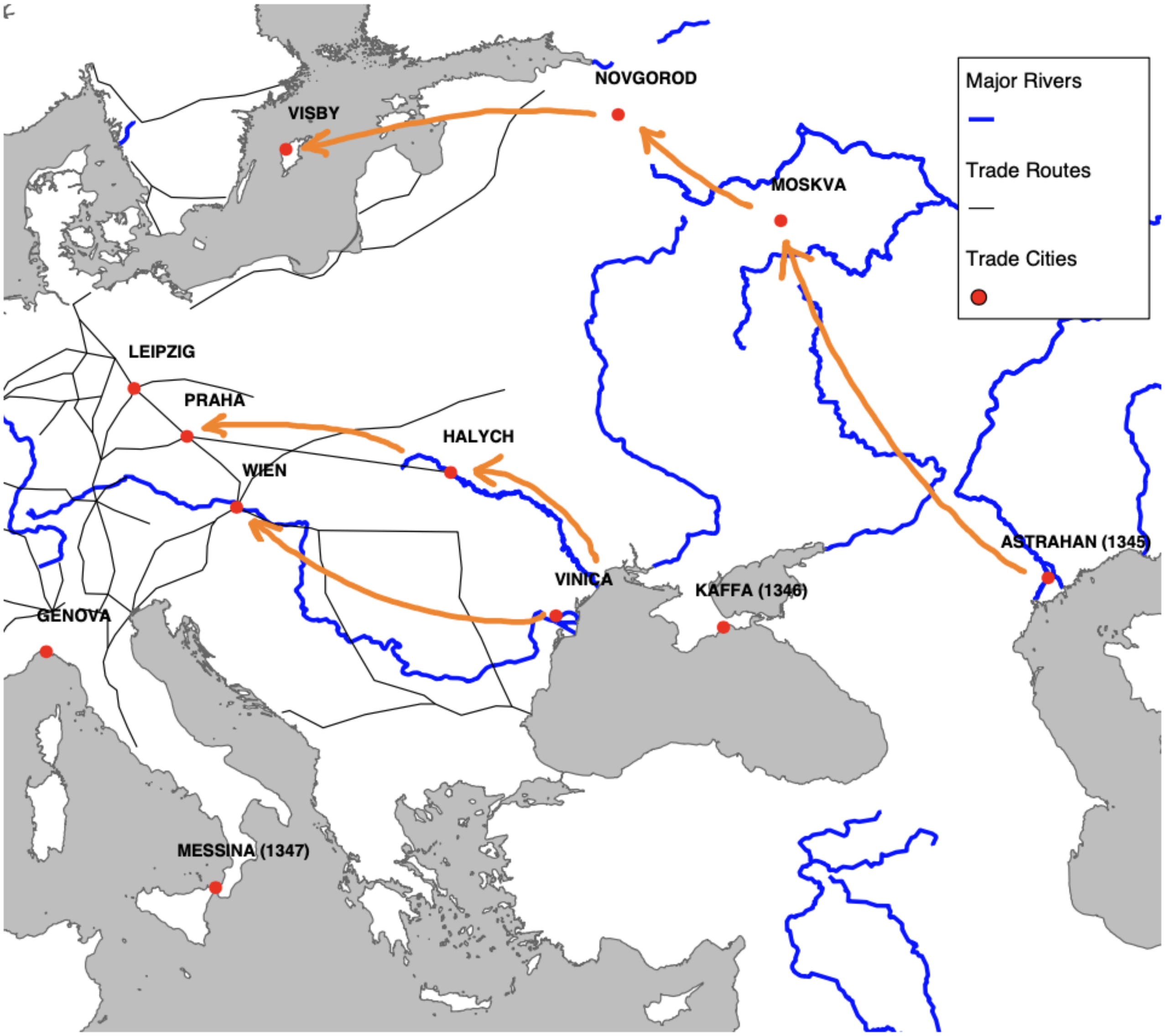
6.1 Additional Figures and Tables

Figure 9: The Spread of the Black Death



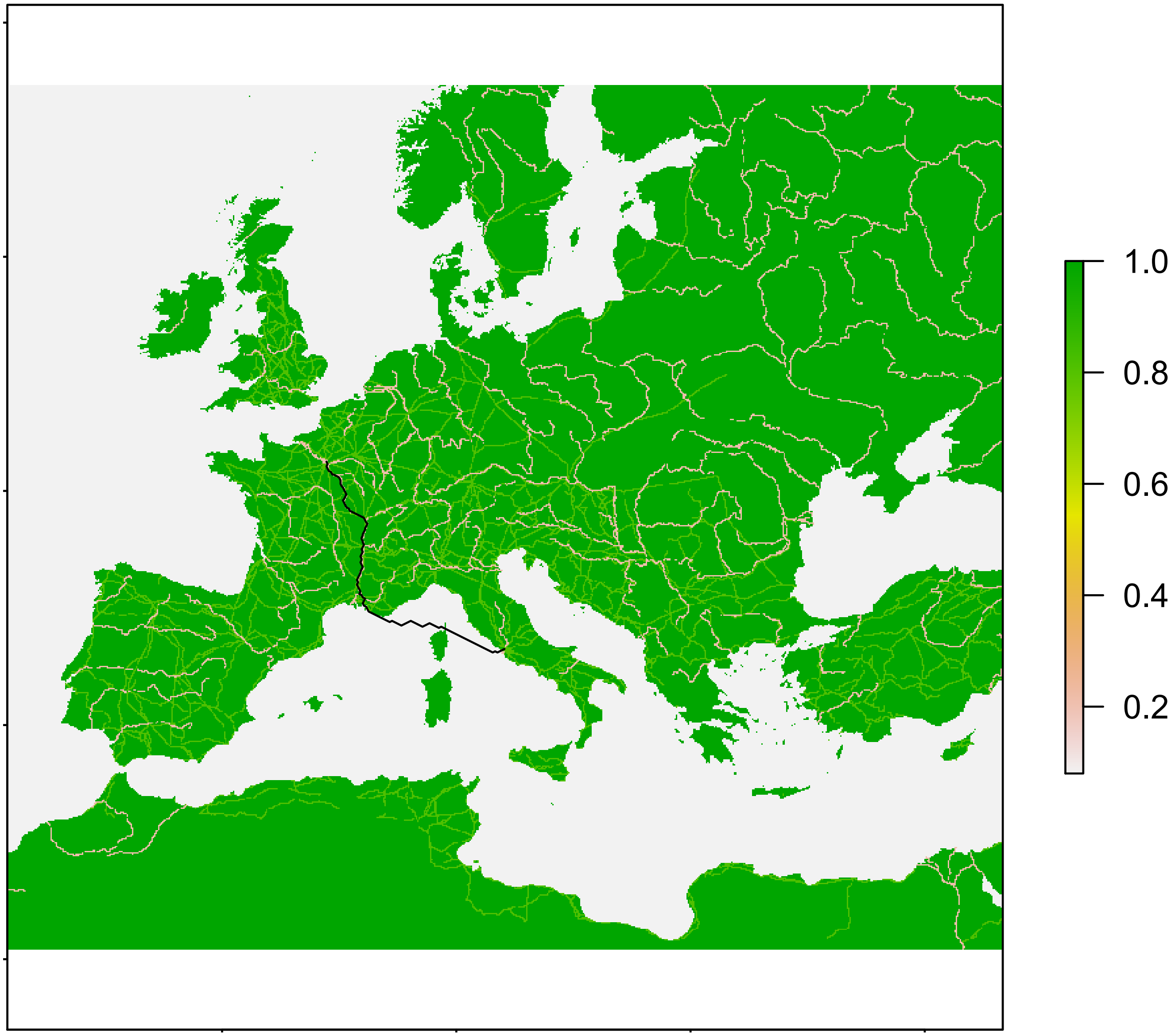
Notes: The spread of the Black Death in Europe, beginning in Messina in 1347. It quickly spread through the rest of Europe. Infected fleas could be transported large distances by either their common black rat hosts or in the clothing of humans, enabling the disease to make large leaps that are difficult to predict. For example, the disease arrived in Messina in October 1347, was in Valencia Spain by May 1348, but already arrived in Brighton England by September 1348. By contrast, it didn't arrive in Köln Germany until December 1349. From Kirsten et al. (1965), accessed in Mengel (2011).

