Creative Destruction in the European State System: 1000-1850

By

David Schönholzer (Stockholm University)
Eric Weese (The University of Tokyo)

February 2020

CREPE DISCUSSION PAPER NO. 64

CENTER FOR RESEARCH AND EDUCATION FOR POLICY EVALUATION (CREPE)
THE UNIVERSITY OF TOKYO
http://www.crepe.e.u-tokyo.ac.jp/

Creative Destruction in the European State System: 1000-1850

David Schönholzer Eric Weese*

December 21, 2019

Abstract

Using newly available data on the universe of boundary changes for all European states over 1000-1850, we argue that competition between states leads to short-term losses and long-term gains in economic growth. In event studies, cities switching between states suffer large transitory losses in population but enjoy sustained population increases under new governance. We then use decomposition techniques to show that improvements in state quality occur both due to improvements of the pool of states over time as well as due to cities gravitating towards higher quality states. Parliamentary activity, fiscal capacity, and protection from predation mediate these effects.

*David Schönholzer (corresponding author), Institute for International Economic Studies, Stockholm University, david.schonholzer@iies.su.se. Eric Weese, Institute of Social Science, Tokyo University, weese@iss.u-tokyo.ac.jp. We are grateful to Davide Cantoni, Alexandre Debs, Mark Dincecco, Mitch Downey, James Fenske, Fred Finan, Walker Hanlon, Philip T. Hoffman, Mark Koyama, Arash Nekoei, Gerard Padró i Miquel, Torsten Persson, Gérard Roland, Jean-Laurent Rosenthal, Jón Steinsson, Hans-Joachim Voth, David Weil, Noam Yuchtman, and Fabrizio Zilibotti. We also thank seminar participants at Caltech, UC Berkeley, Yale, Brown, Tokyo, and Stockholm for helpful comments and discussion. We thank Frank Reed from Clockwork Mapping for preparing the research edition of the *Centennia Historical Atlas*, and Nico Voigtländer and Hans-Joachim Voth for graciously sharing their updated version of the Bairoch data. Finally, we thank Brad DeLong and Jan De Vries for helpful comments and discussions regarding the *Centennia Historical Atlas*. This paper supersedes an earlier draft titled "State Power and Urban Growth: Evidence from the Universe of Boundary Changes in Europe 1000-1850".

1 Introduction

A longstanding tradition in economics and the social sciences places competition between states at the center of Europe's long-run economic and political ascendancy. One popular view is that state competition spurred economic growth by increasing the quality of states, offering a more favorable environment for economic exchange and technological progress (North and Thomas, 1973; Jones, 1981; Hall, 1986; Tilly, 1990; Diamond, 1997; Kennedy, 2010; Acemoglu and Robinson, 2012). In contrast, a newer literature stresses alternative channels of growth such as demographic forces, urbanization, and capital investment (Rosenthal and Wong, 2011; Voigtländer and Voth, 2013a,b; Dincecco and Onorato, 2016).

Evaluating these two alternative explanations of Europe's rise is difficult for at least two reasons. First, it is challenging to separate the effect of state quality from the effect of conflict: competition engenders both the immediate cost of conflict as well as the potential for better governance. Second, even if these effects could be disentangled, data on the quality of historical states is scarce, particularly for the many states that were pushed out of the market for governance and for whom data collection may be nearly impossible.

In this paper, we overcome these obstacles by focusing on an important margin of state competition: territorial changes of competing states. To this end, we use newly available data assembled on the universe of boundary changes of all European states over the years 1000-1850 (Reed, 2016) combined with city population data by Bairoch et al. (1988). We disentangle the costs and benefits of competition by distinguishing cities that experience short-lived switching between states from more permanent changes in governance. We then approach the missing data problem by treating state quality as a latent variable that is to be estimated. Hence, we borrow from the literature in labor economics to estimate the effect of unobserved state quality on city population using panel decomposition methods developed by Abowd et al. (1999).

An immediate concern in estimating the effect of switching on city growth is that affected cities may be systematically different than others. We approach this problem using an event study design, restricting the comparison to cities that experience the same number of switches but at different times in their history. While this allows us to eliminate all unobserved fixed city characteristics, switching cities may still be on a systematic growth trend in the run-up to a switch. Our design allows us to directly evaluate whether this is the case. Finally, concurrent city-specific shocks may be causing both switching and growth. To deal with this concern, we show that results are similar using only on switches affecting many cities simultaneously, which are unlikely to be driven by city-specific shocks. While there are certainly instances in which changes in economic conditions led to changes in governance, our evidence is consistent with a view that changes in territorial control were mainly driven by non-economic motives such as political or religious rivalries (Hoffman, 2015).

Our analysis yields three main results. First, we find that higher-quality states confer large benefits over the long run: cities exposed to better governance become permanently larger. Second, state quality improves along two margins: better states seize control of cities from lower quality rivals, and new entrants tend to be of higher quality than those states exiting the market for governance. And third, we find that these benefits come at substantial costs in the short term, leading to population losses on the magnitude of 5% of urban population in Europe over our study period and setting back cities by almost 40 years relative to their average growth.

We interpret these results as a tradeoff between costs and benefits of state competition in Europe through a model of two states engaged in a process of creative destruction. A challenger state may compete with an incumbent state for governance of a city, and the city bears the cost of war if the challenger succeeds in its conquest. We show that the city benefits from state competition only if the expected quality of a victorious challenger relative to the incumbent overcompensates for the cost of war. Thus, in our case of creative destruction, rather than workers rendered unemployed and capital left obsolete as the costs of progress, cities are pillaged and rulers deposed.

Our model illustrates how a state system may suffer from too little or too much state competition. If there is little entry of high-quality challengers, cities may be spared major losses from conflict, but the states in a region may ossify, leading to low-quality states in power.[1] At the other extreme, if there is frequent entry of low-quality challengers, cities pay a high price from frequent changes in governance with little increase in the quality of governance, as argued by Cosandey (1997).[2] Thus, our model accommodates both views of history: that competition drives growth through improved governance, and that excessive

[1]Diamond (1997) famously made a similar argument comparing the geography of China and Europe, and Landes (2006, p. 8) echoed this sentiment by arguing that "fragmentation and national rivalries compelled European rulers to pay heed to their subjects".

[2]These costs may be so large that they dwarf the benefits from improving states, in line with the view of Rosenthal and Wong (2011, p. 229): they "view European political competition less as the source of economic virtue and more as a vice that reduced the possibility of economic growth".

competition inhibits growth through conflict. We apply the estimating equation derived from the model directly to the data to establish which view has more empirical support.

We begin our investigation into the costs and benefits of state competition by estimating the dynamic response of city populations to switching between states. To separate the costs from the benefits, we distinguish cities that experience switching but ultimately end up with the same state in power as before from those that end up in a new state. If a city reverts back to the original state by the end of the period, it experiences a large, transitory loss to its population of around 11%. But cities respond very differently if their governance changes more permanently: in this case, they undergo an immediate and sustained population increase of a similar magnitude.

To directly compare the costs and benefits of switching, we borrow tools from the literature in labor economics originating from Abowd et al. (1999) on the contributions of unobserved worker and firm characteristics. Specifically, due to the fact that cities are governed by multiple states of varying territory over their lifetime, we can separately identify unobserved city, period, and state characteristics, as well as the costs of switching between states. Using this decomposition, we find that in the short run the costs of war outweigh the benefits of switching to another state but in the long run the quality of states explains about five times as much variation in city population as does switching between states. We can also show that the short-term benefits of switching can be attributed in about equal parts to a state quality effect, a match effect, and a pure change-in-governance effect.

Next, we provide direct evidence for the key proposition that better states may take over cities from worse states by investigating the average state effects estimated in the panel decomposition. Specifically, we show that cities experiencing changes in governance typically end up in a state with a higher average estimated state effect. We quantify the average increase in the quality of states due to switching to be around 9% (in terms of their effect on city growth), half of which is due to a pool of states of increasing quality, and half due to better states taking over cities from worse states in a given pool of states. We also estimate that the returns to changing governance are decreasing, above and beyond the costs of war.

Finally, we explore the mechanisms underlying the growth effects of governance by correlating the estimated state effects with various state characteristics. We show that our estimates of unobserved state characteristics are correlated with parliamentary activity (Van Zanden et al., 2012) as well as fiscal capacity (Dincecco, 2011; Stasavage, 2011) in a sample of historical states. We then present a series of case studies to illustrate how these mechanisms interact with the stability of competitive states: higher-quality economic governance only has a lasting effect on city growth when it is sustained by stable rule.

Our work contributes to the large literature on the causes of Europe's ascent.[3] Although centuries of conflict came with short-run costs, high-quality states won more often than they lost, which was good for long-run growth. While our explanation supports institutional arguments for Europe's success, we show that institutional change came at a price, and in some periods slower institutional change with a lower burden of violent conflict may have been better for European economic development. Our view of institutions is expansive, allowing for features ensuring stability and security to be relevant besides inclusivity. Thus, our evidence on the magnitude of destruction brought on through state competition suggests that political stability itself may have been an important force for economic growth in Europe and beyond.[4]

Our findings also add to the literature studying the long-run impact of state capacity on economic development (Cox, 2017). Dell (2010) documents how pre-colonial institutions affected land tenure and public goods provision in Peru. Michalopoulos and Papaioannou (2014) find no effect of externally imposed boundaries on economic development of African regions, but they find positive effects for pre-colonial states (Michalopoulos and Papaioannou, 2013). Bockstette et al. (2002) document how the age of the state correlates with modern economic development. Dell et al. (2017) show how a strong, centralized historical state contributes to higher living standards. Together, these results make the case for the potential long-run benefits of higher higher quality states. We show how selection pressure may have resulted in such states proliferating in Europe, and confirm the importance of historical state boundaries for economic outcomes in a dynamic setting.[5]

Finally, our work also connects to the literature on state capacity and conflict (Mann,

[3] Broadly, the rise of Europe has been attested to at least four categories of causes: culture (Landes, 2006; Clark, 2008), geography (Diamond, 1997; Nunn and Qian, 2011), technology (McNeill, 1982; Mokyr, 1992), and institutions (North and Thomas, 1973; Acemoglu et al., 2005).

[4] Our results extend work by De Long and Shleifer (1993), who find that regions ruled by states with stronger property rights experience higher growth. In particular, De Long and Shleifer (1993, p. 693, Table 5) find higher growth in a small sample of European countries when governed by nonabsolutist instead of absolutist regimes. In addition to providing more systematic evidence for this phenomenon, we show that regions actually switch to states that offer better conditions for economic growth, and show that in the long run the benefits of switching outweigh the short run adjustment costs.

[5] In related ongoing work, Cervellati et al. (2018) explore the issues of sovereignty and territorial control of European states on economic growth using similar data. Given our focus on political competition and its implications for the quality of the state, our work offers a complementary explanation for the role of the state in European city growth.

1984; Tilly, 1990; Besley and Persson, 2010; Gennaioli and Voth, 2015). Protection against external threats is a key driver of the development of state capacity and lays the foundation of economic development (Queralt, 2018). We provide evidence for this link from security to economic growth. Relatedly, Dal Bó et al. (2018) show theoretically how states can achieve both security and prosperity, a link that we establish empirically for Europe over the second millennium.

The rest of the paper is organized as follows. In Section 2, we present the model of creative destruction through state competition. Section 3 presents the newly available data on the universe of boundaries of all states over 1000-1850 by Reed (2016) and summarizes the extent of switching between states across European cities. In Section 4, we then conduct our event study analysis of city populations around the timing of switching between states. Having established robust systematic responses to switching, we then decompose the contributions of switching and the quality of the state and show how cities gravitate towards better states in Section 5. We demonstrate that estimated state effects are correlated with parliamentary activity and fiscal capacity and offer some historical examples in Section 6. Finally, Section 7 concludes.

2 Theoretical Model

Consider a city whose size is given by $\log Y = \alpha + \psi_j + S\beta$, where α is the city's baseline size and ψ_j is the quality of state j in control of the city. An incumbent state $j = I$ provides public goods, guarantees internal peace, and sets rules for economic exchange. A challenger state $j = C$ with quality drawn from $F(\cdot)$ may take over the city through conquest and expose the city to its institutions instead. We take no stand on whether the Challenger emerges due to a secession from an existing state, a merger with another state, or a challenge by a rival state – all of these may pose a challenge to the Incumbent.[6]

Since our primary interest is in the effect of state competition, we also do not specify the motivation of the Challenger to take over the city, which could be driven by animosity between rulers, for example over religious matters, or by the desire to rule the city. We simply assume that whether the Challenger succeeds in conquering the city is a Bernoulli random variable S with mean $p(\psi_C, \psi_I)$ for a given draw of ψ_C, which places no restrictions on

[6]Our model can involve more than one city; however, it considers competition between states on the intensive margin only. If the Challenger takes over city a, this has no effect on the disposition of city b, and without efficiencies of scale, there is no meaningful exit decision.

the correlation structure between the respective state qualities and the Challenger's success probability. If the Challenger succeeds in the conquest, the city suffers a proportional loss of population $\beta < 0$ due to destruction or looting in the course of the change in governance.[7]

Given these assumptions, the expected city size is

$$\begin{aligned} E[\log Y] &= \alpha + E[\psi_j] + p(\psi_C, \psi_I)\beta \\ &= \alpha + \psi_I + \{(E[\psi_C|S=1] - \psi_I) + \beta\}\, p(\psi_C, \psi_I) \end{aligned} \tag{1}$$

where $E[\psi_C|S=1]$ is the average quality of the Challenger conditional on its victory against the Incumbent.[8] Equation (1) summarizes the conditions through which state competition trades off potential increases in the quality of states against the costs of war. The first two terms, $\alpha + \psi_I$, set the baseline city size in the absence of any competition from other states (that is, when $p(\psi_C, \psi_I) = 0$).

The effect of competition is then captured in two components: first, the terms in curly brackets capture net benefits (or net costs) of state competition. Specifically, they consist of the potential improvement (or deterioration) of the quality of governance – the terms in parentheses within the curly brackets – reduced by the costs β of switching governance to the Challenger state. The potential improvement itself is the difference between the quality of the *victorious* Challenger and the Incumbent, which is positive if Challengers who emerge victorious are typically of higher quality than the Incumbent. The extent of positive (or negative) selection is a product both of the distribution of Challenger quality as well as the correlation structure of qualities with the success of the Challenger.[9]

[7] It is fairly clear that modern warfare is often associated with a high negative β, for example Stalingrad in World War II. Importantly, historical sources indicate that destruction also occurred in earlier conflicts: for example, see De Vries (1997) for a discussion of the sack of Antwerp in 1576 and the resulting reduction of population in that city. Concerning looting, while it has become disreputable in modern conflicts, it was an important benefit of conquest. Hoffman (2015, Ch. 2) argues that "leaders making decisions about war [...] stood to win a disproportionate share of the spoils from victory but avoided a full share of the costs." He offers the particular example of the plunder of parts of India by Nadir Shah (p. 148), which allowed for a three year tax holiday in Persia. Jackson (1999, p. 20) also discusses the valuable plunder obtained from conquest. An additional benefit of (attempted) conquest is that it avoids "the problem of a large inactive standing army" (Jackson, 1999, p. 240) and ensures that troops remain well trained. We restrict looting to the Challenger, consistent with historical behavior of states with regard to cities in their domain. (Jackson, 1999, 1999, p. 282) describes the norm of not plundering already controlled territory, in the context of the Delhi Sultanate.

[8] The second equality in equation (1) holds because we can express the quality of the governing state as $\psi_j = (1-S)\psi_I + S\psi_C$ with $E[\psi_j] = (1 - p(\psi_C, \psi_I))\psi_I + E[S\psi_C]$, and according to the Law of Iterated Expectations, $E[S\psi_C] = E[E[S\psi_C|S=s]] = E[\psi_C|S=1]p(\psi_C, \psi_I)$.

[9] For example, if the unconditional mean of Challenger quality is the same as the Incumbent's, there is

The second component determining the effect of state competition is the Challenger success function $p(\psi_C, \psi_I)$, warping any net benefits or costs of state competition. This component highlights the importance of two aspects of the nature of competition. First, challengers may be more or less likely to succeed in taking the city.[10] Second, the Challenger success function can be biased towards (or against) quality-improving states. In a strongly quality-biased environment, even small net benefits can be magnified by war fortunes skewed towards high-quality Challengers. Conversely, in an environment in which low-quality states have more success in matters of war, even large potential net benefits may rarely translate into economic growth. Thus, the model illustrates how there can be both too little or too much state competition: if potential net benefits are large but few (high-quality) challenges succeed, the economy would benefit from more (and better) challenges to the Incumbent. However, if the net benefits are small or even negative, lower entry and lower success for Challengers would be welfare improving.

Looking ahead to the empirics, note that the net benefits of state competition correspond exactly to the difference in the conditional means of the city population: $E[\log Y|S = 1] - E[\log Y|S = 0] = (E[\psi_C|S = 1] - \psi_I) + \beta$. Thus, we separately identify costs and benefits by looking at two types of city switches: first, controlling for state qualities, any period in which a city switches between states is informative about the cost parameter β, no matter who governs the city at the time of population measurement; and second, once we know β, restricting to periods in which the current governing state is different from the one at a previous instance of population measurement identifies the difference in their respective governing qualities.

Our model gives rise to the interpretation of state competition over territory as a form of creative destruction: better states may enter the market for governance and gradually push worse states out. While the cities undergoing a change in governance may suffer in the short term due to destruction and looting, they may benefit from improved governance in the long term. This is the key proposition we test in the empirical section.

In standard models of creative destruction, firms develop new technology over time to

still scope for improvement as long as there is positive selection. Conversely, if the unconditional mean of the Challenger is high enough, the victorious Challenger can still be better than the Incumbent as long as the extent of negative selection is not too large.

[10]Specifically, for any given set of state qualities mapping into values of $p(\psi_C, \psi_I)$ bounded away from zero and one, the Challenger success function may be shifted up or down by a constant. This reflects the extent of churn or instability in a state system: how easy it is for a Challenger to defeat an Incumbent, independently of their qualities.

gain an edge over their competitors (see Aghion et al. 2014 for an overview of the literature). The average technology used in the market may improve via three possible margins: firms with better technology may be more likely to enter the market, firms with worse technology may be more likely to leave the market, and, among firms already in the market, those firms with superior technology are likely to gain market share at the expense of their less effective rivals. All three of these forces may be at work in competition between states as well. Potential states may be more likely to enter the state system if they are of higher quality (e.g. a republic vs. absolute monarchy). Existing states may be more likely to exit if they are low quality. Finally, when states compete over territory, states with higher quality may gain territory at the expense of those states with lower quality.

3 Data and Historical Background

To empirically investigate the tradeoff between the costs of changing state ownership and the gains of improved governance, we combine two data sources: the *Centennia Historical Atlas* (henceforth: Atlas) by Reed (2016) and city population data by Bairoch et al. (1988). The former dataset is newly available and marks a significant improvement in the quality of European state boundary data, as detailed below. The latter dataset has been widely used and recently updated in Voigtländer and Voth (2013b), with the only addition from our side being a more precise geocoding of all cities, as illustrated in maps described further below. Cities are population agglomerations with at least one thousand inhabitants.[11]

3.1 The Centennia Atlas

The Atlas consists of state boundary data covering Europe, Western Asia, the Middle East and North Africa, although most changes take place in Europe. Instead of a single static map, the Atlas covers ten sets of boundaries each year for every year between AD 1000 and AD 2003, resulting in a total of 10,030 maps (that is, a map covers 5.2 weeks).

The data has been collected and processed over several years by Clockwork Mapping, a cartography business with a focus on historical mapping. It is the basis of a view-only

[11]It should be noted that there may be a censoring issue at the bottom of the city size distribution such that small cities may be underrepresented especially in the early periods of the Bairoch data. We approach this problem by looking at various alternative transformations of the city population data in our robustness checks.

dynamic map-based guide of the history of Europe and the Middle East, which is available at `http://www.clockwk.com/`. The goal of the Atlas is to depict *de facto* territorial control by states, as opposed to claims to territory. Most of the work consisted of assembling various historical maps according to the consensus of the historical cartography community, with some discretion if the consensus reflected claims rather than actual territorial control. In case there was no consensus, an attempt was made to provide a consistent judgement across regions and time with the goal of depicting "boots-on-the-ground" power.[12]

In the Atlas, a state is defined as an entity holding the best claim of *de facto* power over a given territory. In other words, it is the entity most likely to have the capacity to extract taxes, set rules, and hold the monopoly of violence in a territory. States have continuity over time as long as they don't either cease to exist or absorb territory of sufficient size and independence so that the old state identity is too restrictive. For example, while France continues to exist after absorbing Burgundy and Brittany, England is replaced by Britain after the union with Scotland.[13]

To ensure the data was of high quality, we consulted several historians with knowledge of various periods of European history and tested the boundaries shown in the Atlas against their historical timeline of individual places, with special attention to the accuracy of the state in control and the timing of boundary changes. In the two test regions, the Low Countries and the central Holy Roman Empire, territorial control and timing of changes were consistent with the historical narrative.[14] Of course, while the data may accurately depict which state had de facto power over a given territory, some states had a much stronger grip on the territory they claimed sovereignty over than others, and some boundaries were more porous than others.[15] Additionally, European states also took on increasingly large

[12]This information is based on personal communication with the head cartographer of Clockwork Mapping, Frank Reed.

[13]The use of *de facto* power as a criterion to identify states also leads to cases in which an entity may not have officially dissolved but its power has degraded so far that its constituent entities are classified as the state in power. For example, after the abdication of Charles V as the head of the Holy Roman Empire at the Peace of Augsburg in 1555, the remnants of the Empire are reclassified as Lesser Imperial States, a collection of semi-independent duchies, principalities, republics and cities, while other entities such as Bavaria, Bohemia or Switzerland have already achieved sufficient autonomy from the Holy Roman Empire at earlier stages.

[14]The historian Charles W. Ingrao specializing on Early Modern Europe independently evaluated the quality of the Atlas, concluding that he was "impressed by the developer's incredible eye for detail" (on Centennia website and confirmed in personal communication).

[15]This problem becomes most clear in an example discussed with Frank Reed of Centennia: while their assessment was that the medieval French king had very little de facto power over much of his territory, the consensus in historical cartography is to assign most subject territories to the French crown. In the absence

territorial possessions overseas, which we cannot observe.

Figure 1 shows an example of the Atlas. It can be seen that state boundaries have been coded in great detail, sometimes carving out individual cities with special status from larger bodies. Changes over time due to changes in *de facto* power may be due to treaty, conquest, royal inheritance, loss of control, or significant troop movements leading to occupation over months or years. While we cannot distinguish between changes driven by military conquest and contractual changes, the data provides a precise measure of the length of time a city is governed by each state.

3.2 Entry and Exit of States

The Atlas allows for a unified description of state activity over 1000-1850 in great detail. Figure 2 shows the timeline of existence of all 209 states holding at least one city (according to the Bairoch data) in a given year, organized from the oldest states at the top to the newest at the bottom of the figure. Looking at the initial cohort of states in the year 1000, there is substantial variation in their longevity: while some exist all the way to 1850, many others last only a few centuries or even decades. While survival is highly variable, the rate of entry is surprisingly constant: The entry margin from the index of the first entries in 1000 to the last entry in 1849 is nearly linear. The generally constant trend of state entry is perturbed by periods of lulls, such as between 1050 and 1150, often followed by bursts of entries. Major continent-wide conflicts leading to the intermittent or permanent exit of many states are also visible, for example during the Thirty Year War in the mid-17$^{\text{th}}$ century.

The number of states with *de facto* power over cities varies between 20 in the beginning of the Late Middle Ages to around 50 at the end (see figure A.1 in the appendix). The latter number is somewhat smaller than typically described for this period, for which some scholars identify hundreds of states (e.g. Tilly 1990). The reason for this is that the Atlas is more restrictive in assigning *de facto* power to a state. For example, while there existed dozens of duchies and principalities in Central Europe at the time, few of them were strong enough to act independently of the Holy Roman Empire.

---

of a meaningful faction of historical cartographers drawing the territories of local lords as independent states, Centennia decided to stick with the consensus opinion. If this type of measurement error – retrospective assignment of a successful state to territories it had little control over – were systematic, it would likely work against us since we would not be able to pick up the strengthening grip of the French king over his territories, decreasing the power to detect meaningful correlations between state quality and city growth.

3.3 State Competition over Cities

State competition over territory led to frequent changes in state ownership of cities, affecting substantial shares of cities and city populations. Figure 3 shows the annual number and share of cities switching from one state to another, as well as averages per decade for the periods defined by Bairoch et al. (1988). While many switches occur in spikes during major wars or periods of increased state exit and entry, most years see at least some cities switching. As the number of cities grows, so do the number of switches, leading to roughly constant shares of affected cities of around 10-20%, with a period of extreme instability in city ownership by states during the period 1700-1750, as the War of the Spanish Succession led to major territorial changes across the continent.[16] Figure A.2 in the appendix shows the same statistics for city populations. Switches affect around 10-15% of city dwellers. Around six million city dwellers are affected by 1800-1850.

