Abstract This paper challenges the long-standing belief that changes in patterns of war and war making caused the emergence of large territorial states. Using new data describing the universe of European states between 1100 and 1790, I find that small political units continued to thrive well into the “age of the territorial state.” Some scholars have argued that changes in the production of violence led to the dominance of geographically large political units during this era. In contrast, I find evidence that variation in patterns of economic development and urban growth caused fragmented political authority in some places and the construction of geographically large territorial states in others. Exploiting random climatic deviations in the propensity of certain geographical areas to support large populations, I show via an instrumental-variables approach that the emergence of towns and cities caused the formation of small and independent states.

For the nearly forty years since the publication of The Formation of National States in Western Europe questions of the territorial state’s origin have aroused the interest of scholars spanning the social sciences. Yet despite extensive scholarship on the topic, there has been little work aimed at explicitly evaluating the hypotheses generated by theories of state formation. Bringing to bear new data describing the entire universe of European states, I fill this gap by explaining variation in the number and size of states between 1100 and 1790.

Scholarship points to two broad factors that, in some combination, are believed to have determined the number and size of states: capital and coercion. Variation in access to economic resources, on the one hand, and the ability to produce large-scale collective violence, on the other, are thought to have selected large states.
against their smaller competitors. My results indicate that current scholarship has overemphasized the role of war making and underemphasized the consequences of economic change in explaining the eventual dominance of large territorial states. I find that many predictions of those who suggest war’s determinative role are not borne out in data describing the entire set of states. Rather, these effects are manifest only in the relatively small subset of cases that many scholars use to inductively generate theories linking war-making patterns to processes of state formation. Focusing instead upon the entire set of cases, I provide evidence that changes in urban growth and the revival of commerce largely caused observed temporal and spatial variation in the geographic scale of political organization.

My first set of empirical results cast doubt on bellicist approaches to state formation.5 Dating to German sociologists Otto Hintze and Max Weber, this line of scholarship attributes the development of the modern state to changes in patterns of war and war making.6 In these theories large states could most easily raise the manpower and finance required to field the increasingly large standing armies and increasingly dear technologies of coercion necessary to survive an era of endemic warfare. Technological shocks to the production of violence requiring increasing numbers of soldiers and more expensive armaments selected states most capable of adapting to these changes.7 Although a number of possible military innovations are identified, according to this logic the most fit states were those that maintained an advantage in the form of substantial populations, larger tax bases, and greater access to natural resources.8 Bellicists argue that because of these endowments large states were more capable than their smaller counterparts of meeting the demands of war and therefore were more likely to survive.

However, before the French Revolution, before the era of the mass-conscript army, wealth could purchase not only the technologies of violence but also the manpower required to prosecute major wars.9 Rather than being an age when large states dominated militarily, this was a period when the population and natural resource advantages of territorial states provided little benefit in the production of violence. Just as they could use economic resources to purchase the most advanced technologies of coercion like siege artillery or firearms, leaders of states could, for a negotiated price, hire a Hessian colonel or an Italian condottiero10 and retain their men for a campaign season. It was by virtue of their economic capacity that city-states like Genoa and Florence or groups of independent towns like the Swabian league could

9. Onorato, Scheve, and Stasavage (2014) provide evidence that mass armies are the consequence of even later technological developments and did not arise until well after the French Revolution. Focusing upon battle deaths, Cederman, Warren, and Sornette (2011) find a structural break corresponding with the French Revolution.
10. Literally “contractor.”
raise armies that matched or even exceeded those of territorial states like France or England.\(^{11}\)

What is more, if economic resources largely determined the ability to project force, it is not clear that large territorial states held an advantage over their smaller competitors. In fact, the opposite is likely true. The tax income of the Della Scala signoria\(^{12}\) in the fourteenth century, for example, was double that of England.\(^{13}\) The revenue of Venice in the fifteenth century was 60 percent greater than that of France and more than double Spain’s or England’s.\(^{14}\) Moreover, geographically small and urban city-states could far easier and earlier construct financial instruments necessary to purchase the means of defense required to survive interstate competition.\(^{15}\)

If the economic capacity of states best explains their ability to produce force, it follows that the empirical predictions of war-making theories should not be borne out in data. After introducing a new data set encompassing the entire universe of European states between 1100 and 1790, I show just this. Namely, I demonstrate that the military revolution, a sequence of dramatic changes to the production of force expected to select large states, did not significantly alter the typical size of states. In contrast to the predictions of these bellicist theories, the typical state remained small during the period associated with these large-scale changes in the cost of war making. Then I show that the relationship between geographic scale and survival probability is the opposite of what war-making theories predict. Over this span small states were more likely to survive than their larger counterparts. In other words, rather than being “an age of the territorial state,”\(^{16}\) the period between 1500 and 1800 was one in which small political communities not only persisted but remained the typical form of political organization.

My second set of empirical results provides evidence that variation in the economic resources available to states and potential state makers best explains patterns of European state formation before the French Revolution. Economic theories of this sort reemerged with the work of Hendrik Spruyt\(^ {17}\) and build on both the political sociology of Stein Rokkan and the economic history of Henri Pirenne in viewing the development of the territorial state in some places (and its absence in others) as the consequence of variation in the dominant social coalitions that formed from changing patterns of trade and economic development.\(^ {18}\)

The reemergence of the Eastern trade and the revival of urban life during the last half of the tenth century created in some places new commercial classes.\(^ {19}\) The wealth

---

11. For example, Genoa raised an army of 40,000 men in 1295 (Scammell 1981, 161). In contrast, the French army at the end of the Hundred Years War was estimated to be 14,000 (Spruyt 1994a).
12. Territory encompassing Parma, Lucca, and Modena.
that came to exist in these new towns and cities allowed urban social groups, in ways no different from the constructors of territorial states, to project force and assert themselves as independent states within the international system. Where towns formed and burgheers could force their rights upon princes and kings, smaller political units proliferated. Only in the weakly commercialized European periphery did the large territorial states form.

Since the relationship between patterns of state formation and urban growth is likely affected by unobservable confounders and, moreover, because the direction of causality runs in both directions, I take an instrumental variables approach to identify the effect of changes in patterns of urban development on the size and number of states within some defined geography. To show that political fragmentation was caused by the uneven reemergence of urban life, I exploit random climatic variation in propensity of some pieces of geography to sustain large populations. Treating shocks to the optimal growing temperature for cereals like wheat as an instrument for the development of towns and cities, I show that variation in urban growth caused political fragmentation. Where commerce and urban life reclaimed a foothold, where cities reemerged from the Dark-Age nadir, small independent political communities formed. In those places where cities and urban life did not emerge, large territorial states took shape.

What States Are (And Aren’t)

To conduct a rigorous empirical analysis, I need to collect and define systematic data describing my outcome of interest. To do this in a theoretically satisfying manner it is crucial for my object of inquiry, the state, to be clearly defined and carefully operationalized. This section first lays out a minimalist Weberian definition. Then, I outline three observable criteria that allow for the systematic and replicable coding of political units as states or not. Next, I argue that this way of understanding statehood is both reasonable from a historical perspective and, moreover, captures the crucial distinction others have highlighted between alternative institutional forms like empires, leagues of city-states, and territorial states. In other words, the coding scheme I produce captures the ways in which these institutions represent centralized versus fragmented political authority.