Figure 10: Counterfactual Points of Entry for the Black Death



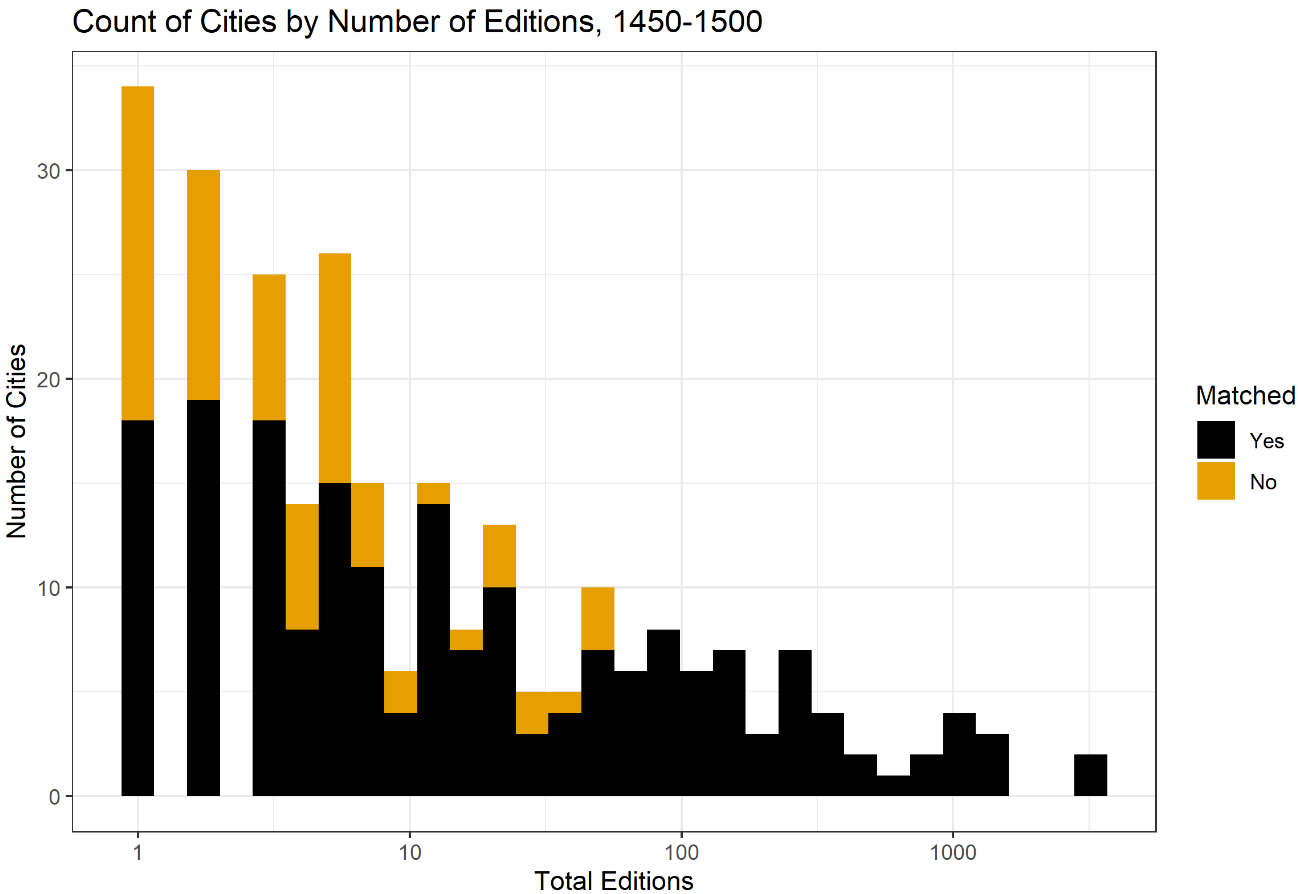
Notes: Potential alternative routes of Black Death spread that do not go through Messina. The trade center of Astrakhan was hit in 1345 and the Genoese colony of Kaffa was infected in 1346, from which Genoese galleys traveled to Messina and brought the plague to Europe. However, the disease could just as easily gone through the popular trade-route from Astrakhan to Moscow, then to Novgorod, and have entered through the Baltic rather than the Mediterranean. Or, it could have easily entered via Prague in central Europe.

Figure 11: Example of a Least Cost Travel Path from Paris to Rome



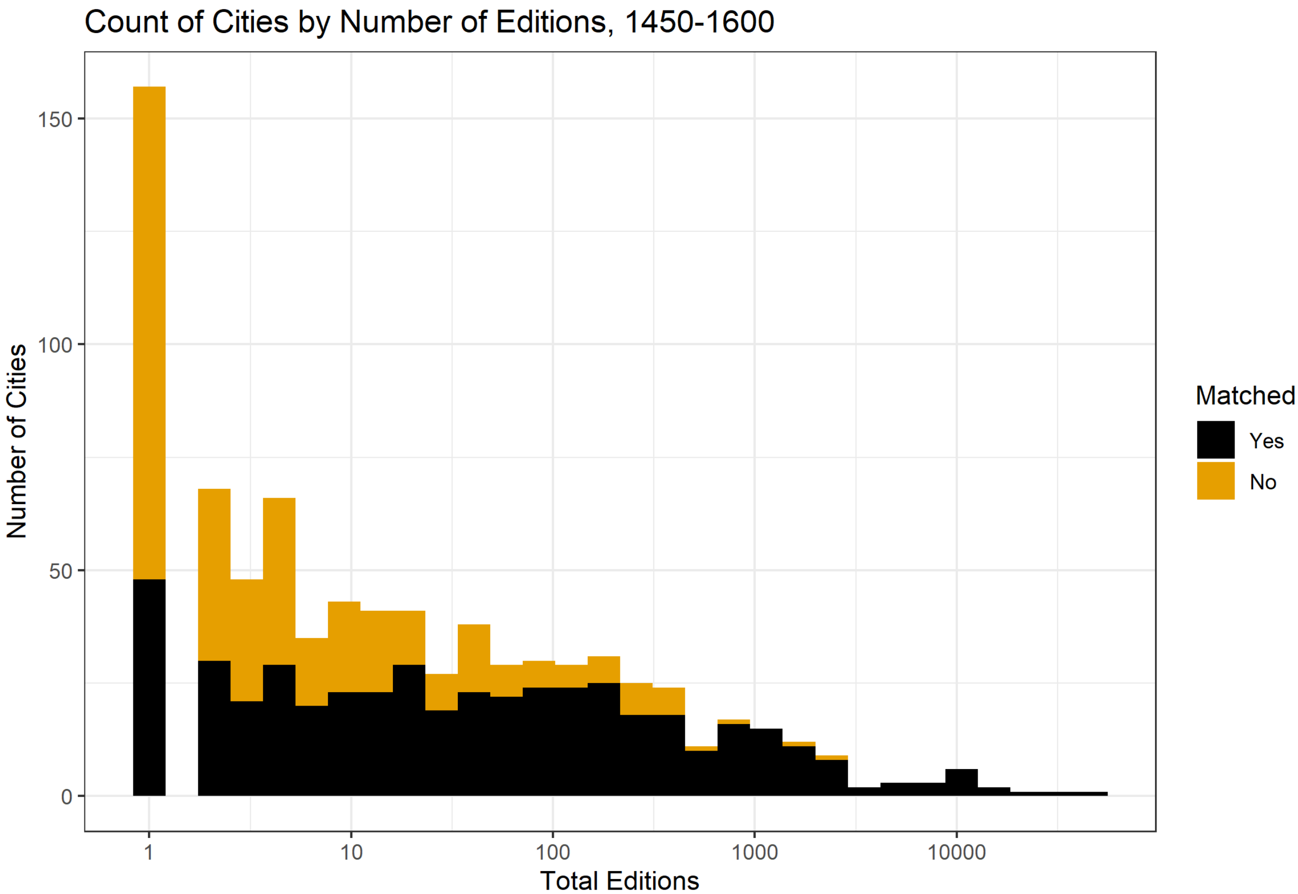
Notes: An example of a least cost travel path from Paris to Rome. To calculate least cost travel paths, we divide Europe into 10km x 10km grids and assign to each grid cell the lowest travel cost associated with the travel technologies located inside them. We then apply Djikstra's algorithm to determine the lowest cost of travel between all city pairs. Colors correspond to the cost of travel for each transportation technology (relative to portage = 1).