There is substantial variation across space in the extent to which cities were exposed to switching. Figure 4 shows the average number of switches per decade each city experienced since its foundation up to 1850. Northern Italy and the Low Countries stand out as areas with intense switching activity, but so do individual cities with a particularly long history of upheaval, such as Warsaw, Naples, or Prague. Other than parts of France and Russia, cities in every region of Europe experienced at least some switching.

Variation in switching across space varies from period to period. Appendix Figure A.3 shows the spatial distribution of switching intensity across Europe during 1800-1850. Some regions of the continent saw no switching at all, such as the British Isles, western France, most of Russia, or Sicily; other regions experienced as much as one switch per decade, such as central Italy, the Low Countries, or northern Germany. While these were areas of intense switching in this period, these spatial patterns shift from period to period: for example, in the period 1700-1750 (Appendix Figure A.4), Sicily and Spain (especially Catalonia) experienced more than one switch per decade, while northern Germany was relatively calm.

Two examples illustrate the two types of switches that speak to the costs and benefits of competition. In the course of the Thirty Year War, Prague switched from the Habsburgs to Sweden and back to the Habsburgs within two years. In contrast, Nancy switched from Burgundy to France for decades. We use the former to identify the costs of switching (since

[16] It should be noted that these shares of affected cities are averages within a period across cities. This means that a city switching multiple times within a period leads to a higher share of average decadal switches.

the Habsburgs governed Prague both before and after), while the latter serves to identify the benefits of French governance over Burgundian governance. Both cities switched states during the 17$^{\text{th}}$ century, so both inform the costs associated with changes in governance; but only Nancy is informative about the benefits, since it allows us to compare Nancy's performance under two different states at the beginning and the end of the century.

How often do cities experience switching in any period over their lifetime, and among switching periods, how often do we observe another state at the end of the period than at the beginning? Figure 5 shows the distribution of the number of switches cities experienced over their entire lifetime (i.e. in periods with nonmissing population). As shown in the gray bars, out of the 2,182 cities in our data, only 419 (19%) did not see a change in state ownership at some point during their lifetime. 17% saw exactly one switch; 19% saw two; and 45% saw three or more.

However, importantly, cities often revert back to the same state as before within the periods defined by the timing of population measurements in Bairoch et al. (1988) (1100-1700 in centuries; 1750-1850 in half-centuries). This is what the transparent bars in Figure 5 show. 886 cities (41%) never experienced a sequence of switches that resulted in another state holding power by the end of the period. 23% experienced exactly one lasting switching period ending up in another state, and 36% experienced two or more. The distinction between intermittent switches with the same state in power by the end of the period and more permanent switches with another state in power by the end of the period is important for us to distinguish the costs and benefits of state competition, as described in detail further below.

We summarize further relevant statistics in Table 1. As expected, city size distributions are highly skewed in every period. Thus, we work mainly with the natural logarithm of city population size. On average, cities switch 7.7 times over their lifetime, so that switching occurs in 2.6 out of up to ten periods. Only half of those switching periods result in a different state in power (as opposed to the same state regaining power by the end). The average city is governed by 2.2 states over its lifetime, taking into account that the same state may be governing it at different times.

While we had 209 states with cities in the annual dataset, once we reduce it to the Bairoch periods (1100 to 1700 in centuries and 1750 to 1850 in fifty-year intervals), we are left with 140 states.[17] While many states control only a few cities, large states with over 400

[17] There are a total of 140 states with cities but due to missing observables we only show statistics for 129

cities appear as early as 1500. We can again see the relatively short lifespan of states: the average state entry year is 1418, and the average exit year is 1611; on average, a state tends to exist for only 3.3 periods with a median of two periods. Every period a state survives, it gains on average 4.1 cities either through city foundations or conquests. Conditional on surviving another period and gaining cities, the average number of cities gained is 13.2; and conditional on surviving and losing, the loss is on average 10.3 cities.

4 Event Study Analysis of Switching Cities

With this background on the extent of competition between states over cities, we now investigate how city population sizes are affected by switching, with two goals in mind. First, we aim to demonstrate that cities respond to switching in a systematic way that allows for a causal interpretation of the switching effect, both with respect to the costs of switching as well as the benefits of switching to another state. And second, in doing so, we provide evidence for the identification assumptions underlying the decomposition of the switching effect and the state effect, which we undertake in Section 5.

We now extend the notation from the model to our empirical design, which is constrained by the dimensions of the Bairoch city data. Thus, our 2,182 cities are indexed by i, and t indexes periods from 1100 to 1700 in centuries and from 1750 to 1850 in fifty-year intervals. y_{it} is log population of city i in year t, which may be missing if the city has not been founded yet or if the population is unknown.

The function $\mathbf{J}(i,t)$ maps city i in period t to the unique state holding power at the time of population measurement. City i is associated with a switch in t if boundaries have moved in such a way that i fell into the territory of another state at some point since $t-1$. Let S_{it} be the number of switches per decade in the period from $t-1$ to t. Correspondingly, $\mathbf{1}[S_{it} > 0]$ is an indicator for whether the city switched at least once in the period. We can also identify the subset of switching instances where $\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)$, that is, when the state governing a city in t is different from the state governing the city in the $t-1$.[18] Notice

in the summary statistics in Table 1.

[18]If a city underwent switching in the period, there are several ways to assign $\mathbf{J}(i,t)$ to a city. Our default option is to assign the state in power exactly at t because we expect cities to adjust fairly rapidly to the trend determined by the state in power, as suggested by the evidence in Davis and Weinstein (2002). Population measurements are also typically exactly from the change of the century or mid-century, as reported by (Bairoch et al., 1988, p. 290). For robustness, we show our main results with alternative assignments of $\mathbf{J}(i,t)$ in the appendix.

the subtle but important way in which $\mathbf{1}[S_{it} > 0]$ differs from $\mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$: clearly, at least one switch is required for the state in power to be different, but there may be many intermittent switches in a period with the state in power at the end being the same as the one in the beginning.

Since we restrict the estimation sample in the event studies to cities that switch an intermediate number of times, it is useful to know how cities differ by the extent of switching they are exposed to. To this end, we summarize statistics for groups of cities by the number of periods in which they experience any switching as well as switching to another state over their lifetime in Table 2. We can see that cities that never switch tend to be younger, smaller in 1850, closer to borders, and more peripheral to the center of urban activity in Europe than those that switch. Those cities that do switch include the modal city category undergoing two switching periods (and one switch to another state). These cities are about 65 years older and 64% larger in 1850 than those that do not switch. Cities having experienced more switches tend to be older, somewhat larger in 1850, and slightly closer to borders and the center of urban activity in Europe. Switching intensity is broadly similar among cities experiencing more switches over their lifetime at around three switches per century. Cities that switch more since their appearance in the data tend to switch slightly more often to another state, rising from 15% among cities switching one or two periods to 45% for cities switching in five periods or more.

4.1 Cost of Switching

To nonparametrically trace out the dynamic behavior of city populations around the timing of switching, we estimate an event study of the following form:

$$y_{it} = \alpha_i + \gamma_t + \psi_{\mathbf{J}(i,t)} + \sum_{s \in \{-2,0,1,2\}} \mathbf{1}[t = e_i^{\text{Switch}} + s]\lambda_s + \varepsilon_{it} \tag{2}$$

where the city effect α_i captures fixed, unobserved city characteristics; the year effect γ_t captures continent-wide, period-specific trends; and the state effect $\psi_{\mathbf{J}(i,t)}$ takes into account unobserved characteristics of states. Finally, e_i^{Switch} is the first period in which a city experienced switching (that is, the first time for a given i when $S_{it} > 0$).

The λ_s terms are our main coefficients of interest, representing the nonparametric dynamics of city population development around the timing of a switch. In particular, since we omit the coefficient for the period before a city switches (that is, we set $\lambda_{-1} = 0$), the coefficient λ_{-2} estimates the presence of a pre-trend. Because we are controlling for state quality (both the state before the switch, and the potentially but not necessarily different one after the switch) by including state effects $\psi_{\mathbf{J}(i,t)}$, we isolate the costs associated with switching. These costs are captured by λ_0, λ_1 and λ_2, allowing us to examine to what extent they are transitory or permanent. We bin the event times at the endpoints to include periods before the lower bound and after the upper bound, as is standard in the event study literature (see e.g. McCrary 2007).[19]

Given that not all cities switch, and most cities switch multiple times, we begin by restricting the estimation sample to cities that switch once or twice (36% of cities). For those cities that switch twice, we choose the first switch as the relevant event. We then gradually expand the set of included cities to those that switch three and four times (61% of cities), again choosing the first switch as the relevant event. Finally, to ensure results are robust to including other switches besides the first, we duplicate cities by the number of switching events they experience and run the event study relative to each of their switching periods.

Estimation results for cities switching once or twice are shown in Figure 6. Examining first the degree to which there is a pre-trend, we can see that city population two periods before switching is indistinguishable from population at our reference point one period before switching. This suggests there is no evidence of reverse causality: on average, switching cities are neither on an upward nor a downward trajectory. It should be noted that this is entirely consistent with our model, where switching is a function of characteristics of states, not cities.

Looking at the coefficient estimates for periods since the first switch, it appears that cities experience an immediate and substantial reduction in their population in the period of switching of around 22% relative to the period before the first switch. This implies large population losses due to switching: there are 1,194 switching events for cities experiencing one or two switching periods with an average population size of 7,321 in the period before switching, implying a total loss of city population of around 1.9 million people for these 778 cities over 1100-1850. However, these results appear to be transitory: one and two periods

[19] This means we abuse notation in equation (2) slightly. The indicator associated with λ_{-2} is $\mathbf{1}[t \leq e_i^{\text{Switch}} + 2]$ (notice the inequality instead of an equality), and similarly for the indicator associated with λ_2. This implies, for example, that λ_{-2} estimates the average population in two periods *or earlier* before the event.

after switching, population sizes are no longer significantly different from the period before switching. This is consistent with the finding by Davis and Weinstein (2002) that long-run city size is robust to large temporary shocks.

These findings are robust to changes in sample composition and switching event selection. Appendix Figure A.9 shows results for cities that experience one, two, or three switching periods. While there is again no pre-trend and a transitory loss of population, the drop is only about 11%, and population is actually significantly larger two periods after switching. This may be due to cities that experience more switches being more resilient to changing boundaries and being able to benefit cumulatively from changes in governance. The findings are also robust to using events other than the first switching period. Appendix Figure A.10 shows event studies for cities switching once or twice, or one to three times, with each city time series duplicated by the number of switching events experienced over its lifetime and assigned one of its possible switching events. Again, we see no evidence for a pre-trend and a transitory drop from switching, again with slightly smaller magnitudes.

Overall, these event study results for switches that may or may not result in another state holding power provide evidence for an economically significant loss of city populations due to territorial competition between states. Any benefits to cities from this competition have to be measured up against these losses incurred from switching. In terms of the model, these estimates provide evidence for the destruction associated with a transfer from one state to another.

4.2 Benefit of Switching to Another State

The preceding event studies allow us to estimate the causal effect of switching, irrespective of whether it led to a change of the state in power, allowing us to identify the costs associated with state competition. To estimate the causal effect of switching when it brings another state to power, opening up the possibility of benefiting from a higher quality state, we estimate the following event study:

$$y_{it} = \alpha_i + \gamma_t + \mathbf{1}[S_{it} > 0]\beta + \sum_{s \in \{-2,0,1,2\}} \mathbf{1}[t = e_i^{\text{AnotherState}} + s]\pi_s + \varepsilon_{it}. \tag{3}$$

with $e_i^{\text{AnotherState}}$ being the first period in which a city is governed by another state than the period before (that is, the first time for a given i when $\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)$).

This regression differs in three important ways from the one described in equation (2). First, we no longer control for state effects $\psi_{\mathbf{J}(i,t)}$ so we can pick up a potential improvement in outcomes due to having switched to another state. Second, we now control for switching, since we are interested in isolating only the benefits of being governed by another state. And third, our event of interest is the first time a state is governed by another state, as opposed to switches that possibly result in the same state holding power by the end of the period. Out of the 5,588 switching city-period observations in our data (which make up about half of our 10,576 city-period observations with positive populations), 2,917 (52%) lead to another state holding power. The π_s coefficients capture the dynamics around the timing of switching to another state, with π_{-2} again providing a direct test of our identification assumption that pre-trends are flat. The coefficients π_0, π_1 and π_2 trace out the average improvement experienced by a city by virtue of being in the state it switched to, as opposed to the state holding power before the switch.

We show results for this event study using cities that switch to another state once or twice over their lifetime in Figure 7. Just like in the event study for any switches, there is no evidence of cities growing or shrinking in the run-up to a switch to another state, as can be seen by the coefficient estimate for π_{-2}. But once a city switches to another state, we see a large, immediate, and lasting effect on city population on the order of about 21%. Interpreting this estimate again in terms of the overall city population increase across all cities that switched once or twice to another state, the implied total gain of city populations is 1.4 million (the average city size before switching to another state is 5,886, with 1,136 instances of switching to another state for these 823 cities).

We again undertake the same robustness exercises as before, showing the effect for cities switching to another state up to three times, both for the first event as well as for all events with duplicated city time series (see Appendix Figures A.11 and A.12). In each case, we find no pre-trends and sustained, large increases after switching, although the magnitude falls slightly as we include cities with higher numbers of switches to another state and switches other than the first.

4.3 Event Study Robustness: Mass Switches

While these event study results address concerns about static selection and diverging trends over the long run, they cannot account for the possibility that individual cities may experience idiosyncratic shocks that drive both switching and growth. To deal with this concern, we now restrict estimation to cities that switch as a group from the same state and/or to the same state. In this way, idiosyncratic city shocks are unlikely to drive switching, but rather are subject to larger geopolitical forces outside their control. Results are presented in Table 3. We can see that, when restricting to cities that switch in a group of at least ten or twenty cities, respectively, the average growth effect in subsequent periods is very similar to the results in Section 4.2. Thus, there is no indication that idiosyncratic shocks to cities drive these effects.[20]

Together, we interpret these results as causal evidence for creative destruction in the European state system: both transitory, negative effects as well as permanent, positive effects of state competition over territory on city populations in Europe. In this sense, it provides evidence for the validity of our model assumptions that there exist costs of switching as well as benefits of another state taking over. However, these benefits may arise from at least three sources: first, as posited in the model, there could be a *state quality effect*, in that the quality of the state taking over may be higher than the state relenting control; second, there could be a *match effect*, so that cities may have idiosyncratic affinities for states; and finally there could be a *change-in-governance effect*, which could arise from generic benefits from changing governance, by providing an opportunity to break up frictions in the city economy grounded in political rents tied to the old regime. We now turn to studying these components using panel decomposition techniques.

5 Panel Decomposition into Switching Cost and State Effect

In order to investigate the contributions of switching and the quality of the state on economic development in greater detail, we decompose the variation in European city size over the period 1100-1850 into two main components of interest: the penalty associated with a city switching due to changing boundaries and the fixed, unobserved characteristics of the state governing a city. In addition to these main components of interest, we continue to tease out fixed city characteristics α_i and continent-wide trends captured in period fixed effects γ_t. Specifically, and along the lines of the theoretical model, we estimate the following regression:

[20] We borrow this strategy from the mass layoffs literature in labor economics originating from Jacobson et al. (1993).

$$y_{it} = \alpha_i + \gamma_t + \psi_{\mathbf{J}(i,t)} + \mathbf{1}[S_{it} > 0]\beta + \varepsilon_{it} \tag{4}$$

where ε_{it} is an error term correlated arbitrarily over time within city i. Importantly, unlike most previous scholarship on European states, this decomposition allows us to separately identify the contribution of fixed geographic characteristics tied to a city from the contribution of the state in power. Intuitively, this is due to the fact that, as boundaries change, cities are governed by multiple states over their lifetime, allowing us to compare the growth of a city under different rulership. This approach originates from work by Abowd et al. (1999) (henceforth AKM) who study the contributions of unobserved worker and firm characteristics and has spurred a rich literature in labor economics (see e.g. Card et al. 2013).

We interpret β as the proportional cost associated with exposure of cities to state competition over territory. The magnitude of β captures the costs both from the uncertainty of the institutional environment potentially changing due to switching as well as the direct cost of warfare through destruction and looting. We interpret the state effect ψ_j as a proportional change in the achievable size for cities i governed by $j = \mathbf{J}(i,t)$ in t due to the institutional environment provided by j. This institutional environment may affect city size through a number of mechanisms, which we explore in Section 6.

To decompose the benefits of switching to another state into a state quality effect, a match effect, and a change-in-goverance effect, we also study variations of the regression model of the form:

$$y_{it} = \alpha_i + \gamma_t + \psi_{\mathbf{J}(i,t)} + \mathbf{1}[S_{it} > 0]\beta + \mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]\delta + \varepsilon_{it} \tag{5}$$

where δ captures any benefits to switching to another state not captured in the state effects.[21] We discuss this in detail in the following section.

Identification of parameters in models such as (4) and (5) has been extensively covered in AKM and the literature following their work. Card et al. (2013) provide an excellent description of sufficient conditions and sources of potential biases, which we briefly summarize here. Parameters are only identified for a connected set of cities and states, linked by switches of cities between states. Due to the long time horizon, 136 states out of 140 are connected

[21] We abuse notation slightly by continuing to denote the parameter associated with switching as β even though it captures a slightly different negative effect than the one specified in (4), as discussed in the following section.

in our data. We require that switching is orthogonal to the error term. Assuming that the error has mean zero for each city in the data, the key issue for identification is whether the error is orthogonal to the state effects. A sufficient condition for this to hold is that the assignment of cities to states obeys a strict exogeneity assumption with respect to the error.[22] To support this assumption, it is useful to show that (a) match effect models do not fit the data much better, which we show further below; (b) there is no systematic trend in the run-up to switches, as established in the event studies; (c) positive shocks are permanent while pre-trends are absent, which is inconsistent with standard proportional growth models of cities (Gabaix 1999, Eeckhout 2004) absent a sustained positive shock such as a higher state effect.

5.1 Basic Decomposition Results

Table 4 shows estimates based on equations (4) and (5). In specification (1), we find that switching is associated with an instantaneous drop in population of 8.2%. Compared to specifications with switching intensity (see Appendix Tables A.1 and A.2), this suggests some mild non-linearity in the effect, as shown in Appendix Figure A.5. The state effect $\psi_{\mathbf{J}(i,t)}$ included in specification (2) increases the switching cost moderately to 11.3%. This point estimate is very close to the median estimate of the switching cost estimated in the event studies in Section 4.1.

In the remaining specifications, we now include an indicator for whether the city switched to another state, as opposed to reverting back to the state in power in the last period, as in equation (5). The switching cost estimate remains unchanged at 11.3%. This cost is largely compensated by a countervailing positive effect of switching to another state, which is 7.4%. As shown further below in column (3), we can reject that the sum of the two coefficients is zero, meaning that the immediate net benefits of switching are negative.

In column (4), we include state effects. As a result, the cost estimate hardly changes while the switching benefits estimate drops from 7.4% to 3.9% and is now only marginally significant. Since we were not controlling for the state effect in specification (3), this means the lower switching benefit estimate in (4) must be due to the state effects absorbing about

[22] Specifically, we need that $P(\mathbf{J}(i,t) = j|\varepsilon_{it}) = P(\mathbf{J}(i,t) = j) = G_{jt}(\alpha_i, \psi_1, ..., \psi_J)$ for all i, t. This assumption is not violated, for example, by the general tendency for cities to move to higher ψ_j states, as we document below; it is also not violated by smaller or younger cities experiencing a different frequency of switching than older or larger ones. Switching frequencies may also be related to any state characteristics such as location or capacity for city conquest without violating this sufficient condition.

half of the instantaneous benefits from switching. We can strongly reject that the short-term benefits besides an increase in the quality of the state compensate for the cost of switching, as can be seen in the F-test in column (4).

Finally, in specification (5), we estimate a fully saturated version of equation (5), that is, we include a full set of city-by-state effects instead of city and state effects separately. The switching benefit falls slightly to 3.4%. Since we include effects for each city-state combination, this remaining small benefit of switching is the change-in-goverance benefit possibly due to the breakup of patronage networks tied to the former state in power, no matter the quality of the state or of the match.[23] Since the saturated model accounts for match effects, the small difference between the 3.9% estimate in (5) and the 3.4% estimate must be due to the match component between cities and states. We can also investigate the change in the R^2 and residual mean squared error (RSME) of this specification compared to (5). While the fully saturated model fits slightly better, the proportional drop in RMSE is as small as in other studies using AKM, suggesting that any violation of the strict exogeneity assumption is modest.

Since these decompositions are in the class of AKM-models, we can also study the decomposition of the city population variance into the variance accounted for by switching and by the state effects (as well as the city and period effects, which we are not directly interested in). In the sixth to the tenth row of the lower section of the table, we show standard deviations for the two variance components of interest. The standard deviation of the switching component, expressed in log population units, is between 5.6% and 6.0%, while the standard deviation of the estimated state effects is typically around five times higher. So while switching explains some of the variation in city population, the state in power is much more important.[24]

Overall, the results in Table 4 confirm the presence of a robust and economically meaningful cost of switching, while also providing evidence for the various components contributing to the benefits of switching. Comparing specifications (3) through (5), the quality of the

[23] An alternative interpretation of this effect is that is estimates the extent of endogeneity in cities switching from one state to another. Given that it is small and only marginally significant the potential bias is likely to be small.

[24] We report standard AKM estimates of variance components, as in Card et al. (2013), which are known to be biased. These are commonly used in the AKM literature and generally considered informative. Kline et al. (2018) develop a leave-one-out estimator that provides unbiased estimates of variance components. Fortunately, in our case, these estimates are almost identical to the original AKM estimates. For example, we estimate the standard deviation for state effects to be 25%, while using the leave-on-out estimator results in a standard deviation of 24%. Other variance components are also similar.

state accounts for 53% and the change-in-governance effect for about 46% of the overall instantaneous benefit of switching to another state, while the match effect is negligible.[25]

The switching cost estimates are robust to a number of alternative specifications, summarized in table A.3. We can control for the distance to the nearest border, the distance percentile to the centroid of the state, the distance to the continent-wide weighted center of urban activity, and the number of plague reports, none of which moves the coefficients of interest by much. We can also control for period-by-country fixed effects (using post-war country boundaries), allowing there to be different city population (non-parametric) trends in regions defined by current countries. We can use the state holding power in the majority of time instead at the time of population measurement, as well as assign shares of state ownership and control for all of these instead of state effects. We also get similar estimates if we estimate β in first differences. Finally, allowing for an effect on the extensive margin (city formation) by transforming population by $\log(1 + \text{population})$ (with missing values set to zero) also produces similar estimates. All estimates for the effect of any switching in a period are statistically significant and around 4.5-8.7%.

5.2 State Effect Improvement Through Switching

We now turn to studying the estimated state effects $\widehat{\psi}_j$ directly to illuminate the pattern of improvement in states that cities are exposed to over time, as posited by the model. The choice of specification to estimate the state effects is not important: the correlation between the estimated effects in specifications (3) and (5) (as well as most estimates from the robustness table) are on the order of 0.95 or higher.[26]

Equipped with a reasonable estimate of state quality, we can now proceed to investigate the key model proposition that cities may benefit from moving up the quality ladder of states as they switch from one state to the next. To explore this option, we investigate the changes in the estimated $\widehat{\psi}_j$ over the course of a city's sequence of governing states. Specifically, evidence consistent with the model would be if $\widehat{\psi}_{\mathbf{J}(i,t)}$ were on average larger than $\widehat{\psi}_{\mathbf{J}(i,t-1)}$ whenever city i transitioned from one state to another.

In Figure 8, we show that this seems to hold. The points in the figure are average state

[25] These shares are simply the results of the following comparisons of coefficient estimates: $(7.4 - 3.9)/7.4$, $3.4/7.4$, and $(3.9 - 3.4)/7.4$.

[26] One exception to this are specifications in which we change the dependent variable to first differences or to $\log(1 + \text{population})$. However, this is to be expected, given that a different aspect of state effects is estimated.

effects for different subsets of cities at each stage in their sequence of switches from one state to the next. We draw a separate sequence of averages for cities that have been controlled by a total of two, three, four, and five states. In each case, average state effects of subsequent states are higher than preceding states.

To quantify the increase in state effects switching cities are exposed to, we proceed as follows. We begin by simply computing $\widehat{\psi}_{\mathbf{J}(i,t)} - \widehat{\psi}_{\mathbf{J}(i,t-1)}$ whenever $\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)$. Taking the average across all city-period observations for whom this is true yields the average increase when switching to another state – in other words, the average slope of each of the segments connecting the average fixed effects at a given point of a city's sequence of governing states. We estimate that, on average, a city finds itself in a state with an 8.7% higher quality after a switch to another state. To test whether this increase is statistically significant – taking into account that two cities transitioning from one state to another will have the exact same change in state fixed effects – we estimate the clustered standard error of this average. We do this in two ways: first, we cluster on the pairs of states involved in the change, which yields a standard error estimate of 0.043; second, we cluster on states, which yields a standard error of 0.046. These estimates are statistically significant on a 5% and 10% significance level, respectively.