A Definition of the State

I begin with Weber’s treatment of states as political communities that “(successfully) claim a monopoly on the legitimate use of physical force within a given territory.”\(^\text{20}\)

However, to reflect empirically observable phenomena rather than a non-existent

\(^\text{20}\). Weber 1972, 1.
ideal type, I alter this definition and treat states as the organizations that maintain a quasi-monopoly of violence over a fixed territory. That is, states are the organizations that have a clear preponderance of the coercive means over some geographically defined unit.

This understanding of states matches observationally that of Tilly who on the very first page of *Coercion, Capital, and European States* defines states as “coercion-wielding organizations that are distinct from households and kinship groups and exercise clear priority in some respects over all other organizations within substantial territories.”21 The utility of this approach is that it is constructed to include “city-states, empires, theocracies, and many other forms of government” but also excludes other forms of social organization like “tribes, lineages, firms, and churches.”22

A potential downside of this definition is that, in removing the requirement of legitimacy, I may discount the importance of legal, constitutive, notions of statehood that many critical scholars of international relations view as having been crucial in shaping the development of the state and state system we observe today. That is, if states are conceived of as like units that are mutually recognized legal entities endowed with exclusive juridical sovereignty and capable of entering into diplomatic relations with each other, then my definition removes this component. For two reasons this is not problematic for my study.

First off, the assumption of juridical sovereignty as a central feature of statehood is the consequence of a common, but nevertheless anachronistic, reading of nineteenth-century legal scholarship by traditional scholars of international relations.23 The inherent problem is that the contemporary legal understanding of sovereignty simply did not exist as we understand it today and did not develop until well into the nineteenth century. In the words of Stephen Krasner, “the conventional view that the Peace of Westphalia of 1648 marks a turning point in history is wrong.”24 Rather, “the principles associated with both Westphalian and international legal sovereignty have always been violated. Neither Westphalian nor international legal sovereignty has ever been a stable equilibrium from which rulers had no incentive to deviate.”25 This is supported by Osiander’s careful reading of treaties and diplomatic texts that shows how contemporary scholars’ imposition of international legal sovereignty is the consequence of a fundamental misreading of a nineteenth-century historiography biased by the desire to promote nationhood. The series of treaties surrounding Westphalia, in fact, did very little to alter how states interacted with each other.26 Rather, in line with my definition, as the concept of sovereignty was introduced in the late sixteenth century, “political

21. Tilly 1990, 1. Later, Tilly defines the state in a substantively identical way as “any organization that commanded substantial means of coercion and successfully claimed durable priority over all other uses of coercion within at least one clearly bounded territory.” Tilly 1994, 14.
22. This approach to defining the state also matches, in a nearly perfect manner, that of Skocpol. On this see Skocpol 1995, 43.
theorists and practicians alike attached more importance to its domestic than to its external side. They were concerned with the power of rulers over their subjects and only marginally with relations among rulers.”

Until the late twentieth century even legal notions of the state were predicated upon the empirical facets of statehood like those I propose. Through the 1930s the legal standard for statehood was considered to be the presence of “effective government.” Even as late as 1979 the textbook legal definition of the state rested on the existence of a stable political community in a territory with an established legal order. Only as weakly empirical states gained independence in the late 1960s did legal, mutual recognition of states become determinative with respect to which units survived in the international arena.

Second, regardless of a mutually recognized legal status, we might care only about units that, in a de facto sense, were capable of entering into meaningful diplomatic relations with each other. That is, regardless of the legal status, IR scholars might only be concerned with polities that behave as if they were like units. Even if this is a concern, it is nevertheless clear that many of the small states that my definition includes and that critical IR definitions might exclude behaved as states in this fashion. In his careful analysis of the behaviors of late-medieval states Fischer finds that while the late-medieval and early-modern political units I describe as states made outward professions of “adherence to such legal institutions as vassalage, fief, feud, and peer court, they really strove for exclusive territorial control, protected themselves by military means, subjugated each other, balanced against power, formed alliances and spheres of influence, and resolved their conflicts by the use and threat of force.” That is, they behaved like states. For example, the units comprising the Catholic League and Protestant Union, distinct formal alliances of states within the Holy Roman Empire, entered into strategic alliances and waged wars with and against actors outside of the Empire. In a more general sense, as I detail later, institutions like leagues operated functionally like contemporary international organizations composed of like units.

In sum, contemporary notions of legal sovereignty are inappropriate for the period covered by this study and, moreover, behaviorally, the units I define as states frequently entered into international politics as equals. As such, the definition of statehood I provide is both accurate and appropriate for empirical study. That said, it still must be operationalized. To accomplish this I provide three necessary observable criteria to distinguish political entities I call states from non-states in the period before the French Revolution. These criteria are:

29. Brownlie 1979, 74–75.
32. The starting date—1100—is chosen for reasons of data availability and accuracy. At this point, records for small—largely princely and ecclesiastic—states begin to document clear evidence, frequently
1. **Direct Military Occupation**

Following the Correlates of War coding scheme,\textsuperscript{33} if a political unit is militarily occupied by a foreign power, according to my coding it ceases to exist as an independent state. Similarly, if a political unit successfully conquers a piece of territory, this newly occupied territory is treated as a part of the conquering state. For example, when the Ezzelino or Pallavicini families were able to effectively wield military control over several Italian city-states I code the amalgamation of these units as a single state. Analogously, when military orders like the Teutonic Knights or the Knights Hospitaller conquered well-defined territories these new units are coded as independent states. Similarly, when the Castilian-Aragonese state drove the Moors from Grenada, the Emirate of Grenada ceases to be coded as an independent state and its territory gets coded as part of Castile.

2. **The Capacity to Tax**

Expropriative power, the ability to take from another what she owns, is the coercive authority most associated with statehood. Formal expositions of states as wealth-maximizing actors, as “stationary bandits”\textsuperscript{34} or organized criminal organizations, underscore this crucial aspect of state violence: states “steal” from those they govern. Taxation is simply on-the-path coercion. Moreover, the ability to extract is the key feature of state power driving several recent and influential theories of political transitions.\textsuperscript{35} In these theories it is precisely the ability of the state to extract that actors—economic classes in these models—enter into conflict to control. In a game-theoretic sense, taxation is simply on-the-path coercion. As such, I take the capacity to tax as evidence of the state’s quasi-monopoly of coercion. So, for example, when Worms (1184) or Lübeck (1226) demonstrably gained rights to collect taxes and tolls within their boundaries, I code them as independent states.

3. **A Common Executive**

Recognizing that many states during the period studied were “composite” entities, composed of political units that maintained semi-independent bureaucracies, parliaments, and other separate political institutions,\textsuperscript{36} I treat those sharing a common executive as a single state. Coding states this way treats the holdings of imperial families as a common state rather than distinct units. So, for example, all of the territory held by the head of the Wittelsbach family—at various points including the Counties of Holland, Hainaut and Zeeland, as well as the Duchies of Jülich and Berg—all get coded as a single state.

---

\textsuperscript{33} Stinnett, et al. 2002.
\textsuperscript{34} Olson 1993, 2000.
\textsuperscript{35} Acemoglu and Robinson 2006; Boix 2003.
\textsuperscript{36} Nexon 2011.
However, as the family split territory among its various component branches—first between the Bavarian and Palatinate and then the numerous further divisions—each is treated as a distinct state. However, when, as in 1777 the Bavarian line died out and merged with the Palatinate branch they again get treated as a single state.

