History's Masters

The Effect of European Monarchs on State Performance*

Sebastian Ottinger (UCLA)

Nico Voigtländer (UCLA, NBER, and CEPR)

First version: 31 August 2020
This version: 14 July 2021

Abstract

We create a novel reign-level dataset for European monarchs, covering all major European states between the 10th and 18th centuries. We first document a strong positive relationship between rulers' intellectual ability and state-level outcomes. To address endogeneity issues, we exploit the facts that i) rulers were appointed according to primogeniture, independent of their ability, and ii) the wide-spread inbreeding among the ruling dynasties of Europe led over centuries to quasi-random variation in ruler ability. We code the degree of blood relationship between the parents of rulers. The 'coefficient of inbreeding' is a strong predictor of ruler ability, and the corresponding instrumental variable results imply that ruler ability had a sizeable effect on the performance of states and their borders. This supports the view that 'leaders made history,' shaping the European map until its consolidation into nation states. We also show that rulers mattered only where their power was largely unconstrained. In reigns where parliaments checked the power of monarchs, ruler ability no longer affected their state's performance. Thus, the strengthening of parliaments in Northern European states (where kin marriage of dynasties was particularly wide-spread) may have shielded them from the detrimental effects of inbreeding.

JEL: D02, D73, N43, P14, P16.

Keywords: Leaders, State Performance, Institutions

*We would like to thank Paola Giuliano, Thilo Huning, Casey Petroff, Marta Reynal-Querol, Romain Wacziarg, Melanie Wasserman, as well as seminar audiences at UCLA, the RIDGE Virtual Forum 2020, the Armenian Economic Association, HSE, and UC Berkeley for helpful comments and suggestions. Thomas Groesbeck, Angelo Krüger, and Logan Spencer provided outstanding research assistance. We are grateful to Scott Abramson for kindly sharing his data on European state borders.

Corresponding author: Nico Voigtländer (nico.v@ucla.edu)

"It was a time ... 'when the destinies of nations were tied to bloodlines'."

– Robert Bartlett ("Blood Royal: Dynastic Politics in Medieval Europe," 2020, p.432)

1 Introduction

A growing literature points to the importance of leaders for the performance of their firms or organizations (c.f. Bertrand and Schoar, 2003; Malmendier and Tate, 2005). Likewise, characteristics of local leaders have substantial effects on public goods provision and conflict in the region or community under their control (c.f. Chattopadhyay and Duflo, 2004; Logan, 2018; Do, Dray, Huillery, and Keene, 2020; Eslava, 2020). However, identifying such effects at the *national* level is difficult. The question whether national leaders can shape their countries' fortunes has been widely debated in the social sciences over the past two centuries. Early advocates proposed the strong view that the "history of the world is but the biography of great men" (Carlyle, 1840, p. 47). Subsequent qualitative analyses of biographies and comparative studies have lent support to an important role played by individual leaders.[1] On the other hand, a literature in the Marxist tradition has argued that underlying structural demographic and economic forces determine both a state's performance and the endogenous emergence of its leaders. Scholars in this strand view leaders has "history's slaves" (Tolstoy, 2007, p. 605); in the words of Braudel and Reynolds (1992, p. 679): "Men do not make history, rather it is history above all that makes men."[2]

Economists have brought identification to this debate. Jones and Olken (2005) show that random leadership transitions due to natural death or accidents are followed by changes in economic growth over the post-WWII period, providing convincing evidence that leaders do indeed matter. Besley, Montalvo, and Reynal-Querol (2011) expand the underlying data to 1875-2004, documenting that random departures of educated leaders cause particularly strong reductions in growth. While these results are an important step forward in identifying a causal effect of leader capability on state performance, some open issues remain: The actual "quality" of leaders is unobserved; it is estimated as average economic growth a few years before and after a random death, and it therefore captures a plethora of other factors. In the same context Easterly and Pennings (2020) point

[1] See for example Kennedy (1989) and Gueniffey (2020). A literature in political psychology has also underlined the importance of leaders' intellectual capabilities (c.f. Simonton, 2006). Horowitz, Stam, and Ellis (2015, p. 11) conclude that "leaders do matter in systematic ways that we can understand."

[2] In his opus magnum 'War and Piece,' Russian writer Lev Tolstoy attested to leaders that "every act of theirs...is...predestined from eternity" (Tolstoy, 2007, p. 605). Karl Marx wrote: "Men make their own history, but they do not make it as they please; they do not make it under self-selected circumstances, but under circumstances existing already, given and transmitted from the past. The tradition of all dead generations weighs like a nightmare on the brains of the living" (Marx, 1907, p. 5). Friedrich Engels elaborated: "But that in default of a Napoleon, another would have filled his place, that is established by the fact that whenever a man was necessary he has always been found: Caesar, Augustus, Cromwell, etc." (Engels, 1968, p. 704). This alternative view, cautioning the interpretation of history through the biography of individuals, is well alive in the modern debate as well. March and Weil (2009, p. 97) assert that "it is not at all clear ... that major differences in the success of organizations reflect differences in the capabilities of their leaders, or that history is the product of leaders' actions."

out that the high volatility in growth makes it difficult to distinguish between random spikes and actual effects of individual leaders. In addition, while the timing of the transition is exogenously determined by death, the appointment of the subsequent leader is endogenous. Finally, the necessary annual GDP data is only available for the modern period, so that the causal role of leaders in history (where it has been debated most intensely) has not been examined. To make progress on these fronts, the ideal experiment would feature a sequence of randomly appointed leaders with varying, observed capabilities who govern over a long horizon. Europe's monarchies over the late medieval and early modern period provide a context that resembles such a setting.

We study European monarchs over the period 990-1795, assembling a novel dataset on ruler ability and state performance at the *reign* level. To identify a causal effect of ruler ability, we exploit two imminent features of ruling dynasties: first, primogeniture – the pre-determined appointment of rulers by birth order, independent of their ability; second, variation in ruler ability due to the widespread inbreeding of dynasties. Importantly, the negative effects of inbreeding were not understood until the 20th century; if anything, rulers believed that inbreeding helped to preserve 'superior' royal traits. In addition, the full degree of consanguinity (genetic similarity) was unknown due to complex, interrelated family trees over generations. Together, these features deliver quasi-random variation in ruler ability.

We collect data on the ability of 336 monarchs from 13 states, building on the work by historian Woods (1906), who coded rulers' intellectual capability and character traits based on hundreds of biographies. While Woods explicitly aimed to assess ruler's intellectual capability and character traits independent of the performance of their state, this coding nevertheless raises endogeneity concerns. We thus instrument for ruler ability with the coefficient of inbreeding of rulers. We collect this variable for all rulers with the necessary information on family lineages from a rich genealogical database. The coefficient of inbreeding is a strong and robust predictor of ruler ability. To assess state performance during a ruler's reign, we use three different outcome variables. First, a coding of state performance that is based on several underlying metrics and summarizes the work by numerous historians (Woods, 1913). Because there are natural concerns with this subjective coding, we use two additional, objective measures. Our second outcome variable measures changes in land area during each ruler's reign. We derive this variable from Abramson (2017), who provides European state borders at 5-year intervals over the period 1100-1795. Finally, we also calculate the change in urban population within the (potentially changing) area ruled by each monarch, combining border changes with the urban population data of Bairoch, Batou, and Chèvre (1988).

We find that ruler ability is strongly associated with all three measures of state performance, and our IV results suggest that this relationship is causal. These findings hold when we use country fixed effects and account for time trends by comparing only rulers whose reigns overlapped. A one standard deviation (std) increase in ruler ability leads to about a one std higher state performance, to an expansion in territory by about 17 percent, and to an increase in urban population by 19 percent. The latter is driven mostly by capable rulers expanding their territory into urbanized areas (as opposed to urban growth within existing borders). We also study the institutional circumstances under which individual rulers mattered particularly strongly. We construct a novel country-year specific measure of historical constraints on rulers, combining definitions of the modern Polity IV score with historical sources on factors such as the power of parliaments. To bypass endogeneity issues, we use constraints on rulers from the period just before they were appointed, and we only focus on our two 'objective' outcome variables because historians' subjective assessment of state performance may be influenced by the state's institutions. We find that the ability of unconstrained leaders had a strong effect on state borders and urban population in the reign, while the capability of constrained rulers made almost no difference.