We can go a step further and estimate this increase in fixed effects in a regression framework to control for various confounders. To this end, we continue to restrict the data to cities whose history of governing states over their switching sequence includes two to five states. Let $h_{it} = \sum_{s \leq t} \mathbf{1}[\mathbf{J}(i,s) \neq \mathbf{J}(i,s-1)]$ be the hth state in a city's sequence of switches from one state to another (that is, h_{it} is the number of states having governed i up to t), and let $H(i) = \max_t h_{it}$ be the total length of the sequence of switches to another state – that is, the number of states ever governing i. Then we run regressions of the following form:

$$\widehat{\psi}_{\mathbf{J}(i,t)} = \omega_t + \kappa_{H(i)} + h_{it}\theta + \nu_{it} \tag{6}$$

where ω_t is a period fixed effect and captures continent-wide trends in the state effect; and $\kappa_{H(i)}$ is a sequence fixed effect and restricts the comparison to cities with the same total number of governing states over their lifetime. Because a given state's effect is constant across the cities it holds, we can clearly not expect the error term ν_{it} to have a standard distribution. To mitigate this problem, we two-way cluster our standard errors both on states as well as cities.

Across several approaches, we confirm that the increase in state quality after a switch to another state is around 7-8%, with slightly diminishing returns to the number of switching. However, the average increase conditional on period fixed effects is only around 3-4%, suggesting that about half of the effect arises from the overall pool of states improving over time, and another half from cities disproportionately switching to better states for a pool of states in a given period. We describe these findings in greater detail in Appendix A.1.

Finally, we show that this result holds even if we allow state effects to vary over time. To this end, we replace ψ_j and γ_t with ψ_{jt} in Equation (5) after residualizing the log population panel with respect to period effects and repeat the city-sequence exercise. As shown in Appendix Figure A.8, cities continue to be exposed to strongly increasing state quality when state quality can change for a given state over time – in fact, the increase is slightly stronger than with fixed state effects.

5.3 Heterogeneity across Space and Time

While Europe as a whole has benefited from state competition, some regions may have reaped the lion's share of benefits while others may have suffered the brunt of the costs. Thus, we now turn to studying regional heterogeneity in the costs and benefits across European cities. In Figure 9, we show each city's average improvement in state effects since its foundation, and we summarize region averages in Table 6. Clearly, cities on the British Isles are winners of this process, and so are parts of Northern Europe, Central Europe, and individual cities spread across the continent. These regions typically experienced large increases in the quality of the state as cities underwent changes in governance, opening the door to sustained growth.

The magnitude of improvement on the British Isles combined with the fact that English cities experienced a potentially crucial switch from the old English state to the union under the Dutch stadtholder and then to the union with Scotland raises the concern that our findings simply reflect competition within states as opposed to between states. However, our findings both on the costs and benefits hold up even when controlling for region-specific time effects, as shown in the specification with country-by-year fixed effects in Appendix Table A.3 for the cost side, and cost estimates also hold up after dropping Britain from the sample. To ensure that the state quality improvements are not driven by Britain alone, we reproduce Figure 8 after dropping all British cities from the data in Appendix Figure A.7, with the same upward trend in the quality of states clearly visible.[27]

The big losers are cities on the Iberian peninsula, Northern Italy, and Southeastern Europe. In some of these cases, as in Northern Italy, this may have been driven by excessive competition, bringing about only marginally better states and exposing cities to repeated looting. This can also be seen through the high average switching cost experienced by cities in Southern Europe. Many cities experienced close to no improvement from state competition (or even negative changes in state quality, as on Iberia), such as those in most of France and parts of Eastern Europe including Russia. In these cases, there may have been insufficient competition over cities by alternative states, leading to relative stagnation.

As shown in Table 6, not only did regions differ in the extent of improvement of state quality but also in the extent of costs incurred from violent conflict between states. While Central Europe saw large and immediate short-term increases in the quality of states with each switch in governance, this came at a high price: the switching intensity associated with Central Europe was so high that these costs were almost twice as large as the short-term benefits. Comparing switching costs to benefits reveals that Southern European cities were struck by an unusually unfortunate bargain: the highest switching costs of all European regions, with no improvement in states at all.[28] Understanding the heterogeneity in the rate of improvement across Europe is an interesting and important margin for future research.

It should be noted that these changes in average state quality capture only the benefits enjoyed by cities affected by changes in governance. However, presumably, a substantial part of the benefits of state competition came in the form of a positive externality on new cities being born into better states compared to a world in which states were not locked into a process of creative destruction. Quantifying the extent of these positive externalities is another area for future research.

We now turn to the question how these costs and benefits have changed over time as reported in Table 7. Looking first at the costs of competition, remarkably, switching costs

[27]Moreover, while there was considerable local support, the removal of James II during the Glorious Revolution was accompanied by an invasion of a foreign state, resulting in a qualitatively new (and arguably superior) state (Pincus, 2009), in line with our model. Pincus (2009, p. 6) writes, in reaction to the view that the new state created after the removal of James was simply a return to the moderate character of the past: “James’s opponents were, by and large, revolutionaries, not reactionaries. They appreciated that only a modernized English state could compete in contemporary Europe.” This line of reasoning is also supported by Cox (2016).

[28]It should be noted that the average change in state effects and the average switching cost only reflect the short-term benefit and cost of competition by region. For a true cost-benefit analysis, it would be necessary to allow the benefits from increased state effects to cumulate over subsequent periods.

are actually negative in the early period 1100-1300 (i.e. associated with larger city-periods), although it is not significant.[29] The point estimate is consistent with strategic switching of cities in an era in which they had more control over their fate: at the time, for city states in Northern Italy and the Lowlands, switching among rulers was a city-specific issue based on deliberate decisions in a strategic setting – a dynamic we exclude in our mass switching design. In later periods, switching costs are positive and become very large after 1700: a single switch is associated with a 7.63% drop in city population.

Turning to the benefits of state competition, we can see that these are rising steadily from the early period to the late period, from about 1.15% per switch to another state to 2.51%.[30] The rise is consistent with entry of higher-quality states over time, relative with the pool of states at the beginning of the period. Correspondingly, we show in Appendix Table A.6 that switching to a new entrant is associated with a considerably larger positive effect on city size compared to switching to another incumbent state. Overall, these results point towards better entry of more competitive states over time.

6 Mechanisms

We have shown that state competition leads to considerable costs and benefits to cities. What mechanisms underly these effects? Scholars have argued that a number of capacities and constraints govern the relationship between states and economic growth (Johnson and Koyama, 2017). While data availability limits the evidence we can bring to bear on this question, we offer three analyses to speak to this issue. First, we inspect the role of constraints on the executive as a way through which higher-quality states may lead to more growth, in line with a broad literature going back to North and Thomas (1973). Second, we provide evidence for higher fiscal capacity mediating the effect of higher-quality states on growth, consistent with work by Dincecco (2009) and Dincecco and Katz (2016). Finally, we turn to how protection from predation interacts with executive constraints and fiscal capacity by examining two case studies.

To analyze the role of executive consraints, we use data from Van Zanden et al. (2012) on

[29]The p-value on the test of the average of $S_{it}\widehat{\beta}_{1,1} + \mathbf{1}[S_{it} > 0]\widehat{\beta}_{2,1} = 0$ is 0.61 (the first subscript on β denotes whether the parameter captures the intensive or extensive margin; the second subscript captures the era, i.e. $\beta 1, 1$ is the intensive margin of the first era 1100-1300). That is, we cannot reject that the switching cost in the era 1100-1300 was zero.

[30]These estimates are somewhat lower than before because we are missing substantial improvement occurring over the period 1600-1700.

average parliamentary activity across 26 states.[31] Parliamentary activity is an index ranging from 0 to 100 counting the average number of years per century the state's parliament was in session. As Van Zanden et al. (2012, p. 835) write, "[e]conomists often assume that constraints on the executive – such as fully functioning parliaments – contribute to the efficiency of the economy via the protection of property rights." Thus, we interpret a correlation between the parliamentary activity index and the state effects as evidence for the state effects presenting plausible proxies for the institutional quality of the state, and in particular its capacity to protect property rights and thereby contribute to city growth.

The correlation can be seen in Figure 10. Clearly, the $\widehat{\psi}_j$ coefficients correlate strongly with the parliamentary activity index. Despite the small number of observations, the bivariate regression coefficient has a p-value of 2.0%. The magnitude suggests that an additional ten years of parliamentary sessions in a century are associated with a 8% higher state effect (in terms of city growth performance). This correlation suggests that state effects may serve as a useful proxy for the underlying quality of the state.

Constraints on the executive alone are unlikely to determine a state's quality in a comprehensive way. In fact, autocratic regimes with efficient means of taxation, such as Prussia, may have higher quality and stronger growth than more democratic states.[32] To provide evidence for this channel, we show that state quality is correlated with several proxies of fiscal capacity from Dincecco (2011); Stasavage (2011) in Figure 11. In particular, states with higher estimated ψ_j have higher tax revenue per capita, are able to maintain a higher deficit ratio, are charged a lower interest rate, and were able to obtain long-term loans sooner than others. While the individual correlations are not statistically significant, a joint test using methods from meta-analysis yields a p-value of 0.02.[33] While each of these measures

[31] We do not include states for which Van Zanden et al. (2012) do not provide estimates for parliamentary activity. While they have done admirable work covering most states for a long period of European history, it is possible that some parliamentary activity has received insufficient historical treatment and has thus not made it into their data. For example, the Estates General of Burgundy met every three years and had considerable influence in matters of legislation and taxation (Richard, 1984), and the Republic of Novgorod featured fairly democratic institutions for its time (Sixsmith, 2011); neither of these two is included in their data.

[32] This point is made powerfully by Rosenthal and Wong (2011, p. 226): "The level of economic growth in Wilhelmine Germany was remarkably robust even though by English or French standards it was an incomplete democracy. Equally problematic, the levels of economic achievement of England had few echoes in Ireland (although it was formally part of the same polity) during the 120 years in which the union between the two countries prevailed."

[33] Specifically, we construct joint tests by combining the point estimates and standard errors of the four regressions associated with the correlations shown in Figure 11 using a number of meta-analysis methods. The median p-value across eight methods is 0.007, and the maximum is 0.021.

is imperfect, together they form a picture of higher-quality states on more stable fiscal footing, reducing uncertainty about the ability to protect from competing states or the need for excessive extraction.

Finally, we illustrate how executive constraints and fiscal capacity interacted with protection from predation in determining how state competition affects city growth. We do so through the lens of two city histories shown in Appendix Figure A.13. Gent and the Low Countries more generally demonstrate the costs of instability and poor governance. The Spanish Habsburgs refused to give up the Low Countries but lacked the financial resources to finance an army to secure them. The result was the Eighty Years' War, where unpaid troops repeatedly mutinied, massacring civilians and looting goods: there were at least 45 separate mutinies according to Parker (2004, Ch. 8). Severe taxes levied locally in an attempt at properly funding these troops instead led to further chaos, motivating an anti-Spanish alliance to drive out the "hard government of the Spaniards and their adherents", as described in the treaty on the Pacification of Ghent 1576 (Koenigsberger, 2001, p. 272). The population of Ghent recovers as control shifts away from Spain.

In contrast, Leipzig serves as an illustration of the benefits from constrained governance, preventing inefficient appropriation by the state. After attaining autonomy from the Holy Roman Empire during the 15^{th} century, the Saxon state developed under a restrained nobility and court, whose rulers supported unusually modern civil law and commercial ordinances, sometimes in opposition to special interests (Blaschke, 1990; Matzerath, 2006). As a result, Leipzig hosted one of the largest trade fairs in Central Europe for almost two centuries, supported by legal institutions that protected merchants from expropriation and fraud and guaranteed access to credit (Beachy, 1999). Thus, Leipzig experienced substantial population growth on the basis of flourishing industries, especially trade and manufacturing.

7 Conclusions

We argue that competition between European states caused substantial short-term losses compensated by meaningful long-term gains to the economic development of cities in the region. In a process reminiscent of creative destruction operating through technological innovation on the quality of firms, European states compete for territory through institutional innovation. We provide evidence for both of these forces by separating cities that experience competition without any change in governance from those that end up with a new state in power. The former cities suffer from large transitory losses in population, while the latter enjoy sustained population increases. Decomposing the contributions of changes in governance into a switching effect and a state effect, we show that improvements in the quality of states occur both due to improvement of the pool of states over time as well as due to cities gravitating towards higher quality states.

One simple way to summarize the key tradeoff between benefits and costs of creative destruction is to study the average gains for cities of a given length of switching sequence against the costs of switching, as we do in Figure 12. Total gross benefits of state competition are shown in the solid line, illustrating the concavity in the estimated returns to switching states. Looking at net benefits, which take the cost of switching into account, we can see that these benefits rise from slightly negative (even cities that never switch to another state are exposed to some switching) to almost 10% per period for cities that switch states up to two times; but once a city switches state three times or more, due to the decreasing marginal benefits of switching, the costs of switching begin to outweigh the benefits. This is consistent with the model prediction that there is a positive, optimal number of switches cities should undergo for welfare to be maximized. It should also be noted that any positive externalities on cities not directly affected by switching are not captured in our approach, which may be valuable to attempt to estimate in future research.

In light of Europe's ascendancy over other world regions, this tradeoff seen within Europe may have acted globally across a much wider domain. While Europe may have experienced close to the optimal level of creative destruction in its state system, China and South Asia present two cases that may have been exposed to insufficient and excessive creative destruction, respectively. Due to China's early unification under a single state after the establishment of the Qin dynasty and especially the Han dynasty around 200 BC, the Chinese state faced only infrequent threats from other states (Finer, 1997; Rosenthal and Wong, 2011). As a result, Chinese cities were exposed to much less upheaval than European ones, but in turn the Chinese state may have ossified and the quality of governance may have failed to improve.

On the other hand, South Asia experienced nearly one hundred "supra-regional" or "pan-Indian" states that have ruled over substantial portions of the region (Schwartzberg, 1978). Figure 14.3 of Schwartzberg (1978) shows that the only city that was reliably a center of political power in South Asia was Delhi, and that numerous other capitals appear randomly scattered about the sub-continent. This contrasts substantially with the state system in

Europe, where Rome, Constantinople, Vienna, Lisbon, Paris, and London served as capitals of European states for most of the last thousand years. It seems plausible that the extreme variability in the number and location of states on the Indian subcontinent, perhaps caused by the lack of easily defended peninsulas and protective mountains, led to the repeated capture and looting of cities and thereby delayed capital deepening. If this were indeed the case, then South Asia would be an example of a state system where control of cities was too ephemeral, perhaps causing slower economic progress due to an excess of creative destruction.

Ultimately, whether a state system can benefit from competition depends crucially on its tendency to engender states that are both high quality and highly competitive in the contest with existing states. While this seems to have been true in Europe over 1100-1850, there is no guarantee for this to be true in other places or at other times – for instance, ancient Greece may offer a cautionary example, where states with seemingly high quality institutions succumbed to more powerful states that were detrimental to long-run growth. Understanding the features of a state system in which high quality states survive and thrive is an important avenue for future research in political economy.

References

Abowd, John M., Francis Kramarz, and David N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, Vol. 67, pp. 251–333.

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

Acemoglu, Daron and James A Robinson (2012): *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*: Crown Publishing Group, Random House Inc., New York.

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt (2014): "Chapter 1 - What Do We Learn From Schumpeterian Growth Theory?" in Philippe Aghion and Steven N. Durlauf eds. *Handbook of Economic Growth*, Vol. 2 of Handbook of Economic Growth: Elsevier, pp. 515 – 563.

Bairoch, Paul, Jean Batou, and Pierre Chevre (1988): *The Population of European Cities. Data Bank and Short Summary of Results: 800-1850*: Geneva Switzerland Librairie Droz 1988.

Beachy, Robert (1999): "Reforming Interregional Commerce: The Leipzig Trade Fairs and Saxony's Recovery from the Thirty Years' War," *Central European History*, Vol. 32, pp. 431–452.

Besley, Timothy and Torsten Persson (2010): "State Capacity, Conflict, and Development," *Econometrica*, Vol. 78, pp. 1–34.

Blaschke, Karlheinz (1990): *Hof und Hofgesellschaft im Königreich Sachsen während des 19. Jahrhunderts*, 1st edition, pp. 177–206: De Gruyter.

Bockstette, Valerie, Areendam Chanda, and Louis Putterman (2002): "States and Markets: The Advantage of an Early Start," *Journal of Economic Growth*, Vol. 7, pp. pp. 347–369.

Büntgen, Ulf, Christian Ginzler, Jan Esper, Willy Tegel, and Anthony J. McMichael (2012): "Digitizing Historical Plague," *Clinical Infectious Diseases*, Vol. 55, pp. 1586–1588.

Card, David, Jörg Heining, and Patrick Kline (2013): "Workplace Heterogeneity and the Rise of West German Wage Inequality*," *The Quarterly Journal of Economics*, Vol. 128, pp. 967–1015.

Cervellati, Matteo, Sara Lazzaroni, Giovanni Padrolo, and Paolo Vanin (2018): "Sovereign Polities, Territorial Control and City Growth in Pre-Industrial Europe, 1000-1800," *Working Paper*.

Clark, Gregory (2008): *A Farewell to Alms: a Brief Economic History of the World*: Princeton University Press.

Cosandey, David (1997): *Le secret de l'Occident : du miracle passé au marasme présent*: Arléa.

Cox, Gary W (2016): *Marketing Sovereign Promises: Monopoly Brokerage and the Growth of the English State*: Cambridge University Press.

Cox, Gary W. (2017): "Political Institutions, Economic Liberty, and the Great Divergence," *The Journal of Economic History*, Vol. 77, pp. 724–755.

Dal Bó, Ernesto, Pedro Dal Bó, and Erik Eyster (2018): "The Demand for Bad Policy when Voters Underappreciate Equilibrium Effects," *The Review of Economic Studies*, Vol. 85, pp. 964–998.

Davis, Donald R. and David E. Weinstein (2002): "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review*, Vol. 92, pp. 1269–1289.

De Long, J. Bradford and Andrei Shleifer (1993): "Princes and Merchants: European City Growth before the Industrial Revolution," *The Journal of Law and Economics*, Vol. 36, pp. 671–702.

De Vries, J. (1997): *The First Modern Economy: Success, Failure, and Perseverance of the Dutch Economy, 1500-1815*: Cambridge University Press.

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

Dell, Melissa, Nathaniel Lane, and Pablo Querubin (2017): "The Historical State, Local Collective Action, and Economic Development in Vietnam," Working Paper 23208, National Bureau of Economic Research.

Diamond, Jared M (1997): *Guns, Germs, and Steel: A Short History of Everybody for the Last 13,000 Years*: W. W. Norton.

Dincecco, Mark (2009): "Fiscal Centralization, Limited Government, and Public Revenues in Europe, 1650–1913," *The Journal of Economic History*, Vol. 69, pp. 48–103.

——— (2011): *Political Transformations and Public Finances: Europe, 1650–1913*: Cambridge University Press.

Dincecco, Mark and Gabriel Katz (2016): "State Capacity and Long-run Economic Performance," *The Economic Journal*, Vol. 126, pp. 189–218.

Dincecco, Mark and Massimiliano Onorato (2016): "Military conflict and the rise of urban Europe," *Journal of Economic Growth*, Vol. 21, pp. 259 – 282.

Eeckhout, Jan (2004): "Gibrat's Law for (All) Cities," *American Economic Review*, Vol. 94, pp. 1429–1451.

Finer, Samuel Edward (1997): *The History of Government from the Earliest Times: Ancient Monarchies and Empires*, Vol. 1: Oxford University Press.

Gabaix, Xavier (1999): "Zipf's Law for Cities: An Explanation," *The Quarterly Journal of Economics*, Vol. 114, pp. pp. 739–767.

Gennaioli, Nicola and Hans-Joachim Voth (2015): "State Capacity and Military Conflict," *The Review of Economic Studies*, Vol. 82, pp. 1409–1448.

Hall, John A (1986): *Powers and liberties: The causes and consequences of the rise of the West*: Univ of California Press.

Hoffman, Philip T (2015): *Why did Europe Conquer the World?*: Princeton University Press.

Jackson, Peter (1999): *The Delhi Sultanate: a Political and Military History*: Cambridge University Press.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan (1993): "Earnings Losses of Displaced Workers," *The American Economic Review*, Vol. 83, pp. 685–709.

Johnson, Noel D. and Mark Koyama (2017): "States and economic growth: Capacity and constraints," *Explorations in Economic History*, Vol. 64, pp. 1 – 20.

Jones, E. L. (1981): *The European Miracle: Environments, Economies, and Geopolitics in the History of Europe and Asia*: Cambridge University Press.

Kennedy, Paul (2010): *The Rise and Fall of the Great Powers*: Vintage.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten (2018): "Leave-out estimation of variance components," *ArXiv e-prints*.

Koenigsberger, Helmut Georg (2001): *Monarchies, states generals and parliaments: the Netherlands in the fifteenth and sixteenth centuries*: Cambridge University Press.

Landes, David S. (2006): "Why Europe and the West? Why Not China?" *Journal of Economic Perspectives*, Vol. 20, pp. 3–22.

Mann, Michael (1984): "The Autonomous Power of the State: Its origins, Mechanisms and Results," *European Journal of Sociology / Archives Européennes de Sociologie / Europäisches Archiv für Soziologie*, Vol. 25, pp. 185–213.

Matzerath, Josef (2006): *Adelsprobe an der Moderne: sächsischer Adel 1763 bis 1866; Entkonkretisierung einer traditionalen Sozialformation*, No. 183: Franz Steiner Verlag.

McCrary, Justin (2007): "The Effect of Court-Ordered Hiring Quotas on the Composition and Quality of Police," *American Economic Review*, Vol. 97, pp. 318–353.

McNeill, William H (1982): "The Pursuit of Power: Technology, Armed force, and Society."

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

——— (2014): "National Institutions and Subnational Development in Africa," *The Quarterly Journal of Economics*, Vol. 129, pp. 151–213.

Mokyr, Joel (1992): *The Lever of Riches: Technological Creativity and Economic Progress*: Oxford University Press.

North, Douglass C and Robert Paul Thomas (1973): *The Rise of the Western World: A New Economic History*: Cambridge University Press.

Nunn, Nathan and Nancy Qian (2011): "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment," *The Quarterly Journal of Economics*, Vol. 126, pp. 593–650.

Parker, Geoffrey (2004): *The Army of Flanders and the Spanish road, 1567-1659: the logistics of Spanish victory and defeat in the Low Countries' Wars*: Cambridge University Press.

Pincus, Steven CA (2009): *1688: The First Modern Revolution*: Yale University Press.

Queralt, Didac (2018): "The Legacy of War on Fiscal Capacity," *Working Paper*.

Reed, Frank (2016): "The Centennia Historical Atlas: Academic Research Edition," Clockwork Mapping: Jamestown RI.

Richard, Jean (1984): "Les Etats de Bourgogne dans la politique des ducs Valois," *Publications du Centre Européen d'Etudes Bourguignonnes*, Vol. 24, pp. 11–16.

Rosenthal, Jean-Laurent and R. Bin Wong (2011): *Before and Beyond Divergence: The Politics of Economic Change in China and Europe*: Harvard University Press.

Schwartzberg, Joseph E. (1978): *A Historical Atlas of South Asia*: University of Chicago Press.

Sixsmith, Martin (2011): *Russia: A 1,000-Year Chronicle of the Wild East*: Random House.

Stasavage, David (2011): *States of Credit: Size, Power, and the Development of European Polities*, Vol. 35: Princeton University Press.

Tilly, Charles (1990): *Capital, Coercion and European States*: Cambridge, MA: Basil Blackwell.

Van Zanden, Jan Luiten, Eltjo Buringh, and Maarten Bosker (2012): "The Rise and Decline of European Parliaments, 1188–1789," *The Economic History Review*, Vol. 65, pp. 835–861.

Voigtländer, Nico and Hans-Joachim Voth (2013a): "Gifts of Mars: Warfare and Europe's Early Rise to Riches," *Journal of Economic Perspectives*, Vol. 27, pp. 165–86.

——— (2013b): "The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe," *The Review of Economic Studies*, Vol. 80, pp. 774–811.