*Relation to Other Definitions of the State*

The definition of states as quasi-monopolists of violence recognizes the fact that political communities that reasonable coders would identify as states existed before juridical notions of sovereignty. This is not a claim that juridical statehood is unimportant for the study of politics but, rather, that such a coding scheme would fail to capture political organizations like France, Venice, and England, let alone older polites like the Roman or the Han Empires that existed as coercion monopolizing entities long before 1648 or 1555. In the words of H.J.M. Claessen, we have no reason to consider “the realm of the Aztecs, the Mongol Empire … or the late Roman empire qua political structure as qualitatively different from, say, France, Spain, or England in the fifteenth century. They were all states, varying from early to mature.”

Because this coding scheme treats the de facto distribution of power as central in determining which units get coded as states, I treat as distinct political units some states that more traditional historiography might consider unified. For example, when Boleslaw III of Poland divided his kingdom between his sons, thereby creating the Masovian, Seniorate, and Sandomierz provinces along with Greater Poland, my coding scheme treats each as distinct units. Similarly, feudal territories like Toulouse, Provence, or Brittany, although seigniorial dependents of the French king, get coded as independent until they are integrated into France proper. Likewise, imperial city-states, prince bishoprics, free cities, and imperial abbeys that effectively demonstrate independence as I outlined get coded as independent units. For example, when Fredrick Barbarossa attempted to assert imperial rule in Lombardy, the various units that composed the Lombard League and successfully resisted begin to be coded as independent.

With this minimalist definition I account for institutional configurations that others treat as distinct from territorial or “national” states. Without rejecting the notion that political entities like the Holy Roman Empire or the Swabian and Hanseatic leagues were unique constellations of political institutions, based upon my scheme we can consider many of their constituent units as states while still capturing the relevant differences drawn by the literature between these alternatives and territorial states. Unlike territorial states, they preserved fragmented political authority.

Consider the Holy Roman Empire. Whether or not it represents a true alternative to the state, using my coding scheme, I arrive at a conclusion similar to those who, like

Spruyt, view the empire as a fundamentally unique institutional form. That is, from either perspective it is apparent that the Holy Roman Empire was marked by substantial fragmentation of political authority. Indeed, by the end of the thirteenth century an increasing number of units within the empire acquired de facto independence from imperial rule. The extent of this was such that even petty magnates who previously swore “fealty to only God and Emperor eschewed themselves equally of both powers” and maintained “full jurisdiction … rights of legislation, privileges of coining money, levying tolls and (collecting) taxes.” They were, by my definition, independent states. Indeed, “along the Rhine even the Lord of a single tower was often almost an independent prince.”

The same is true for leagues of city-states. Although leagues like the Hansa were far more than loosely bound affiliations of towns centered upon the regulation of trade, when considered in light of my coding scheme they do not represent an alternative to the state. Consider the difficulties leagues faced at creating compliance. Leagues facilitated cooperation among members largely through reputational mechanisms. With respect to generating revenue, like modern international institutions, they faced great difficulty in directly taxing their members. Instead they relied upon the voluntary compliance of individual member cities to obtain revenue. Typical of these organizations and lacking a third-party enforcement mechanism, the Hansa could at most expel member cities who failed to comply with calls for revenue.

The absence of third-party enforcement is perhaps most evident in leagues’ conduct of military affairs. The Hansa and other leagues were certainly capable of projecting military force, fielding armies able to combat large territorial states like Sweden, England, Denmark, and Holland. For example, at the height of its powers the Swabian league could support armies rivaling those of any major power, in 1385 raising an army of more than 12,000 infantrymen and 1,200 cavalry. But the ability to field large armies belied their true capacity to project force. When, for example, the Hansa waged war against the Danish Crown in 1360 it could not compel all of its member states to participate in the conflict. Similarly, it was in part because the consent of the forty-odd commissioners (representatives of the individual cities) was necessary to execute tactical maneuvers that the Swabian league was defeated at Doffingen in 1388; coordination on the battlefield was made so difficult that the allied lords the league opposed were able to emerge victorious despite their numerical inferiority.

40. Ibid.
41. Ewert and Salzer 2006; Greif, Milgrom, and Weingast 1994.
42. Fink 2012.
43. Laffan 1957.
44. Dollinger 1970, 70.
In other words, although leagues were certainly institutional responses to changing patterns of war and trade, they were organizations made up of units I call states and not a fundamentally distinct alternative. Perfectly in line with other interpretations, leagues were institutions that, by providing quasi-public or club goods like collective security or enlarged markets for traded goods, allowed fragmented political authority to persist, an outcome captured in my operationalization.

In sum, forms of political organization like leagues or universal empires preserved and even enhanced local autonomy over coercion. My measure of the state captures precisely this and yields empirical measures of the number and size of states that reflect the degree to which political authority was fragmented across geography.

**Constructing the Data**

Following this coding scheme the data are constructed by manually geo-referencing several sets of historical maps. Two of the main sources from which the base GIS boundaries are constructed are the Centennia Historical Atlas\textsuperscript{46} and the Euratlas\textsuperscript{47} digital atlas. The Nussli data are measured in one-hundred-year panels whereas the Reed atlas utilizes a much more high-frequency approach, recording observations in tenths of years. I use the boundaries as defined by both data sets, aligning them at every hundred-year mark based upon the coding scheme defined earlier. The Nussli data match the Reed data nearly perfectly at these points. Where there are discrepancies it is usually because the Nussli data set takes observations from a window surrounding each panel and not a snapshot exactly at the one-hundred-year point. Because the Reed data are not geo-referenced, I construct shape files that are compatible with GIS analysis by manually constructing the boundaries from reprojected images provided by the atlas and then referencing each observation using the European Albers Equal-Area projection system.

The Nussli data have been used in several recent publications and are considered highly accurate.\textsuperscript{48} Nevertheless, even after combining the data from these digital sources there are still a number of imperfections. For example, units I code as independent states are absent from the reconstructed shape files. These tend to be small independent principalities, ecclesiastical units, and city-states that were not picked up by the historical geographers who created the digital reproductions from which my maps are constructed. In order to rectify these flaws and prevent the ensuing selection problems that would plague the subsequent statistical analysis, I turn to a number of historical and contemporary primary source maps to create high-frequency boundary changes for these missing units.

This combination of secondary and primary cartographic sources allows me to project the boundaries for political units that meet the coding criteria. Using known

\textsuperscript{46} Reed 2008.
\textsuperscript{47} Nussli 2010.
\textsuperscript{48} Blaydes and Chaney 2013; Stasavage 2011a, 2011b.
physical geography, known political boundaries, and the location of cities and towns to properly reference these maps, I create shape files that, with a high degree of accuracy, reflect the geographic scale of each unit. For each unit I track the history of their boundary changes—expansions and contractions—and adjust the shape-files accordingly. To provide a more detailed description of the procedure, a walk through the creation of the shape file for part of Nassau between 1159 and 1328 may be useful.

The town of Nassau dates to at least 915 and was founded by Robert, the son of Dudo-Henry of the House of Laurenburg. The Laurenburgs built Nassau Castle in 1125 and established the County of Nassau in 1159—effectively claiming rights of taxation, toll collection, and justice. As such, Nassau enters the data set as an independent state as of 1160. The County (later principality) of Nassau exists in the digital base maps from this point, giving an accurate measure of its boundaries and size. However, the digital source maps fail to record the dissolutions and mergers of various component units of Nassau, nearly all of which meet my coding as independent states. I manually make these corrections as described in what follows.