Figure 12



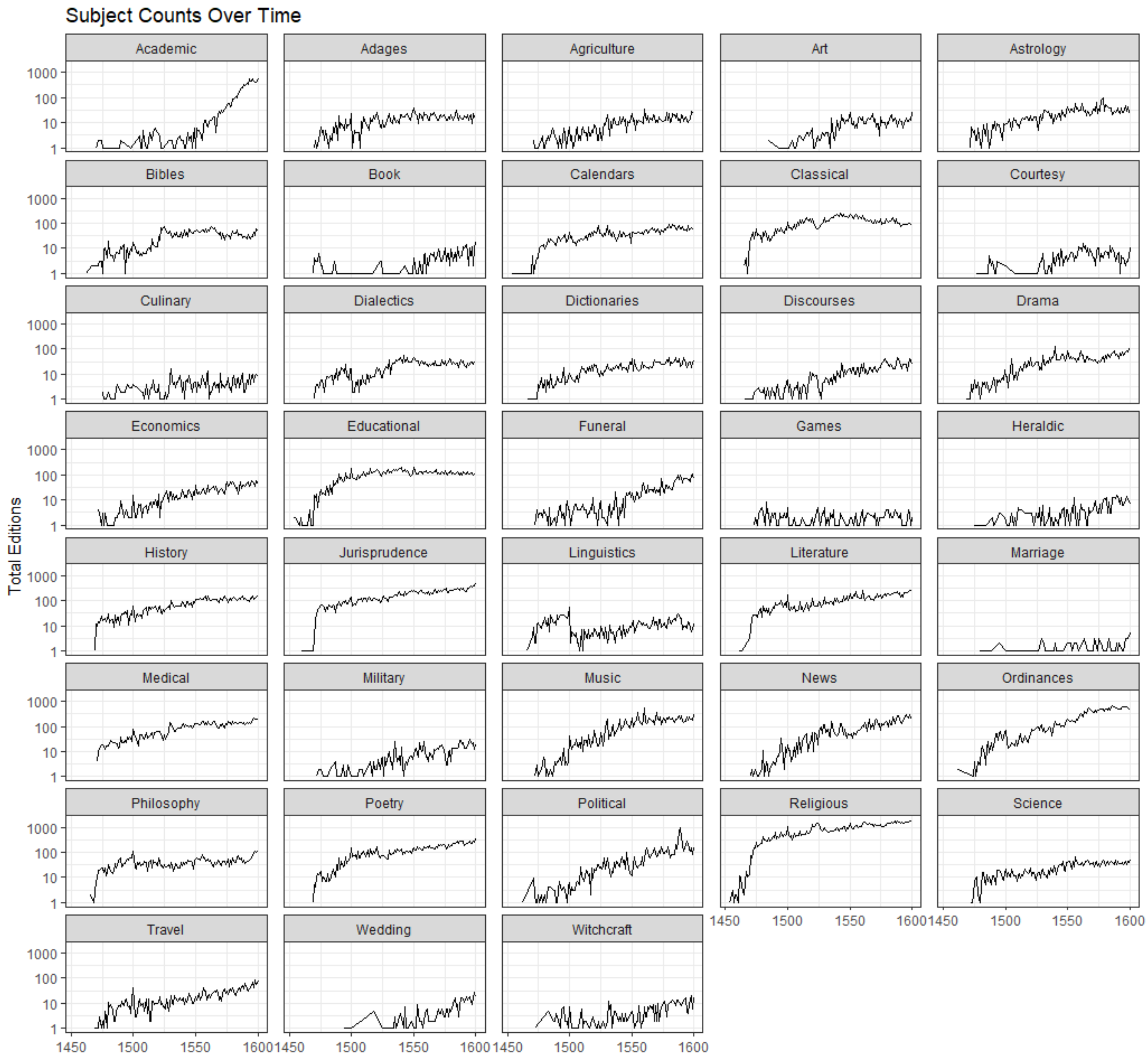
Notes: The count of *Universal Short Title Catalogue* cities matched to the 169 cities in both Bairoch (1988) and the Black Death mortality data sorted by number of editions printed from 1450-1500. The x-axis is log-scaled. A majority of unmatched cities had little print output (¡ 10 editions). All cities that printed ¿ 100 editions were matched to the Bairoch data. As shown in Table 2, 74% of cities printing 98% of editions for this time period are matched to Bairoch.

Figure 13



Notes: The count of *Universal Short Title Catalogue* cities matched to the 169 cities in both Bairoch (1988) and the Black Death mortality data sorted by number of editions printed from 1450-1600. Similar to Figure 12, a majority of unmatched cities had low levels of printing and the most active cities were matched to the Bairoch data. As shown in Table 2, 58% of cities printing 96% of editions for this time period are matched to Bairoch.

Figure 14: Counts of Editions Published by Subject over Time



Notes: The number of editions printed in different subjects from 1450—1600, as classified by the *Universal Short Title Catalogue*. The y-axis is log-scaled. Religious materials were the most widely printed and Marriage materials (largely announcements) the least widely printed. These data are used in the construction of our Herfindahl-Hirschman Index.