We run a battery of checks to confirm the robustness of our results and the validity of our IV strategy. Our findings are robust to numerous alternative specifications such as using dummies for different levels of ruler ability, using ordered Probit, as well as clustering at the country, dynasty, and century level. By controlling for country fixed effects, our baseline results capture state-specific features. Our findings are unaffected when we exclude episodes of governments by regents (for example, when rulers were minor at the time of their appointment), when excluding episodes of foreign rule, or those when the same monarch governed more than one state. We also verify and extend Woods' (1906) and (1913) coding of ruler ability and state performance, showing that our results are robust to using only our own assessments, to using an extended sample covering Poland and Hungary additionally, and until WWI, and even to a conservative coding that specifies ambiguous cases so that they work against a positive association between state performance and ruler ability. Finally, we confirm the robustness of our results in alternative pair-level regressions that compare concurrent rulers *across* countries within shorter time periods.[3]

Our IV results, in particular, are robust to excluding cases of high inbreeding coefficients, and to restricting the sample to those rulers for whom historical sources explicitly confirm that they rose to power via primogeniture; for example, these specifications exclude all cases where a ruler from a new dynasty came to power. We also discuss potential threats to the exclusion restriction (i.e., that inbreeding affected ruler ability but was not related to state performance via other channels). For instance, such a threat would arise if royals tended to marry their kin when state performance was low, and if, in addition, low state performance during parents' reign led to low performance during the reign of the offspring. We show that this is not the case – past state performance predicts

[3]We identify for each monarch all rulers from other countries that had at least a one-year overlap in their reigns. We then run pair-level regressions in differences, also controlling for reign-specific fixed effects. We find that differences in inbreeding across concurrently ruling monarchs are a strong predictor of differences in their ability, which in turn drives differences in country performance.

neither current state performance nor ruler ability, and our IV results are robust to controlling for lags in the coefficient of inbreeding.[4] Another threat to our identification would arise if monarchs made strategic decisions on kin marriage for reasons that are correlated with the *prospects* for future state performance.[5] We address this possibility by exploiting only the *hidden* component of inbreeding that was due to kin marriage over previous generations – and which could only be assessed with methods in genetics that emerged in the early 20th century.[6] We confirm our IV results based on this restrictive measure of inbreeding.

Our paper makes novel contributions both in terms of data collection and empirical results. We are the first to track the performance of all major European states at the *reign* level over a horizon of several centuries, allowing for fluid changes in their borders. In contrast, previous seminal papers have typically used today's country borders as their unit of analysis, and they have relied on (half-) century level outcomes such as GDP per capita or urbanization (c.f. Acemoglu, Johnson, and Robinson, 2005; Nunn and Qian, 2011; Dittmar, 2011). Our dataset thus opens a new dimension to study Europe's economic history. Using this novel dataset, we contribute to a large literature that has debated the role of rulers for nationwide outcomes. We analyze a period that has been at the center of this debate since its beginning in the 19th century.[7] Our paper is the first to provide causal identification of the importance of European rulers over the late medieval and early modern period. State performance during this period had long-lasting consequences, as the foundations for the modern nation states were laid across Europe. Our findings suggest that the territorial organization of Europe as we know it is at least in part the result of chance, embodied in the ability of individual rulers.

We also contribute to a strand of the literature that has underlined the importance of individual

[4]For our first, subjective, measure of state performance, the exclusion restriction could also be violated if inbreeding affected the *assessment* of state performance by historians – for example, if they hypothesized negative effects of inbreeding on rulers, and in turn of bad rulers on states. This is unlikely because Woods was a proponent of 'Social Darwinism,' viewing history as a process of natural selection. Woods' (1913) hypothesis was that moral and intellectual ability is inheritable, so that kin marriage among successful dynasties would produce *better* rulers. This introduces a bias *against* our findings. In addition, the negative effects of inbreeding on fitness were not accepted in biology until the second half of the 20th century. Conclusions such as the following were common: "Inbreeding as such does not cause degeneration; the testimony of biologists is conclusive on this point" (White, 1948, p. 417). See Wolf (2005) for detail on this debate. Correct measures of inbreeding were first developed by Wright (1921). When these measures eventually became available, Asdell (1948) showed that Woods' hypothesis was wrong, using Woods' (1906) own coding of ruler ability.

[5]Note that in this case, controlling for past state performance would not necessarily address the endogeneity issue because the potential for future performance may be uncorrelated with current performance.

[6]Becker, Ferrara, Melander, and Pascali (2020) similarly exploit variation in the pedigree of nobility that was not a direct choice of the nobles themselves (changes of individual positions in the nobility network) to study the effect of conflict on state capacity in the German lands.

[7]For proponents of the "rulers matter" view see for example Carlyle (1840), Weber (1921), William (1880), and Spencer (1896). For the opposite view that "history makes men" see Marx (1907), Engels (1968), Braudel and Reynolds (1992). More recent contributions to this theoretical and empirical debate include March and Weil (2009), Simonton (2006), and Xuetong (2019), as well as Acemoglu and Jackson (2015), Alston (2017), and Alston, Alston, and Mueller (2021).

characteristics of leaders in both managerial and political settings.[8] In the managerial literature, Clark, Murphy, and Singer (2014) have documented that CEOs matter less when they are constrained by a well-defined governance structure, echoing the findings on constrained politicians by Jones and Olken (2005) and Besley et al. (2011). Similarly, Besley and Reynal-Querol (2017) document higher economic growth under hereditary (as compared to non-hereditary) leaders when constraints on them were weak, using data from 1875 onwards. Besley and Reynal-Querol (2017) interpret these correlations as evidence that hereditary leaders have a longer time horizon, improving policy choices.[9] Our results focus *only* on hereditary leaders, showing that their ability (which is not observed by Besley and Reynal-Querol, 2017) had strong effects on state performance – unless it was checked by institutional constraints. This latter finding is particularly interesting because the detrimental effects of inbreeding became more severe in the 17th and 18th century, after centuries of inbreeding.[10] By that time, parliaments across Northern Europe had expanded their power (Van Zanden, Buringh, and Bosker, 2012). Thus, our results suggest that parliaments protected (some) European states from the adverse effects of their ruling dynasties' inbreeding.

The paper is organized as follows. Section 2 introduces the historical background of European monarchs and Section 3 discusses our data sources and coding. Section 4 shows our main empirical results and discusses our identification strategy. Section 5 examines heterogeneity by institutional constraints on rulers. Section 6 concludes.

2 Historical Background: Europe under Dynastic Rule

This section briefly reviews the historical background of European monarchs in the late medieval and early modern period. We pay particular attention to those features that render the setting a rich testing ground for identifying the causal effect of national leaders on state performance.

2.1 Rulers and Country Performance

A plethora of studies in a variety of fields have argued that national leaders affect the fortunes of their countries. For example, the literatures in historiography and political science are full of cases linking the fate of countries to their rulers' actions and abilities.[11] One often-cited case is the series

[8]C.f. Bertrand and Schoar (2003), Malmendier and Tate (2005), Bloom and Van Reenen (2007), and Becker and Hvide (2013) for the importance of managerial traits; and Ferreira and Gyourko (2014), Yao and Zhang (2015), Logan (2018), Dippel and Heblich (2021), and Assouad (2020) for results on traits of political leaders.

[9]A related literature studies political dynasties in modern democracies, where some prominent families repeatedly have members *elected* to important offices (c.f. Dal Bó, Dal Bó, and Snyder, 2009; George and Ponattu, 2018). In contrast, in our setting, succession was guaranteed by law, and dynasties were the central governing bodies over the course of centuries.

[10]The average coefficient of inbreeding increased by 80% between the 15th and the 18th century. In Northern Europe (comprising the countries of England, Scotland, the Netherlands, Denmark, and Sweden), this increase was particularly pronounced, with 180% as compared to 42% in the remaining countries.

[11]Biographies published by historians consistently emphasize the importance of certain individuals and their leadership qualities in shaping the nations they ruled – e.g. for the U.K. see Roberts (2018) and MacCulloch (2018) for the effects of Cromwell's and Churchill's actions and convictions upon their native England. Nicholas (2021) writes:

of able rulers accompanying Prussia's rise from small polity to great power.[12] Similarly, Kennedy (1989) notes that one of the factors aiding Sweden's "swift growth from unpromising foundations" was "a series of reforms instituted by Gustavus Adolphus and his aides," increasing the efficiency of administration and allowing Sweden under Gustavus to play an outsized role in the Thirty Years Wars, which, "militarily and economically [...] was a mere pigmy" when he ascended to the throne. Conversely, the *shortcomings* of individual monarchs have been linked to political failures, such as in the case of John I of England, whose personal incapability in military matters resulted in Britain losing most of its continental possessions.[13] Similarly, the German naval buildup aiming to contest British dominance at sea, and the break-up of the intricate system of alliances designed by Chancellor Bismarck are all linked to individual decisions of Emperor Wilhelm II of Germany. Röhl (1996) emphasizes Wilhelm's character's role and these decisions in paving the way to World War I.