For ninety-six years Nassau existed as a single independent state. Upon the sale of Weilburg to the Count of Nassau in 1255, the territory was split between the two sons of Henry II, with Otto I receiving the territory north of the river Lahn and his older brother Walram II receiving the rest. Using this geographic boundary as the dividing line, I create the Counties of Nassau-Dillenburg and Nassau-Weilburg, respectively. Dillenburg remained a single state until Otto’s death in 1303 after which Nassau-Dillenburg was divided into three units, splitting off Siengen and Hadamar from the initial unit. The boundaries of these new states are constructed using the boundaries as they exist on several historical reproductions and one primary source map. Using the known latitudes and longitudes of the cities of Siengen and Hadamar, I then can reference the projected images from the historical reproductions using points representing the locations of these cities. From here, using these points and the boundaries of pre-existing Nassau-Dillenburg, we can create the boundaries and subsequent shape files for each of these new units. These three states remained independent until Siengen conquered Dillenburg in 1328 and then Hadamar in 1394.

War Making and the European System of States

The mainstream view of state formation has been that “until recently only those states survived that held their own in war with other states … and that over the long run the changing character of war gave the military advantage to states that could draw large, durable military forces from their own populations, which were increasingly national states.” That is, in the face of technological and tactical changes to the production of violence, large, territorial states (in Tilly’s language, national states) came to hold a

50. Tilly 1990, 63.
clear advantage over political units that relied upon “capital intensive” strategies, typically small city-states.

These changes are described by historians as an early-modern “military revolution,” a notion first promulgated by Roberts51 who viewed changes in infantry tactics devised in the mid-sixteenth century by Maurice of Orange and Gustavus Adolphus as fundamentally altering the manner in which armies were trained, raised, and paid for. Others view alternative technological and tactical changes as having marked the sea change between premodern and modern ways of waging war. Parker and McNeill for example, see as crucial the late-fifteenth-century development of siege artillery, a technological innovation that engendered an even more costly defensive response, the trace italliene.52

These accounts are not exhaustive. For example, Black places a greater significance on developments that took place between 1660 and 1710, namely, the invention of the flintlock musket.53 Rogers, on the other hand, dates an almost continuous series of military innovations to the beginning of the Hundred Years War.54 Whether or not the exact timing of the military revolution is clearly defined, a recurrent theme in the study of early-modern military history is that a fundamental change in the production of force took place sometime between the mid-fifteenth and the end of the eighteenth centuries.

These changes have then been thought of as crucial in selecting the type of state that could survive. For example, Bean provides a technologically deterministic theory wherein the introduction of weapons like powerful siege artillery drove up the fixed costs of defense, thus making local political rule untenable and leading to the selection of geographically large states.55 Less economistic in his theorizing, McNeill arrives at a similar conclusion, finding that “the major effect of the new weaponry was to dwarf the Italian city-states and to reduce other small sovereignties to triviality.”56

Even in the more nuanced theory of Tilly where a “capital intensive” path undertaken by small city-states is presented, at the very least, as a viable option for survival, military change is fundamental. For example, with respect to the survival probability of small units, Tilly finds that “war eventually changed in a direction that made their small scale and fragmented sovereignty a clear disadvantage and they lost to large states.”57 Zeroing in on the consequences of military changes on the Italian peninsula, he argues that the “Florentine and Milanese republics crumbled under the weight of the fifteenth and sixteenth century’s military requirements … That era of large armies, heavy artillery, and extensive fortifications relegated all the Italian city-

56. McNeill 1984, 89.
57. Tilly 1990, 65.
states to extinction, subordination, or perilous survival in the interstices of great powers.”

In sum, social scientists have interpreted the large-scale tactical and technical changes in the production of violence that took place in the second half of the last millennium as having been associated with related changes in the nature of the state. This literature predicts that the type of state expected to survive structural changes in war making was geographically large.

**The Size of States Across Time**

With a simple description of the data I show that, during the period associated with the military revolution, the typical state declined in size. The top panel of Figure 1 plots the trend in state size across time measured in square kilometers. If one were to consider only the mean, it appears that the bellicist hypothesis matches the general trend. Between 1400 and 1790 the average size of states more than doubled from approximately 33,000 to 71,000 square kilometers. Although a proponent of the military revolution might view this as confirmation, in the presence of extreme values the mean is a poor indicator of central tendency. This point is made apparent not only by the large spread between the median of state size and its mean but by the relationship between the mean and the third quartile of the distribution. For nearly all of the period for which there are data, the size of the state at the seventy-fifth percentile was less than the mean size.

The reason for this is clear. There are several extremely large states distorting what we might otherwise view as typical. That is, the distribution of the untransformed data is heavily skewed, with far more large states relative to what one would expect if size of states took a more symmetric distribution. An oft-used and simple corrective to this type of problem is to log-transform the data—taking the natural log of each state’s size—and then drawing inferences from the transformed distribution. In doing so, we can make better descriptive inferences about the trend in state size for the typical state in the presence of large outliers.

In other words, when data are skewed (as with state size), we might consider them to be drawn from a distribution like the log-normal that would generate outlying values like those we actually observe. In fact, state size is most likely distributed log-normally. To gain a sense of this, Figure 2 displays the untransformed distribution of state size on the left and the log-transformed on the right (pooling all years). It is clear that the untransformed data are heavily skewed. However, once we transform the entire data, we see a distribution of state sizes that is roughly normal. In the online appendix I provide more detailed evidence that in each cross-section (in every year we observe) the size of states is distributed log-normally.

58. Ibid.
59. Using a similar approach but examining only 1500 and 1998, Warren, Cederman, and Schutte (2011) also find that in both years state sizes are log-normally distributed.
Although a naive interpretation of the untransformed mean trend would indicate a revolution in state size coinciding with known changes in military technology, once we consider the log-transformed distribution this upward trend disappears. Rather, the typical state between 1100 and 1790 declined in size. The lower panel of Figure 1 indicates that both the mean and median of the log-transformed distributions of state size are declining over time and in near perfect tandem. The Engle-Granger two-step procedure indicates that the two series are cointegrated. Estimating the following relationship \( \text{Mean}_t - \beta \cdot \text{Median}_t = \mu_t \) where \( \beta \) is estimated to be 1.04, a Dickey-Fuller test yields a test statistic of \(-4.30\) allowing us with a high degree of confidence to reject the null hypothesis that \( \mu_t \) is a non-stationary series.

**FIGURE 1. The trend in state size**
measures is substantial; the mean and median logged state size decreases between 1100 and 1790 from 9.03 to 6.32 and 9.62 to 5.67, respectively. Re-transforming these results gives declines of 7,818 and 14,816 square kilometers from initial values of 8,372 and 15,106. By these measures the “typical” state, though quite small in 1100, became even smaller over time.61

To summarize, although literature describes radical changes to the production of military violence between 1450 and 1700, a main prediction of the consequence of these changes fails to materialize when the data are examined systematically. During the period associated with large systemic changes in the scale and cost of military force, the typical size of political organization in Europe remained small. In other words, during the period that historians describe as the “age of the territorial state,” large territorial states were not the norm. Rather, states of small geographic scale remained the modal form.