Second Stage Bin Scatters for 1450-1600

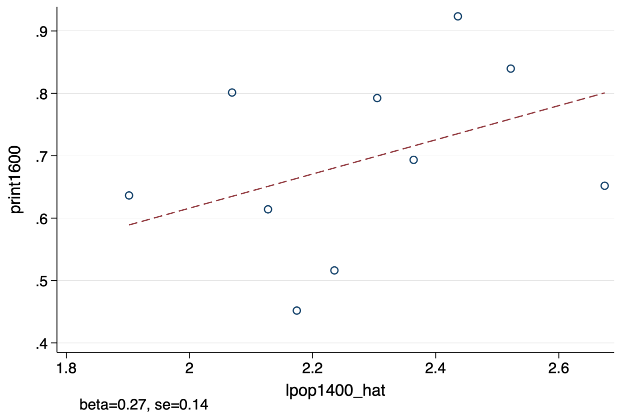


Figure 15: 2SLS Second Stage:
1400 Population vs. 1600 Print Dummy

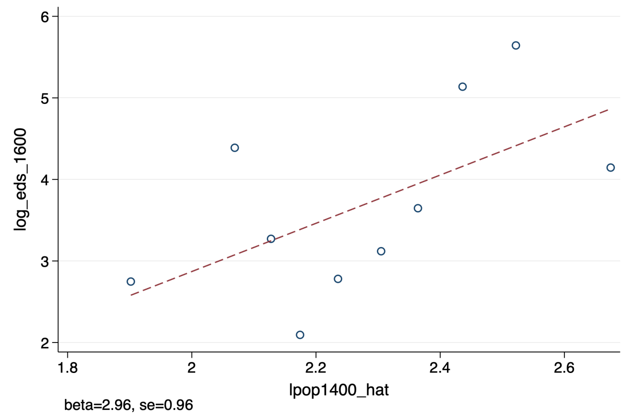


Figure 16: 2SLS Second Stage:
1400 Population vs. 1600 No. Editions

Bin Scatter Plots for 100k Regressions in Tables 8 and 9

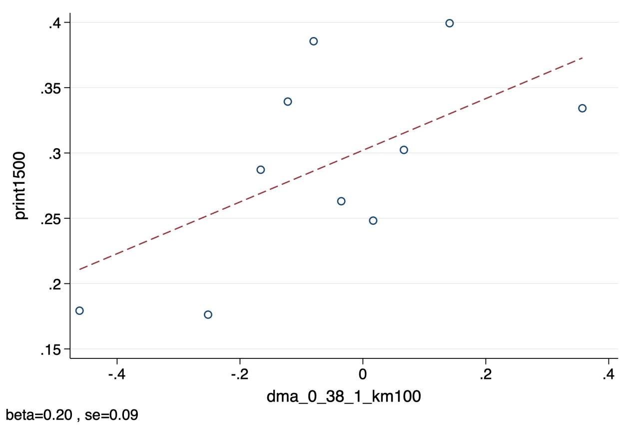


Figure 17: MA Change (100k) vs. Print Dummy

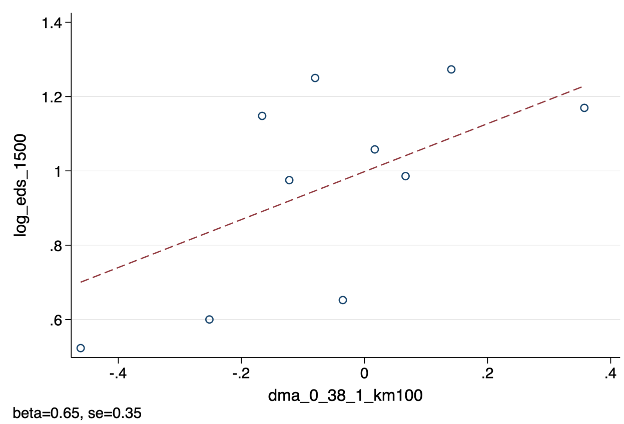


Figure 18: MA Change (100k) vs. No. Editions

Table 10: Summary Statistics for City Level Observables

Variable	N	Mean	Std. Dev.	Min	Max
Mortality Rate	169	40	17	0	93
Average Temperature 1500-1600	169	18	3.2	9.3	24
Elevation	169	148	226	1	1999
Cereal Suitability 25km	169	4.6	0.96	1.3	7.6
Closest Potato Suitability	169	5.7	1.1	3	8
Closest Pastoral Suitability	169	0.91	0.29	0	1
Distance to AMNB	169	0.23	0.42	0	1
Proximity to Rivers	169	0.31	0.47	0	1
Longitude	169	4.7	6.3	-9.1	18
Latitude	169	47	5.1	37	60
Log Population 1300	169	2.4	1	0.41	5.4
Market Access 1300	169	3.1	2.7	-1.8	10
Distance to Major Roman Road	169	0.64	0.48	0	1
Distance to Any Roman Road	169	0.72	0.45	0	1
Distance to Major Roman Intersection	169	0.36	0.48	0	1
Distance to Any Roman Intersection	169	0.49	0.5	0	1
Distance to Land Route	169	0.34	0.47	0	1
Distance to Land Route Intersection	169	0.1	0.3	0	1
Proximity to Fairs	169	0.2	0.4	0	1
Hanseatic League Influence	169	0.1	0.3	0	1
Proximity to Aqueduct	169	0.14	0.34	0	1
Presence of University	169	0.083	0.28	0	1
Bishopric Capital	169	0.12	0.32	0	1
Representative Assembly 1300	169	0.26	0.44	0	1
Parliament 1300	169	0.54	0.5	0	1
Distance to Parliament	169	4.4	3.2	-12	6.7
Distance to Battle 1300-1350	169	0.37	0.48	0	1

Table 11: Summary Statistics for Mortality Cities

Variable	N	Mean	Std. Dev.	Min	Max
Printing Press in 1500	169	0.43	0.5	0	1
Printing Press in 1600	169	0.69	0.46	0	1
Log Unique Editions in 1500	169	1.6	2.2	0	8
Log Unique Editions in 1600	169	3.7	3.2	0	11
Mortality Rate	169	40	17	0	93
Log Population 1300	169	2.4	1	0.41	5.4
Log Population 1400	169	2.3	1.1	0	5.6
Presence of Bishopric	169	0.56	0.5	0	1
Presence of University	169	0.083	0.28	0	1
Cereal Suitability 25km	169	4.6	0.96	1.3	7.6
Distance to Mainz	169	727	407	0	1949
Least Cost Distance to Mainz	169	16	6	0	33
Market Access 1300	169	3.1	2.7	-1.8	10
Year of First Printing Press	118	1498	31	1454	1600
Topic HHI in 1600	117	0.48	0.23	0.16	1

Table 12: Summary Statistics for Market Access Cities

Variable	N	Mean	Std. Dev.	Min	Max
Printing Press in 1500	470	0.29	0.45	0	1
Printing Press in 1600	470	0.54	0.5	0	1
Log Unique Editions in 1500	470	0.96	1.8	0	8
Log Unique Editions in 1600	470	2.5	2.9	0	11
Change in Market Access 1300-1400	470	0.043	0.69	-2.3	5
Change in Market Access 1300-1400, 50km	470	-0.036	0.32	-1.6	1.6
Change in Market Access 1300-1400, 100km	470	-0.041	0.24	-0.94	0.93
Change in Market Access 1300-1400, 150km	470	-0.055	0.2	-0.81	1
Change in Market Access 1300-1400, 200km	470	-0.059	0.17	-0.62	0.73
Change in Market Access 1300-1400, 250km	470	-0.074	0.15	-0.51	0.43

Table 13: Balance Table for Mortality Rate

Dependent Variable:	Mortality Rate			
Model:	(1)	(2)	(3)	(4)
Variables				
Average Temperature 1500-1600	0.5043 (0.7373)			1.100 (0.8870)
Elevation	0.0049 (0.0054)			0.0084 (0.0071)
Cereal Suitability 25km	1.789 (1.923)			2.104 (1.696)
Closest Potato Suitability	0.3681 (1.130)			0.0201 (1.745)
Closest Pastoral Suitability	1.214 (3.287)			1.293 (4.478)
Distance to AMNB	4.894 (3.543)			5.355 (4.104)
Proximity to Rivers	-4.617* (2.560)			-4.790 (3.510)
Longitude	-0.1405 (0.2227)			-0.0152 (0.2946)
Latitude	-0.5705 (0.5340)			-0.1786 (0.5967)
Log Population 1300		-0.9998 (1.448)		-2.685 (1.662)
Market Access 1300		0.3308 (0.5281)		-0.1057 (0.5056)
Distance to Major Roman Road		-2.989 (7.643)		-2.467 (6.185)
Distance to Any Roman Road		6.514 (8.697)		4.532 (6.185)
Distance to Major Roman Intersection		4.797 (4.025)		6.568* (3.501)
Distance to Any Roman Intersection		-2.824 (4.141)		-1.612 (3.375)
Distance to Land Route		2.018 (3.237)		2.685 (3.705)
Distance to Land Route Intersection		-5.441 (4.372)		-6.692 (4.043)
Proximity to Fairs		-5.313 (4.538)		-1.431 (4.342)
Hanseatic League Influence		0.8562 (4.029)		3.559 (6.096)
Proximity to Aqueduct		2.577 (3.196)		0.6488 (3.411)
Presence of University		6.819* (3.561)		5.735* (3.245)
Bishopric Capital			4.276 (4.439)	0.9003 (3.617)
Representative Assembly 1300			-5.292 (3.453)	-0.1325 (2.987)
Parliament 1300			3.040 (2.712)	1.239 (3.356)
Distance to Parliament			0.7069* (0.4065)	0.0178 (0.3631)
Distance to Battle 1300-1350			-4.191 (2.648)	-2.450 (3.529)
Fit statistics				
Observations	169	169	169	169
R^2	0.15095	0.07718	0.06486	0.22269
Adjusted R^2	0.10289	0.00620	0.03617	0.08037

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 14: IV Regressions with Mainz Controls

Dependent Variables:	Printing Press in 1500		Log Unique Editions in 1500	
Model:	(1)	(2)	(3)	(4)
Variables				
Log Population 1400	0.3017**	0.3474**	1.166	1.435**
	(0.1503)	(0.1461)	(0.7307)	(0.6663)
Least Cost Distance to Mainz		-0.0022		-0.0335
		(0.0099)		(0.0423)
Distance to Mainz	-9.25×10^{-5}		-0.0008	
	(0.0001)		(0.0005)	
Presence of Bishopric	0.0488	0.0284	0.3571	0.2579
	(0.0698)	(0.0689)	(0.3598)	(0.3119)
Presence of University	0.0269	-0.0169	0.8121	0.5659
	(0.1140)	(0.1013)	(0.8774)	(0.8227)
Market Access 1300	-0.0275**	-0.0300**	-0.0831	-0.1033*
	(0.0138)	(0.0135)	(0.0587)	(0.0583)
Cereal Suitability 25km	0.0472	0.0490	0.1667	0.1748
	(0.0339)	(0.0300)	(0.1951)	(0.1587)
Fit statistics				
F-test	3.6027	3.3914	4.9804	4.4024
Observations	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Least Cost Distance to Mainz is calculated using our preferred set of travel costs. Distance to Mainz is the straight line distance in km. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and trade elasticity. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 15: Threshold Regressions