A Tale of two Carloses

In the empirical analysis, we compare rulers of the same country. In what follows, we provide an illustrative example of such a comparison. Carlos II was king of Spain from 1665 to 1700. Hailing from a line of successive marriages of relatives from the Spanish and Austrian Habsburgs, he was highly incestuous and commonly described as an incapable ruler with little effective power. While his parents technically were 'merely' uncle and niece, the build-up of consanguinity over previous generations due to marriage among relatives resulted in Carlos' parents sharing as many genes as siblings would. As the pedigree in Figure 1 shows, all of Carlos II's grandparents descended from Joanna and Philip I of Castile (Alvarez, Ceballos, and Quinteiro, 2009). Repeated marriage between cousins and uncles and nieces ultimately led to most of the inbreeding in Carlos II being 'hidden' in the deeper layers of the pedigree, and not merely resulting from one (potentially strategic) kin marriage in the one generation prior. Carlos II's coefficient of inbreeding was 25.36, of which 12.5 was due to his parents being uncle and niece, with the remainder being a 'hidden' component due to accumulated inbreeding over previous generations. The degree of inbreeding was of no concern (not even the 'visable' uncle-niece dimension) when Carlos II's parents married

"In any age and time a man of Churchill's force and talents would have left his mark on events and society."

[12]In particular, Frederick William I. (the "Soldier King," who reigned 1713-1740) and his son, Frederick II (the "Great," 1740-1786), facilitated the rise of Prussia into the rank of a Great Power of Europe with their administrative reforms and military decisiveness. And even if – by his father's achievements – "Frederick the Great came into a rich inheritance, [...] the favorable circumstances do not in the least explain his great success." (Woods, 1913, p. 159). The often idiosyncratic decisions of earlier rulers also shaped Prussia, as for instance that of Elector John Sigismund to convert to Calvinism in 1613 (Clark, 2007, p. 115).

[13]"John was little, if at all, lagging behind Philip [his adversary] in wealth and resources. The explanation of the defeat [which lead to losing England's continental possessions] does not reside in economics. It rests between John's fault as a commander and his faults as a man" (Bradbury, 1999, p. 349).

in 1649.[14]

The "inbreeding depression" resulting from intermarriage over generations left Carlos II hostage to physical and mental fragility.[15] Carlos II only started talking at age 4, and walking at age 8. Alvarez et al. (2009) describe him as "physically disabled, mentally retarded and disfigured." As Carlos II became king of Spain when he was 4 years old, his mother Mariana became regent and initially influenced his policies. The resulting power struggles between factious rivals to influence Carlos II did not aid in solving the domestic and foreign challenges Spain faced (Mitchell, 2013).[16]

The power struggles that followed Carlos II's death brought a new dynasty to the Spanish throne – the Spanish Bourbons. The ranks of the Bourbon dynasty first led to two relatively undistinguished monarchs.[17] Thereafter, the highly capable Carlos III came to inherit the throne in 1759 through the rules of primogeniture from his half-brother, who had left no heirs. Carlos III's parents were cousins of third degree, and the accumulated 'hidden' component of inbreeding was also small, resulting in a degree of inbreeding of only 3.9, slightly more than half of that of first cousins (6.25). Spain flourished under Carlos III's reign, and contemporaries and historians hold him in high regards: He "was probably the most successful European ruler of his generation. He had provided firm, consistent, intelligent leadership [...and] had chosen capable ministers" (Payne, 1973, p. 371). Consequently, Carlos III's reign saw the "continued improvement in financial and commercial conditions, including agriculture and the useful arts" (Woods, 1913, p. 331).

2.2 Dynastic Rule and Primogeniture

The vast majority of European monarchs came to power according to fixed rules of accession. While these rules differed across countries and time, primogeniture became increasingly common. Primogeniture determines that the eldest living offspring of the current ruler becomes the country's next ruler. This practice was common on the Iberian peninsula early on, from where it spread to other countries quickly (to England in 1066, and France in 1222). It gradually replaced the two other common forms of successions – by siblings and other relatives of the current ruler, and

[14] As we discuss below, restrictions on cousin marriage were not enforced among the European nobility, and knowledge about the adverse effects of inbreeding only emerged in the early 20th century and was not widely accepted even in academic circles until the second half of the 20th century. In addition, the 'hidden' degree of inbreeding in Carlos II's pedigree was, if anything, interpreted as a positive feature, signaling a 'clean' royal bloodline (Van Den Berghe and Mesher, 1980; Scheidel, 1995).

[15] While population biology strongly suggests that inbreeding was responsible for Carlos II's mental fragility, such assertions cannot be proven for historical cases, because genetic samples are not available.

[16] "Diseased in mind and body from infancy, and constantly preoccupied with his health and eternal salvation, Charles II was incapable not only of governing personally but of either selecting his ministers or maintaining them in power. From the assumption of the regency by the Queen Mother, Mary Anne of Austria (...) to the death of Charles II not one of the many individuals who rose to power displayed genuine ability" (Hamilton, 1938).

[17] Philipp V (ruled from 1700 to 1745) and Ferdinand IV (1745-1759) "both were undistinguished rulers frequently incapacitated by near lunacy (Philip V dined at 5 a.m. and went to bed at 8 a.m., refusing to change his clothes)"(Carr, 1991). Philipp V's coefficient of inbreeding was 9.27, and that of his successor, Ferdindand VI, was 9.55 – both were thus more inbred than first-degree cousins (6.25), but significantly less than Carlos II. In both reigns, Spain's economic fortune improved moderately, starting off from the low levels left behind by Carlos II.

election of rulers by feudal elites.[18] In most cases, agnatic primogeniture was practiced, implying that the eldest living male offspring was heir apparent. In the absence of an heir (for instance due to premature death of the current ruler), the reign passed on to close relatives.[19]

Due to primogeniture, dynasties often stayed in power for centuries. For example, until the French Revolution, all kings of France were direct ancestors of Hugh Capet, who had ruled eight centuries earlier (from 987 to 996) and founded the "Capetian dynasty."[20] For more than half of the rulers in our dataset, there is unambiguous information for ascension to the throne by primogeniture. More than three quarters of successors in Austria, Prussia, and France ascended through explicit primogeniture, but less than half did in England, Denmark, Russia, and Sweden.[21]

2.3 Intermarriage Among Dynasties

Intermarriage among ruling dynasties was common, even *across* the countries of Europe. The leaders of the Spanish and Austrian Habsburgs, for instance, practiced cousin marriage over multiple generations in the 16th century, culminating in Carlos II, as described above. Alvarez et al. (2009) argue that the frequent dynastic marriages ultimately resulted in the extinction of the Spanish Habsburgs. While the Catholic Church had formal restrictions on cousin marriage, these were rarely enforced for European monarchs.[22] The pope could – and usually did – grant "dispensations" (exceptions) from the ban for Catholic rulers. As a result, intermarriage among royal dynasties actually *increased* throughout the early modern period (Benzell and Cooke, 2018), aided also by Protestantism lifting the ban entirely.

[18] Tullock (1987) describes theoretically that both current monarchs and elites favor primogeniture over other forms of succession, as it delivers political stability. Kokkonen and Sundell (2014) provide empirical evidence for this theory during our sample period. Often, kings crowned their sons while they were still alive to ensure a stable succession (Bartlett, 2020, p. 93).

[19] In general, the reign passed on to those individuals with the closest genealogical distance to the last male monarch. Whether this included female lines of succession as well as the exact definition of genealogical distance differed by ruling dynasty according to their "house law." In some cases, such laws of ascension where incomplete and left multiple potential claimants to the throne. As in the case of the heirless death of Carlos II, such cases often resulted in succession crises, sparked conflicts, and, later, amendments to succession laws.

[20] While the direct line of succession broke twice when kings died heirless, the title always passed to someone related to Hugh Capet. This happened first in 1328 (triggering a succession crisis that resulted in the Hundred Years War), when the Valois dynasty came to power, and again in 1589 with the rise of the Bourbon dynasty.

[21] The remaining cases have either no information, ambiguous information, or deviations from primogeniture – most prominently because of heirless rulers. In many such cases, monarchs ascended to the throne by agnatic succession (e.g., the monarch's younger brother becoming king), the decision by former rulers or by parliaments, or by usurpation of the throne.

[22] Restrictions on cousin marriage had been put in place starting from the 8th century – but *not* because of concerns about the physical or mental effects of inbreeding. Instead, these restrictions weakened the political power of closed kinship networks (Ausenda, 1999) and inhibited their further formation (Schulz, 2016); they also increased the likelihood that bequests would fall to the Church (Goody, 1983).