The Relationship Between Size and Survival

Thus far I have shown that in the period associated with the military revolution the number and size of states fail to change as hypothesized in war-making theories.

61. In the appendix I detail a series of statistical tests to see if the data generating process that determined number of states changed over time and find no evidence that it was altered over the same period.
The mechanisms proposed in these theories operate by altering the type of state capable of surviving in an anarchic international arena. That is, the costs associated with mass armies consisting of professionally trained riflemen, the purchase of increasingly powerful cannons, and the development of defensive ramparts made small states unviable. This section provides two findings. First, over the entire time period of study small states were more likely to survive, the opposite of that hypothesized in bellicist theories. Second, I find that the relationship between geographic scale and the survival of states did not change across time. That is, the relationship between size and failure was the same before, after, and during the period associated with the military revolution.

Since the theories I am evaluating concern the capacity of states to survive, I utilize duration analysis which requires a coding of failure or “state death.” I treat state death as any instance in which an existing state ceases to appear as an independent political unit according to my coding scheme. Thus, if a state is conquered it is treated as “dying.” If two states merge I treat the new state as either a new unit (and the pre-existing states as being censored) or, if it is clear that one subsumed the other, the subsumed state as having died. In the few cases where this is ambiguous I alternate codings and re-estimate the model with each possible alternative. The treatment of these ambiguous cases does not substantively alter the results.

The relationship between the hazard rate and the geographic size of states is estimated via a mixed-effects Cox proportional hazards model of the basic form:

$$\lambda_i(t) = \lambda_0(t) \times \exp(\delta_p \cdot \ln(\text{Size}_{it}) + \epsilon_i + \eta_v)$$

Time is described in three ways. First, $t$, indexes the time in years since a state came into existence. Second, $v$ indexes chronological time, for example, 1445 or 1750, and third, $p$ captures a multi-year period in chronological time, for example, 1450 to 1500.

The baseline hazard rate is captured by $\lambda_0(t)$. The relationship between size and the hazard rate is captured by the set of parameters $\delta_p = \mu + \gamma_p$ where it is assumed that $\gamma_p$—the period-varying effect—is distributed $\mathcal{N} \sim (0, \sigma^2_\gamma)$ and where $\mu$, captures the time invariant, mean relationship between state size and death. The magnitude of each $\gamma_p$ signifies the deviation for each period $p$ from the time-invariant mean effect $\mu$. I present results allowing the effect of state size to vary by one-hundred- and fifty-year intervals, respectively. Since the data include repeated observations—because some states die and then reappear only to die again—I follow convention and include a unit-specific random effect, $\epsilon_i$. Again, because state deaths might be clustered by chronological time, I include a time random effect, $\eta_v$.

First I estimate the model without the time-varying component and find that, in contrast to the conclusions of bellicist theories, there is a robust negative relationship between the probability of survival and the size of states. That is, I find that

62. The definition of failure I provide is most similar to that of Fazal (2004, 2011) who calls this “state death.”
geographically large states are more likely to fail than their smaller counterparts. I then estimate the same model, allowing the effect of size to vary across period. The magnitude of these effects is roughly uniform across models. Figure 3 plots these coefficient estimates. In the left panel I plot $\mu$, the parameter capturing the time-invariant relationship between size and the hazard rate, which is positive and statistically significant in all specifications. To gauge the magnitude of this effect, I have plotted in Figure 4 survival curves from the most conservative model, manipulating state size from the first to the third quartile. A substantial difference is apparent. For example, having survived up to 100 years the probability of surviving in the next period is roughly one-third less for the state at the seventy-fifth percentile of state size than the state at the twenty-fifth.

To examine the hypothesis that the relationship between size and failure changed during the period associated with the military revolution I compare the time-varying effect of size $\gamma_p$ across periods. Since each $\gamma_p$ captures the period-specific deviation from $\mu$, if $\gamma_p$ differs in a statistically significant way from 0 we can say that for period $p$ the relationship between geographic scale and survival differed from the average effect. Plotted in the right-hand panel of Figure 3 we see that this is not true for any time period; none of the time-varying effects differ in a statistically significant way from 0.

Matching my previous results, I find no evidence in favor of the notion that the military revolution affected the size of states. The survival probability of small states was greater than that of large states. Moreover, there is no evidence of a change in the relationship between geographic scale and failure during the expected period. Indeed, I find that the positive relationship between size and the failure rate of states did not change in a statistically significant way across time.

**Commerce and the Origins of the Modern State**

A number of scholars have drawn the causal link between the revival of commerce and patterns of European state formation. Despite this, few systematic empirical tests of this relationship have been undertaken. This section briefly outlines the relevant literature tying the re-emergence of commerce and urban life to political fragmentation.

In an early incarnation, Stein Rokkan argued that the existence of a “city belt” running through central Europe was the crucial determinant explaining why the modern territorial state developed in places peripheral to this productive core.63 The existence of a large number of prosperous urban centers prevented the rulers of any one from consolidating rule over the others. In peripheral England and France, for example, the absence of many urban centers allowed monarchs, by force or diplomacy, to establish rule over expansive territories.

FIGURE 3. Cox proportional hazards parameter estimates of the relationship between state size and failure

Notes: The left panel plots the parameter estimates of the time-invariant relationship between size and the hazard. Models 1 and 2 do not include time-varying estimates of this effect. Model 2 is a standard Cox-proportional hazards model and does not include any random effects. Models 3 and 4 include period-varying effects of size. Model 3 allows the effect of size to vary by one-hundred-year periods and model 4 allows the effect of size to vary by fifty-year periods. These period-varying effects, none of which are statistically distinguishable from 0, are plotted in the right panel.
Spruyt views the creation of new social groups as both an outcome of the revival of commerce and a catalyst for premodern innovation in the organization of the state.\textsuperscript{64} In those places where trade resumed, new towns and cities formed as hubs of economic life, allowing a class of burghers to establish effective independence.\textsuperscript{65} By virtue of these groups’ economic power, monarchs, particularly the Holy Roman Emperor, were impeded in their attempts to create geographically large states and were forced into political bargains accepting the de facto independence of these states.\textsuperscript{66} Moreover, because of their wealth and ability to establish institutions like those promoting private property, they could accumulate economic resources and invest them in urban development.

\textsuperscript{64} Spruyt 1994b.

\textsuperscript{65} Even Tilly who clearly stressed the centrality of war, came to recognize that his early work (Tilly 1975; 1985) “emphasized the extractive, coercive, war-making side of state activity so strongly as to substitute implicitly a new unilinear model” of state formation (Tilly 1994; p. 5.) In later work he sought to elucidate the role of “capital,” for example, the effects of uneven access to financial resources – typically concentrated in cities – on the process of state formation Tilly (1990) and towards the end of his career he had more fully developed a theory of state-formation that placed uneven urban development at the fore Tilly and Blockmans 1994, Tilly 1994.

\textsuperscript{66} Blockmans (1989) similarly finds that the material advantages maintained by social groups residing in cities allowed them to obstruct the “voracious” centralizing efforts of monarchs seeking to construct large, “national” states.