	(1) print1500 OLS	(2) print1500 OLS	(3) print1500 IV	(4) print1500 IV
Region1				
Log Population in 1400	-0.0237	0.166***	0.258	0.585**
	(0.134)	(0.0614)	(0.197)	(0.250)
University	-0.180	-0.188	0	0
	(0.252)	(0.188)	(.)	(.)
Bishopric	0.236**	0.102	-0.0264	-0.00539
	(0.117)	(0.0796)	(0.113)	(0.149)
Market Access in 1300		-0.0436***		-0.0497***
		(0.0138)		(0.0186)
Cereal Suitability		0.0233		0.0181
		(0.0419)		(0.0628)
Constant	0.0603	-0.0482	-0.183	-0.711
	(0.155)	(0.239)	(0.384)	(0.637)
Region2				
Log Population in 1400	0.243***	0.140	0.0275	0.720**
	(0.0516)	(0.109)	(0.456)	(0.296)
University	0.168	0.172	0.0736	-0.178
	(0.140)	(0.157)	(0.254)	(0.233)
Bishopric	0.00913	-0.0118	0.714*	0.173
	(0.0838)	(0.132)	(0.374)	(0.190)
Market Access in 1300		0.00991		0.00489
		(0.0228)		(0.0207)
Cereal Suitability		0.0593		0.0615
		(0.0616)		(0.0823)
Constant	-0.126	-0.00841	-0.108	-1.822**
	(0.152)	(0.401)	(1.136)	(0.888)
N	169	169	169	169
Predicted Threshold	5,000	18,000	10,000	10,000

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: These regressions are run using the "threshold" command in Stata. Each column represents a single regression. In all regressions the program picked out a single threshold break (we allow the program to test for both a change in the slope of the parameter estimated on Log Population in 1400 and y-intercept changes). There are two sets of parameter estimates reported for each regression corresponding to the two sides of the estimated threshold. In columns 1 and 2 we report OLS regressions. In column 1 we control for university or bishopric presence. In column 2 we add controls for market access in 1300 as well as potential cereal suitability. In columns 3 and 4 we approximate IV regressions by first predicting population in 1400 using mortality and the controls and then including the predicted mortality in the threshold regression.

Table 16: Cities > 10,000 in 1300 but < 10,000 in 1400

City	**Mortality Rate**	**Print1500**
REGENSBURG	16.66667	1
BAEZA	50	0
BEZIERS	50	0
CARCASSONNE	40	0
LAON	50	0
NARBONNE	40	1
PERPIGNAN	60	0
NORWICH	42.5	0
PLYMOUTH	51.85185	0
WINCHESTER	50	0
AREZZO	60	0
ORVIETO	50	0
PISTOIA	40	0
PRATO	40	0

Table 17: Cities Projected to be > 10,000 in 1400 but < 10,000 in 1400

City	Mortality Rate	Print1500
FRANKFURTANDERODER	16.66667	0
HALBERSTADT	44.16667	0
KIEL	25	0
KONSTANZ	12.5	1
LANDSHUT	50	0
LUNEBURG	36	1
MUNCHEN	50	1
OSNABRUECK	4.166667	0
PADERBORN	33.33333	0
PASSAU	16.66667	1
WISMAR	42	0
ATH	16	0
BAEZA	50	0
CASTELLON-DE-LA-PLANA	0	0
GERONA	66	1
LERIDA	23	1
MADRID	22.5	1
PAMPLONA	54	1
TARRAGONA	50	1
ALBI	55	1
ANNECY	50	0
APT	52	0
CARCASSONNE	40	0
CARPENTRAS	5	0
CHAMBERY	59	1
GRENOBLE	45	1
LIMOGES	25	0
LODEVE	53	0
MONTAUBAN	40	0
PERPIGNAN	60	0
ST-FLOUR	34	0
TARBES	50	0
TOULON	50	0
ABERDEEN	50	0
BATH	44.44444	0
EDINBURGH	33	0
LEICESTER	50	0
NOTTINGHAM	33.33333	0
PLYMOUTH	51.85185	0
SHREWSBURY	58.33333	0
DROGHEDA	45	0
KILKENNY	50	0
AREZZO	60	0
PRATO	40	0
SIRACUSA	50	0
TORINO	33	1
OSLO	50	0
SANTAREM	66	0
SILVES	80	0
MALMOE	50	0
STOCKHOLM	50	1
ZUERICH	60	1
PARCHIM	40	0
INCA	20	0

Table 18: Impact of Change in Market Access on Print Adoption up to 1600

Dependent Variable:	Printing Press in 1600					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Change in Market Access 1300-1400	0.0764**					
	(0.0343)					
Change in Market Access 1300-1400, 50km		0.2632**				
		(0.1173)				
Change in Market Access 1300-1400, 100km			0.3802**			
			(0.1570)			
Change in Market Access 1300-1400, 150km				0.2609**		
				(0.1310)		
Change in Market Access 1300-1400, 200km					0.2266	
					(0.1946)	
Change in Market Access 1300-1400, 250km						0.3209
						(0.2827)
Market Access 1300	-0.0227*	-0.0208*	-0.0199*	-0.0201	-0.0217*	-0.0195
	(0.0120)	(0.0110)	(0.0114)	(0.0122)	(0.0118)	(0.0119)
Presence of Bishopric	0.2908***	0.3035***	0.3091***	0.2950***	0.2903***	0.2920***
	(0.0687)	(0.0641)	(0.0610)	(0.0665)	(0.0699)	(0.0690)
Presence of University	0.2949***	0.3090***	0.3057***	0.3067***	0.3051***	0.3045***
	(0.0940)	(0.0949)	(0.0971)	(0.0989)	(0.0997)	(0.0956)
Cereal Suitability 25km	-0.0909***	-0.0851***	-0.0849***	-0.0904***	-0.0926***	-0.0917***
	(0.0240)	(0.0236)	(0.0231)	(0.0220)	(0.0206)	(0.0220)
Fit statistics						
R^2	0.18125	0.19711	0.20145	0.18057	0.17597	0.17809
F-test	20.544	22.783	23.411	20.449	19.818	20.108
Observations	470	470	470	470	470	470

Conley (240km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample consists of 470 cities for which a population estimate is recorded in 1300. Printing Press in 1600 takes on a value of 1 if a USTC record appears in that city up to the year 1600, or 0 otherwise. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details on market access construction. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 19: Impact of Change in Market Access on Log Unique Editions up to 1600

Dependent Variable:	Printing Press in 1600					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Change in Market Access 1300-1400	0.0764**					
	(0.0343)					
Change in Market Access 1300-1400, 50km		0.2632**				
		(0.1173)				
Change in Market Access 1300-1400, 100km			0.3802**			
			(0.1570)			
Change in Market Access 1300-1400, 150km				0.2609**		
				(0.1310)		
Change in Market Access 1300-1400, 200km					0.2266	
					(0.1946)	
Change in Market Access 1300-1400, 250km						0.3209
						(0.2827)
Market Access 1300	-0.0227*	-0.0208*	-0.0199*	-0.0201	-0.0217*	-0.0195
	(0.0120)	(0.0110)	(0.0114)	(0.0122)	(0.0118)	(0.0119)
Presence of Bishopric	0.2908***	0.3035***	0.3091***	0.2950***	0.2903***	0.2920***
	(0.0687)	(0.0641)	(0.0610)	(0.0665)	(0.0699)	(0.0690)
Presence of University	0.2949***	0.3090***	0.3057***	0.3067***	0.3051***	0.3045***
	(0.0940)	(0.0949)	(0.0971)	(0.0989)	(0.0997)	(0.0956)
Cereal Suitability 25km	-0.0909***	-0.0851***	-0.0849***	-0.0904***	-0.0926***	-0.0917***
	(0.0240)	(0.0236)	(0.0231)	(0.0220)	(0.0206)	(0.0220)
Fit statistics						
R^2	0.18125	0.19711	0.20145	0.18057	0.17597	0.17809
F-test	20.544	22.783	23.411	20.449	19.818	20.108
Observations	470	470	470	470	470	470