Table 1: Summary statistics for cities and states with cities

Panel A: Cities	Mean	P(50)	SD	Min	Max
City population					
1100 (N = 117)	20.3	16	25.6	1	250
1200 (N = 185)	17.5	13	17.5	1	150
1300 (N = 521)	11.6	6	16.2	1	150
1400 (N = 631)	9.6	5	17.3	1	275
1500 (N = 834)	9.5	5	14.1	1	225
1600 (N = 1114)	10.4	6	19.0	1	300
1700 (N = 1409)	9.9	5	26.0	1	575
1750 (N = 1595)	11.1	6	28.2	1	675
1800 (N = 2107)	12.0	7	31.1	1	948
1850 (N = 2063)	19.7	10	63.8	1	2236
Other characteristics					
Year of city foundation	1574.0	1600	221.4	1100	1850
Total number of switches	7.7	6	7.3	0	44
Number of periods with switching	2.6	2	2.0	0	10
Number of changes in governing state	1.3	1	1.5	0	8
Number of states ever governing city	2.2	2	1.3	1	8
N	2,181				
Panel B: States with cities	Mean	P(50)	SD	Min	Max
Number of cities in state					
1100 (N = 20)	6.1	4	7.4	1	30
1200 (N = 30)	6.5	1.5	13.8	1	70
1300 (N = 47)	11.6	3	24.1	1	138
1400 (N = 43)	14.7	7	21.5	1	114
1500 (N = 39)	21.4	8	33.6	1	149
1600 (N = 41)	27.8	12	67.0	1	417
1700 (N = 38)	38.1	11	79.4	1	409
1750 (N = 39)	40.9	16	62.2	1	230
1800 (N = 41)	54.0	6	99.2	1	461
1850 (N = 41)	50.3	10	85.9	1	340
Other characteristics					
Period of state entry	1418.2	1400	245.0	1100	1850
Period of state exit	1610.9	1700	236.1	1100	1850
Number of periods in existence	3.3	2	2.5	1	10
Average net change in number of cities per period	4.1	0.3	12.7	-21	71
Average net gain in number of cities per period	13.2	5.8	20.3	1	107.5
Average net loss in number of cities per period	-10.3	-5.5	13.5	-61	-1
N	129				

Notes: P(50) is the median. **Panel A**: Population counts are in thousands. The "N" in parentheses next to the year refers to the number of cities with nonmissing population in that year (note there are observations with zero population in years 1100 and 1200, which is due to linear interpolation between 1000 and 1200 in our source data). "Total number of switches" is the sum of switches over 1000-1850, that is, $\sum_t \#[\text{Decades in period } t] \times S_{it}$. "Number of periods switching" is $\sum_t \mathbf{1}[S_{it} > 0]$. "Number of changes in governing state" is $\sum_t \mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$. "Number of states ever governing city" is $\#[j : j = \mathbf{J}(i,t) \text{ for all } t]$. **Panel B**: Only states with at least one city included. The "N" in parentheses next to the year refers to the number of existing states with cities in that year. Changes in number of cities can be territorial or foundations/disappearances. Average net gains (losses) are averages over periods when the state experiences any gains (losses) at all.

Table 2: City summary statistics by number of switching periods

	Number periods switching				Switching to another state			
	0	1-2	3-4	5≥	0	1-2	3-4	5≥
First year in Bairoch et al. (1988)	1748	1683	1487	1320	1727	1549	1345	1279
	(109)	(171)	(173)	(142)	(144)	(182)	(162)	(135)
Population at first appearance in data	6.29	5.49	5.58	9.23	6.24	5.29	7.90	11.31
	(3.83)	(4.93)	(5.06)	(16.23)	(4.36)	(5.98)	(15.57)	(13.13)
Population in 1850	11.14	18.22	20.41	29.01	12.69	19.42	31.44	35.91
	(10.70)	(50.11)	(99.98)	(51.79)	(37.86)	(39.13)	(125.71)	(52.28)
Latitude	47.96	46.77	45.62	47.18	46.50	46.93	46.61	48.66
	(5.45)	(5.65)	(6.15)	(5.51)	(5.20)	(6.28)	(5.97)	(5.09)
Longitude	13.96	9.49	6.99	10.45	13.07	7.04	9.84	6.00
	(19.14)	(10.95)	(9.35)	(7.64)	(15.12)	(9.81)	(8.55)	(7.53)
Distance to state centroid (percentiles)	0.45	0.48	0.52	0.52	0.46	0.51	0.54	0.54
	(0.27)	(0.27)	(0.20)	(0.16)	(0.28)	(0.21)	(0.18)	(0.13)
Distance to borders (km)	323	136	106	71	209	127	93	61
	(344)	(121)	(95)	(61)	(270)	(117)	(78)	(54)
Distance to continent urban center (km)	1,254	938	897	761	1,076	904	831	736
	(846)	(473)	(479)	(397)	(695)	(477)	(417)	(427)
Switches per decade	0	0.29	0.30	0.35	0.16	0.30	0.31	0.36
	–	(0.25)	(0.19)	(0.15)	(0.24)	(0.20)	(0.17)	(0.17)
Per period probability of switching	0	0.54	0.62	0.76	0.30	0.60	0.68	0.82
	–	(0.26)	(0.18)	(0.16)	(0.34)	(0.22)	(0.18)	(0.15)
Per period probability of new state	0	0.15	0.31	0.45	0	0.29	0.46	0.68
	–	(0.18)	(0.17)	(0.20)	–	(0.12)	(0.13)	(0.13)
Number of cities	419	778	552	432	886	823	365	107

Notes: Means by groups of cities switching the number of times indicated in the columns; standard deviations in parentheses. Populations in thousands. State centroids weighted by the populations of all cities in them. Continental urban center weighted by the populations of all cities on the continent. "Number of periods switching" is $\sum_t \mathbf{1}[S_{it} > 0]$, and "Switching to another state" is $\sum_t \mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$. Columns 1-4 (number of periods switching) and 5-8 (number of periods switching to another state) are two different groupings of the 2,181 cities in the data.

Table 3: Benefits of switching using "mass switches"

	Dependent variable: log(population)							
			Mass switches					
	Any number of cities		At least 10 cities together			At least 20 cities together		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pre-switching size	0.024	0.007	-0.013	0.015	-0.004	-0.024	0.001	-0.023
	(0.037)	(0.035)	(0.038)	(0.038)	(0.042)	(0.041)	(0.041)	(0.048)
Post-switching to another state	0.164***	0.185***	0.184***	0.163***	0.184***	0.171***	0.187***	0.198***
	(0.024)	(0.022)	(0.024)	(0.023)	(0.026)	(0.026)	(0.025)	(0.031)
City FE	X	X	X	X	X	X	X	X
Period FE	X	X	X	X	X	X	X	X
Switch to same state	—	—	X		X	X		X
Switch from same state	—	—		X	X		X	X
Sequence lengths $H(i)$	{1,2}	All	All	All	All	All	All	All
Switching event type	First	First	First	First	First	First	First	First
City N	1,188	2,063	1,846	1,868	1,665	1,663	1,680	1,415
Total N	7,107	10,458	8,785	9,040	7,648	7,448	7,782	5,983
Adjusted R^2	0.69	0.69	0.66	0.67	0.65	0.65	0.64	0.60

Notes: **Specification**: Event study estimates via OLS according to equation (3). **Samples**: Sequence length $H(i) = \max h_{it}$ where $h_{it} = \sum_{s \le t} \mathbf{1}[\mathbf{J}(i,s) \neq \mathbf{J}(i,s-1)]$. Thus, for example, sequence lengths $\{1,2\}$ denotes all cities that switched states once or twice, while "all" denotes the inclusion of cities of any number of switches over their lifetime. Columns 3-8 restrict the sample to the inclusion of cities that switch together with at least 10 (columns 3-5) or 20 (columns 6-8) other cities, either *from* the same state (columns 3 and 6), *to* the same state (columns 4 and 7), or *both* from and to the same state (columns 5 and 8). **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching to another state once, the event time is the period of switching to another state; for cities switching to another state multiple times, the event time is the first period of switching to another state. **Control variable**: controlling for $\mathbf{1}[S_{it} > 0]$, i.e. all periods in which the city switched (ending up in a different state or the same state). **Fixed effects**: included are city and period fixed effects. **Standard errors**: clustered on the city level.

Table 4: Basic panel decomposition results

	Dependent variable: log(population)				
	(1)	(2)	(3)	(4)	(5)
Switching indicator {1}	-0.082***	-0.113***	-0.113***	-0.128***	-0.112***
	(0.014)	(0.014)	(0.018)	(0.017)	(0.016)
Switch to another state {2}			0.074***	0.039*	0.034*
			(0.022)	(0.021)	(0.020)
City FE	X	X	X	X	
Period FE	X	X	X	X	X
State FE		X		X	
City × state FE					X
F-test [p-value]	–	–	5.59	22.4	18.6
H_0: {1} + {2} = 0			[.018]	[.000]	[.000]
Std. dev. of **Xb**	.041	.056	.053	.06	.
Std. dev. of $\widehat{\psi}_j$	–	.327	–	.327	–
Adjusted R^2	0.69	0.73	0.69	0.73	0.78
RMSE	0.51	0.48	0.51	0.48	0.43
City N	1,949	1,949	1,949	1,949	1,949
State N	–	89	–	89	–
Total N	8,629	8,629	8,629	8,629	8,629

Notes: **Specification**: Regression estimates according to specification (4) for 2,182 cities and ten periods (1100 to 1700 in centuries; 1750 to 1850 in half-centuries). The number of cities is lower than the total number due to collinearity created by the fixed effects (i.e. some city-period observations become collinear with a combination of fixed effects). **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Independent variables**: "Switching intensity per decade" is the average number of switches per decade over the period, i.e. normalized to take into account varying period length. It corresponds to S_{it}. "Switching indicator" is an indicator variable for $\mathbf{1}[S_{it} > 0]$. "Switch to another state" is the indicator for $\mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$. Together, the independent variable(s) of a given model are denoted as **Xb**. **Standard errors**: clustered on the city level.

Table 5: Testing the increase in state fixed effect

	Dependent variable: $\widehat{\psi}_j$									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Number of state in sequence	0.068*	0.036*					0.161*	0.132**		
	(0.039)	(0.020)					(0.092)	(0.054)		
Post-switching to another state			0.087	0.044*						
			(0.058)	(0.023)						
Fraction of switching sequence					0.265*	0.162*				
					(0.157)	(0.085)				
(Number of state in sequence)^2							-0.019*	-0.019**		
							(0.011)	(0.008)		
State in sequence of length 2									0.131	0.069
									(0.112)	(0.059)
State in sequence of length 3									0.126*	0.092**
									(0.072)	(0.045)
State in sequence of length 4									0.052	0.026
									(0.042)	(0.034)
State in sequence of length 5									0.038*	0.019
									(0.021)	(0.016)
Sequence length FE	X	X	X	X	X	X	X	X	X	X
Period FE		X		X		X		X		X
City N	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194	1,194
State N	115	115	115	115	115	115	115	115	115	115
Total N	3,535	3,535	4,690	4,690	3,535	3,535	3,535	3,535	3,535	3,535
Adjusted R^2	0.06	0.15	0.03	0.13	0.07	0.16	0.07	0.16	0.07	0.16

Notes: **Specifications**: OLS estimation according to equation (6). **Sample**: All cities that are governed by two to five states over their lifetime. **Dependent variable**: Estimated state effects, that is $\widehat{\psi}_j$. **Independent variables**: "Number of state in sequence" is $h_{it} = \sum_{s \leq t} \mathbf{1}[\mathbf{J}(i,s) \neq \mathbf{J}(i,s-1)]$. Post-switching to another state is $\mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$ in one duplicate of duplicated city time series. "Fraction of switching sequence is $h_{it}/H(i)$, where $H(i) = \max_t h_{it}$. "State in sequence of length s" is $h_{it} \times \kappa_{H(i)}$, where $\kappa_{H(i)}$ is a set of indicators for the length of the switching sequence being s. **Fixed effects**: included are period and sequence length fixed effects. **Standard errors**: clustered both on the state and the city level.

Table 6: Switching and state improvement by European region

	Regions						
	Britain	Northern	Central	Eastern	Western	Southern	Iberia
First year in Bairoch et al.	1545	1529	1488	1651	1572	1575	1577
Distance to borders (km)	275	164	26	309	97	82	170
Population in 1850	48.35	12.62	17.31	18.96	17.38	16.07	14.88
Switches per decade	0.17	0.09	0.33	0.18	0.20	0.38	0.21
Per period prob. of switching	0.38	0.34	0.67	0.43	0.39	0.64	0.49
Per period prob. of new state	0.26	0.18	0.32	0.15	0.23	0.21	0.20
Number periods in data	5.38	5.60	5.93	3.90	4.96	4.66	4.80
Total periods switching	2.14	1.92	3.77	2.00	2.15	2.96	2.59
Total periods new state	1.58	0.96	1.87	0.90	1.38	1.31	1.32
Avg. change state effect (%)	16.24	5.54	3.83	0.55	0.07	-0.39	-1.10
Avg. switching cost (%)	-4.07	-3.18	-7.38	-4.53	-4.34	-7.46	-5.15
Number of cities	187	48	281	406	474	469	316

Notes: Means by region. Population in 1850 in thousands. Per period probabilities of switching and switching to a new state are estimated $P[S_{it} > 0]$ and $P[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$, respectively. Change in state effect is the average of $\widehat{\psi}_{\mathbf{J}(i,t)} - \widehat{\psi}_{\mathbf{J}(i,t-1)}$ for all city-periods with positive population. Switching cost is estimated cost coefficients times the number/frequency of occurrence. Specifically, we run $y_{it} = \alpha_i + \gamma_t + \psi_{\mathbf{J}(i,t)} + S_{it}\beta_1 + \mathbf{1}[S_{it} > 0]\beta_2 + \varepsilon_{it}$ and then compute the switching cost the per-region average of $S_{it}\widehat{\beta_1} + \mathbf{1}[S_{it} > 0]\widehat{\beta_2}$. Regions are comprised as follows (using current countries):

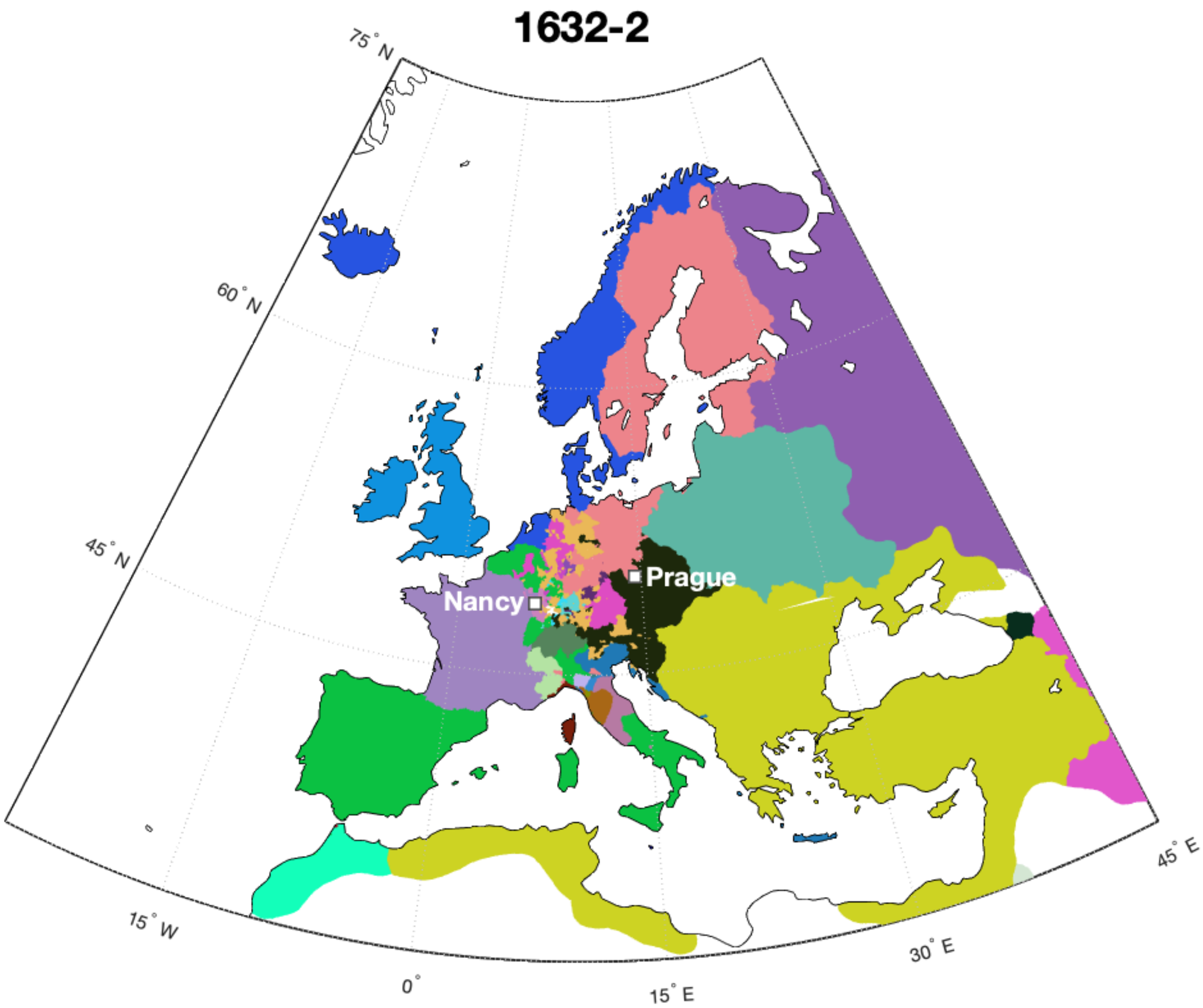
- Britain: Great Britain and Ireland.
- Northern: Norway, Sweden, Denmark, and Finland.
- Central: Austria, Germany, and Switzerland.
- Eastern: Czech Republic, Slovakia, Bulgaria, Hungary, Poland, Romania, and Russia.
- Western: France, Belgium, Netherlands and Luxembourg.
- Southern: Italy, Greece, Malta, and Yugoslavia.
- Iberia: Spain and Portugal.

Table 7: Switching and state improvement by era

	Eras		
	1100-1300	1400-1600	1700-1850
Distance to borders (km)	97	96	150
Population	14.14	9.92	13.61
Switches per decade	0.12	0.14	0.29
Per period prob. of switching	0.49	0.62	0.50
Per period prob. of new state	0.15	0.40	0.25
Avg. change state effect (%)	1.15	2.37	2.51
Avg. switching cost (%)	1.46	-2.13	-7.63
Number of cities	523	1129	2177

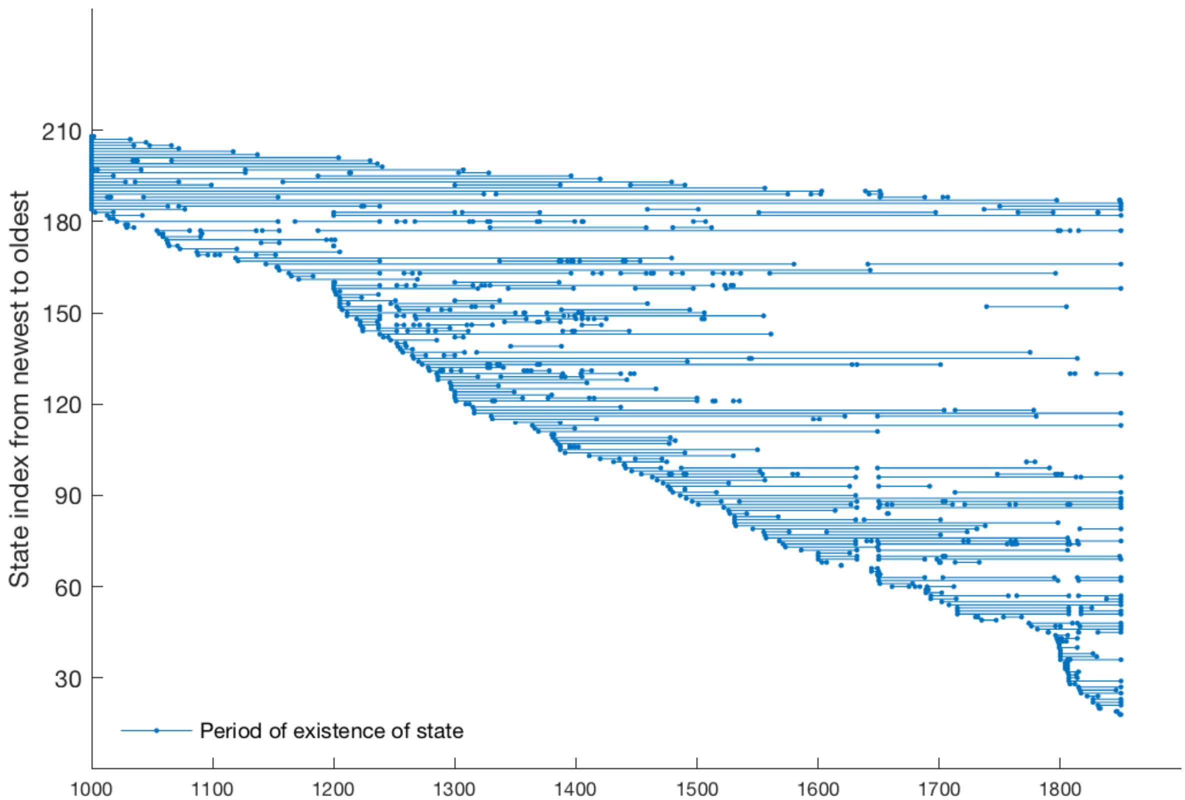
Notes: Means by era. Population in 1850 in thousands. Per period probabilities of switching and switching to a new state are estimated $P[S_{it} > 0]$ and $P[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$, respectively. Change in state effect is the average of $\widehat{\psi}_{\mathbf{J}(i,t)} - \widehat{\psi}_{\mathbf{J}(i,t-1)}$ for all city-periods with positive population. Switching cost is estimated cost coefficients for each era separately multiplied times the number/frequency of occurrence in the era. Specifically, we run $y_{it} = \alpha_i + \gamma_t + \psi_{\mathbf{J}(i,t)} + \sum_{k=1}^{3}\{S_{it}\beta_{1,k} + \mathbf{1}[S_{it} > 0]\beta_{2,k}\} + \varepsilon_{it}$ and then compute the switching cost the per-era average of $S_{it}\widehat{\beta}_{1,k} + \mathbf{1}[S_{it} > 0]\widehat{\beta}_{2,k}$.

Figure 1: Example of the *Centennia Historical Atlas*: year 1632, period 2



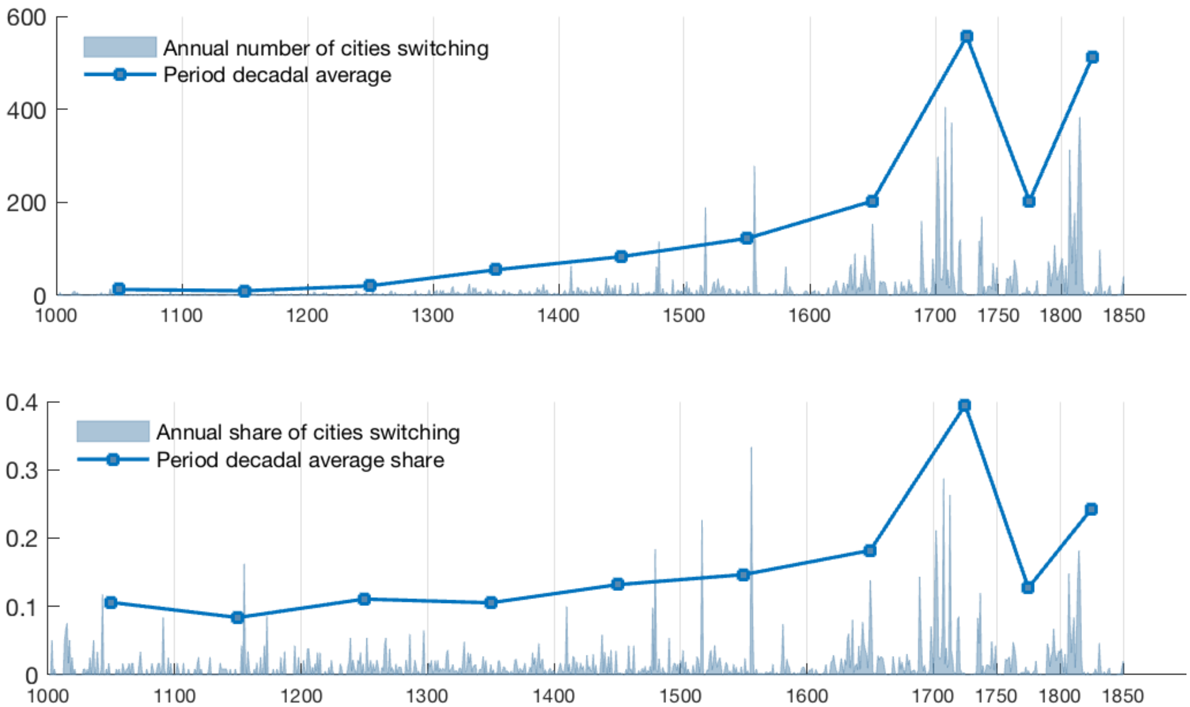
Notes: Territorial control of European states in early 1632, period 2 out of $\{0, 1, ..., 9\}$ based on the data from Reed (2016) in the *Centennia Historical Atlas.* The map shows *de facto* territorial holdings during the early stage of the Thirty Year War. Highlighted are two cities that exemplify the two types of switches we use to estimate costs and benefits. In the course of the conflict, Prague switched from the Habsburgs to Sweden for less than two years, while Nancy switched from Burgundy to France for decades – we use both cases to identify the costs of switching, while only the latter serves to identify the benefits of French governance over Burgundian governance.