2.4 The Negative Effects of Inbreeding on Capability

A crucial feature of our identification strategy is that offspring of repeated dynastic marriages were less likely to become capable monarchs. Inbreeding reduces genetic diversity and evolutionary fitness (Robert, Toupance, Tremblay, and Heyer, 2009; Ceballos and Álvarez, 2013; Royuela-Rico, 2020). Thereby, it systematically increases the risk of genetic disorders, affecting physical and mental capability. Children of first cousins have a five times higher risk of intellectual disability (Morton, 1978), and their intelligence is reduced by 10% (Afzal et al., 1993). Inbreeding further results in lower height and weight (Fareed and Afzal, 2014b), and it decreases fertility while raising child mortality (Fareed, Ahmad, Anwar, and Afzal, 2017), thus lowering the probability of successfully producing heirs for the dynasty (Alvarez et al., 2009). Most important for our context, inbreeding depresses many individual physical and psychological traits that are associated with successful leadership.[23] European royal families did not defy the laws of biology. After the methodology for computing coefficients of inbreeding became available, Asdell (1948) showed that more inbred rulers had been assessed by Woods (1906) as systematically less capable – despite the fact that Woods had the opposite hypothesis (see footnote 4).[24]

In sum, our setting features rulers who ascended to power by pre-defined rules, independent of their inherent ability for office. In addition, the frequent intermarriage and negative effects of incest gave rise to variation in monarchs' ability that was unrelated to state performance.

3 Data

In this section we describe our dataset of ruler ability, country performance, inbreeding, and constraints on ruler power at the reign level.

3.1 Ruler Ability and State Performance

Our measure of ruler ability comes from the work of Frederick Adam Woods. A lecturer in biology at MIT at the beginning of the 19th century, Woods took an interest in heredity and, ultimately, history. To understand the heredity of moral and mental status across generations, Woods turned to the royal families of Europe.[25] In his 1906 publication on "Mental and Moral Heredity in Royalty"

[23] The literature on leadership traits has emphasized the importance of cognitive capabilities for leadership (c.f. Judge, Colbert, and Ilies, 2004). At the same time, there is a large literature documenting that inbreeding negatively affects these traits (Afzal et al., 1993; McQuillan, Eklund, Pirastu, Kuningas, McEvoy, Esko, Corre, Davies, Kaakinen, Lyytikäinen, et al., 2012; Fareed and Afzal, 2014b,a). More directly, Adams, Keloharju, and Knüpfer (2018) show that cognitive and non-cognitive ability, measured during military tests in Sweden, are strong positive predictors of individuals assuming leadership roles – becoming CEO's – later in life.

[24] A recent literature has argued that – in modern data – the size of the negative effects of marriages among first cousins may be confounded by poverty (Hamamy, Antonarakis, Cavalli-Sforza, Temtamy, Romeo, Ten Kate, Bennett, Shaw, Megarbane, van Duijn, et al., 2011; Bittles, 2012; Mobarak, Chaudhry, Brown, Zelenska, Khan, Chaudry, Wajid, Bittles, and Li, 2019). In contrast, our results do not depend on first-cousin marriage, as consanguinity due to complex intermarriage over generations (and beyond first cousins) drives our first stage.

[25] The appeal of this group of people to study heredity was manifold to Woods: The pedigrees of royal families were (and are) comparably well-documented over multiple generations. Further, for most of these individuals, their life,

(Woods, 1906), he "graded" more than 600 individual members of royal families based on their mental and moral qualities. This grading was based on adjectives used in written sources that describe these individuals. Based on these data, Woods concluded that mental and moral status was heritable.[26]

Subsequently, in his endeavor to test for the heredity of mental and moral status, Woods ventured beyond the realm of biology to the "great men" debate in history (Carlyle, 1840). Woods noticed a correlation between able rulers and favorable political and economic conditions in the country they ruled. In Woods' (1913) publication "The Influence of Monarchs," he extended his 1906 tabulation of the ability of rulers and also added a systematic coding of their states' performance for 13 states, ranging from their foundation until the French Revolution. This publication is a central data source for our empirical analysis. It contains the ability of rulers and state performance for more than 300 European reigns. Figure 2 shows the borders of the covered states in different time periods.[27]

Similar to Woods's earlier work, this grading is largely based on the assessment of historians and contemporaries, as distilled by Woods from reference works and country-specific histories. For each of the reigns, Woods provided a brief summary underlying his assessment and references. For instance, Woods' assessment of Carlos II is brief, characterizing him as an "imbecile" with negative virtues. In contrast, Carlos III is described as "enlightened, efficient, just, and sincere. Not brilliant, but had a very well-balanced mind." In terms of state performance, Spain under Carlos II is characterized by "misery, poverty, hunger, disorders, decline, especially in agriculture, finances and strength of the army" while a century later, Carlos III's reign saw "continued improvement in financial and commercial conditions, including agriculture and the useful arts." As an additional example, consider Maria Theresa, who reigned over Austria from 1740 to 1780, and was judged by Woods as "able and very industrious." Under her reign, "the various portions of the kingdom [were] unified and centralized" and "Austria gained slightly in territory and greatly in prestige," while "industry, commerce, and agriculture improved."

character, and achievement was documented from letters, court biographies, or other written sources.

[26]Woods was part of a (then active) research agenda in biology on heredity sparked by the publication of Darwin's "Origin of Species" in 1859 and Galton's "Hereditary Genius" in 1869. Social Darwinism, foremost that of Grant (1919), had an influence on the eugenics crusade in the United States and on the US Immigration legislation after World War I (Saini, 2019). After World War II, Social Darwinism was largely discredited, as was the concept of heritability of traits such as mental or moral qualities (at the level of societies). While heritability of intelligence at the individual level is sizable (Neisser, Boodoo, Bouchard Jr, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg, et al., 1996; Devlin, Daniels, and Roeder, 1997), differences between population groups are resulting from other environmental differences (Lewontin, 1970). Simonton (1983) used Woods' data to analyze the intergenerational transmission of individual differences.

[27]Note that many states started out small and come to dominate the map over time. Thus, territorial gains were positive on average over time (as opposed to a zero-sum game). The states covered are Castile, Aragon (Spain), Portugal, France, Austria, England, Scotland, Holland, Denmark, Sweden, Prussia, Russia, and Turkey. Appendix Figure A.1 provides a timeline of coverage for each state.

Woods assigned a "+" to rulers with high ability, a "-" to incapable ones, and "±" to those not clearly capable or incapable. In his coding of state performance, Woods covered the following dimensions: "finances, army, navy, commerce, agriculture, manufacture, public building, territorial changes, condition of law and order, general condition of the people as a whole, growth and decline of political liberty, and the diplomatic position of the nation, or its prestige when viewed internationally," while purposefully *excluding* "literary, educational, scientific, or artistic activities" (Woods, 1913, p. 10). Woods coded a three-valued variable summarizing the political and economic performance of the country during each reign, using again the three-tier scale "+, ±, -." We transform these into "1," "-1," and "0" and create the variables *Ruler Ability* and *State Performance*, respectively. Out of 336 reigns for which we have information on both the monarch's ability and the performance of the state, 127 rulers are rated as clearly incapable, 122 as clearly capable, and 87 as neither; regarding state performance, 112 reigns are rated as clearly bad, 143 as clearly good, and 98 are neither.[28]

The fact that both ruler ability and state performance were coded by the same historian gives rise to obvious endogeneity concerns. We address these in a multitude of ways, including extensive checks of Woods' coding, our IV strategy, as well as the use of alternative outcome variables. We also note that if Woods did have a bias in coding, then this should have worked against our IV strategy: Woods believed that more kin marriage among "successful" dynasties produced *better* rulers.

We discuss the quality and reliability of Woods' coding in Appendix A.2. For example, Thorndike (1936) had numerous research assistants "grade" the morality and intellect of more than 300 rulers.[29] This data quality assessment resulted in correlations of the intellectual grade across different graders (including Woods) ranging from 0.73 to 0.82. We similarly asked research assistants to assess the capability of individual rulers, as well as state performance, on a three-point scale based on articles in online encyclopedias (and without reference to Woods' coding). This exercise also largely confirmed Woods' data (see Figure A.2 in the appendix).

In principle, our extensive checks of Woods' data would us to run our empirical analysis based on our own coding. Nevertheless, we use Woods' original coding as our baseline, because Woods' hypothesis works against our IV strategy (as explained above), providing a conservative baseline.

[28]Woods collected information for 368 reigns in total. Especially for early and short reigns, Woods did not provide an assessment. In instances of co-reign, as for Ferdinand and Isabella of Castile from 1479 to 1504, we generally take the assessment of one individual if it is only available for one of two rulers. When both are available, we use the assessment of the individual working against our hypothesis. In cases where Woods expressed a doubt by, say, "+ or ±," we use the average (in this example, 0.5). In a robustness check, we recode all these cases conservatively so as to work against our baseline findings.