Conley (240km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Sample consists of 470 cities for which a population estimate is recorded in 1300. dma variables note the change in market access for a city from 1300 to 1400 based on full population data. dma variables appended with ‘km‘ exclude cities within a 50, 100, 150, 200, and 250 km radius of each observation are excluded from the market access calculation. Log Unique Editions in 1600 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1600. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details on market access construction. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 20: First Stage, Travel Costs

Dependent Variable:	Log Population 1400			
Model:	(1)	(2)	(3)	(4)
Variables				
Mortality Rate	-0.0137***	-0.0138***	-0.0140***	-0.0135***
	(0.0043)	(0.0042)	(0.0043)	(0.0045)
Presence of Bishopric	0.5714***	0.5785***	0.5772***	0.5776***
	(0.1736)	(0.1763)	(0.1745)	(0.1783)
Presence of University	0.7990***	0.7978***	0.8022***	0.7971***
	(0.2463)	(0.2517)	(0.2447)	(0.2584)
Cereal Suitability 25km	-0.1506	-0.1460	-0.1500	-0.1444
	(0.0982)	(0.0962)	(0.0987)	(0.0954)
Market Access 1300, Cost 1	0.0216			
	(0.0266)			
Market Access 1300, Cost 2		0.0185		
		(0.0340)		
Market Access 1300, Cost 3			0.0220	
			(0.0274)	
Market Access 1300, Cost 4				0.0046
				(0.0438)
Fit statistics				
R^2	0.18145	0.18016	0.18174	0.17850
F-test	7.2264	7.1640	7.2408	7.0836
Observations	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details. Cereal 25k is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 21: Second Stage, Travel Costs

Dependent Variables:	Printing Press in 1500				Log Unique Editions in 1500			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Log Population 1400	0.3695**	0.3306**	0.3188**	0.4010**	1.770**	1.639**	1.585**	1.861**
	(0.1548)	(0.1356)	(0.1332)	(0.1665)	(0.7646)	(0.6640)	(0.6394)	(0.8338)
Presence of Bishopric	0.0152	0.0281	0.0369	-0.0055	0.0568	0.0985	0.1374	-0.0125
	(0.0663)	(0.0663)	(0.0661)	(0.0755)	(0.2733)	(0.2592)	(0.2462)	(0.3337)
Presence of University	-0.0400	-0.0069	-0.0055	-0.0574	0.2162	0.3278	0.3420	0.1645
	(0.1021)	(0.0808)	(0.0817)	(0.1112)	(0.9020)	(0.8265)	(0.8136)	(0.9379)
Cereal Suitability 25km	0.0503*	0.0401	0.0444	0.0455	0.1949	0.1589	0.1738	0.1757
	(0.0290)	(0.0279)	(0.0298)	(0.0278)	(0.1248)	(0.1109)	(0.1239)	(0.1197)
Market Access 1300, Cost 1	-0.0302**				-0.1070*			
	(0.0136)				(0.0593)			
Market Access 1300, Cost 2		-0.0474***				-0.1612**		
		(0.0135)				(0.0629)		
Market Access 1300, Cost 3			-0.0393***				-0.1425**	
			(0.0129)				(0.0574)	
Market Access 1300, Cost 4				-0.0326**				-0.0828
				(0.0146)				(0.0806)
Fit statistics								
R^2	0.27814	0.33541	0.33889	0.23634	0.34706	0.38433	0.39415	0.31601
F-test	4.0921	5.1602	4.8544	4.2032	5.1317	5.6931	5.6165	5.0493
Observations	169	169	169	169	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and a trade elasticity of 3.8. See Appendix for further details. Cereal 25k is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 22: First Stage, Trade Elasticity

Dependent Variable:	Log Population 1400		
Model:	(1)	(2)	(3)
Variables			
Mortality Rate	-0.0137***	-0.0139***	-0.0138***
	(0.0043)	(0.0042)	(0.0042)
Presence of Bishopric	0.5714***	0.5553***	0.5417***
	(0.1736)	(0.1626)	(0.1573)
Presence of University	0.7990***	0.7871***	0.7782***
	(0.2463)	(0.2446)	(0.2493)
Cereal Suitability 25km	-0.1506	-0.1494	-0.1506
	(0.0982)	(0.0978)	(0.0963)
Market Access 1300, TE 3.8	0.0216		
	(0.0266)		
Market Access 1300, TE 2		0.1118	
		(0.0962)	
Market Access 1300, TE 1			0.4204
			(0.4130)
Fit statistics			
R^2	0.18145	0.18748	0.18846
F-test	7.2264	7.5219	7.5707
Observations	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and trade elasticity iterations of 3.8, 2, and 1. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 23: Second Stage, Trade Elasticity

Dependent Variables:	Printing Press in 1500			Log Unique Editions in 1500		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Log Population 1400	0.3695**	0.3576**	0.3671**	1.770**	1.713**	1.746**
	(0.1548)	(0.1487)	(0.1571)	(0.7646)	(0.7008)	(0.7225)
Presence of Bishopric	0.0152	0.0322	0.0321	0.0568	0.1390	0.1539
	(0.0663)	(0.0607)	(0.0636)	(0.2733)	(0.2213)	(0.2160)
Presence of University	-0.0400	-0.0195	-0.0228	0.2162	0.3066	0.3040
	(0.1021)	(0.0920)	(0.0936)	(0.9020)	(0.8408)	(0.8280)
Cereal Suitability 25km	0.0503*	0.0444	0.0457	0.1949	0.1744	0.1811
	(0.0290)	(0.0277)	(0.0287)	(0.1248)	(0.1142)	(0.1157)
Market Access 1300, TE 3.8	-0.0302**			-0.1070*		
	(0.0136)			(0.0593)		
Market Access 1300, TE 2		-0.0975**			-0.4086**	
		(0.0430)			(0.1938)	
Market Access 1300, TE 1			-0.2932*			-1.361*
			(0.1610)			(0.7472)
Fit statistics						
Observations	169	169	169	169	169	169

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: Log Population 1400 is constructed as the natural log of the city population (in thousands) plus 1 to include three cities with population 0 in 1400. Printing Press in 1500 takes on a value of 1 if a USTC record appears in that city up to the year 1500, or 0 otherwise. Log Unique Editions in 1500 is the natural log of one plus the number of unique editions present in the USTC for a city up to the year 1500. Bishopric or University is a set of controls coded as 1 if a city contains a bishopric or archbishopric, or if the city contains a university, and 0 otherwise. Market Access in 1300 is calculated using our preferred set of travel costs and trade elasticity iterations of 3.8, 2, and 1. See Appendix for further details. Cereal Suitability is the average value of GAEZ rainfed cereal suitability for a 25km radius around each city.

Table 24: MA Shock, Travel Costs

Dependent Variables:	Printing Press in 1500				Log Unique Editions in 1500			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Change in Market Access 1300-1400, Cost 1	0.0783***				0.3545**			
	(0.0257)				(0.1721)			
Change in Market Access 1300-1400, Cost 2		0.0586**				0.2842*		
		(0.0230)				(0.1691)		
Change in Market Access 1300-1400, Cost 3			0.0611***				0.2959*	
			(0.0225)				(0.1682)	
Change in Market Access 1300-1400, Cost 4				0.0814***				0.3842**
				(0.0274)				(0.1736)
Market Access 1300, Cost 1	-0.0099				-0.0231			
	(0.0093)				(0.0342)			
Market Access 1300, Cost 2		-0.0194**				-0.0512		
		(0.0091)				(0.0338)		
Market Access 1300, Cost 3			-0.0186**				-0.0537*	
			(0.0080)				(0.0290)	
Market Access 1300, Cost 4				-0.0044				0.0083
				(0.0121)				(0.0472)
Presence of Bishopric	0.2393***	0.2311***	0.2315***	0.2408***	0.8840***	0.8595***	0.8603***	0.8945***
	(0.0638)	(0.0665)	(0.0667)	(0.0617)	(0.2442)	(0.2544)	(0.2554)	(0.2337)
Presence of University	0.2732***	0.2837***	0.2761***	0.2755***	1.870***	1.905***	1.882***	1.867***
	(0.0552)	(0.0517)	(0.0537)	(0.0544)	(0.2746)	(0.2780)	(0.2811)	(0.2667)
Cereal Suitability 25km	-0.0398***	-0.0405***	-0.0358**	-0.0447**	-0.1236	-0.1275	-0.1115	-0.1337
	(0.0152)	(0.0153)	(0.0143)	(0.0185)	(0.0943)	(0.0958)	(0.0925)	(0.1034)
Fit statistics								
R^2	0.12310	0.12841	0.13183	0.12049	0.14729	0.14766	0.15144	0.14860
F-test	13.028	13.672	14.091	12.713	16.030	16.077	16.562	16.197
Observations	470	470	470	470	470	470	470	470