Figure 2: Entry, existence and exit of European states, 1000-1850



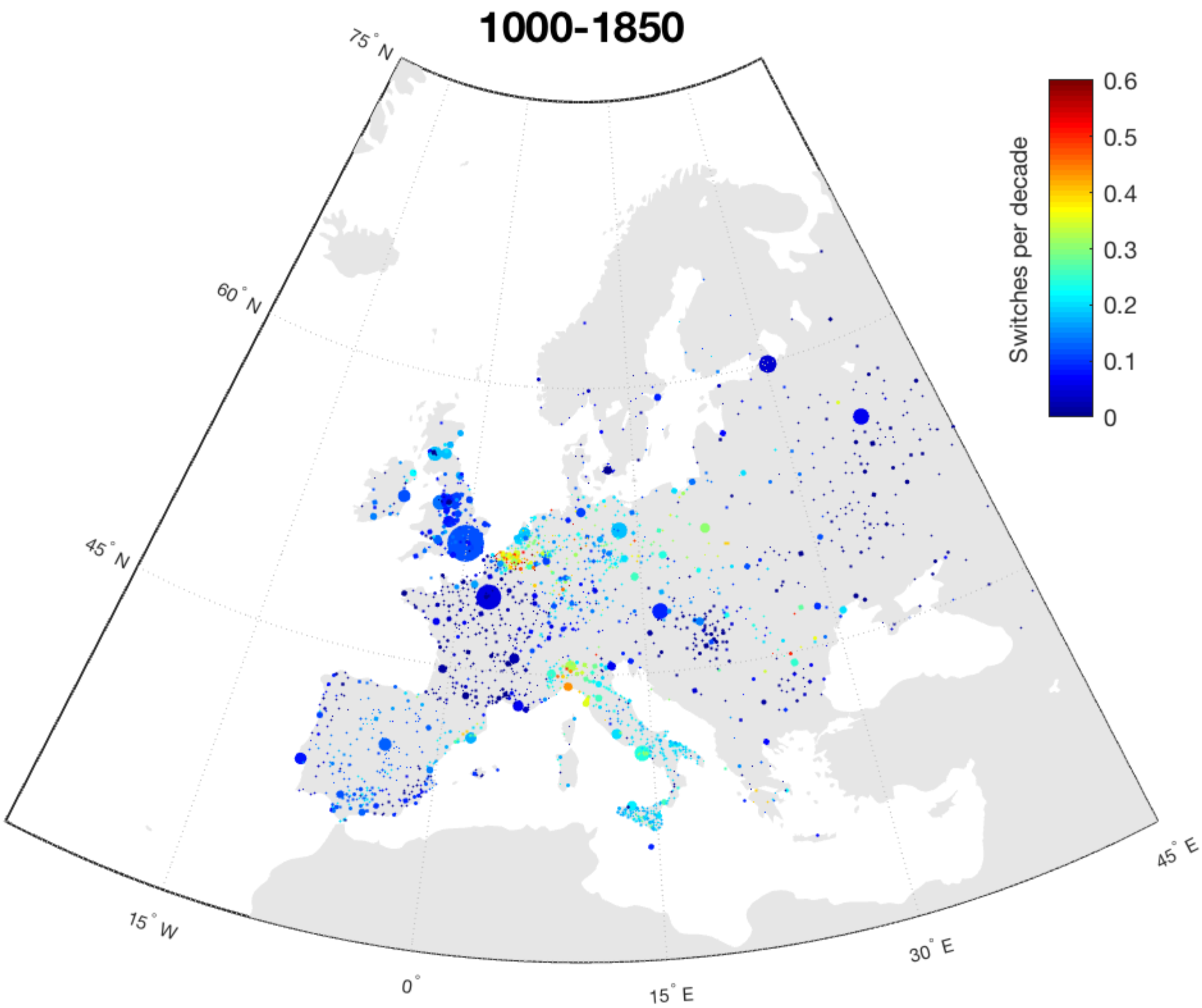
Notes: Entry, existence and exit of European states with at least one city as defined by Bairoch over 1000-1850 according to the coding in Reed (2016). States are ordered by index from top to bottom by year of entry, and within those, year of exit.

Figure 3: Frequency of city switches over 1000-1850



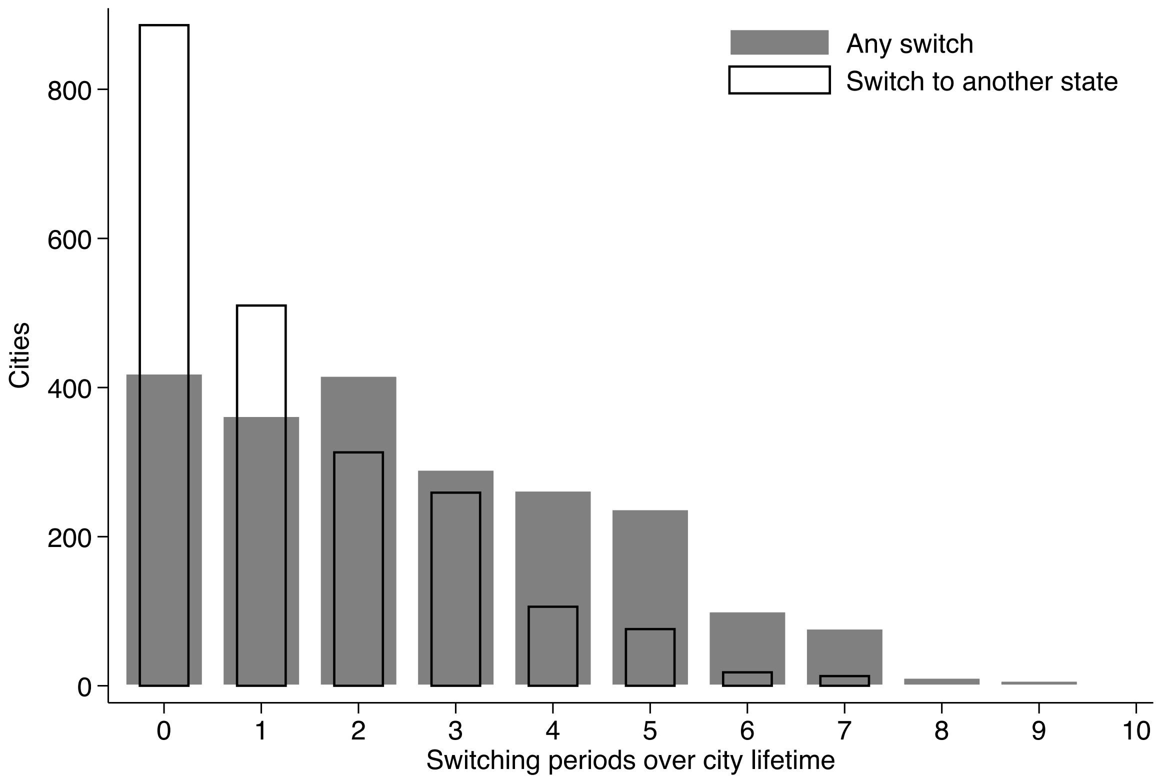
Notes: Annual number and share of cities switching in each year as well as decadal averages per Bairoch period. A switch is defined as a boundary change that leads to another state holding the city.

Figure 4: Geocoded city population data and average number of switches per decade over 1000-1850



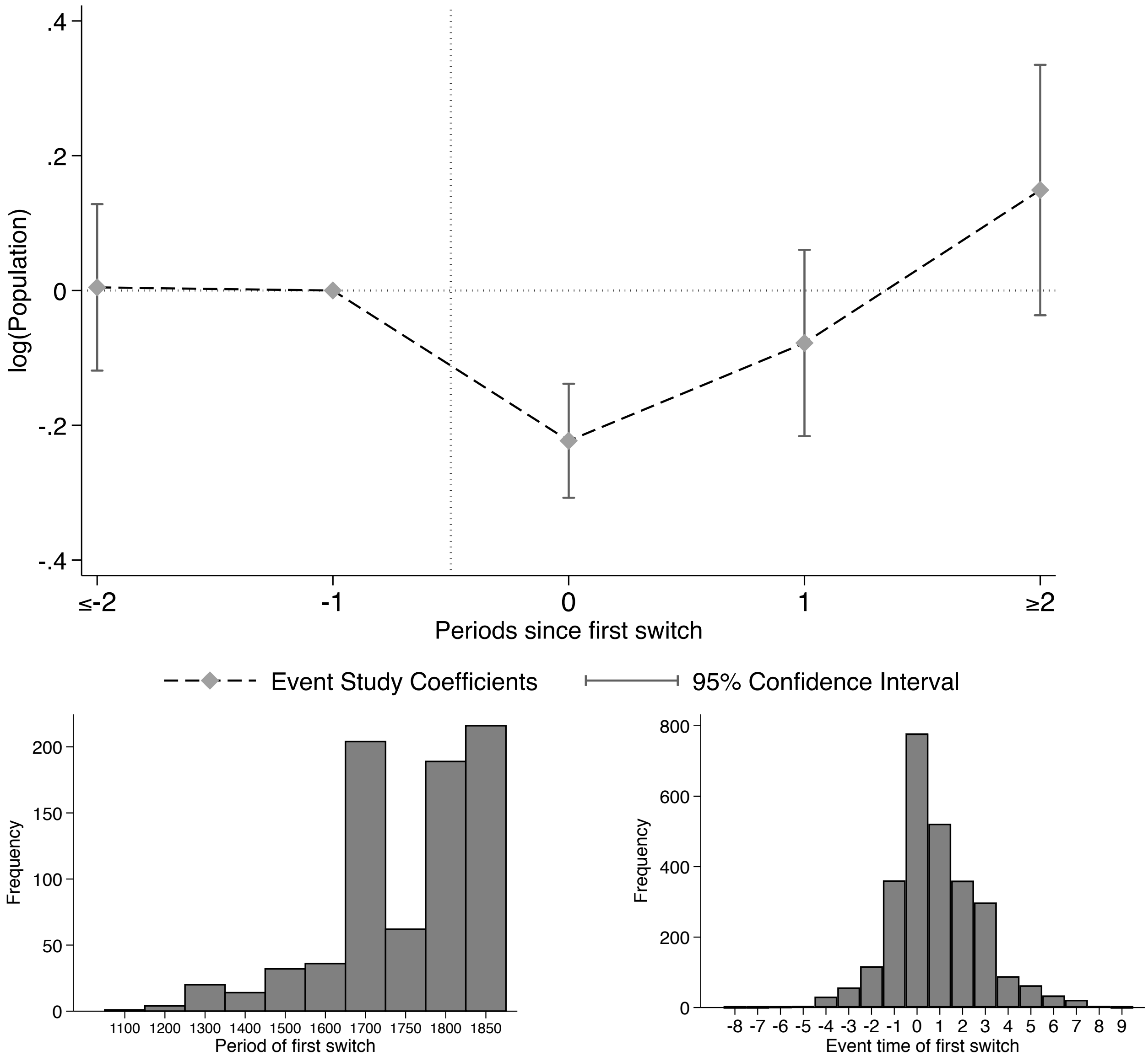
Notes: City switches per decade since city foundation averaged over 1000-1850 across Europe based on data from Bairoch et al. (1988) combined with Reed (2016). Markers scaled to population size in 1850.

Figure 5: Distribution of number of switching periods by city



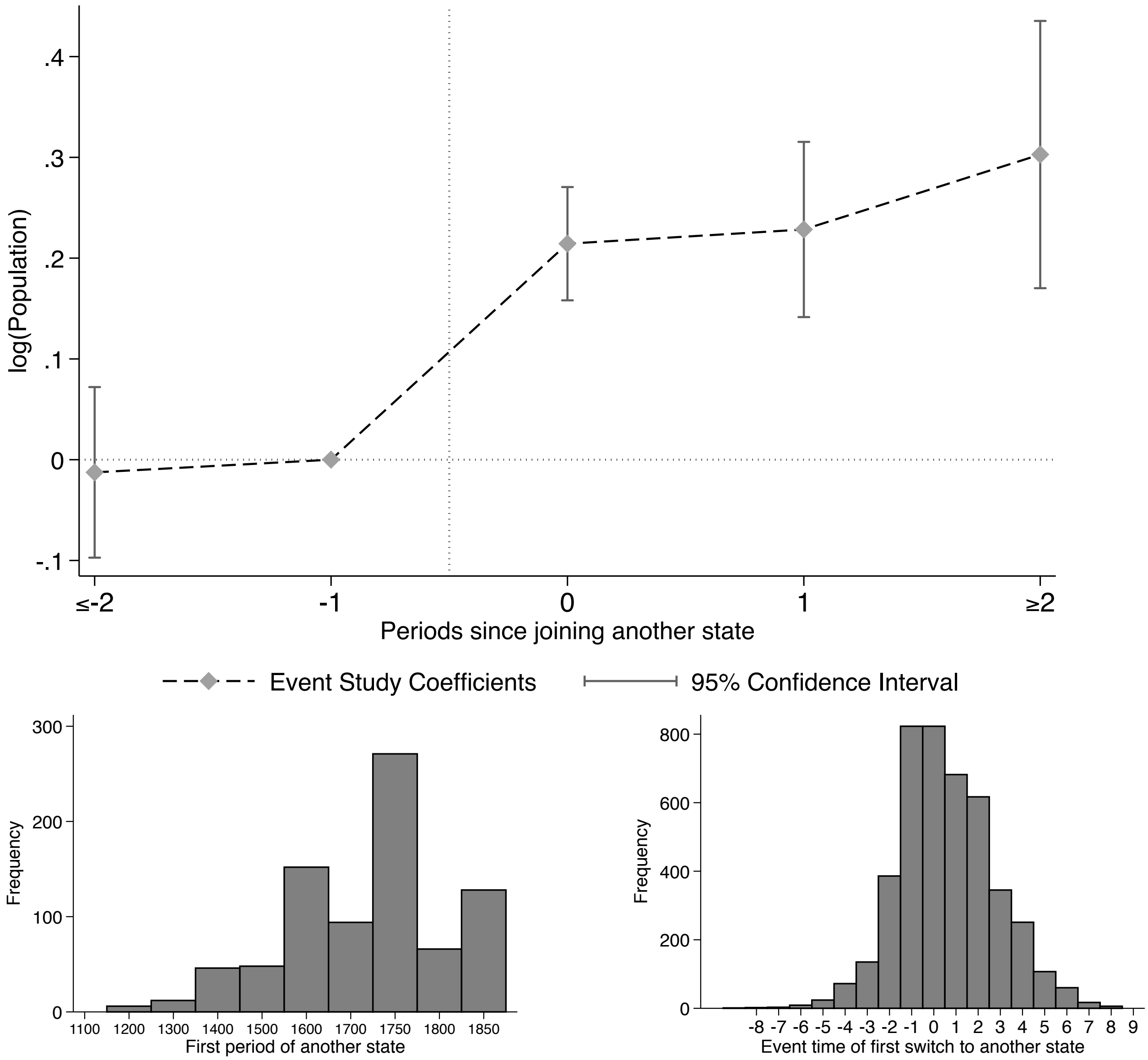
Notes: "Any switch": Distribution of number of periods in which a city experiences a switch over its entire existence. That is, the distribution of $\sum_t \mathbf{1}[S_{it} > 0]$. "Switch to another state": Distribution of number of periods in which a city experiences a switch to another state over its entire existence. That is, the distribution of $\sum_t \mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$.

Figure 6: Event study estimates of switching cost



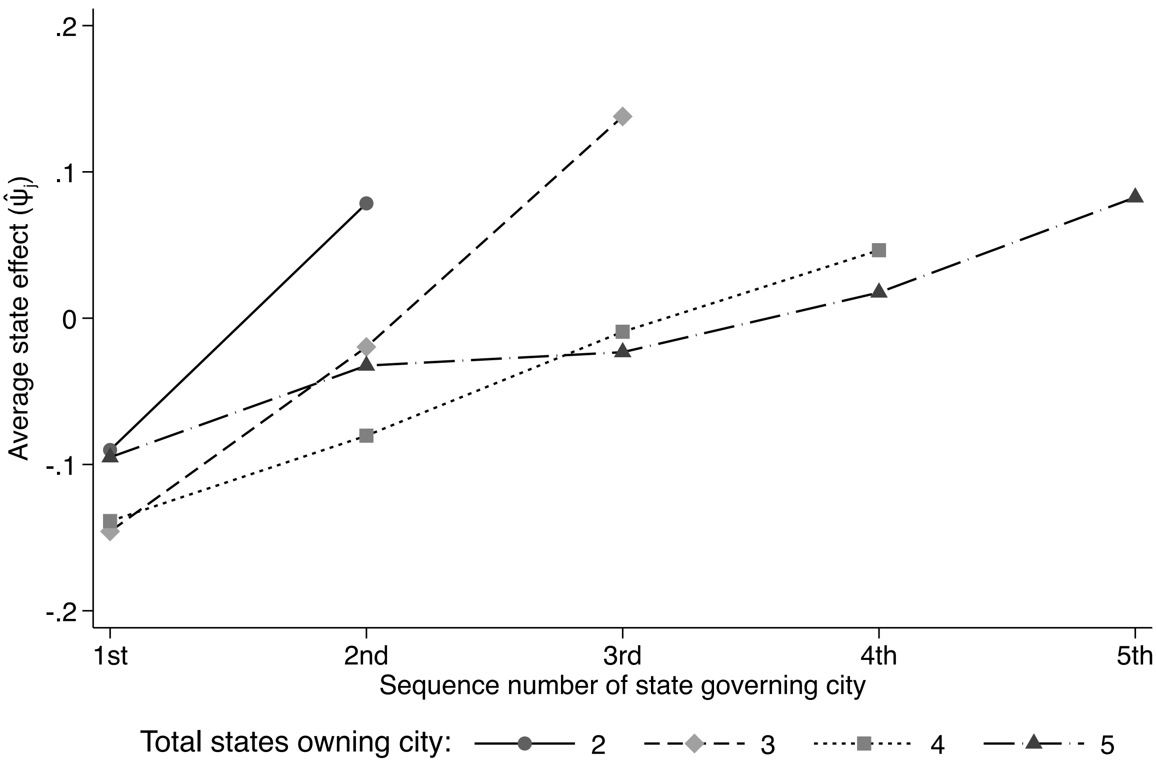
Notes: **Specification**: Event study estimates via OLS according to equation (2). **Sample**: All cities that experience one or two switching periods. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching once, the event time is the period of switching; for cities switching twice, the event time is the first period of switching. **Fixed effects**: included are city, period and state fixed effects. **Standard errors**: clustered on the city level.

Figure 7: Event study estimates of improved state quality



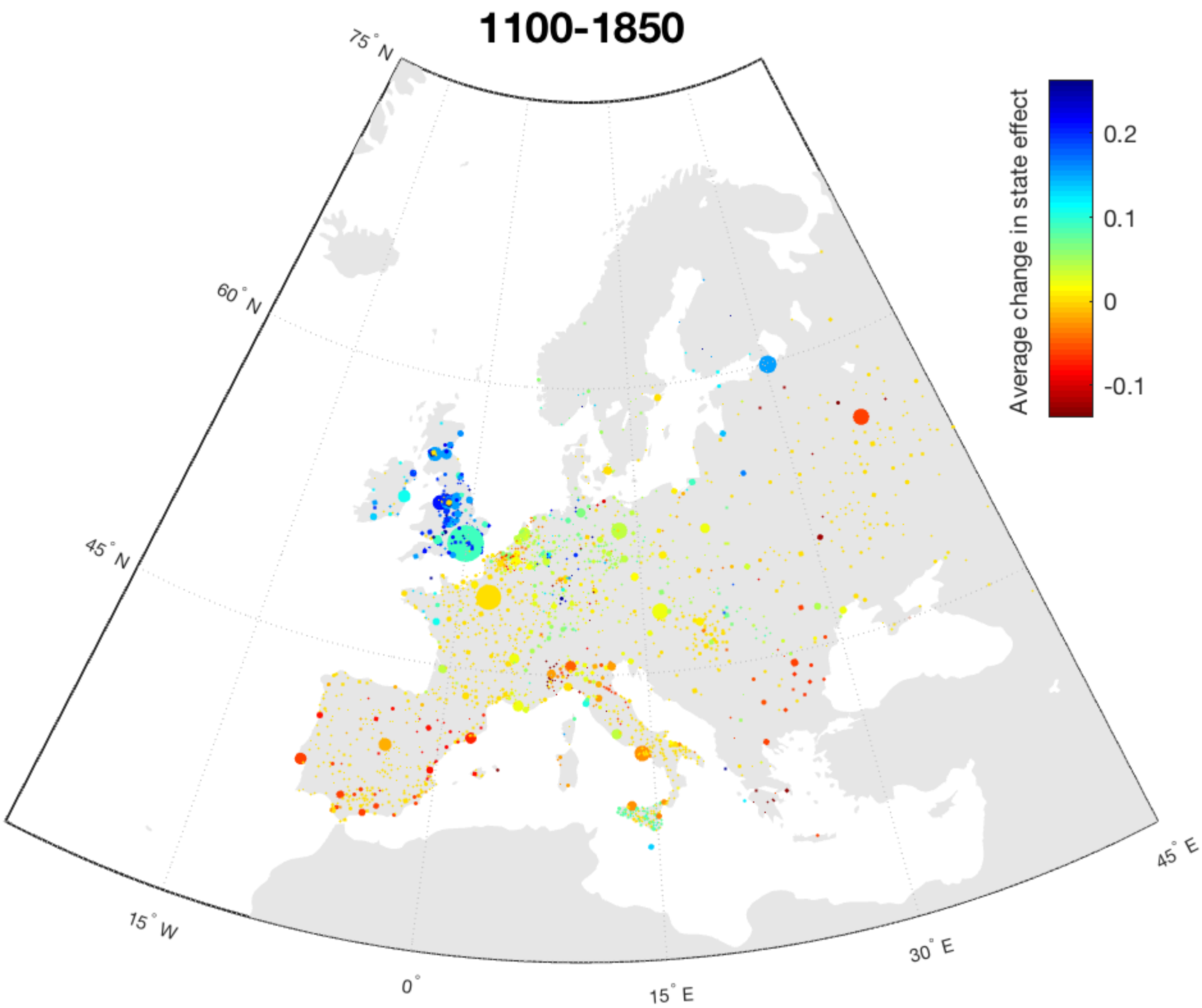
Notes: **Specification**: Event study estimates via OLS according to equation (3). **Sample**: All cities that experience one or two switches to another state. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching to another state once, the event time is the period of switching to another state; for cities switching to another state twice, the event time is the first period of switching to another state. **Control variable**: controlling for $\mathbf{1}[S_{it} > 0]$, i.e. all periods in which the city switched (ending up in a different state or the same state). **Fixed effects**: included are city and period fixed effects. **Standard errors**: clustered on the city level.

Figure 8: Improvement in state fixed effects over switching sequences



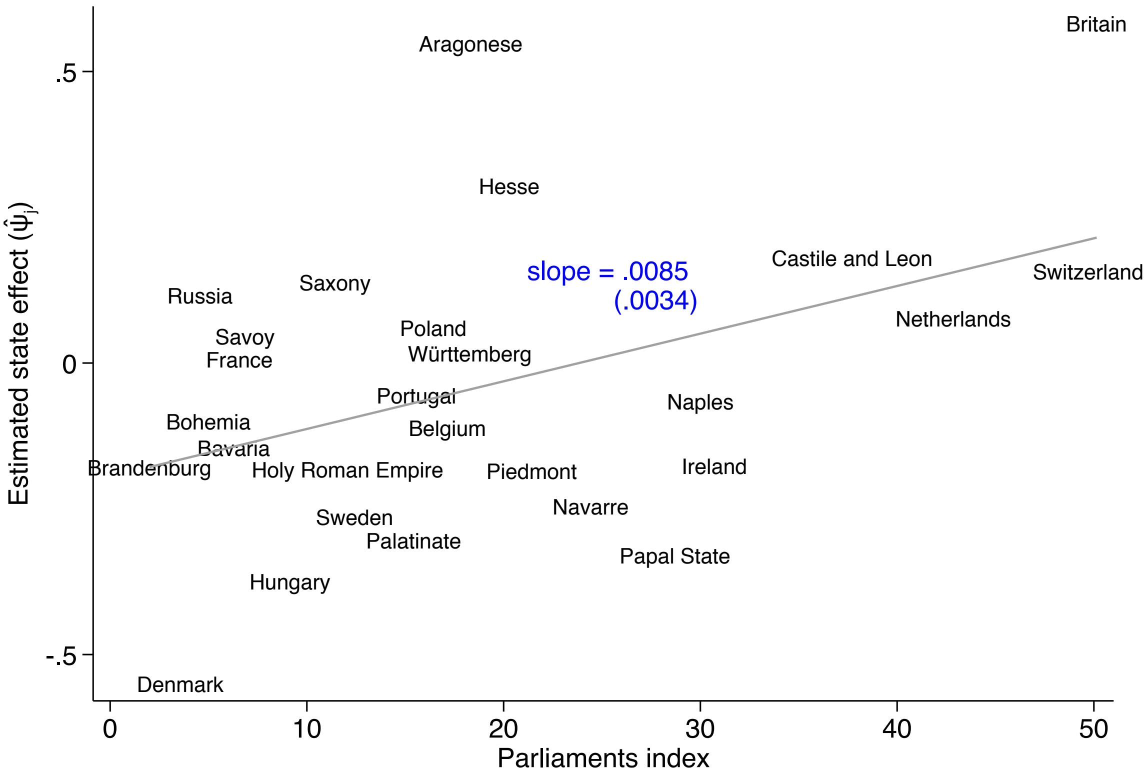
Notes: Average of state level fixed effect estimates ($\widehat{\psi}_j$) for cities of a given length of switching sequence over all city-period observations. We include cities that are in two to five different state over their lifetime (i.e. switch states one to four times).