[29]Thorndike's student, Dr. Edith E. Osburn "read what was printed about each of about four hundred of the persons studied by Woods, in each of the six biographical dictionaries used by him. This occupied her about forty hours a week for about eight weeks. She then read through the entire set of references again" (Thorndike, 1936, p. 322). At the same time, Thorndike had five more research assistants independently do the same coding.

In Appendix B.1 we report results based on our own coding of Woods' sample coverage. In addition, we also provide robustness checks with an extended sample – both in terms of time period (until World War I) and countries covered (adding Hungary and Poland). We coded this extended sample using Woods' original sources, as well as modern encyclopedias. The two countries and additional century of data add a total 95 reigns to our baseline sample (see Appendix A.4 for detail).

3.2 State Border Changes and Urbanization

Because our outcome variable 'state performance' is ultimately a subjective measure, we collect two additional outcome variables. First, we calculate changes in the size of a state's territory during the reign of each monarch. Abramson (2017) provides borders and the area of the independent polities of Europe at five-year intervals from 1100 to 1795. We link these to the beginning and end of each reign and calculate the percentage change in area ruled during a reign, $\Delta log(Area)$.[30] For example, Austria during the reign of Maria Theresa (1740 to 1780), Silesia was lost to Prussia, while Austria gained areas from Poland (see Appendix Figure A.3). In net terms, Austria increased its area by 7%.

Territorial expansions do not necessarily and unambiguously imply better state performance. In fact, "overexpansion" might weaken the power of a state (Kennedy, 1988). However, Baten, Keywood, and Wamser (2021) show that expansions of territory go hand-in-hand with increases in taxes per capita, and thus use territorial expansions as a proxy for state capacity.[31] In order to address potential shortcomings of this measure, we also code changes in urban population within the territory ruled by each monarch. We impute the total urban population within the boundaries of each state by combining the borders provided by Abramson (2017) with city population data from Bairoch et al. (1988).[32] For each reign, we calculate the total urban population within the state borders at the beginning and at the end of each reign. We then calculate the percentage change in total urban population $\Delta log(UrbPop)$. For an additional check, we also decompose this measure into an intensive (changes in population in existing cities) and extensive margin (conquering/losing cities).

[30] We link the end year of each reign to the subsequent five-year observation, and start dates to the preceding five-year observation. For reigns shorter than five years, we check historical sources to confirm that the implied territorial changes indeed occurred during the respective reigns.

[31] We mainly use territorial changes as an objective (i.e., not coded by historians) and high-frequency measure of (an admittedly narrow aspect of) state performance, while our main measure of state performance (as assessed by Woods) is very broad. In fact, we find that our results are driven by many economic and political aspects of state performance (see Appendix D.2) and do not depend on the consideration of territorial expansions in assessing state performance (see Appendix C.3). Most other available measures of state performance, such as GDP per capita, are only available for few states (with changing borders) and are typically measured at the century or half-century level.

[32] We use linear interpolation to obtain city population in 5-year intervals (corresponding to the border data frequency) from Bairoch et al.'s century- and half-century data, assuming a linear growth rate. We geocode the location of each city to determine the polity that it belonged to at each 5-year interval.

3.3 Coefficient of Inbreeding for European Monarchs

The first correct measure of the degree of similarity in the genes of offspring due to common ancestors was developed by Wright (1921). This "coefficient of inbreeding" is the probability that both gene copies at any locus in an individual are identical by descent, i.e., from a common ancestor. Higher F thus means lower diversity in an individuals gene pool. Diversity has a positive effect because humans are diploid, i.e., they have two chromosomes copies each (one from each parent), and for recessive disorders to appear, both need to be deleterious. Hence, the lower diversity, the higher the risk of recessive gene disorders.[33] Offspring of siblings or of parent-offspring couples have a coefficient of inbreeding of $F = 25$, while offspring of uncle-niece couples would have $F = 12.5$, and first cousins couples would have $F = 6.25$.[34]

We collect F for 265 monarchs from http://roglo.eu/, a crowd-sourced online data source of the genealogy of European noble families. For 243 of these monarchs, Woods assessed both state performance and ruler ability. We start by identifying each monarchs' parents. For these, in turn, http://roglo.eu/ calculates the coefficient of inbreeding for their offspring, relying on rich data on relationships between their ancestors.[35] Figure 4 shows a histogram of the coefficient of inbreeding for all monarchs in our dataset. The figures also provides two illustrative examples. Carlos II is the individual with the highest coefficient of inbreeding. With $F = 25.36$, he is more inbred than offspring of siblings would be. Yet, his parents were "merely" uncle and niece (which in itself would imply $F = 12.5$). This points to an important feature of our setting: A sizable amount of the observed inbreeding is not the result of just one generation of consanguineous mating, but rather driven by a "build up" of inbreeding over previous generations.[36] We will use this 'hidden' component of inbreeding explicitly in robustness analyses later.

3.4 Constraints on Ruler Power

We collect data on the legal and de facto constraints on the power of monarchs from a variety of sources. Our baseline variable refines and extends the measure "constraints on the executive"

[33] This "dominance hypothesis" is the prevailing explanation for "inbreeding depression"in genetics (c.f. Charlesworth and Willis, 2009).

[34] The coefficient of inbreeding ranges from 0 to 100 (%). Humans inherit one allele at each locus from each parent. Because humans carry two alleles at each locus, the probability to pass on a particular allele to a particular offspring is 0.5. Hence, the offspring of self-fertilization would have $F = 50$, as there is a one-half chance for each locus that the entire pair of alleles was passed on. Hypothetically, with repeated self-fertilization, F would approach 100. Offspring of completely unrelated parents have $F = 0$. We provide more detail on the calculation in Appendix A.5.

[35] We cross-checked and validated the coefficients we obtained from http://roglo.eu/ extensively with other publications, among them Asdell (1948) and Alvarez et al. (2009). Turkey is not covered by this source and is thus not included in our IV results. For 43 rulers, no known relationship link was recorded. This could either imply that they were unrelated, or simply that the information on distant family relationships did not survive. We thus exclude these cases from our baseline, but we show robustness to their inclusion in Appendix Table A.14.

[36] Consider again the pedigree of Carlos II (Figure 1). While Philipp IV, the father of Carlos II, married his niece, past consanguineous marriage weighted heavy in opening up many pathways for the common ancestors Joanna, "The Mad" and her husband Philip generations earlier.

from Acemoglu et al. (2005), which is available between 1000 CE and 1850 (first at the century level and after 1700 CE in fifty-year intervals). Acemoglu et al.'s measure was coded following the approach of the Polity IV project (Marshall, Jaggers, and Gurr, 2017) at the level of today's countries. Using the same coding approach, we refine the coding of "constraints on the executive" on a *year-by-year* basis at the *historical state* level, guided by the Polity IV rating and the same primary sources used by Acemoglu et al. (2005). Appendix E explains our methodology in detail.

Figure 6 illustrates our annual measure, using England during its turbulent seventeenth century. The black solid line shows the institutional score by Acemoglu et al. (2005), which is constant at 3, indicating "slight to moderate limitation on executive authority" from 1600 to 1700 CE. Our measure (the dashed green line) is much more finely grained, reflecting the variability of constraints on the monarch during that century. Consider 1629, when the English parliament was dissolved and "Charles [I] governed without a parliament, raising money by hand-to-mouth expedients, reviving old taxes and old feudal privileges of the crown and selling mentarians contrary to the spirit of the constitution" (Stearns and Langer, 2001). This is reflected by a sharp drop of our measure from "substantial limitations on the monarch's authority" (a score of 5) to "no regular limitations on the executive's actions" (score of 1). Constraints became stronger again during the "Long Parliament" from 1640-1660, as a consequence of the "Triennial Act [of 1641], requiring the summoning of parliament every three years without an initiative of the crown. [This was] followed by [... a] bill to prevent the dissolution or proroguing of the present parliament without its own consent" (Stearns and Langer, 2001).

Based on our year-reign specific measure for constraints on the executive, we define the variable *Constrained* if the constraints on the ruler in the year prior to the beginning of the reign were above a score of 5, indicating "substantial limitations on executive authority." This cutoff is further defined as follows: "The executive has more effective authority than any accountability group but is subject to substantial constraints by them." In our sample, this applies to 28 monarchs, of which 10 were rulers of England. Appendix E provides further detail and a list of all seven cutoffs.

As an alternative measure for constraints on the executive, we use parliamentary activity from Van Zanden et al. (2012), who compile the frequency of parliamentary meetings across European countries from the 12th to the 18th century (at the century level). We use the measure of parliamentary activity in the century before the start of a reign and code rulers as *Constrained* if parliamentary activity was above the 95th percentile of the entire sample.[37] Unfortunately, this measure is not available at the sub-century (let alone annual) level. We thus use it as a (rough) consistency check of our findings.