\textbf{FIGURE 4. The estimated survival curves from Model 1}
the Hansa capable of providing collective security, these states could wield sufficient military force to not merely claim independence but to sustain it over time.67

What is common to these economic theories is that changes in patterns of economic activity, principally those associated with changes in commerce and urban life, altered the balance of political power, causing political fragmentation in the urban core of Europe. This section examines trends in the number and size of states at the regional level, providing evidence that only in the most productive places in Europe, a central regional band extending, roughly, in an arc from the Low Countries, through the Rhineland and into Northern Italy, could small political communities persist. That is, in the places where the “commercial revolution” of the first half of the previous millennium took hold political fragmentation ensued. However, in less-productive regions, in the absence of dense urban and commercial growth, large territorial states formed.

![Figure 5](https://doi.org/10.1017/S0020818316000308)

**FIGURE 5.** The average size and number of states separating out the urban European from the rest of the continent

67. The political economy literature on state size has constructed several related models. Friedman (1977) is the closest in explicitly linking changes patterns of trade—via changes in appropriable revenue—to the size of states. Bolton and Roland (1997) explore the optimal size of states in a democratic setting where a voter faces a tradeoff between the economic gains from a large state and from possibly lower levels of redistribution following secession. Alesina and Spolaore (1997, 2005) construct a series
As initial evidence Figure 5 plots both the trend in logged state size and number of states across time, dividing the map into two broad regions: the urban European core (the area I described earlier), and the remainder of the map, or what I will call the periphery. In both regions the number of states is increasing before the early thirteenth century. After this period the upward trend in the number of states continues in Central Europe whereas in the periphery it plateaus and then begins to decline in the early fifteenth century. Similarly the average size is initially declining in both groups. However, in the periphery, beginning in the early sixteenth century, the average size starts to increase whereas it continues to decline in the center.

These patterns coincide with the reemergence of trade and a general revival of commercial and urban life during the first third of the millennium. Like the regional pattern in the number and size of states, these economic trends were not uniform across space but were geographically concentrated in a manner nearly identical to the patterns of state formation I described. The emergence of towns and cities—a product of the geographically concentrated revival of commerce and economic development—created groups in these places with the resources necessary to maintain or claim independence, thereby preventing the consolidation of territorial states.

Urban Growth and State Formation

This section directly estimates the relationship between development and state formation. As the unit of analysis I employ arbitrarily defined pieces of geography, 10,000 square-kilometer grid squares, and estimate the effect of changes in urban population on the number of states forming within the bounds of a given unit. The data on urban population come from the widely used data set of Bairoch, Batou, and Chevre which describes the population sizes of the about 2,200 towns which ever had 5,000 or more inhabitants at some time between 800 and 1800.

of models where a decisive actor selects the number and size of states in the face of a tradeoff between local tastes in public goods provision and the fixed costs associated with their provision. In these models access to trade (or a world in which barriers to trade are limited) results in smaller states as it decreases the returns (relative costs of public provision) associated with large size.

68. The “core" here consists of grid-units whose centroid falls within contemporary Germany, The Netherlands, Belgium, and Northern Italy (Vall de’Aosta, Piemonte, Liguria, Trentino Alto Adige, Veneto, Friuli-Venezia Giulia). The periphery are those units whose centroid falls outside of the core.

69. I have estimated all of the subsequent analysis using an alternative dependent variable, a Herfindahl like index of fragmentation, defined for each grid-square j as $H_j = \sum_i (\frac{\text{Area}_i}{\text{Area}_j})^2$ where Area, is the total area held by state i in grid-square j and Area, is the total area in grid-square j. This measure is highly correlated with the number of entrants ($\rho = -.72$) and all of the results remain qualitatively the same if this measure is used. I treat grid-units that have no state, but which by 1790 come under the control of one, as having a count of 0 in the periods of no state control. Estimating all models excluding these “zero” units yields qualitatively similar results.

I estimate the following model:

\[
\text{Number of States}_it = \alpha + \beta \cdot \log(\text{Urban Population}_it) + \epsilon_{it} \quad (1)
\]

where the parameter of interest \( \beta \) captures the effect of urban population on the number of states on a given grid-square. I estimate this relationship via pooled OLS and, since the number of entrants is by definition a non-negative integer, negative binomial regression. The results are summarized in Table 1. I estimate three models, the marginal effects of which are extremely close in size. For each model I successively add in year effects and latitude and longitude as controls. The estimated effects of a 100 percent change in urban population range from 14 to 20 percent, a substantial effect when the differences between the maximum urban population in any given year range from forty-seven to sixty-six times the mean.

**TABLE 1. The relationship between urban population and political fragmentation**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \log(\text{Urban Population}_it) )</td>
<td>0.19</td>
<td>0.11</td>
<td>0.20</td>
<td>0.11</td>
<td>0.14</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>[0.15, 0.24]</td>
<td>[0.10, 0.13]</td>
<td>[0.15, 0.24]</td>
<td>[0.10, 0.13]</td>
<td>[0.10, 0.19]</td>
<td>[0.05, 0.08]</td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>OLS</td>
<td>Neg Bin</td>
<td>OLS</td>
<td>Neg Bin</td>
<td>OLS</td>
<td>Neg Bin</td>
</tr>
<tr>
<td>Year Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lat/Long</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
<td>0.22</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
</tr>
<tr>
<td>T</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.13</td>
<td>0.14</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Coefficient estimates of the effect of urban population on the number of entrant states on a given grid-square. Ninety-five percent confidence intervals in brackets estimated from robust standard errors clustered by grid-square.

These results provide some initial evidence that the most urban quadrants were those that had the most states form on them. As a consequence, the average size of these states in these places was by definition smaller. Nevertheless, these results need not signify a causal relationship. That is, the assumptions necessary for this regression analysis to consistently estimate \( \beta \) from Equation 1 are unlikely to be satisfied. The size of the urban population and the number of states that exist on a given piece of territory are likely driven by a number of common but unobservable confounders. Moreover, simultaneity bias, the fact that state-formation processes are likely a cause of urban development as well as its consequence, may plague estimates of this sort.71

71. See Dincecco and Onorato 2014.
To show that the relationship between changes in patterns of urban development and state formation is causal, I take an instrumental variables approach, exploiting random climatic variation in the ability of arbitrary pieces of geography to support large urban populations. Using a set of paleoclimatological sources I construct an estimate of the propensity of land to feed large populations by growing cereals like wheat. I show that the ease with which some places could produce calorically dense and easily storable foods of this sort is a strong cause of urban population and, I argue, has no direct effect on state-formation processes. It follows that this measure is an ideal instrument, satisfying the necessary exclusion restriction needed to estimate causal effects.