Conley (240km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 25: MA Shock, Trade Elasticity

Dependent Variables:	Printing Press in 1500			Log Unique Editions in 1500		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Change in Market Access 1300-1400, TE 3.8	0.0783***			0.3545**		
	(0.0257)			(0.1721)		
Change in Market Access 1300-1400, TE 2		0.1767**			0.5983*	
		(0.0795)			(0.3400)	
Change in Market Access 1300-1400, TE 1			0.8863**			3.125**
			(0.3650)			(1.349)
Market Access 1300, TE 3.8	-0.0099			-0.0231		
	(0.0093)			(0.0342)		
Market Access 1300, TE 2		-0.0269			-0.0736	
		(0.0271)			(0.0954)	
Market Access 1300, TE 1			-0.0060			0.0115
			(0.1045)			(0.3522)
Presence of Bishopric	0.2393***	0.2438***	0.2567***	0.8840***	0.8876***	0.9335***
	(0.0638)	(0.0644)	(0.0609)	(0.2442)	(0.2488)	(0.2418)
Presence of University	0.2732***	0.2909***	0.3042***	1.870***	1.932***	1.982***
	(0.0552)	(0.0553)	(0.0551)	(0.2746)	(0.3011)	(0.2952)
Cereal Suitability 25km	-0.0398***	-0.0361*	-0.0308	-0.1236	-0.1210	-0.0982
	(0.0152)	(0.0191)	(0.0244)	(0.0943)	(0.1044)	(0.1175)
Fit statistics						
R^2	0.12310	0.11889	0.12328	0.14729	0.13692	0.14216
F-test	13.028	12.522	13.049	16.030	14.723	15.378
Observations	470	470	470	470	470	470

Conley (240km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 26: Horserace Regressions

Dependent Variables:	Printing Press in 1500		Log Unique Editions in 1500	
Model:	(1)	(2)	(3)	(4)
Variables				
Mortality Rate	-0.0051***	-0.0017	-0.0242***	-0.0076
	(0.0019)	(0.0019)	(0.0082)	(0.0098)
Log Population 1400		0.2452***		1.213***
		(0.0386)		(0.2077)
Presence of Bishopric	0.2263**	0.0862	1.068***	0.3752**
	(0.0907)	(0.0561)	(0.3790)	(0.1781)
Presence of University	0.2553***	0.0593	1.630**	0.6614
	(0.0690)	(0.0557)	(0.6533)	(0.5586)
Market Access 1300	-0.0222	-0.0275**	-0.0688	-0.0950*
	(0.0179)	(0.0131)	(0.0666)	(0.0513)
Cereal Suitability 25km	-0.0053	0.0316	-0.0717	0.1110
	(0.0452)	(0.0309)	(0.2440)	(0.1778)
Fit statistics				
Observations	169	169	169	169
R^2	0.11153	0.33575	0.13600	0.40352
Adjusted R^2	0.08427	0.31115	0.10950	0.38143

Conley (260km) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

6.2 Control Variables and Other Variables

Market Access. Market access for city i in 1300 is defined as $MA_i = \Sigma_j \frac{P_j}{D_{ij}^{\sigma}}$, with P_j being the population of the other 470 - 1 = 469 cities $j \neq i$ in 1300 for which populations are listed in Bairoch, D_{ij} the travel time between city i and city j, and σ is a trade elasticity (we use 3.8 in our baseline regressions and check for robustness using 1 or 2). We use the gdistance package in R to compute the least cost travel paths between cities via four transportation modes — by sea, by river, by road and by walking (portage). We use four different sets of travel costs, normalizing the speed of portage to 1. Our main parameterization, Cost 1, is from Bairoch (1988) which uses: porters = 1; roads = 0.81; rivers = 0.21; seas = 0.08. Three alternative parameterizations of travel costs from the literature are Cost 2 from Boerner and Severgnini (2014) which uses: porters = 1; roads = 0.50; rivers = 0.50; seas = 0.13; Cost 3 from Masschaele (2008) which uses: porters = 1; roads = 0.81; rivers = 0.51; seas = 0.10; and Cost 4 from Galloway et al. (1996) which uses: porters = 1; roads = 0.81; rivers = 0.59; seas = 0.06.

More on the gdistance package is available for download at this link: `https://cran.r-project.org/web/packages/gdistance/vignettes/gdistance1.pdf`.

Average Temperature 1500—1600. We use temperature data from Luterbacher et al. (2004). They reconstruct seasonal European temperatures (celsius degrees) since 1500 using proxy data from ice cores, tree rings, and written records. The data cover 0.5° x 0.5° grids which is approximately 50km x 50km at European latitudes. The data extend from 25° W to 40° E and 35° N to 70° N which includes all of the cities in our full sample. We extract the growing season (summer) temperature for each of our cities during the 16th century as this is the closest century to the Black Death period for which we have data. No comparable data exist for earlier centuries.

Elevation. City elevation data come from Jarvis et al. (2008) which is available at `http://srtm.csi.cgiar.org`. These data report elevation in meters with a spatial resolution between 1 and 3 arc-seconds. Where there are missing data we have supplemented it using Wikipedia.

Cereal Suitability. Our soil suitability data are from the FAO Global Agro-Ecological Zones (GAEZ) dataset as described in Fischer et al. (2002). We use these in preference to Ramankutty et al. (2002) as the latter does not have full coverage for all of western Europe. We use the GAEZ's cereal suitability data assuming low inputs and rain-fed irrigation. We extract the average soil suitability within 25 km radius circles around each city. Overall, cereal suitability is scaled from 1-9 where 1 is best, 8 is unsuitable and 9 is water (seas and oceans are treated as missing values).

Potato Suitability. The potato suitablility numbers are constructed using the Global Agro-Ecological Zones (GAEZ) data. We specifically use the data on white potatoes grown under conditions using low inputs and rain-fed irrigation for the baseline period 1961-1990. The raster file for the data along with support documentation are available for download from: `http://www.fao.org/nr/gaez/newsevents/detail/en/c/141573/`. These data are constructed in two stages. First the Food and Agriculture Organization (FAO) compiles information on the nutrients, soil, irrigation, and climatic conditions under which the potato grows best. Then the FAO compiles data on the physical environment for the entire world at a resolution of 5 arc minutes x 5 arc minutes ($\approx$ 10 x 10 Km). These characteristics include soil type, slope, average water availability, humidity, temperature, wind speed, etc. Then these two types of data are combined in order to create a value for "potential suitability for potato cultivation" for each raster cell. These values run from 1 to 9, where 1 is most suitable, 8 is least suitable, and 9 is water (or impossible to cultivate). See Monteduro (2012) for more details on the construction of the suitability raster. We use the GAEZ data to construct our city-level measures of potato suitability by extracting the average value of the raster cells within a 10 km radius of each city.

Pastoral Suitability. We control for the potential suitability of a region surrounding a city for pastoral farming with a variable measuring grazing suitability. This variable come from Erb et al. (2007) who create land use measures at a resolution of 5 arc minute cells ($\approx$ 10 km X 10 km). They record how land is used in each cell in 2000. The five categories they code for are: cropland, grazing, forest, urban, and areas without land use. Their grazing category is calculated as a residual after accounting for the percentage of area taken up by the other four uses. As part of this analysis they also generate a variable measuring the suitability of each cell for grazing (as opposed to actual present-day use). The suitability measure is created by first separating grazing land into three categories based on cover: 'high suitability of cultivated and managed areas, medium suitability of grazing land found under tree cover, and low suitability if shrub cover or sparse vegetation is detected in remote sensing' (Erb et al., 2007, 199). They then further subdivide the first two of these categories into areas with a net primary productivity of Carbon per meter squared is greater than 200 grams and those in which it is less than 200 grams. This results in five categories which they regroup into four categories with 1 = most suitable and 4 = least suitable. There is a fifth category which is 'no grazing' which we re-code as 5. We then create a dummy equal to 1 if the cell is most or moderately suitable. Finally, we extract the average suitability of the region around a city for grazing using circles of 10 km's.