[37] Van Zanden et al. (2012) collect the information on the relative frequency of meetings of parliaments from a variety of sources. For all countries except Turkey, we can link this to our data set. We link Prussia to the "Brandenburg Diet" and the "Generallandtag" of Austria to the Habsburgs. The data are separately available for Scotland and England, for Castile (and Leon) – which we match to Castile, and for Aragon. All other matches are straightforward.

4 Main Empirical Results

In this section we first document a strong association between the capability of European monarchs and the performance of their countries. We show that this association is robust to measurement, specification and in different samples. We then discuss our identification strategy and present our IV results.

4.1 Baseline OLS Results

Our baseline regressions are at the state-reign level:

$$y_{r,s} = \beta RulerAbility_{r,s} + \delta_s + \varepsilon_{r,s} \ , \tag{1}$$

where $y_{r,s}$ is one of the three the outcome variables for state s in reign r, as defined in Sections 3.1 and 3.2: *State Performance*$_{r,s}$, $\Delta log(Area)_{r,s}$, or $\Delta log(UrbPop)_{r,s}$. *Ruler Ability*$_{r,s}$ is the assessment of the monarch's ability. For a straightforward interpretation of coefficients, we standardize the assessments of *State Performance* and of *Ruler Ability* so that both variables have mean zero and standard deviation one.[38] We include state fixed effects δ_s, so that we effectively compare rulers of the same state over time.[39] Throughout, we report standard errors clustered at the state level.

Table 1 shows that *Ruler Ability* is strongly associated with *State Performance*. Column 1 reports the raw correlation. The coefficient of interest, β, is highly significant and sizable: A one standard deviation increase in *Ruler Ability* is associated with a 0.62 standard deviation (std) increase in *State Performance*.[40] Column 2 shows that this association is unchanged when we add state fixed effects, thus comparing only monarchs who ruled the same state. The outcome variable *State Performance* is subject to concerns about biased coding by Woods. We address this by using 'objective' (and also continuous) outcome variables in the next columns. For the reign-specific percentage change in state area, we document a significant and sizable association with ruler ability (column 3). Again, these results are stable when we include state fixed effects (column 4). A one std increase in ruler ability in the same state and century is associated with land area expanding by about 10%. Finally, columns 5 and 6 use the change in urban population during a

[38]Note that while both are categorical variables, we treat them as continuous variables for ease of estimation throughout the paper. We provide a robustness check using ordered Probit below.

[39]Accounting for time trends is not straightforward in our main regressions because there is no clear-cut time variable: Reigns begin and end at different times in different states, and they also often span across centuries. Below we present a flexible method to filter out time effects: regressions at the ruler pair level, comparing monarchs in different states who ruled contemporaneously.

[40]Woods (1913) himself had also manually computed the (not standardized) correlation coefficient of 0.6 in his raw data. He asserted a causal direction from monarch ability to country performance: "Only very rarely has a nation progressed in its political and economic aspects, save under the leadership of a strong sovereign." While Woods was well aware of reverse causality concerns, he provided descriptive evidence in favor of this conclusion. We go beyond Woods' findings by exploring richer specifications and, in particular, by providing an identification strategy.

reign as outcome variable. We document a sizable association: A one std increase in ruler ability in the same state is associated with total urban population in the state expanding by about 9%.

4.2 Robustness of OLS Results

Next, we examine the robustness of our baseline OLS results. Beginning, with the baseline sample in column 1, Table 2 successively reduces the sample until column 6. In column 2, we focus on reigns in which the ruler was linked to a dynasty. Thereby, we exclude cases of interregna, regencies in which non-royal individuals exerted power, and instances of non-monarchical governance (as in the Netherlands).[41] The coefficient increases slightly and remains highly significant. Column 3 excludes all regencies, independent of whether the regent was a dynasty member or not. The coefficient again increases slightly. Note that the variation explained (R^2) actually increases in columns 2 and 3, indicating that indeed monarchs hailing from dynasties are crucial to the relationship between ruler ability and state performance. Column 4 excludes the few instances of foreign rule.[42] Column 5 excludes all individuals who appeared as rulers in more than a single reign. These are either monarchs that repeatedly came to power in the same country, or who ruled in more than one country contemporaneously. In both columns 4 and 5, the coefficient remains significant and comparable in size to the baseline. Finally, column 6 applies all restrictions of the preceding columns simultaneously. With only about 75% of the initial sample left, the coefficient remains almost unchanged, and the variation in country performance explained by the regression is actually higher than in the baseline.

Columns 7 and 8 present extensions of our baseline sample. Woods' data stops around 1790, before Napoelon, and he excluded states in Eastern Europe – most prominently, Poland(-Lithuania) and Hungary. In column 7, we extend the coding of states covered by Woods (1913) until World War I based on internet encyclopedias, which in turn draw on historical sources (see Appendix A.4 for detail and Appendix B.4 for additional results). In column 8 we also add data for Poland(-Lithuania) and Hungary from their foundation until 1914. Both extensions yield results that are very similar to those in the baseline sample. In Table A.3 in the appendix we show that our results also hold when we use only our own coding of state performance and ruler ability for the core sample of Woods.

[41] Interregna are periods between the rule of two monarchs when no monarch is present. Regencies are periods of government by others (regents) in lieu of the designated ruler. Usually, these are close relatives such as the mother of an underage monarch, but sometimes these can be officials or members of the elite. In column 2, we exclude all rulers during whose reign regents from outside their dynasty governed. We still include cases of rule by relatives of the designated heir until the heir assumed office. For example, Mariana was regent for Carlos II of Spain until he reached adulthood, and then tried to regain regency by arguing that he was unfit for office.

[42] Foreign rule refers to instances when monarchs of one country temporarily ruled over another country. For instance, Philipp II of Spain ruled Spain and Portugal from 1580 to 1598, and James VI of Scotland also reigned over England from 1603 to 1625. When excluding episodes of foreign rule, we drop the corresponding observations for Philipp in Portugal, but keep his reign in Spain. When excluding monarchs who governed in more than one country (column 5), we drop both observations, his reign in Spain and that in Portugal.

We document further robustness checks in Appendix B. Table A.4 shows robustness to measurement. We find that our results based on Woods' (1913) original coding are highly robust when we exclude cases that Woods coded with intermediate values for state performance or ruler ability, indicating that he felt a clear judgment was not warranted by the underlying information. In fact, our results are even robust when we recode all those middling values to work *against* a positive association between ruler ability and state performance (Table A.4, col 5). Finally, Table A.5 shows robustness to different specifications, such as using dummies for different values of state and ruler performance, ordered probit, as well as clustering at the state, century, and dynasty levels.

4.3 Heterogeneity

How does the association between state performance and ruler capability vary across time, space, and personal characteristics of rulers? In Table 3, we include interaction terms between ruler ability and numerous characteristics. We collect the variables used in this section from encyclopedias and biographies, as explained in Appendix A.1. For column 1, we define a dummy indicating whether a monarch was female, which was the case for 40 of the 338 reigns to which a gender was assignable.[43] The small and statistically insignificant coefficient suggests that there the relationship between ruler ability and state performance did not vary by the ruler's gender. In column 2, we interact *Ruler Ability* with a dummy indicating whether a monarch ascended to the throne before the median age of ascension (28 years). While the interaction term is quantitatively somewhat larger, it remains small compared to the coefficient on *Ruler Ability*, and it is also statistically insignificant. Column 3 uses a dummy indicating whether a ruler was raised as designated heir. This means that a ruler was raised as monarch, rather than ascending to the throne because another designated heir unexpectedly died earlier, or ascending to the throne by other means.[44] The interaction term is positive but minuscule and again statistically insignificant. Finally, in column 4, we interact with a dummy indicating that the prior ruler was executed after trial or murdered (Kokkonen and Sundell, 2014). We find no difference in our baseline association for those reigns. This speaks against the possibility that our result is primarily driven by able monarchs deposing of their incapable predecessor, as for instance Catherine the Great, who ascended to power through the murder of her husband. We also note that of the dummies in Table 3 only one is itself statistically significant – in column 1, indicating that female rulers were associated with somewhat lower state performance. One possible explanation is that states led by queens were more frequently involved in warfare (Dube and Harish, 2020). In contrast, age at ascension, being raised as a designated hair, or regicide of the previous ruler are by themselves not associated with state performance.

Did the relationship between monarchs' ability and state performance change over time? The

[43] For 28 reigns we cannot assign a gender. As explained in Appendix A.1, these other instances are interregna or reigns by councils.