Figure 9: Improvement in state effect over city lifetime across Europe



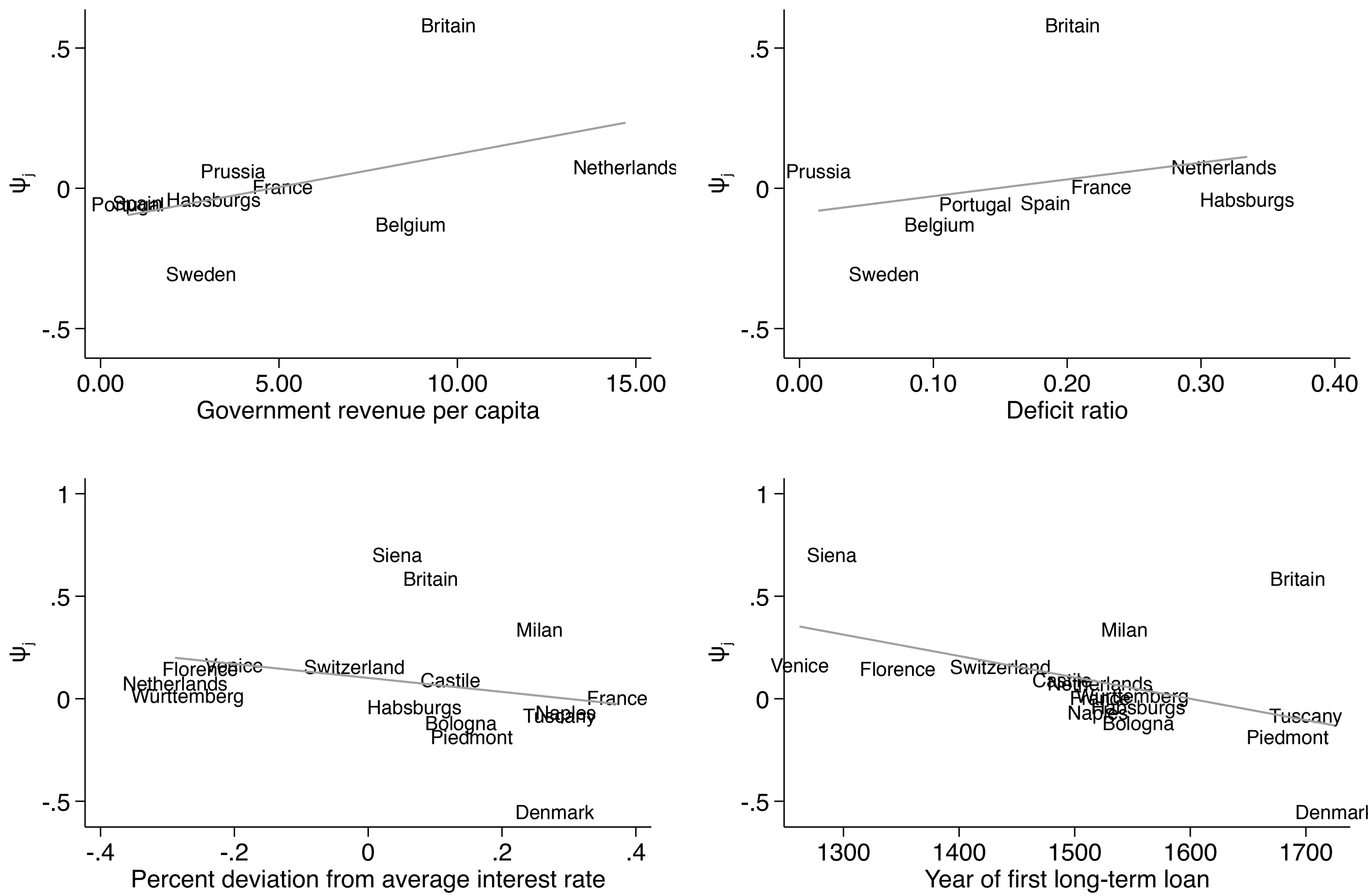
Notes: Average improvement in estimated state effects since city foundation. Estimates windsorized at 1st and 99th percentiles. Markers scaled to population size in 1850.

Figure 10: Correlation of state effects with parliamentary activity



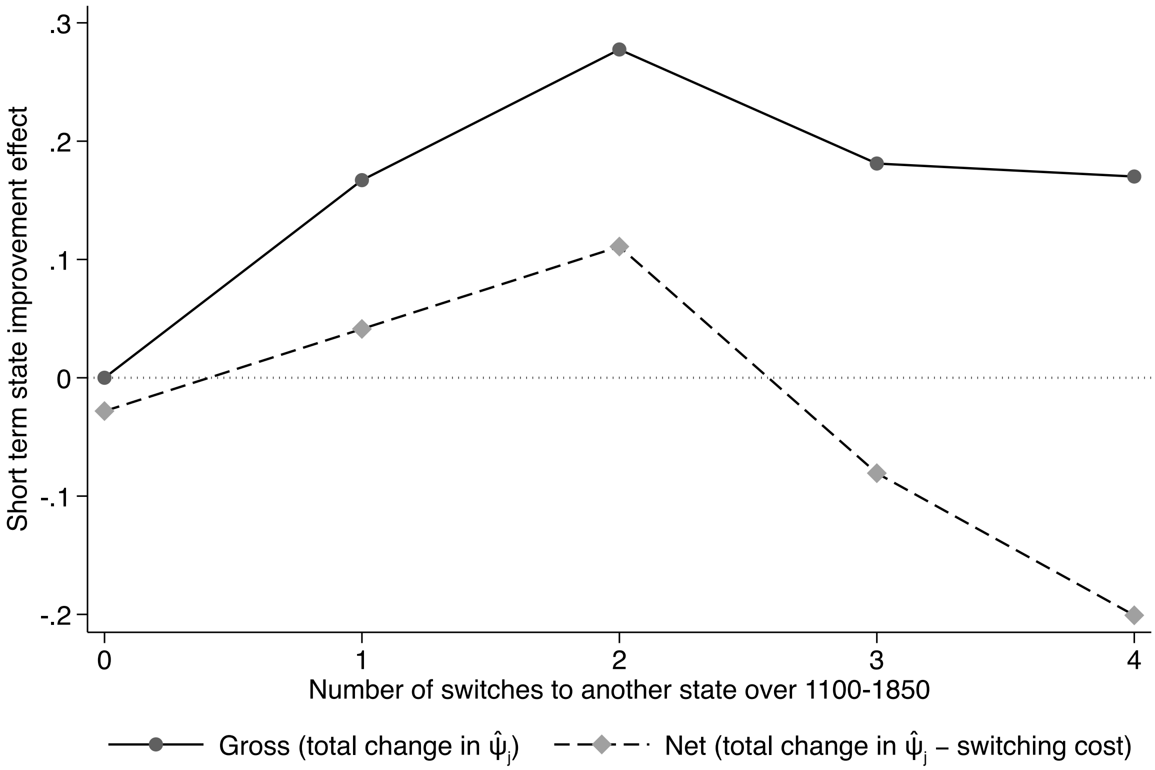
Notes: Scatter plot of estimated state fixed effects ($\widehat{\psi}_j$) against the average parliamentary activity index (Van Zanden et al., 2012) as a proxy of constraints on the executive of the state. We proceed as follows to match the states in the two datasets – the Atlas used to estimate the $\widehat{\psi}_j$ coefficients, and the Van Zanden et al. parliaments data: Aragonese is the average of Catalonia, Aragon, and Valencia. The Holy Roman Empire is the average of Germany and Austria. Britain is the average of England and Scotland. We match to Sicily to Naples and Sardinia to Savoy since they were the states in charge of each place for the longest period.

Figure 11: Correlations of estimated state effects with fiscal capacity proxies



Notes: Scatter plot of estimated state fixed effects ($\widehat{\psi}_j$) against four proxies of fiscal capacity of a state: (a) tax revenue per capita; (b) deficit ratio; (c) interest rate on long-term loans; (d) year of first long-term loan. Data for (a) and (b) are from Dincecco (2011); data for (c) and (d) are from Stasavage (2011).

Figure 12: Tradeoff between average improvement in states and cost of switching



Notes: Gross and net average improvement of state effects for cities switching zero to five times over the course of their history. The benefits are computed from the total improvement shown in Figure 8. The costs are estimated as the cumulative population loss due to switching (using both the indicator for switching and switches per decade) for each category of city (those that switch to another state a certain number of times).

A Appendix

A.1 Details on Quantifying State Improvement

Here, we describe in greater detail the regression results in Table 5. Our first approach to estimating a rise in the state effect is represented by θ. This parameter captures the average increase in the state effect when a switch to another state occurs across the entire switching sequence. Graphically, θ is the slope of the best linear fit through the microdata underlying Figure 8. Estimation results corresponding to equation (6) are displayed in Table 5. Columns (1) and (2) show estimates from our first approach. We estimate that a switch to another state along the sequence of switches, after controlling for period fixed effects, is associated with a state effect that is about 2.1% to 3.6% higher, significant at the 5% and 10% level, respectively. These estimates are smaller than the simple average increase in state effects documented earlier, suggesting that about half of the increase in state effects is due to states existing in later periods being better than those in earlier periods.[34] However, even when conditioning on the secular rise in the quality of states, cities that switch seem to benefit more strongly from the increase in state quality.

In our second approach, we duplicate all city-period observations that do not belong to either the first or the last state in a sequence of switches, assigning a post-indicator to one of the duplicated observations in each pair of duplicated city-periods as well as the last city-period observation in a city's sequence. This means we are essentially stacking all the segments from Figure 8 to estimate their average slope. This slope corresponds to the average increase in the state effect irrespective of what part of a sequence it is. In columns (3) and (4) we show that the second approach yields similar results as the first. The point estimate is 4.4% in both cases, significant at 5% and 10% significance levels.

The third approach consists of replacing h_{it} on the right-hand side of equation (6) with $h_{it}/H(i)$. In this way, we normalize the length of all sequences to one, meaning we are estimating the increase in the state effect over the entire sequence, no matter the length of the sequence. Graphically, we are compressing all sequences in Figure 8 to range from 1 to 2, estimating the slope over all these compressed sequences. In (5) and (6), we estimate this average increase in the average $\widehat{\psi}_j$ over the whole switching sequence. The estimates suggest that the average switching sequence undergone by cities being governed by two to five states

[34] In a descriptive sense, this is true only when state effects are weighted by the number of cities or the city population they hold; across states without weighting, there is no secular trend in the state effects.

over their lifetime is about 16.2%.

It may be possible that switching states may have diminishing returns. To this end, we first include a quadratic term of h_{it}, which is shown to be negative and significant in columns (7) and (8). Second, we test whether sequences of different lengths see different increases in the state effect for a given switch to another state. We do this by replacing h_{it} with the interaction term $h_{it} \times \kappa_{H(i)}$. Graphically, we are now estimating the slope of each sequence in the figure separately. Columns (9) and (10) show that longer switching sequences tend to experience smaller increases in fixed effects: as we move from a total of two switches to a total of five, the increase per switch falls from and 6.9% to 1.9% in the specification with sequence fixed effects and period fixed effects. While we only have enough power to reject an equal increase for cities with sequence length of three from those with sequence length of five, these results are suggestive of decreasing returns to switching.

Each of these sets of results provides evidence for exposure to systematically better states through switching. Interestingly, the unconditional average of the improvement in state effects through switching is about twice as high as the average conditional on period fixed effects, suggesting that about half of the effect arises from the overall pool of states improving over time, and another half from cities disproportionately switching to better states for a pool of states in a given period.

A.2 Further Descriptive Statistics

A.3 Further Results

Table A.1: Decomposition results with intensive margin of switching

	Dependent variable: log(population)				
	(1)	(2)	(3)	(4)	(5)
Switching intensity per decade	-0.151***	-0.127***	-0.215***	-0.126***	-0.112***
	(0.021)	(0.019)	(0.022)	(0.021)	(0.020)
Switch to another state			0.083***	-0.004	-0.006
			(0.015)	(0.019)	(0.019)
City FE	X	X	X	X	
Period FE	X	X	X	X	X
State FE		X		X	
City × state FE					X
Std. dev. of **Xb**	.043	.036	.057	.036	.
Std. dev. of $\widehat{\psi}_j$	–	.322	–	.323	–
Adjusted R^2	0.69	0.73	0.68	0.73	0.78
RMSE	0.51	0.48	0.52	0.48	0.43
City N	1,949	1,949	2,063	1,949	1,949
State N	–	89	–	89	–
Total N	8,629	8,629	10,458	8,629	8,629

Notes: **Specification**: Regression estimates according to specification (4) for 2,182 cities and ten periods (1100 to 1700 in centuries; 1750 to 1850 in half-centuries). The number of cities is lower than the total number due to collinearity created by the fixed effects (i.e. some city-period observations become collinear with a combination of fixed effects). **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Independent variables**: "Switching intensity per decade" is the average number of switches per decade over the period, i.e. normalized to take into account varying period length. It corresponds to S_{it}. "Switch to another state" is the indicator for $\mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$. Together, the independent variable(s) of a given model are denoted as **Xb**. **Standard errors**: clustered on the city level.

Table A.2: Decomposition results with intensive and extensive margin

	Dependent variable: log(population)				
	(1)	(2)	(3)	(4)	(5)
Switching intensity per decade	-0.123***	-0.038	-0.122***	-0.038	-0.036
	(0.027)	(0.026)	(0.027)	(0.026)	(0.025)
Switching indicator	-0.029	-0.096***	-0.061***	-0.111***	-0.096***
	(0.018)	(0.019)	(0.022)	(0.021)	(0.020)
Switch to another state			0.073***	0.039*	0.034*
			(0.022)	(0.021)	(0.020)
City FE	X	X	X	X	
Period FE	X	X	X	X	X
State FE		X		X	
City × state FE					X
Std. dev. of **Xb**	.045	.055	.053	.059	.
Std. dev. of $\widehat{\psi}_j$	–	.326	–	.325	–
Adjusted R^2	0.69	0.73	0.69	0.73	0.78
RMSE	0.51	0.48	0.51	0.48	0.43
City N	1,949	1,949	1,949	1,949	1,949
State N	–	89	–	89	–
Total N	8,629	8,629	8,629	8,629	8,629

Notes: **Specification**: Regression estimates according to specification (4) for 2,182 cities and ten periods (1100 to 1700 in centuries; 1750 to 1850 in half-centuries). The number of cities is lower than the total number due to collinearity created by the fixed effects (i.e. some city-period observations become collinear with a combination of fixed effects). **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Independent variables**: "Switching intensity per decade" is the average number of switches per decade over the period, i.e. normalized to take into account varying period length. It corresponds to S_{it}. "Switching indicator" is an indicator variable for $\mathbf{1}[S_{it} > 0]$. "Switch to another state" is the indicator for $\mathbf{1}[\mathbf{J}(i,t) \neq \mathbf{J}(i,t-1)]$. Together, the independent variable(s) of a given model are denoted as **Xb**. **Standard errors**: clustered on the city level.

Table A.3: The cost of switching: alternative specifications

Dependent variable:	log(population)						population growth		log(1 + population)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Switching indicator	-0.072***	-0.064***	-0.067***	-0.072***	-0.067***	-0.054***			-0.053***	-0.087***
	(0.019)	(0.018)	(0.014)	(0.014)	(0.017)	(0.016)			(0.011)	(0.012)
Switch in period (first diff.)							-0.045**	-0.046**		
							(0.020)	(0.020)		
City FE	X	X	X	X	X	X	X	X	X	X
Period FE			X	X	X	X	X	X	X	X
Period-country FE	X	X								
State FE	X	X					X	X	X	X
Majority state FE			X	X						
Fractional state FE					X	X				
Control variables		X		X		X		X		X
City N	2,058	2,058	2,060	2,060	2,063	2,063	1,583	1,583	2,182	2,060
Total N	10,418	10,418	10,444	10,444	10,458	10,458	7,882	7,882	21,810	10,439
Adjusted R^2	0.75	0.75	0.72	0.73	0.73	0.73	0.16	0.17	0.74	0.73

Notes: **Specification**: Regression estimates according to specification (4) for up to 2,182 cities and ten periods (1100 to 1700 in centuries; 1750 to 1850 in half-centuries). The number of cities is lower than the total number due to collinearity created by the fixed effects or control variables (i.e. some city-period observations become collinear with a combination of fixed effects). **Dependent variables**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. Population growth is the growth rate over the period, i.e. $(Y_{it} - Y_{it-1})/Y_{it-1}$. log(1 + population), where population was set to zero in case it was missing, jointly captures the extensive (e.g. city foundation) and intensive margins. Note that the constructed distance controls (see below) are still missing when populations are zero since the distance percentile is meaningless when a number of cities of zero population, so specification (10) has less observations than (9). **Independent variables**: "Switch in period" is an indicator variable for whether the city switched at least once in the period, i.e. $\mathbf{1}[S_{it} > 0]$. "Switch in period (first difference)" is $\mathbf{1}[S_{it} > 0] - \mathbf{1}[S_{it-1} > 0]$. **Additional fixed effects**: Country-period fixed effects are period fixed effects interacted with Cold War country borders, as coded by Bairoch et al. (1988). Majority state fixed effects use the state that holds the city over the majority of the period as opposed to the state that holds the city at the time of population measurement by Bairoch (our default). Fractional state "fixed" effects are shares of states in power over the period, which sum up to one. **Control variables**: these include the number of reported plague outbreaks (Büntgen et al., 2012), the distance to the nearest border, the percentile distance to the state's weighted centroid (the latter two are time-varying because borders move over time), and the distance to the continent-wide center of city population (which is time-varying because of differential urban development in different parts of Europe). The latter three control variables are constructed based on Reed (2016). **Standard errors**: clustered on the city level.

Table A.4: Switching to new entrant

	Dependent variable: log(population)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post-switching to another state	0.205***	0.205***	0.126***	0.198***	0.199***	0.120***
	(0.029)	(0.024)	(0.027)	(0.030)	(0.024)	(0.027)
Post-switching to new entrant				0.107	0.093	0.175*
				(0.110)	(0.083)	(0.093)
City FE	X	X	X	X	X	X
State FE			X			X
Period FE	X	X	X	X	X	X
Sequence lengths $H(i)$	{1,2}	All	All	{1,2}	All	All
Switching event type	First	First	First	First	First	First
City N	823	2,063	2,060	823	2,063	2,060
Total N	4,363	10,458	10,437	4,363	10,458	10,437
Adjusted R^2	0.69	0.68	0.72	0.69	0.68	0.72

Notes: **Specification**: Generalized difference-in-difference estimates via OLS according to the following equation: $y_{it} = \alpha_i + \gamma_t + \mathbf{1}[S_{it} > 0]\beta + \mathbf{1}[t \geq e_i^{\text{AnotherState}}]\pi + \mathbf{1}[t \geq e_i^{\text{AnotherState}}] \times \mathbf{1}[\mathbf{J}(i,t) \text{ is new entrant}]\eta + \varepsilon_{it}$. The parameter η captures how much larger the benefit is to switching to a new entrant compared to some other state. **Samples**: Sequence length $H(i) = \max h_{it}$ where $h_{it} = \sum_{s \leq t} \mathbf{1}[\mathbf{J}(i,s) \neq \mathbf{J}(i,s-1)]$. Thus, for example, sequence lengths $\{1,2\}$ denotes all cities that switched states once or twice, while "all" denotes the inclusion of cities of any number of switches over their lifetime. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching to another state once, the event time is the period of switching to another state; for cities switching to another state multiple times, the event time is the first period of switching to another state. **Control variable**: controlling for $\mathbf{1}[S_{it} > 0]$, i.e. all periods in which the city switched (ending up in a different state or the same state). **Fixed effects**: included are city and period fixed effects. **Standard errors**: clustered on the city level.

Table A.5: Event study estimates of switching cost

	Dependent variable: log(population)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[t \leq e_i^{\text{Switch}} - 2]$	0.005	-0.017	-0.006	-0.030	0.008	0.013
	(0.063)	(0.057)	(0.053)	(0.026)	(0.015)	(0.011)
$\mathbf{1}[t = e_i^{\text{Switch}}]$	-0.223***	-0.106***	-0.073**	-0.144***	-0.069***	-0.050***
	(0.043)	(0.038)	(0.036)	(0.016)	(0.012)	(0.009)
$\mathbf{1}[t = e_i^{\text{Switch}} + 1]$	-0.078	0.012	0.056	-0.059**	-0.025*	0.008
	(0.070)	(0.053)	(0.047)	(0.023)	(0.014)	(0.012)
$\mathbf{1}[t \geq e_i^{\text{Switch}} + 2]$	0.149	0.263***	0.235***	0.099***	0.068***	0.042***
	(0.095)	(0.069)	(0.059)	(0.029)	(0.016)	(0.010)
City FE	X	X	X	X	X	X
Period FE	X	X	X	X	X	X
State FE	X	X	X	X	X	X
Sequence length $H(i)$	{1,2}	{1,2,3}	{1,2,3,4}	{1,2}	{1,2,3}	{1,2,3,4}
Switching event type	First	First	First	All	All	All
City N	706	996	1,262	711	1,002	1,264
Total N	2,660	4,265	6,000	4,462	9,286	16,198
Adjusted R^2	0.71	0.70	0.70	0.75	0.74	0.74

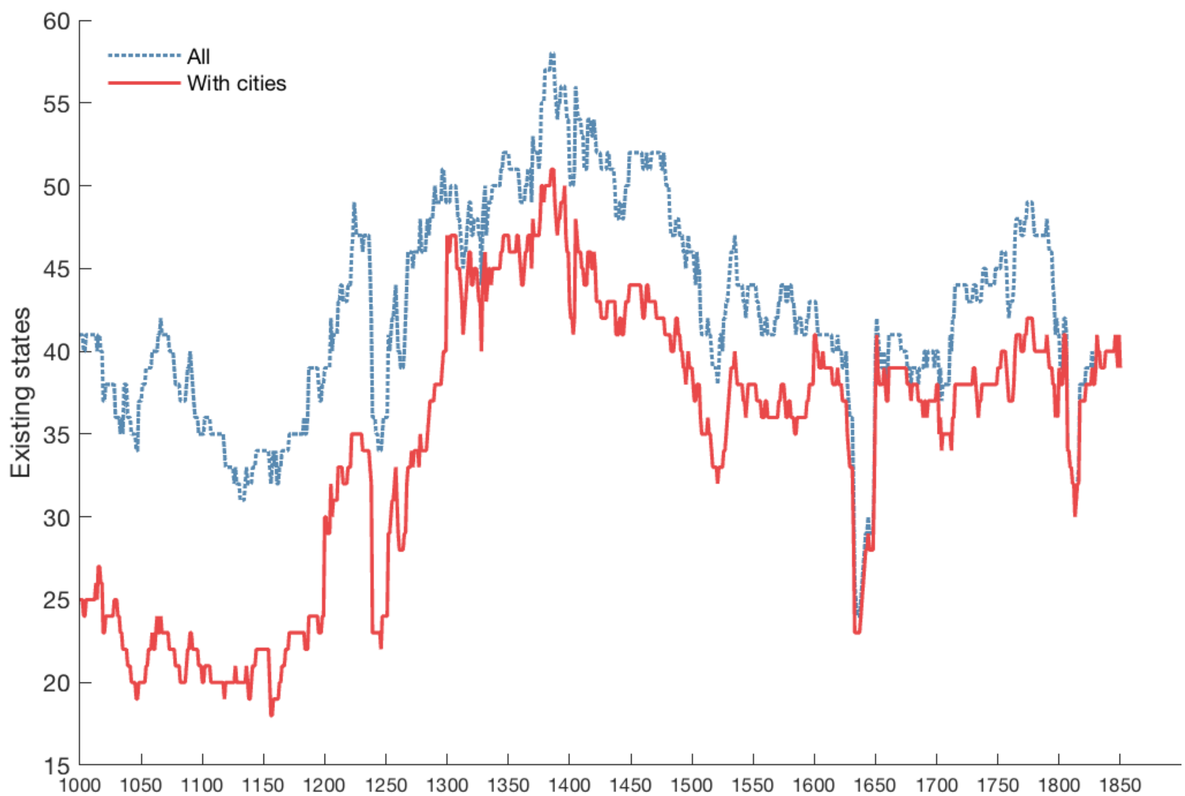
Notes: **Specification**: Event study estimates via OLS according to equation (2). **Samples**: Sequence length $H(i) = \max h_{it}$ where $h_{it} = \sum_{s \leq t} \mathbf{1}[\mathbf{J}(i,s) \neq \mathbf{J}(i,s-1)]$. Thus, for example, sequence lengths $\{1,2,3\}$ denotes all cities that switched once, twice, or three times over their entire history. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching once, the event time is the period of switching; for cities switching multiple times, the event time is the first period of switching in columns (1)-(3); and all periods of switching in columns (4)-(6), assigned across the duplicate time series. **Fixed effects**: included are city, period and state fixed effects. **Standard errors**: clustered on the city level.