Coasts and Rivers. We create variables to measure distances to the coast and major rivers using ArcGIS. We base these distances on the 1300 shape file downloaded from Nussli (2011). We then create two dummies for whether each city is within 10 km from the coast or a river.

Roman Roads. Data on Roman roads is provided by the *Digital Atlas of Roman and Medieval Civilizations*. We use this shape file to create two distances: (1) distance to all Roman roads and (2) distance to 'major' Roman roads. Since major settlements often formed along intersections of the road network, we also create variables for distances to Roman road intersection. We then create four dummies if the city is within 10 from any Roman road, a major Roman road, any Roman road intersection, or a major Roman road intersection.

Medieval Trade Routes. We use Shepherd (1923) to create a map of major medieval land trade routes. We create a GIS file that allows us to measure the distance to major medieval land trade routes or the intersection of two of them. We then create dummy variables that take the value of 1 if a city is within 10 kilometers of a trade route or an intersection of two of them.

Market Fairs. We obtain data on the location of important medieval fairs from two sources. The main source is Shepherd (1923). The second source we use is the *Digital Atlas of Roman and Medieval Civilizations*. The original source for this information is: Ditchburn, David and MacLean, Simon (eds.) 2007, Atlas of Medieval Europe, 2nd edn, London and New York, p. 158. We drop fairs that they cannot be matched with cities in the Bairoch dataset.

Hanseatic League. We document whether or not a city was a member of the Hanseatic League. We do this by matching where possible the city data with available lists of cities which belonged to the League. We include only cities which were members of the League and do not include cities with Hansa trading posts or communities. The source we use is from Dollinger (1970) and is the most comprehensive list of Hanseatic cities available. Unfortunately, Dollinger does not provide details on when each city became a member of the Hanseatic league. However Wikipedia provides information on a subset of Hanseatic cities. Using this data, we estimate that approximately 75% of these cities were likely members of the league prior to the time of the Black Death, thus giving us confidence that our Hansa dummy mostly captures pre-plague conditions.[32]

Aqueducts. We use GIS to create a shape file for whether or not a town was within 10 km from a Roman aqueduct using the map provided by Talbert (2000) as well as information from two Wikipedia webpages: `https://en.wikipedia.org/wiki/List_of_aqueducts_in_the_Roman_Empire` and `https://fr.wikipedia.org/wiki/Liste_des_aqueducs_romains`.

Medieval Universities. Bosker et al. (2013) provides data on the presence of medieval universities for European cities with populations greater than 10,000 (at some point between 800 and 1800). We consulted Wikipedia and other sources to find evidence of medieval universities with smaller populations. There are five medieval universities missing from the list in Bosker et al. (2013): Angers, Greifswald, Ingolstadt, Tuebingen, and Uppsala. However, as none of these were established prior to the Black Death we do not include them in our analysis.

Bishoprics and Archbishoprics. Our data on which cities have a bishopric or archbishopric come from Bosker et al. (2013).

State Capital in 1300. We use the data provided by Bosker et al. (2013) who collect data on capital cities from McEvedy and Jones (1978).

Representative Body in 1300. Bosker et al. (2013) provide information on the existence of communes for a subset of the cities in the Bairoch dataset. Bosker et al. (2013) create a variable"commune" that takes a value of 1 if there is indication of the presence of a local urban participative organization that decided on local urban affairs. Stasavage (2014) provides data on 169 cities that were autonomous at some point between 1000 and 1800. We use the variable for 1300-1400. Stasavage (2014) defines autonomous cities in the following terms:

> 'I have defined an "autonomous city" as being one in which there is clear evidence that such institutions of self-governance existed, and in addition there is also clear evidence of exercise of prerogatives in at least one of the policy areas referred to above. In the absence of such evidence the default is to code a city as non-autonomous (6).'

As Stasavage (2014) notes, his definition of city autonomy is stricter than the definition of commune used by Bosker et al. (2013). We create a dummy equal to one if the city is a commune in the Bosker et al. (2013) data set or a self-governing city according to Stasavage (2014).

Parliamentary Activity and Distance to Parliament 1300—1400. Our data on parliamentary activity is from van Zanden et al. (2012). This measures the number of times that Parliaments met at a regional level in 1300–1400. We create a dummy variable based on whether or not a town is in a region/country which had above the median number of parliamentary meetings. We also obtain a list of whether the parliaments were held for each region/country. We then use GIS to compute for each city the minimal Euclidean distance to a parliament.

Battles. As our main source we use Wikipedia's list of all battles that took place between 1300 and 1600. `https://en.wikipedia.org/wiki/List_of_battles_1301-1800`. This is a highly reliable source for the most important battles of the period. We are not concerned about sample selection

[32]Data available on request.

here as Wikipedia's coverage of European history is extensive; battles not listed on Wikipedia are likely to have been extremely small. For each battle we assign geo-coordinates based on either the location of the battle or the location of the nearest city mentioned in the entry. Note that we exclude naval battles.

Appendix References

Bairoch, Paul (1988), *Cites and Economic Development, from the dawn of history to the present*, University of Chicago Press, Chicago. translated by Christopher Braider.

Bosker, Maarten, Eltjo Buringh and Jan Luiten van Zanden (2013), 'From Baghdad to London: unravelling urban development in Europe and the Arab world 800-1800', *Review of Economics and Statistics* **95**(4), 1418–1437.

Dollinger, Philippe (1970), *The German Hansa*, MacMillan, London. translated and edited by D.S. Ault and S.H. Steinberg.

Erb, Karl-Heinz, Veronika Gaube, Fridolin Krausmann, Chritoph Plutzar, Alberte Bondeau and Helmut Haberl (2007), 'A comprehensive global 5 min resolution land-use data set for the year 2000 consistent with national census data', *Journal of Land Use Science* **2**(3), 191–224.

Fischer, Gunther, Harrij van Nelthuizen, Mahendra Shah and Freddy Nachtergaele (2002), *Global Agro-Ecological Assessment for Agriculture in the 21st Century: Methodology and Results*, Food and Agriculture Organization of the United Nations, Rome.

Jarvis, A., H.I. Reuter, A. Nelson and E. Guevara (2008), Hole-filled seamless srtm data v4. International Centre for Tropical Agriculture (CIAT).

Kirsten, Ernst, Ernst Wolfgang Buchholz and Wolfgang Kollmann (1965), *Raum und Bevolkerung in der Weltgeschichte: Bevolkerungs-Ploetz*, Verlag Herder.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean and H. Wanner (2004), 'European seasonal and annual temperature variability, trends, and extremes since 1500', *Science* **303**, 1499–1503.

McEvedy, Colin and Richard Jones (1978), *Atlas of World Population History*, Facts of File, New York.

Mengel, David C (2011), 'A plague on bohemia? mapping the black death', *Past and Present* **211**(1), 3–34.

Monteduro, Patrizia (2012), Agro-climatically attainable yield for low input level rain-fed white potato for baseline period 1961-1990, Technical report, FAO/IIASA.

Nussli, Christos (2011). URL: http://www.euratlas.com/about.html.

Ramankutty, N.J., J. Foley, J. Norman and K. McSweeney (2002), 'The global distribution of cultivable lands: Current patters and sensitivity to possible climate change', *Global Ecology and Biogeography* **11**, 377–392.

Shepherd, William R. (1923), *Historical Atlas*, Henry Holt and Company, New York.

Stasavage, David (2014), ‘Was Weber right? the role of urban autonomy in Europe’s rise’, *American Political Science Review* **108**, 337–354.

Talbert, Richard J. A, ed. (2000), *Barrington Atlas of the Greek and Roman World*, Princeton University Press, Princeton.

Zanden, Jan Luiten Van, Eltjo Buringh and Maarten Bosker (2012), ‘The rise and decline of European parliaments, 1188–1789’, *Economic History Review* **65**(3), 835–861.