[44] Note that we were only able to assess whether monarchs were raised for particular roles for 155 observations, of which 121 where raised as monarchs.

left panel in Figure 3 depicts the coefficient on *Ruler Ability* for different time periods, showing a statistically highly significant correlation throughout.[45] After 1600, the coefficient size decreases. This period also coincides with the rise of parliaments in Western Europe (Van Zanden et al., 2012). Below, we examine whether this trend may have affected the role of ruler ability in their states' performance. The right panel of Figure 3 shows the correlation between ruler ability and state performance for all states in our sample.[46] The coefficients are relatively similar across states, and they are statistically highly significant for all states except Denmark.[47] The coefficient is strongest for Prussia, implying that this state fared particularly well under good rulers and/or suffered particularly strongly under bad ones. Prussia's institutional setting featured few if any constraints on the monarchs' executive power. The other extreme is England, where the association between ruler ability and country performance is less pronounced. This is particularly true after 1600, when the English Parliament gained power vis-à-vis the Crown. For this period, we observe no more relationship between ruler ability and the country's performance.[48]

4.4 IV Results

In what follows we provide evidence for a causal relationship between ruler ability and country performance. We first discuss our identification strategy based on primogeniture and inbreeding. Then we introduce our instrument – the coefficient of inbreeding – and document that it is a strong predictor of ruler ability. The corresponding IV results reveal a positive causal effect of ruler ability on country performance in the second stage.

Identification

An causal interpretation of our OLS estimates is subject to numerous concerns. Omitted variables could influence both the performance of a country and the ability of the ruler in power, and reverse causality is also a possibility – for example, better country performance driving the selection of more capable rulers. In addition, historians may have assessed rulers of better-performing states more favorably (or vice-versa).

Our identification strategy (in combination with our 'objective' outcome variables) enables us to address these concerns. We rely on the combination of two features. First, primogeniture resulted in pre-determined ruler succession, independent of ability. Second, we leverage the varia-

[45] As before, reigns are allocated to time periods according to the start year of each reign. Table A.8 in the appendix provides point estimates for these broad periods and also in a more disaggregate fashion, by century.

[46] Table A.7 provides the corresponding regression estimates.

[47] A possible explanation is that Danish crown had not fully transitioned to a hereditary monarchy. Danish kings were de jure elected by the nobility. However, de facto the oldest son of a ruler was usually elected as his successor (Bartlett, 2020, p. 398). Therefore, Danish monarchs may have been impeded by relatively strong constraints on their executive power.

[48] Woods (1913, p. 245) also noted that the positive association between state performance and ruler ability disappeared for England after 1600. In Section 5 we provide systematic evidence that this is linked to the English monarchs becoming constrained by ever stronger parliaments.

tion in ruler capability due to the wide-spread inbreeding within and between European dynasties. Centuries of intermarriage resulted in a sizable degree of genetic closeness between the potential marriage partners of Europe's monarchs. The exclusion restriction is that inbreeding was not related to state-level outcomes via channels other than ruler ability. There are two potential paths for violating this condition: First, potential historical factors that linked inbreeding to state performance. This is unlikely – at least in terms of concerns about inbreeding being related to state outcomes. As we discussed above, the negative effects of inbreeding were unknown to the royal families. If anything, they believed that marrying within dynasties strengthened their noble traits, which would tend to work against our results. However, inbreeding may have been the result of strategic marriage for other reasons. To address this concern, we show that our results also hold when we focus only on the 'hidden' degree of inbreeding – the dimension beyond the parents' relatedness that was embedded in the intertwined family trees of previous generations and thus not observable (and impossible to compute prior to the 1920s).

Second, the exclusion restriction may be violated because of a bias in the coding of historical data. Again, this is unlikely because of Woods' (1906; 1913) hypothesis that intermarriage among what he considered the superior stock of royal families led to *more capable* rulers.[49] Further, the correct measurement of inbreeding was unknown when Woods was writing in 1913, and the fact that inbreeding has negative consequences in humans was only accepted in academic circles decades later.[50] Therefore, the timing of scientific progress on inbreeding renders a violation of the exclusion restriction in Woods' (or other underlying historians') assessment of monarchs unlikely.

First Stage

Our first stage shows that monarchs with a higher coefficient of inbreeding are significantly less capable rulers. This is in line with the fact that genetic closeness between partners carries an increased risk of genetic disorders for their offspring. and that these disorders, in turn, increased the probability that rulers were incapable and could not effectively fulfill the duties of their offices. Formally, our first stage is:

$$RulerAbility_{r,s} = \gamma F_{r,s} + \delta_s + \varepsilon_{r,s} \ , \tag{2}$$

[49] "The very formation of royal families was thus a question of selection of the most of able in government and war. From their intermarriage with their own kind, in connection with the force of heredity, we find an explanation in their relative superiority" (Woods, 1906, p. 302). See Section 3.1 for further detail.

[50] Darwin was the first to show experimentally that inbreeding depression exists in plants, and then worried that his own offspring might be affected (his wife was his first cousin, cf. Berra, Alvarez, and Ceballos, 2010). It took decades for researchers to become convinced that humans are similarly negatively affected by inbreeding. In 1927, Bronislaw Malinowski, one of the "founding father[s] of social anthropology" (Young, 2004), stated that "biologists are in agreement that there is no detrimental effect produced upon the species by incestuous unions" (Malinowski, 1927). See also Wolf (2005).

where $F_{r,s}$ is the coefficient of inbreeding of the ruler of state s in reign r (as described in Section 3.3), *Ruler Ability*$_{r,s}$ is the capability of said ruler, and δ_s are state fixed effects. Again, we cluster standard errors at the state level.

Column 1 in Table 4 documents a negative and statistically highly significant raw relationship between a ruler's coefficient of inbreeding and her or his capability. We obtain a similar result in column 2, where we add country fixed effects (our preferred specification). The effect is sizable: Increasing the coefficient of inbreeding by one standard deviation decreases ruler ability by 0.3 standard deviations (standardized beta, unreported). Figure 5 shows a binned scatter plot of the variation underlying column 2, illustrating that the first-stage relationship is not driven by outliers. The next two columns in Table 4 exclude individuals with high coefficients of inbreeding. If anything, the first-stage coefficient increases when we exclude Carlos II, whose parents were as related as siblings (column 3), or when excluding all individuals whose parents were at least as related as uncle-niece pairs, corresponding to $F \geq 12.5$ (column 4). In column 5, we focus on cases of documented primogeniture, i.e., those cases for which we could explicitly confirm from historical sources that the ruler ascended to power according to the laws of primogeniture. The coefficient remains similar in terms of both magnitude and statistical significance. This is reassuring for our use of the IV strategy in the full sample, which includes cases where primogeniture is not historically documented (although it was the norm), or cases of heirless rulers where other close relatives were appointed as the next king or queen (see footnote 21 for detail).

Finally, note that the monotonicity assumption required for IV is likely fulfilled. In our setting, this assumption requires that the instrument does not trigger "defiers," i.e., that inbreeding does not (by accident) lead to ingenious leaders. The literature in genetics documents that "inbreeding depression" only has negative effects on fitness (c.f. Robert et al., 2009; Ceballos and Álvarez, 2013), and therefore in all likelihood on leader ability. It is also not the case that inbreeding increases variance in ability; that is, it is essentially impossible that inbreeding leads to "genius" by accident.

Second Stage Results

Table 5 presents our second stage results in Panel A. In column 1 we use Woods' assessed state performance as the outcome. Note that the instrument is strongly relevant (effective F-statistic of 42).[51] The IV coefficient is positive and strongly significant, suggesting that the ability of monarchs had a positive causal effect on the performance of the countries they reigned. In column 2 we exclude monarchs who were at least as inbred as the offspring of uncles and nieces. We obtain a very similar 2SLS coefficient on ruler ability.

[51] We follow the recommendation by Andrews, Stock, and Sun (2019) and report the effective F-statistic by Montiel Olea and Pflueger (2013), which can be compared to the Stock and Yogo (2005) critical values in our case with one endogenous regressor and one instrumental variable (Andrews et al., 2019). The corresponding critical value for max. 10% relative bias is approximately 16.4 for all three 2SLS specifications.

Our IV strategy addresses reverse causality and some, but not all, omitted variable biases. For example, our IV would not address the possible issue that Woods' (1906) initial assessment of ruler ability may have influenced his subsequent coding of state performance in Woods (1913). To speak to this concern, columns 3 to 6 in Table 5 turn to our second and third outcome variables – the changes in land area and in urban population during the tenure of each monarch. Again, both OLS and IV coefficients point to a positive, large, and significant effect of ruler ability, and the results are robust to excluding rulers with relatively high coefficients of inbreeding above 12.5.