Table A.6: Event study specifications of improved state quality and switching cost

	Dependent variable: log(population)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbf{1}[t \leq e_i^{\text{AnotherState}} - 2]$	0.014	0.023	0.025	0.038	0.018	0.027**
	(0.040)	(0.038)	(0.037)	(0.024)	(0.015)	(0.011)
$\mathbf{1}[t \geq e_i^{\text{AnotherState}}]$	0.205***	0.198***	0.188***	0.152***	0.120***	0.043***
	(0.029)	(0.027)	(0.026)	(0.020)	(0.015)	(0.013)
Switching penalty: $\mathbf{1}[S_{it} > 0]$	-0.127***	-0.127***	-0.118***	-0.075***	-0.085***	-0.086***
	(0.021)	(0.019)	(0.018)	(0.021)	(0.020)	(0.019)
City FE	X	X	X	X	X	X
Period FE	X	X	X	X	X	X
Sequence lengths $H(i)$	{1,2}	{1,2,3}	{1,2,3,4}	{1,2}	{1,2,3}	{1,2,3,4}
Switching event type	First	First	First	All	All	All
City N	823	1,082	1,188	823	1,082	1,185
Total N	4,363	6,269	7,107	6,381	12,099	15,442
Adjusted R^2	0.69	0.68	0.69	0.71	0.69	0.76

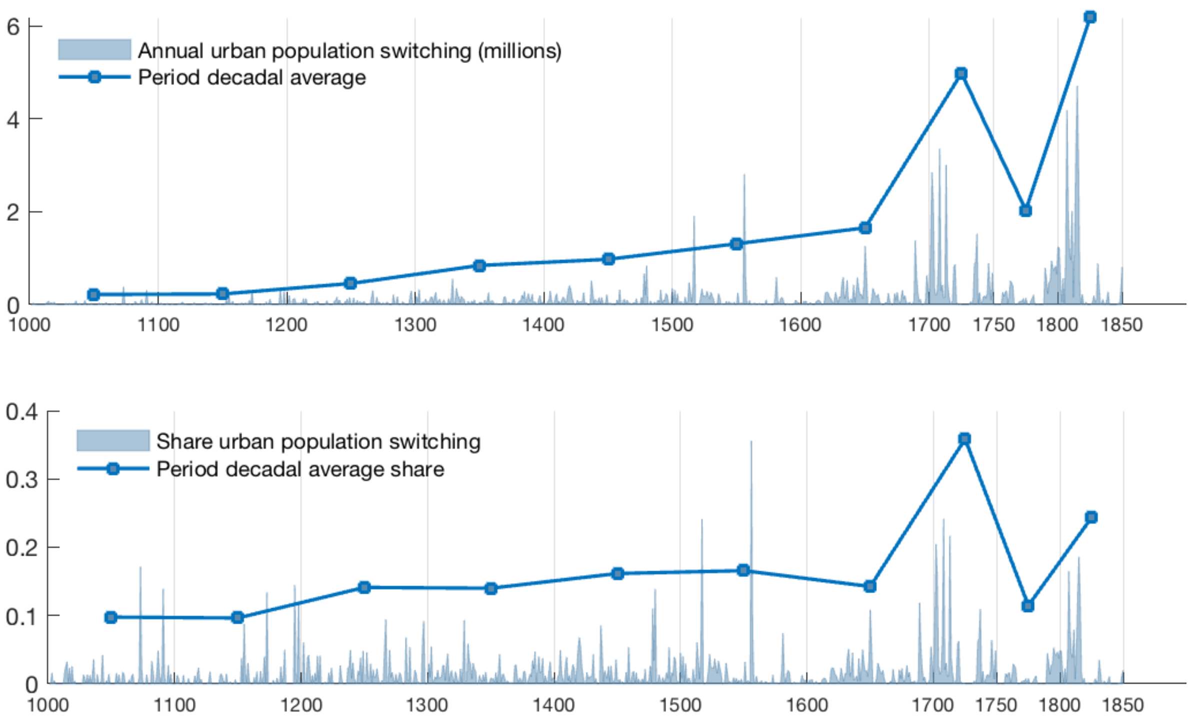
Notes: **Specification**: Event study estimates via OLS according to equation (3). **Samples**: Sequence length $H(i) = \max h_{it}$ where $h_{it} = \sum_{s \leq t} \mathbf{1}[\mathbf{J}(i,s) \neq \mathbf{J}(i,s-1)]$. Thus, for example, sequence lengths $\{1,2,3\}$ denotes all cities that switched states once, twice, or three times over their entire history. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching to another state once, the event time is the period of switching to another state; for cities switching to another state multiple times, the event time is the first period of switching to another state in columns (1)-(3); and all periods of switching to another state in columns (4)-(6), assigned across the duplicate time series. **Control variable**: controlling for $\mathbf{1}[S_{it} > 0]$, i.e. all periods in which the city switched (ending up in a different state or the same state). **Fixed effects**: included are city and period fixed effects. **Standard errors**: clustered on the city level.

Figure A.1: Number of existing states with and without cities, 1000-1850



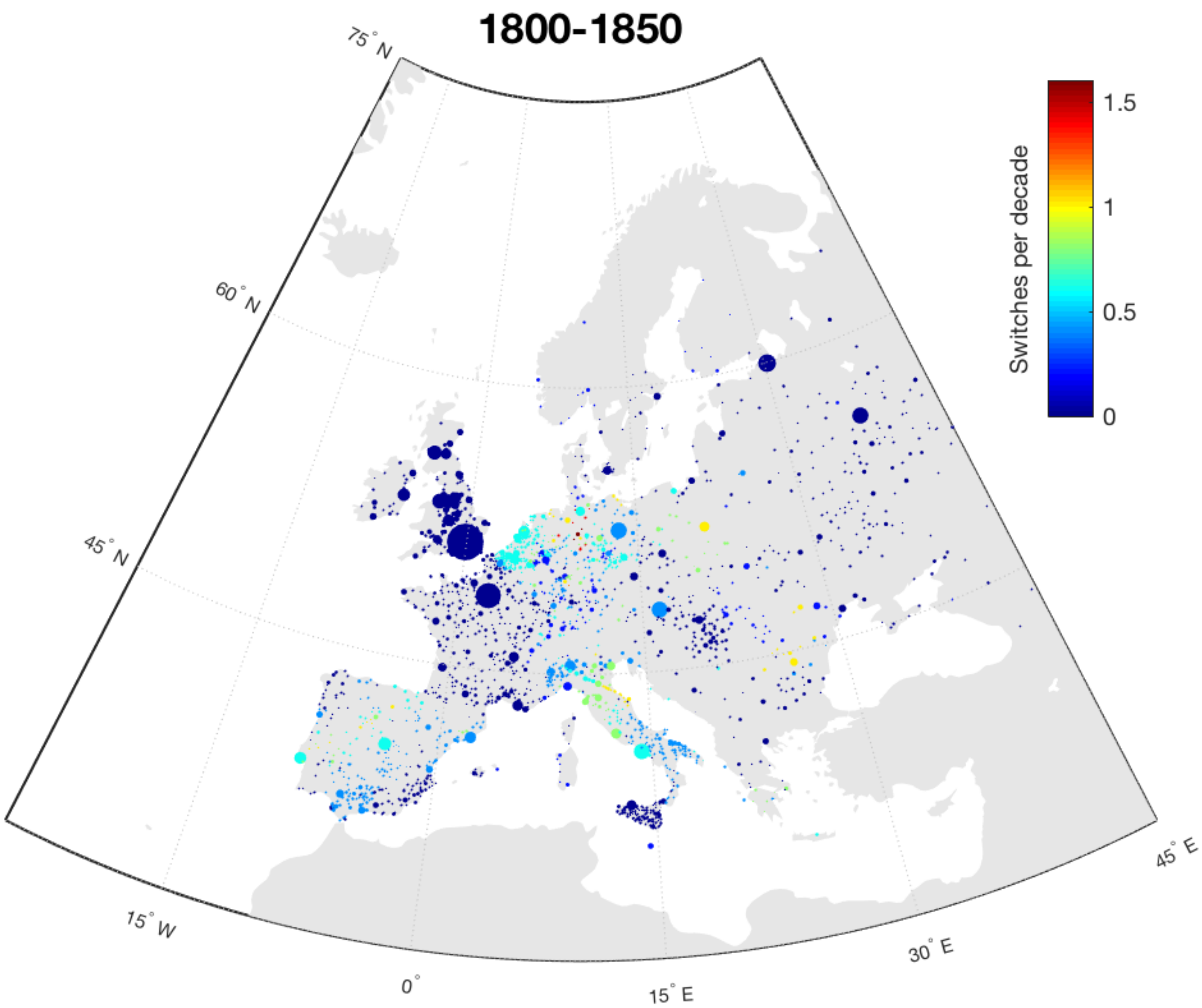
Notes: Number of existing states with and without cities each year between 1000-1850.

Figure A.2: Number and share of city population switching over 1100-1850



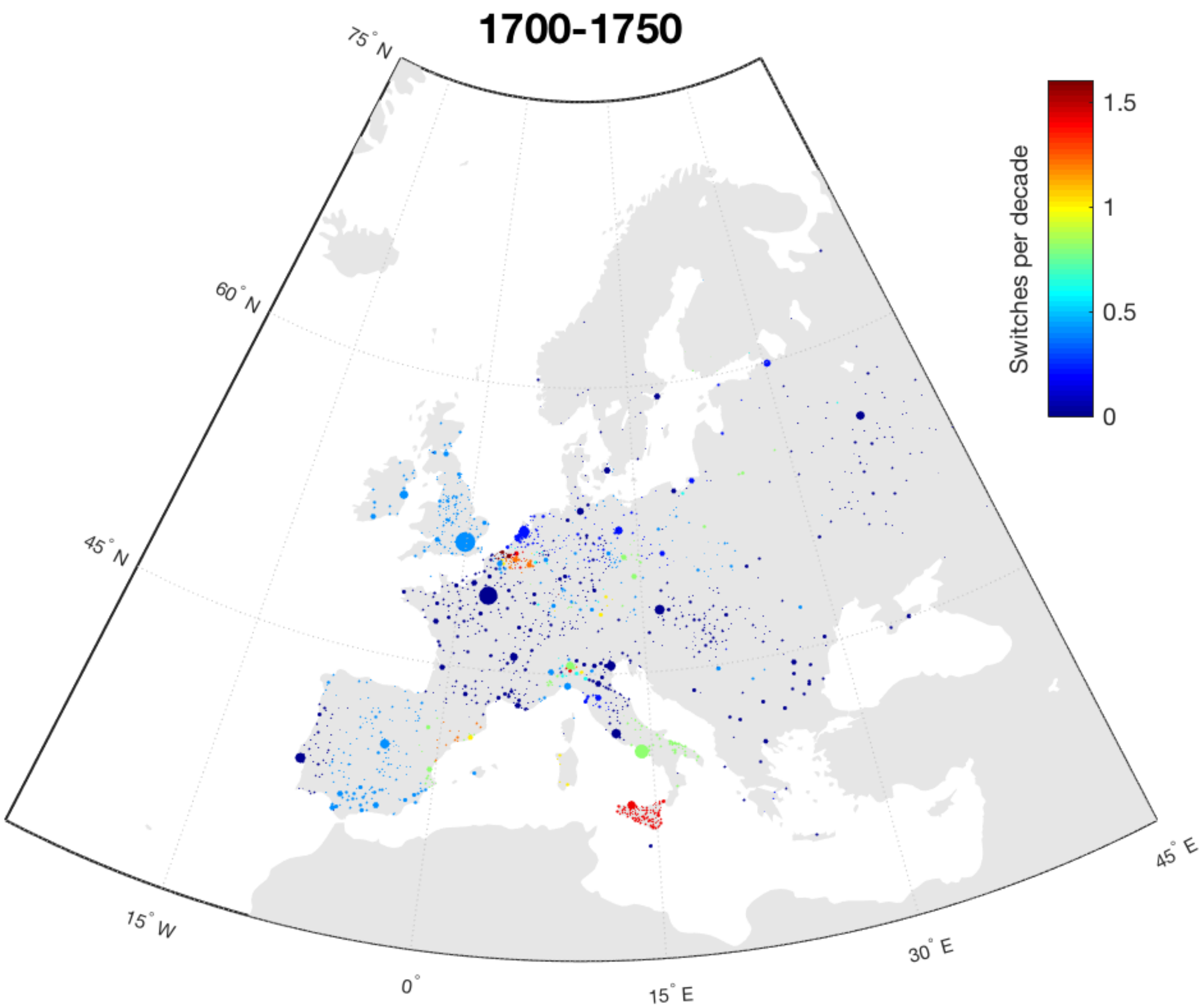
Notes: Annual number and share of city dwellers switching in each year as well as decadal averages per Bairoch period.

Figure A.3: Geocoded city population data and switching intensity: 1800-1850



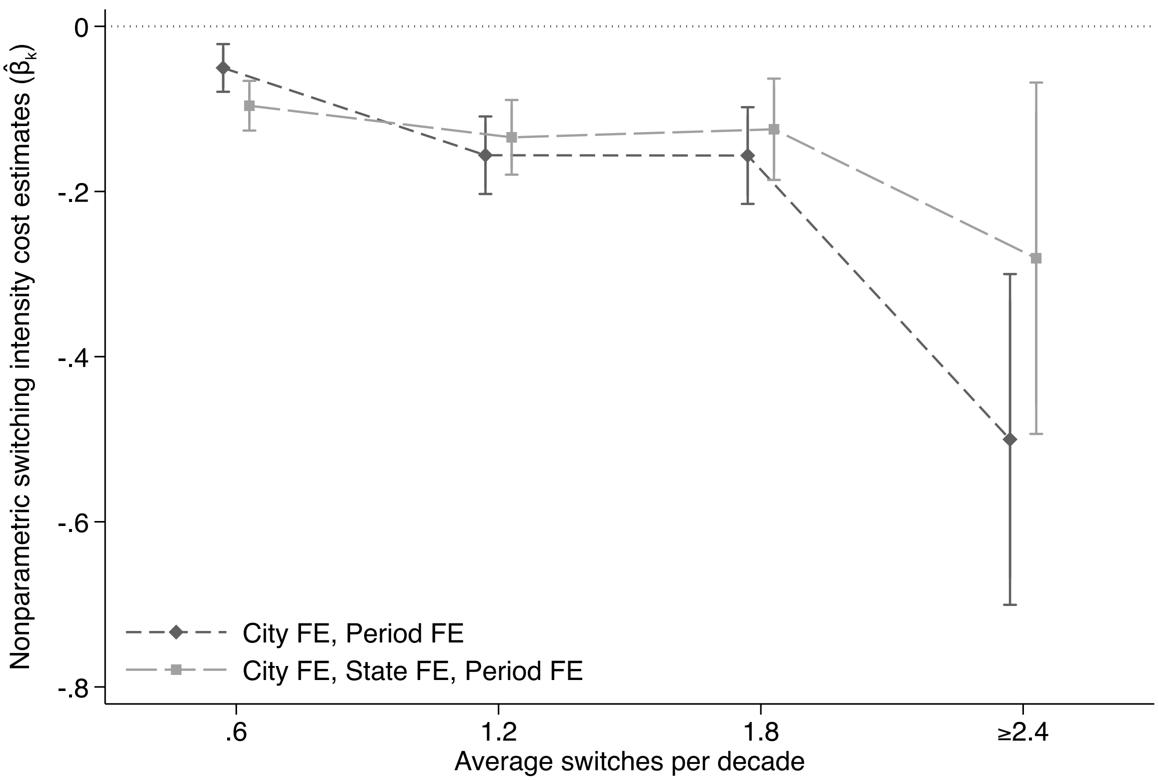
Notes: Number of cities switching per decade (i.e. switching intensity) in 1800-1850 across Europe based on data from Bairoch et al. (1988) combined with Reed (2016). Markers scaled to population size.

Figure A.4: Geocoded city population data and switching intensity: 1700-1750



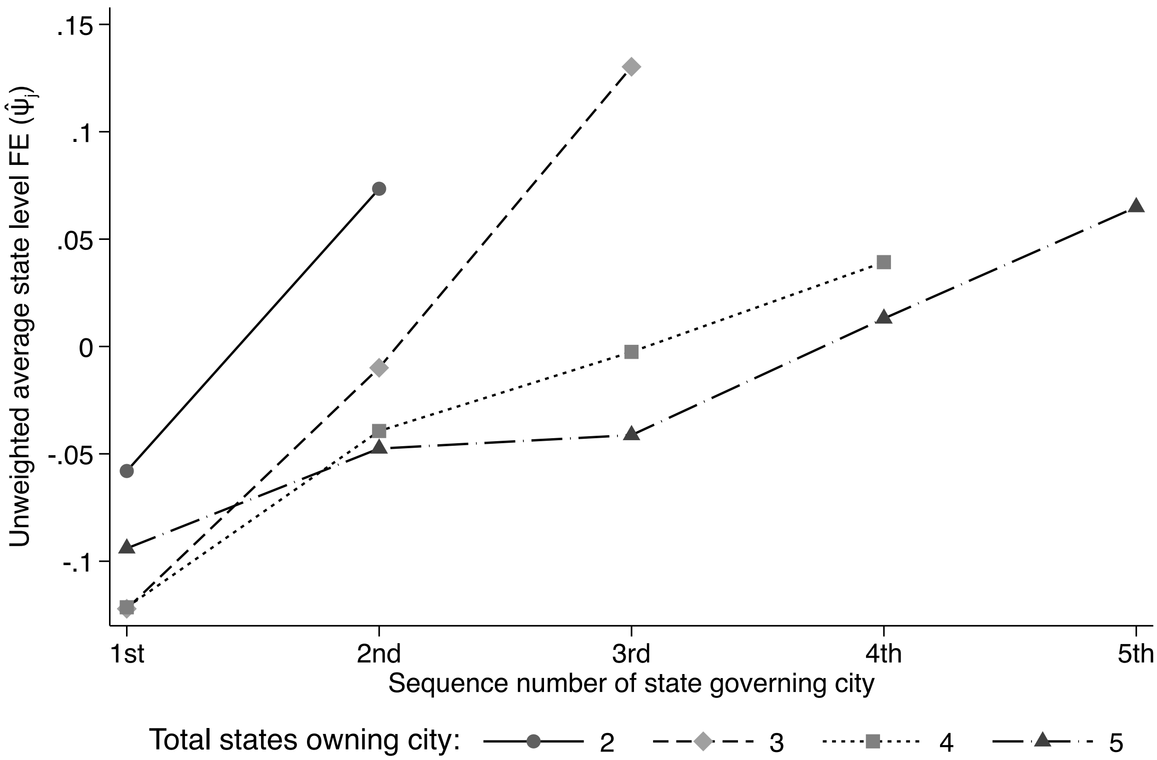
Notes: Number of cities switching per decade (i.e. switching intensity) in 1700-1750 across Europe based on data from Bairoch et al. (1988) combined with Reed (2016). Markers scaled to population size.

Figure A.5: Nonparametric switching intensity cost estimate



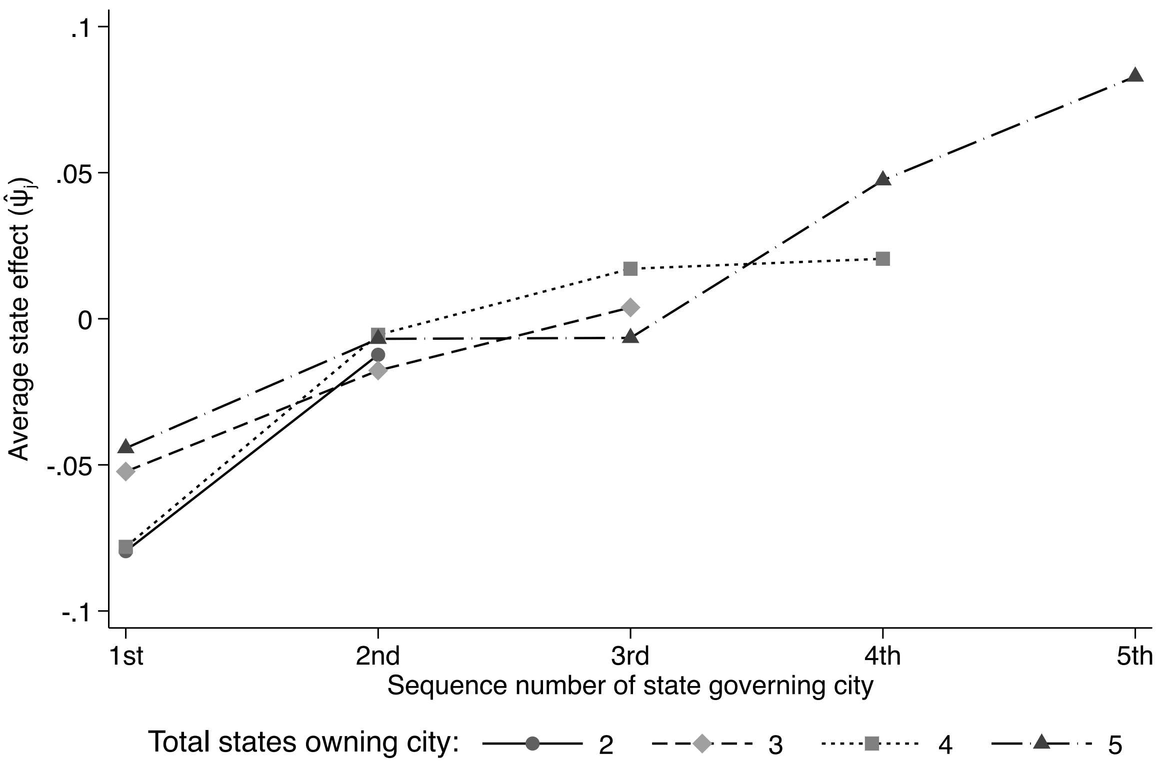
Notes: Nonparametric Estimates of switching intensity on city population according to regression (4), but with $\mathbf{1}[S_{it} > 0]\beta$ replaced with $\sum_{k=2}^{4} S_{it} \in [s_{k-1}, s_k)\beta_k$, where $s_k \in \{0.6, 1.2, 1.8, 2.4, \infty\}$ for $k \in \{1, ..., 4\}$, respectively.

Figure A.6: Improvement in state fixed effects over switching sequences (only switches)



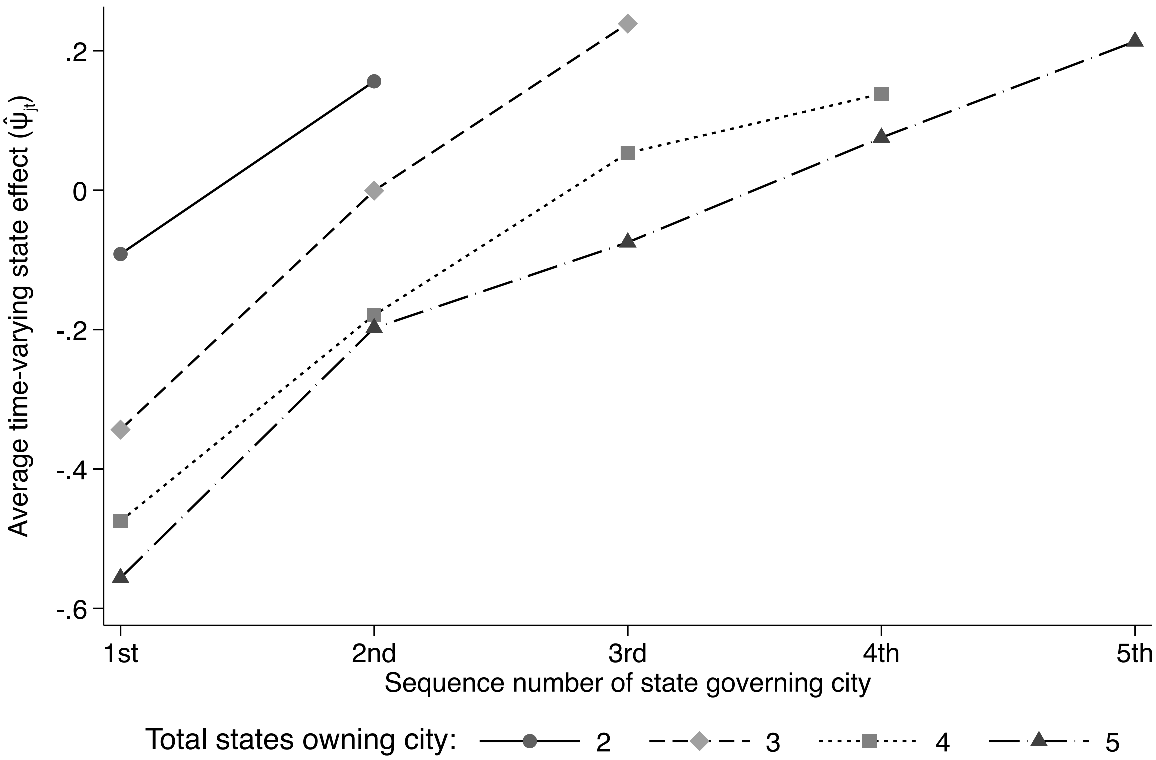
Notes: Average of state level fixed effect estimates ($\widehat{\psi}_j$) for cities of a given length of switching sequence, using only city-period observations that switched from one state to another. We include cities that are in two to five different state over their lifetime.