Urban Growth and Agricultural Productivity

Cities as centers of economic specialization can exist only once populations are able to devote effort to activities other than subsistence. Places that could produce certain foods most easily were also those more likely to develop as urban centers. In premodern Europe these locations could grow a specific set of crops that were superior to other alternatives in terms of their ability to support sizable populations over an extended period.72

To identify the causal effect of urban development on state formation in an instrumental variables framework I exploit random climactic variation in the ability of a given location to produce key agricultural outputs necessary to support large populations. A number of economic historians have directly related the natural predisposition for some places to feed large groups to the development of urban life and the revival of commerce. Pirenne, for example, argues that the location of towns in premodern Europe was a function of natural geography, that “in a more advanced era, when better methods would permit man to conquer nature and to force his presence upon her despite handicaps of climate or soil, it would doubtless have been possible to build towns anywhere the spirit of enterprise and the quest of gain might suggest a site.” This was, however, not the case. Rather, “the first commercial groups were formed in neighborhoods which nature had disposed to become … the focal points of economic circulation.”73

I focus on the ability of some places to produce cereals like wheat for two reasons. First, the European diet of the premodern era centered on the consumption of complex carbohydrates derived from cereals. Economic historian Robert Lopez notes that “in the form of bread, porridge, or mush, cereals were almost everywhere the basis of human alimentation.”74 Cereals were central to diets not only across

72. See Bairoch and Braider 1991; De Vries 1984; and Nicholas 1997 for three prominent examples of the view that increases in agricultural productivity were a necessary precondition for the growth of cities.
74. Lopez 1976, 37.
European geography but across classes as well and were integral to the consumption of the aristocracy and peasantry alike, although certainly in unequal proportions.\footnote{Duby 1974.}

Second, the ability to grow cereals has been directly linked to the support of large populations. Cereals like wheat, unlike other plants, are most capable of feeding large populations with minimal effort; cereal crops, unlike fruits, pulses, or nuts, are extremely fast growing, high in calories from carbohydrates, and have extremely high yields per hectare.\footnote{Diamond 1997.} Moreover, unlike other crops, cereals can be stored for long periods of time enabling communities to smooth consumption over extended periods. The ability to feed large populations was key to the development of cities. Since in premodern Europe the principle component of diets was cereal like wheat, foods that are particularly good at supporting large populations, climatic variation across time and space in the ability to grow these crops serves as a good encouragement for urban growth.

Three related facts allow me to summarize the logic of my identification strategy. 1) We have strong theoretical reason to view cities as the locus of the socioeconomic groups capable of resisting the construction of centralized territorial states. 2) The size of cities, moreover, is a strong measure of these groups’ economic power. And 3) before the industrial revolution the size of the urban population a given piece of territory could support relied upon the local biogeography, particularly the capacity to grow cereals. I propose to exploit random perturbations in the ability to sustain large urban populations as an instrument to understand the effect of city-size—our measure of the strength of commercial groups—on the number of states forming on a given piece of territory. In the appendix I detail how the instrument is constructed from paleo-climatic sources. Furthermore, to show that the logic of the instrument holds, in the appendix I use twentieth-century data on temperature and two measures of the ability to grow wheat to demonstrate the robust relationship between this measure and actual production of cereals.

\textit{The Validity of the Instrument}

In an instrumental variables framework several assumptions must be satisfied to consistently estimate $\beta$, the parameter describing the relationship between urban population and the number of states forming on a given piece of territory. The first assumption, random assignment, is likely satisfied since until the nineteenth century there was little to no human effect on climate. The second assumption that the instrument is strong is also satisfied. In each of the first-stage regressions all tests against weak instrumentation meet conventional levels of statistical significance.

The last assumption, the exclusion restriction, is also likely satisfied. For this to be violated the instrument, deviation from the optimal growing temperature for cereals,
would have to affect the number of states entering on a given quadrant in some way other than through my measure of economic development, urban population. It seems unlikely that changes in this measure of climate will affect the choice to form a new state or for states to fail other than through their effect on economic incentives as proxied by urban development.

To see this we can broadly divide the constraints facing state makers into two components. The first are the costs of statehood. Substantively these can be taken to mean the size of militaries and bureaucracies necessary to maintain statehood. Shocks to the propensity for a given piece of geography to support urban populations should have no direct affect on this particular incentive to form or dissolve as an independent state. The instrument perturbs the major component of the second constraint, the economic surplus available for latent or existing states to claim, here measured by urban population. The question then becomes do changes in this measure of temperature affect economic incentives not captured by changes in urban population?

Here, we must recognize two facts about the preindustrial economy. The first is that transportation costs were extremely high such that long-distance trade was concentrated in luxury goods.\(^\text{77}\) As such, markets for agricultural products were, for much of this period, local.\(^\text{78}\) Second, the market for surplus agricultural product was concentrated in towns and cities. That is, those who specialized in non-agricultural sectors — located in towns — exchanged their goods for the surplus agricultural product produced in the hinterland. Because of the limited ability to trade these goods across extreme distances, changes in the productivity of agriculture affect these local markets either by shifting labor from the countryside to towns or by allowing a greater number of individuals to specialize in non-agricultural activity, both of which result in a greater number of people living in towns and cities. Since my measure of urban population includes towns as small as 1,000 individuals, I should pick up these dynamics.

Moreover, since the changes in optimal growing temperature over a hundred-year panel are extremely slow, the long-term trend would be very difficult to perceive at any given point in time in an era before meteorological data was systematically collected. Because of this these changes would only affect the economic incentive to form states through their long term effects—the ability to sustain large populations.

Nevertheless, in case the growing climate for wheat has some direct effect on urban population, I attempt to control for alternative channels through which the optimal growing temperature might affect state formation. I control for the ways in which climate, other than through the deviation from this optimal temperature for cereals, might affect the number of states on a given square. I do this first by controlling for both latitude and longitude. Since climate is strongly correlated with geographic location, controlling for the position in space should similarly control for the effects of climate other than through the optimal growing temperature.

\(^\text{77}\) Findlay and O’Rourke 2007.

\(^\text{78}\) A number of recent econometric studies find little to no apparent market integration in Europe before the modern period. See Allen 2001; Bateman 2011; Özmucur and Pamuk 2007; and Unger 2007.
By including grid-square fixed effects I get similar results. Here identification is most plausible and comes from within-unit variation, for example, shocks from the mean of each unit’s distance from the optimal growth temperature for cereals. In this sense I am again controlling for long-term climatic conditions and identification is coming from the random changes from this long-term value.

In the online appendix I present results where I construct and control for grid-square-level estimates of per capita income. By controlling for income at this level, I identify the effect of urban population after accounting for the possible violation of the exclusion restriction as it might operate through other economic channels captured by per-capita wealth. Across these specifications the results remain positive and statistically significant. Moreover, they are very close in magnitude to those presented here.

Furthermore, in the online appendix I include specifications that consider additional possible confounders and violations of the identifying assumptions. First, in case my results are being driven by my coding of small states within the Holy Roman Empire I divide the data and estimate the same specifications separately for the core region of Europe and the remaining periphery. While the resulting effects are slightly smaller in the periphery than the core the main result holds. Similarly, in case there are confounders that vary by region I include a series of regional dummies. Again, the results remain roughly constant to those presented here.79 Similarly, in the cross-sectional estimates I account for access to over-land travel routes by including a measure of river density. Again, the results remain unchanged.

It could be that my results are driven by units where the city and territory were perfectly coterminous (e.g., city-states). That is, were states small because they were simply city-states and not because of changes in the material wealth and capabilities of the social actors living in cities? To assess whether political fragmentation is being driven by city-states or by changes in the economic potential of territory, in the appendix I rerun the main set of statistical models controlling for the number of cities in each territory. By controlling for the number of cities on a given territory we are comparing changes in urban population across geographic units that have the same number of cities within them. Across specification the results are positive and statistically significant, indicating that the effect of changes in urban population is not operating through the creation of more city-states. In fact, across specification when I control for the number of cities on a given unit the effect of changes in urban population on the number of states in a given grid-square increases by over 30 percent.