Figure A.7: Improvement in state fixed effects over switching sequences (without Britain)



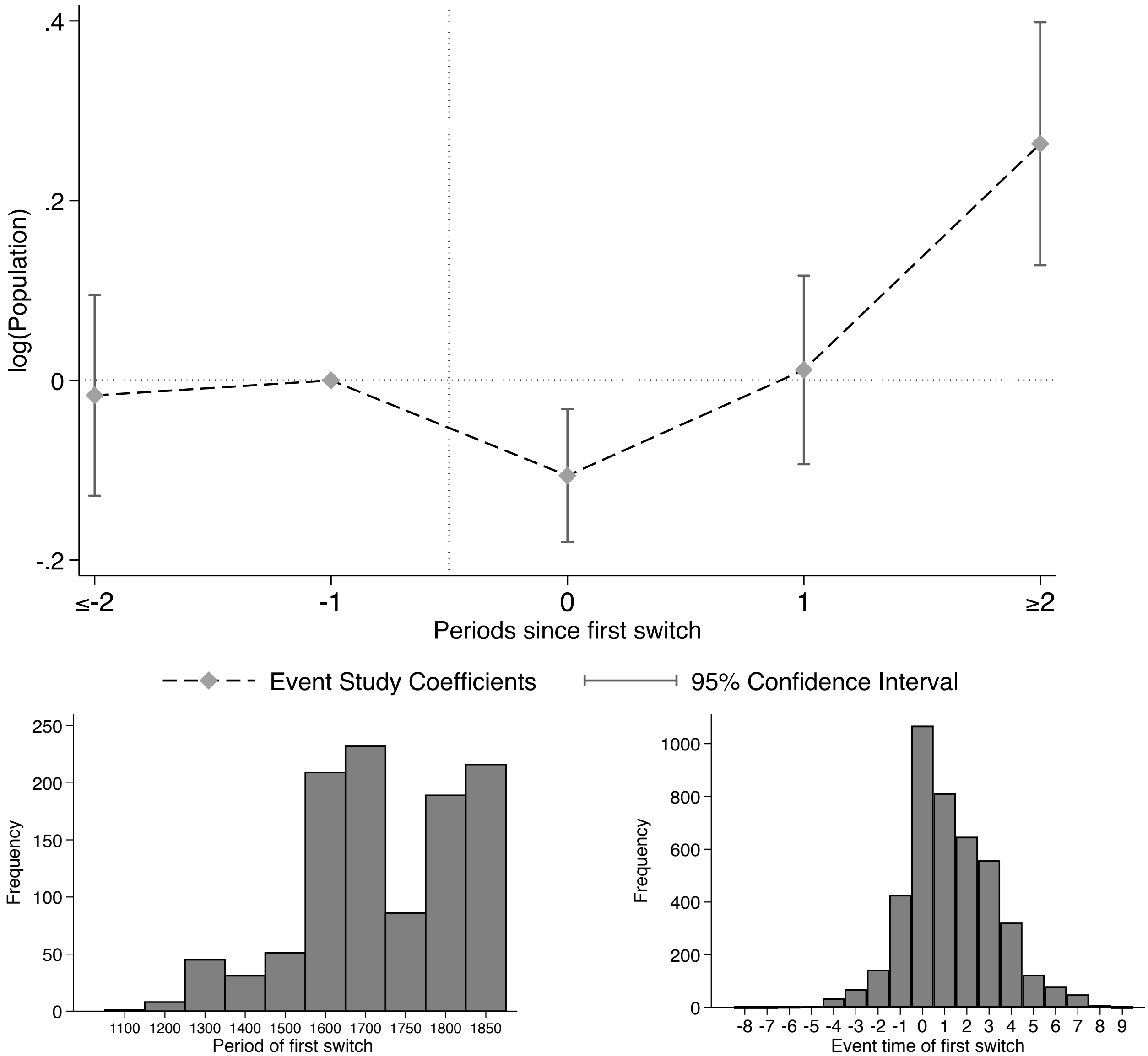
Notes: Average of state level fixed effect estimates ($\widehat{\psi}_j$) for cities of a given length of switching sequence, dropping all cities in Great Britain. We include cities that are in two to five different state over their lifetime.

Figure A.8: Improvement in state fixed effects over switching sequences (time-varying ψ_{jt})



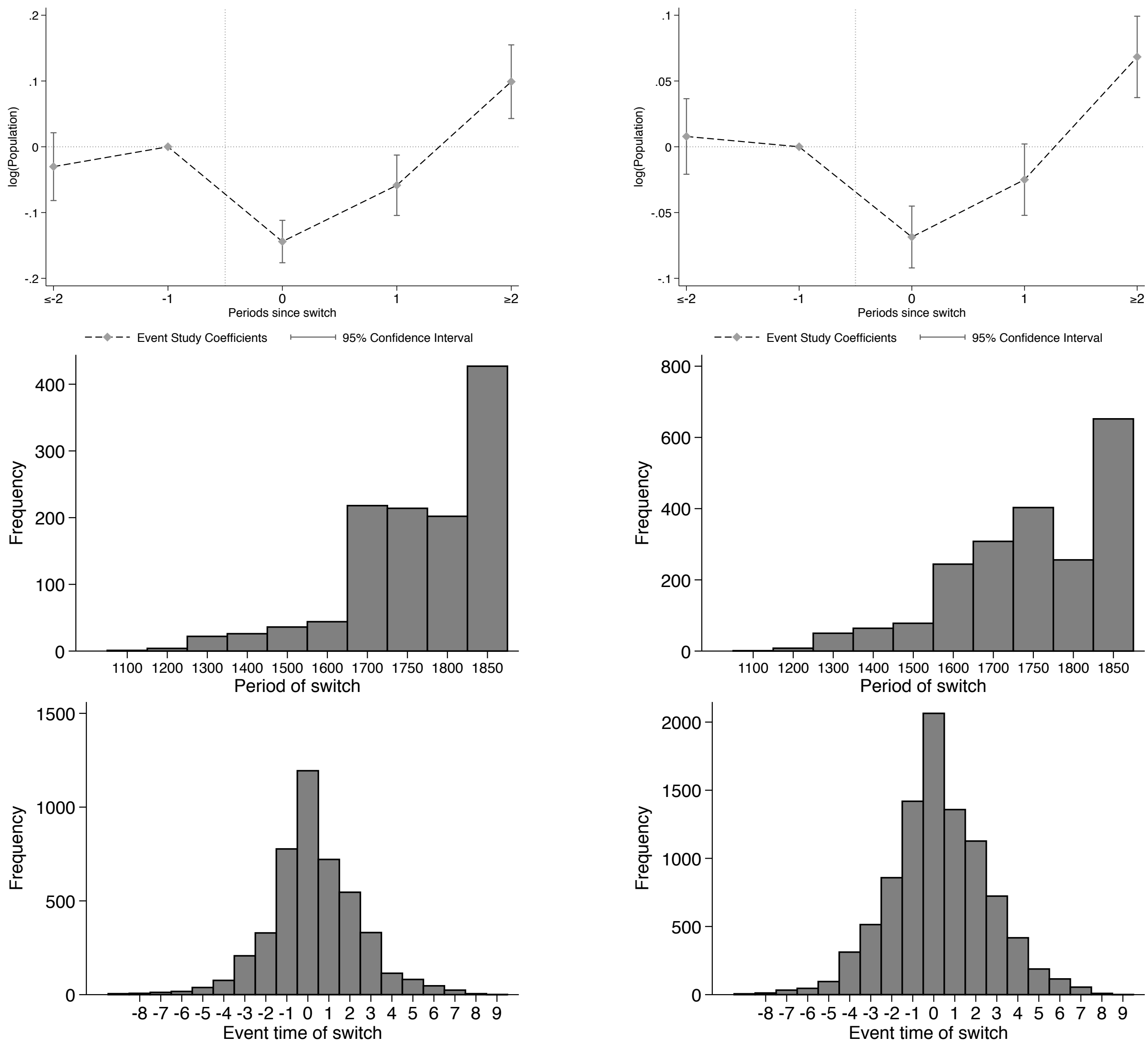
Notes: Average of state level fixed effect estimates ($\widehat{\psi}_j$) for cities of a given length of switching sequence over all city-period observations. We include cities that are in two to five different state over their lifetime (i.e. switch states one to four times). To compute the time-varying state effects ψ_{jt}, we first regress population on a full set of period dummies and then regress the residuals on $\mathbf{1}[S_{it} > 0]$, city effects, and state-by-period fixed effects.

Figure A.9: Event study estimates of switching cost: up to three switches



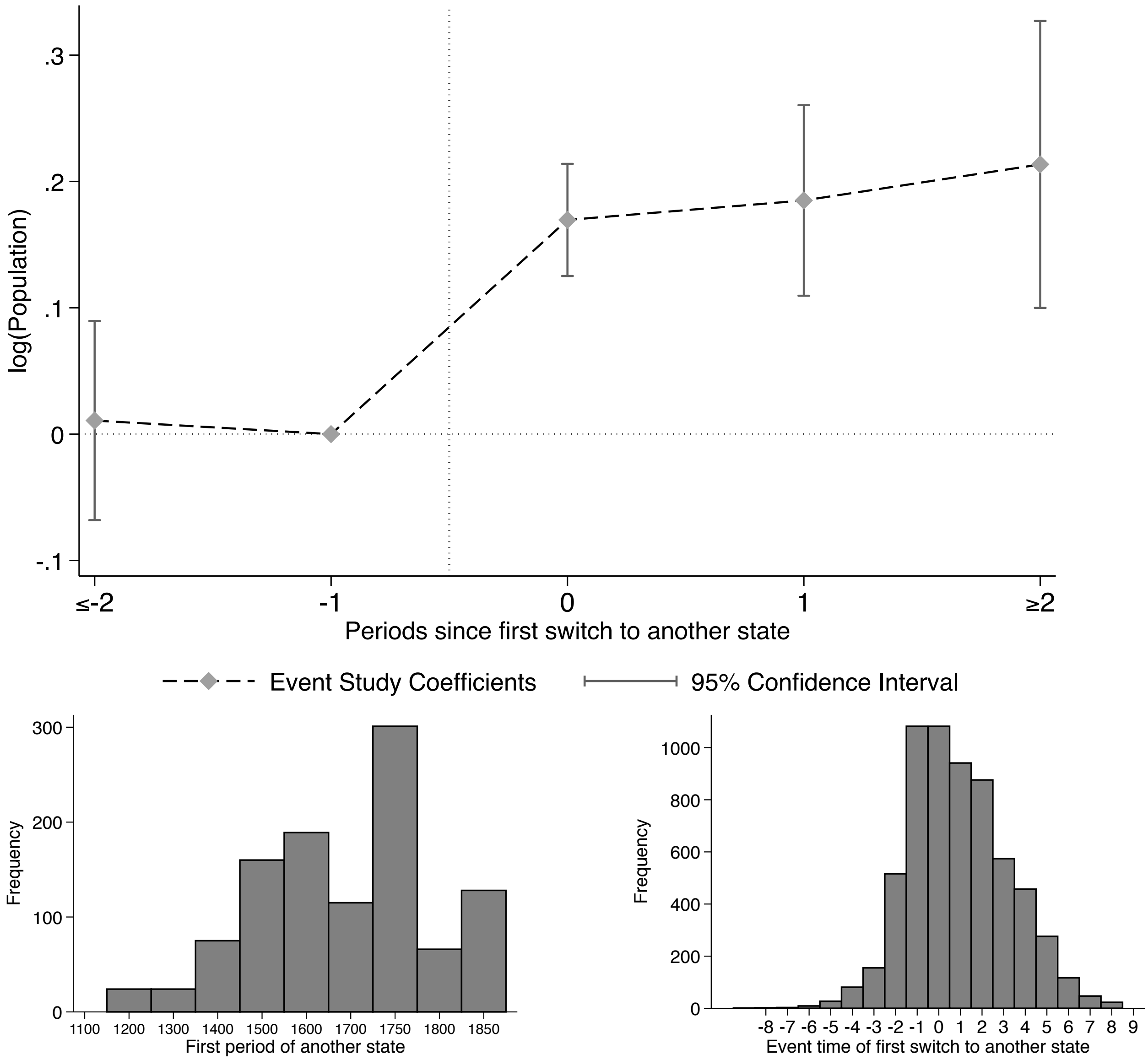
Notes: **Specification**: Event study estimates via OLS according to equation (2). **Sample**: All cities that experience one, two, or three switching periods. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching once, the event time is the period of switching; for cities switching twice, the event time is the first period of switching. **Fixed effects**: included are city, period and state fixed effects. **Standard errors**: clustered on the city level.

Figure A.10: Event study estimates of switching cost: duplicated, up to two/three switches



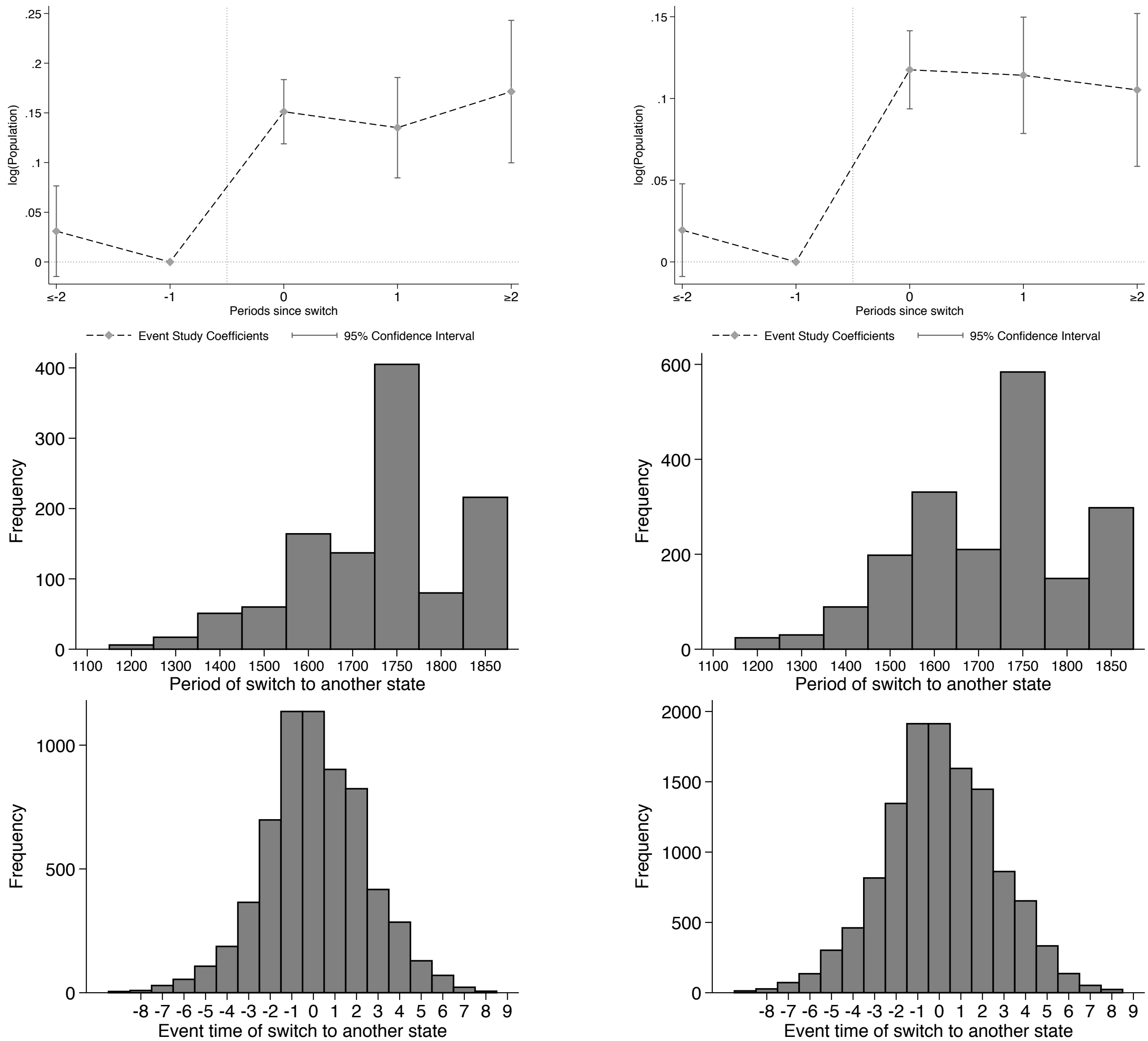
Notes: **Specification**: Event study estimates via OLS according to equation (2). **Sample**: All cities that experience one or two (left panel) or one, two, or three (right panel) switching periods, with cities duplicated by the number of switching events. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching once, the event time is the period of switching; for cities switching multiple times, the event time is any period of switching, assigned across the duplicate time series. **Fixed effects**: included are city, period and state fixed effects. **Standard errors**: clustered on the city level..

Figure A.11: Event study estimates of improved state quality: up to three switches



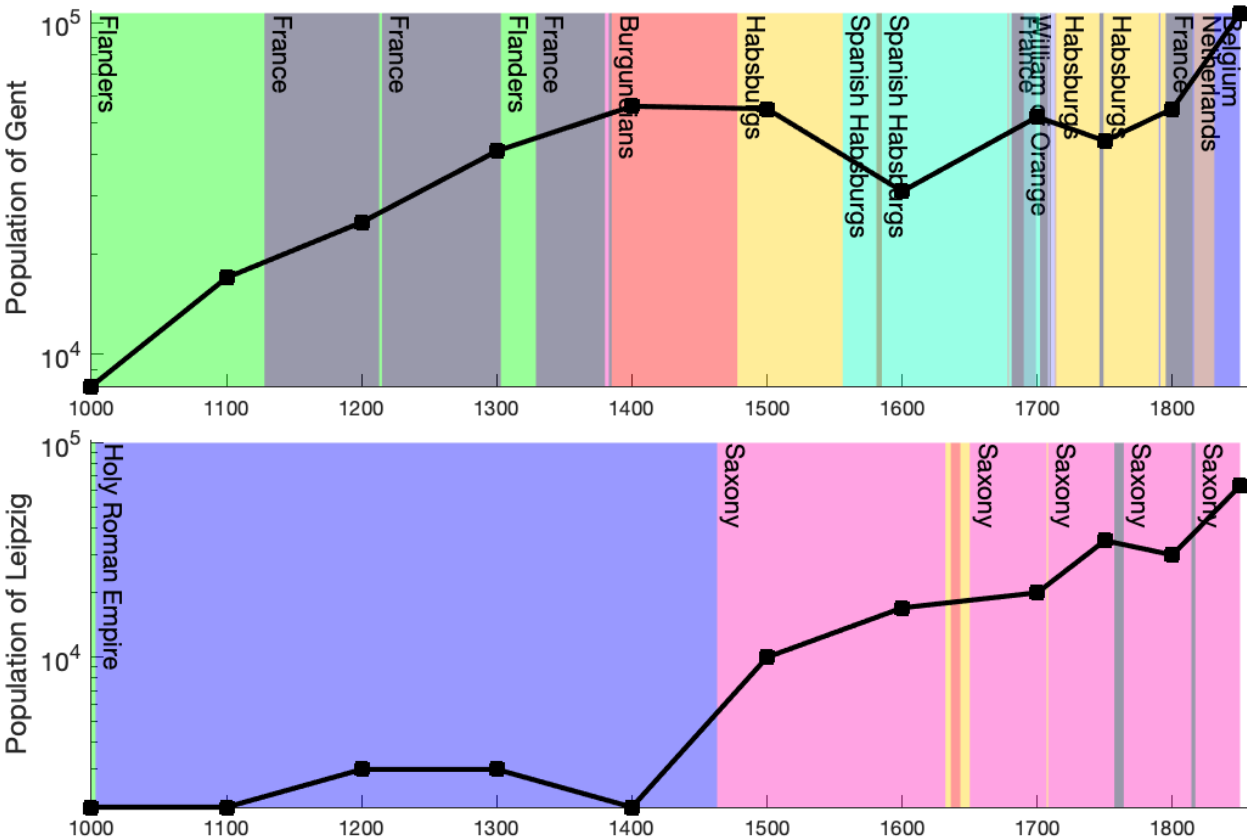
Notes: **Specification**: Event study estimates via OLS according to equation (3). **Sample**: All cities that experience one, two, or three switches to another state. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching to another state once, the event time is the period of switching to another state; for cities switching to another state multiple times, the event time is the first period of switching to another state. **Control variable**: controlling for $\mathbf{1}[S_{it} > 0]$, i.e. all periods in which the city switched (ending up in a different state or the same state). **Fixed effects**: included are city and period fixed effects. **Standard errors**: clustered on the city level.

Figure A.12: Event study estimates of improved state quality: duplicated, up to two/three switches



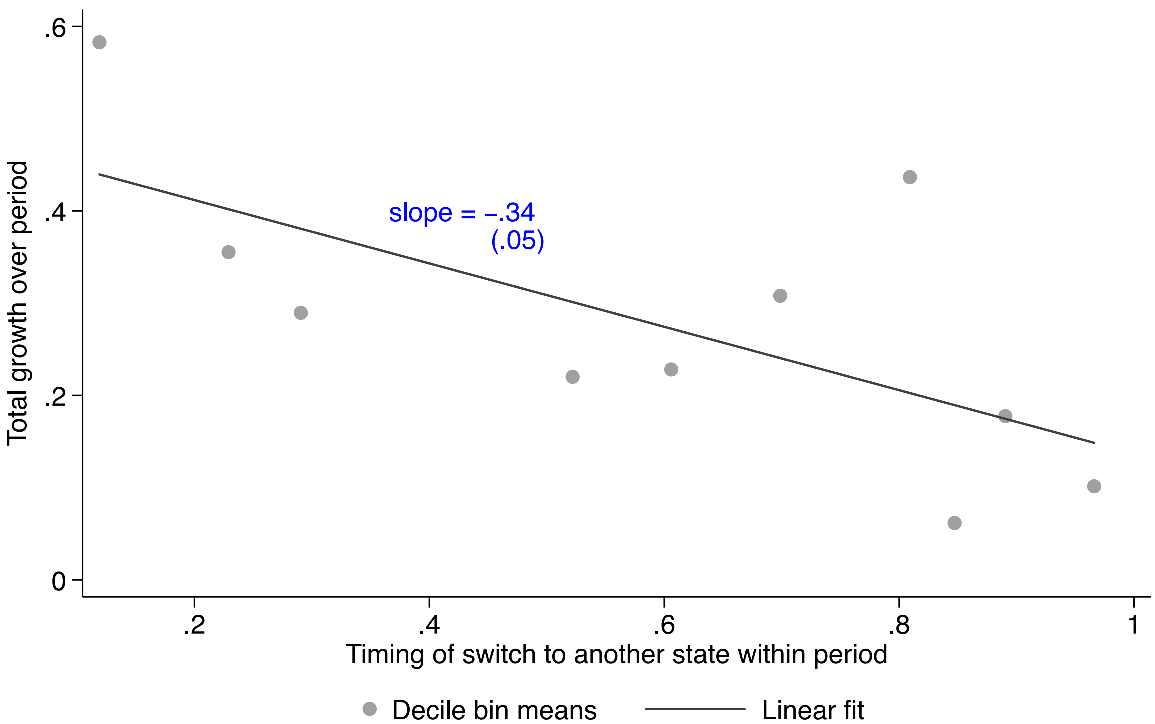
Notes: **Specification**: Event study estimates via OLS according to equation (3). **Sample**: All cities that experience one or two (left panel) or one, two, or three (right panel) switches to another state, with cities duplicated by the number of switches to another state. **Dependent variable**: log(population) is the natural logarithm of city population from Bairoch et al. (1988), which may be missing when no population data is reported. **Event of interest**: For cities switching to another state once, the event time is the period of switching to another state; for cities switching to another state multiple times, the event time is any period of switching to another state, assigned across the duplicate time series. **Control variable**: controlling for $\mathbf{1}[S_{it} > 0]$, i.e. all periods in which the city switched (ending up in a different state or the same state). **Fixed effects**: included are city and period fixed effects. **Standard errors**: clustered on the city level.

Figure A.13: Two examples illustrating mechanisms: Gent and Leipzig



Notes: Two examples of city population and the various states in control over 1000-1850. Top panel: Gent, illustrating the costs of low fiscal capacity (Spanish Habsburgs) and frequent switching between states. Bottom panel: Leipzig, illustrating the benefits of constrained rule and protection from expropriation. Random colors change with state in power; name of state is printed if it rules the city for at least 10 years.

Figure A.14: Intra-period analysis: timing of switch to another state within period



Notes: Empirical relationship between timing of switch and growth after switching to another state. Points are mean growth deciles over the preceding period before switching to another state conditional on deciles of the timing of switching, controlling for switching intensity over the period. Average growth over a period is around 20%, which is very close to predicted growth right if the switch occurs right before the end of the period. In contrast, if the switch happens in the beginning of the period, growth of the period is about twice as large. We control for average switches per decade to take into account the intensity of switching. Without this control, the coefficient on the timing of switching is $-.32$ with standard error .05. We can also control for the switching intensity in quartiles of the timing, which yields a coefficient of $-.25$ and standard error .06.