On top of these robustness checks, following Conley, Hansen, and Rossi80 I conduct a sensitivity analysis of the assumption that the instrument perfectly satisfies the exclusion restriction. Detailed in the appendix, I show that my results are robust at the 5 percent significance level up to a direct negative effect of the instrument on the

79. I divide the map into the following regions coterminous with contemporary boundaries: the British Isles and France, Iberia, Italy, Scandinavia and the Baltics, Eastern Europe including Russia, and the Ottoman Territories which includes the Balkans.

outcome equivalent to 36 percent of the estimated effect of urban population. This would require a substantively large violation of the identifying assumption to obtain a null effect of urban population on the number of states formed on a given geographic unit.

**Instrumental Variables Results**

The instrumental variables estimates of $\beta$ are shown in Table 2. The effect sizes are rather large, a 100 percent increase in total urban population on a given grid-square is expected to increase the number of states locating within that same unit by between just under four- and just over eight-tenths of a new state depending upon specification. Again, this is a rather large effect size because the inequality between geographic units is quite large. Taking the smallest 2SLS estimated effect size, .37, it would only take slightly more than a two-and-a-half-fold increase in the total urban population to increase the number of states on a given unit by one.

**TABLE 2. 2SLS estimates of the effect of urban population on political fragmentation**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Urbanization$_{it}$)</td>
<td>0.37</td>
<td>0.36</td>
<td>0.67</td>
<td>0.79</td>
<td>0.58</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>[0.31, 0.42]</td>
<td>[0.31, 0.42]</td>
<td>[0.39, 0.95]</td>
<td>[0.40, 1.19]</td>
<td>[0.15, 1.01]</td>
<td>[−0.05, 1.82]</td>
</tr>
</tbody>
</table>

**Controls**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lat/Long</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Entrants$_{it-1}$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>First Differencing</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>FirstStage</td>
<td>$\betaITEM_{it-10.5l}$</td>
<td>$\betaITEM_{it-10.5l}$</td>
<td>$\betaITEM_{it-10.5l}$</td>
<td>$\betaITEM_{it-10.5l}$</td>
<td>$\betaITEM_{it-10.5l}$</td>
<td>$\betaITEM_{it-10.5l}$</td>
</tr>
<tr>
<td>F</td>
<td>351.06</td>
<td>352.49</td>
<td>24.18</td>
<td>22.31</td>
<td>9.76</td>
<td>3.81</td>
</tr>
<tr>
<td>N</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
<td>1684</td>
</tr>
<tr>
<td>T</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

**Notes:** 2SLS estimates of the effect of urban population on the number of states existing on a given 10,000 square-kilometer piece of geography. Controls for latitude and longitude, year effects, and fixed effects are continued in both the first- and second-stage regressions. Ninety-five percent confidence intervals estimated from robust standard errors clustered by grid-square are in brackets.

The estimates where identification is coming from changes in optimal growing temperature after accounting for other possible climatic channels, those controlling for latitude and longitude as well as for unit-specific heterogeneity (Models 3 through 6) are slightly larger, ranging from .67 to .88. Of these, the fixed effects (Models 4 and 5) and first-differencing (Model 6) estimates have a ready interpretation. Because they are identified from within-unit changes, they tell us how the
number of states on a given geographic unit changes over time with changes in urban population. As an additional specification Model 5 includes the lag of the number of states. Controlling for the number of states in a territory during the last century, a 100 percent change in the urban population still causes .58 of a new state to form on the same piece of geography in the next century.81 Interpreted this way these results provide an explanation for why the central European core became simultaneously more urbanized and politically fragmented. Distance from this central corridor was highly predictive of the size of the urban population living within the bounds of a given geographic unit.82

In the mechanism I propose, urban growth causes fragmented political rule by creating new social groups associated with the revival of commerce and economic development. The wealth embodied in cities, the only real location of economic specialization in the pre-industrial world,83 made these groups capable of using their newly acquired material capabilities to assert independence and preventing the construction of would-be territorial states. To assess the claim that the effect of urban development on political fragmentation operates through the presence of new social actors associated with commercialism, in the appendix I exploit data on proto-industrial (textile and metallurgical) activity as a proxy for these actors’ presence. First, I simply describe the statistical relationship between proto-industry and political fragmentation and show a substantial positive relationship between the two. Next, using the methodology proposed by Imai, Keele, and Tingley,84 I treat the presence of the actors proxied by the existence of proto-industry as a mediating variable affecting the number of states forming on a given piece of geography. That is, I estimate the relationship between urbanization in 1200 and political fragmentation in 1500 as it operates through the existence of proto-industrial activity. In all, I find that about 40 percent of the effect of urban population can be explained by my rough proxies for the presence of commercial social actors.

Where new towns and cities formed, the urban groups that inhabited them were able to construct and maintain independent states. In the absence of city life, the makers of territorial states won out, consolidating large swaths of geography into comparative few large states.

81. Additional specifications, including controls for possible region effects, geographic features like river density, and estimating separate effects for the core and periphery are presented in Table A2 in the appendix.
82. On this see Abramson and Boix 2016.
84. See Imai, Keele, and Tingley 2010; Imai, Keele, Tingley, and Yamamoto 2011; and Imai, Tingley, and Yamamoto 2013.
Conclusion

The empirical results presented in this paper indicate that observed changes in the number and size of states before the French Revolution were driven by economic factors. In those regions where urban life reclaimed a foothold, where new cities formed and the new social classes that emerged within them were capable of asserting themselves, political life fragmented. These new political actors, by virtue of their material wealth, were capable of resisting the centralizing efforts of the leaders of would-be territorial states. In the continent’s periphery, where these groups were weak and incapable of subverting the construction of large states, the earliest national states formed. Even so, these territorial states—before the nineteenth century—did not dominate as an institutional form. Small types of political organization not only persisted but flourished right up until the French Revolution. Rather, the modern state as a geographically centralized entity that monopolized violence over large territories is a fundamentally contemporary institution asserting its dominance only in the nineteenth century.

Although the notion of the state I have presented is founded upon Weberian coercion rather than constitutive principles that some scholars of IR view as crucial, my findings, nevertheless, speak to claims of fundamental early-modern changes in the organizing principles of the state system. In line with the work of Krasner and Osiander, the “Westphalian Myth” is just that—a widely held belief that fails to materialize when confronted with the realities of data. That is, if the very constitution of the international system was fundamentally altered by events surrounding the Thirty-Years war, the consequences, at least in terms of the type of state that continued to survive, were not immediate and are not borne out in data.

In emphasizing the importance of economic change in the development of the territorial state my findings run counter to much of the existing literature that places substantial theoretical weight on changes in patterns of war and war making. The specific effects attributed to the changes in war making, believed to have selected large states against their smaller counterparts, do not appear when the entire distribution of cases is considered. Still, the system of states we observe today is characterized, almost exclusively, by large territorial states. This paper highlights the tendency of contemporary scholars to conflate scale with power. Power manifest in an observed ability to persist—the main objective of states according to much of international relations theory—was, according to my findings, orthogonal to size. Thus, those who focus upon the history of “great powers” to understand the development of the state and state system build in an anachronistic bias. Rather, as this study shows, the execution of power between states should be studied by focusing upon the entire distribution of types, small and large alike.

86. Waltz 2010.
Supplementary Material

Supplementary material for this article is available at https://doi.org/10.1017/S0020818316000308.

References