The IV estimates tend to be somewhat larger than the corresponding OLS coefficients. For instance, the OLS estimate corresponding to Table 5 column 1 is 0.618 (Table 1 column 2). Thus, the IV coefficients of 0.794 is 28% larger. A plausible explanation is that – as discussed above – Woods had a bias in favor of rulers hailing from old dynasties, which may have led him to assign better grades to inbred rulers, and correspondingly, worse grades to less inbred rulers. Our IV strategy corrects for this biased assessment of ruler ability and uncovers larger effects.

Panel B shows the corresponding reduced-form relationships between each monarch's coefficient of inbreeding and the three state-level outcomes during the reign. We find sizeable and statistically highly significant coefficients. A one-std increase in inbreeding leads to a 4.7% increase in land area and to a 4.6% increase in urban population, according to the estimates in columns 3 and 5, respectively. An alternative interpretation of the magnitudes is that on average, monarchs with $F < 6.25$ (less inbred than the offspring of first cousins) saw a 6% larger increase in their territory then more inbred monarchs with $F > 6.25$.

'Hidden' Inbreeding

In what follows, we present IV results for the 'hidden' degree of inbreeding. These specifications can address endogeneity concerns that link inbreeding to state performance – either via strategic royal marriage decisions (e.g., to expand the territory) or via preferences for or against marrying within dynasties that are in turn related to state performance. We first identify the degree of inbreeding that resulted directly from each ruler's parents' family ties (e.g., parents being first cousins or uncle and niece). Then, we deduct this 'naive' degree of inbreeding from the 'full' coefficient of inbreeding. The resulting 'hidden' degree of inbreeding reflects the more remote layers of the pedigree, beyond the parent generation (see Appendix A.6 for further detail on the calculation). Table 6 replicates our previous IV and reduced-form results, using only the 'hidden' component of the coefficient of inbreeding as an instrument for ruler ability. Throughout, we obtain very similar results, both in terms of magnitude and statistical significance. Importantly, the first stage also remains strong, with the effective F-statistic either exceeding or being close to the critical value for max. 10% IV bias.

Other Potential Threats to Identification

In Appendix C we present results that can address further potential threats to our identification strategy. Here, we provide a brief overview. First, the exclusion restriction would be violated if royals married kin when state performance was low, *and* past bad state performance lowered state-level outcomes during the reign of their offspring. Our results on 'hidden inbreeding' can already alleviate this concern – by effectively excluding marriage decisions at the generation of rulers' parents. Nevertheless, we also account for this possibility more directly in Appendix C.1. We show that controlling for state performance over the previous two reigns, or for lags in the coefficient of inbreeding, does not affect our 2SLS results (Table A.13). In addition, neither lagged state performance nor lagged inbreeding predicts current state performance or current ruler ability (Tables A.11 and A.12). Second, we show that our results are similarly not driven by strategic marriages *outside* of the kin network. Marriage between completely unrelated parents would imply rulers with zero inbreeding ($F = 0$). These are excluded from our baseline dataset (see footnote 35). Table A.14 in Appendix C.2 shows that our IV results are almost identical when we include the 43 ruler with $F = 0$. Third, in Appendix C.3 we account for a possible confounding role of conflict, which may have been related to dynastic networks (Benzell and Cooke, 2018). We show that our results are robust to controlling for conflict during reigns (Table A.15), and to residualizing our *State Performance* measure with respect to territorial changes (Table A.16). Fourth, in Appendix C.4 we account for a possible role of founders vs. descendants in dynasties (George and Ponattu, 2018), whereby founders of dynasties may be at the same time more capable and less inbred than later descendants. Table A.17 documents that our IV results hold when we include fixed effects for rulers' order within dynasties. Fifth, a related concern is that monarchs may have selected the most able leaders among their offspring as successor (even to the point of 'ridding themselves' of incapable offspring that came earlier in the birth order), or that offspring who were more affected by their parents' consanguineous relationships died in young age, leaving more capable surviving successors. However, both these mechanisms would work against our first stage: Siblings share the same coefficient of inbreeding, and 'eliminating' the least capable ones would reduce the variation in ruler ability that is due to inbreeding.

4.5 Ruler-Pair Regressions

So far, our regressions have compared rulers from the same country over time. As we noted in footnote 39, accounting for variation over time is not straightforward because reigns begin and end at different points in different states. Yet, rulers and their states' performance might be affected by continent-wide shocks, such as the Black Death, the Reformation, or long-lasting wars.

In order to account for potential confounding factors over time, we introduce a flexible approach that compares leaders in different countries who ruled contemporaneously. For instance, while Carlos III of Spain (assessed as a capable ruler by Woods (1913)) oversaw the "continued improvement" of many aspects of the performance of Spain from 1759 to 1788, Louis XV ruled over France from 1731 to 1774. Described by Woods (1913) as "weak, indolent" and of "inferior capacity," Louis XV oversaw the (for France) "disastrous Seven Years War" and domestically a "decline in commerce [...] Under excessive taxes, the peasantry were reduced to extreme misery."

We identify – for each ruler i – all those rulers j who overlapped in their reign in different countries for at least one year. Then, we calculate pairwise differences in their ability, in the performance of their countries, and in their coefficients of inbreeding. Based on these variables, we estimate regressions at the ruler pair-level:

$$\Delta_{ij}\text{State Performance} = \beta\Delta_{ij}\text{ Ruler Ability} + \mu_{c(i)} + \mu_{c(j)} + \gamma X + \varepsilon_{ij} \quad , \tag{3}$$

where Δ_{ij} indicates the difference in a variable between ruler i and j. For ease of interpretation of coefficients we again standardize the differences in the assessments of *State Performance* and of *Ruler Ability* so that both variables (in differences) have mean zero and standard deviation one. We further estimate IV regressions in this setting using the difference in the coefficient of inbreeding. For the above example of Carlos III and Louis XV, this difference is negative (-5.65) from the perspective of Carlos III, as he had a lower coefficient of inbreeding (3.9) compared to Louis (9.55). In all regressions we further include country fixed effects for both rulers ($\mu_{c(i)}, \mu_{c(j)}$), and we introduce the following additional fixed effects successively: country-pair fixed effects, ruler fixed effects (of ruler i), and country-pair times century fixed effects. Throughout for the ruler-pair regressions, standard errors are clustered at the country-pair level. In total there are 5,512 pairs of overlapping rulers in our sample with data on ability and state performance for both rulers. For 3,070 of these we also know the coefficient of inbreeding for both rulers.

We present the results for our main outcome variable (*State Performance*) in Table 7; the corresponding regressions for territorial changes and urban population are shown in Appendix Table A.10. Panel A presents the OLS results. Differences in the ability of contemporaneous rulers are strongly positively associated with differences in *State Performance* (column 1). The coefficient is similar to our result in levels (see Table 1). While this is unsurprising, given that we standardized both the differences and the levels in ruler ability and *State Performance*, the almost identical coefficient also implies that our results are not driven by time trends. Next, in column 2 we introduce country-pair fixed effects, absorbing features that are specific to country pairs (such as the frequent wars between England and France). In column 3, we further use ruler fixed effects, which prevents that our results may be dominated by individual rulers. Column 4 uses country-pair $times$ century fixed effects, thus absorbing for example differences between England and France that were specific to the 17th century. In all three increasingly restrictive specifications we obtain very similar results. Finally, our results also holds when we compare each ruler only with the *one* ruler with whom (s)he shared the largest overlap in reign (column 5).

Next, we turn to our IV results in the ruler-pair regressions. Panel B in Table 7 presents the first stage, showing strong negative and highly significant coefficients for all specifications. Comparing contemporaneous rulers, those who were more inbred had lower ability. Building on this strong first stage, Panel C shows our IV results at the ruler-pair level. We document sizable effects of pair-wise differences in ruler ability on differences in *State Performance*. The coefficients are similar to our baseline IV specification in Table 5 (column 1), suggesting that aggregate time trends do not confound our identification and causal estimates. Finally, Panel D presents reduced-form results, documenting highly significant and negative relationship between differences in inbreeding across contemporaneous rulers and differences in the corresponding *State Performance*. Appendix Table A.10 shows that the findings for ruler-pair regressions also extend to our other measures of state performance, namely the change in territory and in total urban population during each reign.

4.6 Mechanisms

Why did capable monarchs boost their states' performance? In Appendix D we provide several pieces of suggestive evidence that we summarize here. First, we examine whether physical (as opposed to intellectual) ability could explain our findings, as both are affected by inbreeding. In particular, heavily inbred rulers often had a short life span and lacked reproductive success. Appendix D.1 (Table A.18) shows that our IV results are unchanged when we control for the age at death and the number of offspring.[52] Second, we examine whether which aspects of Woods' (1913) broad *State Performance* measure drive our results. To this end, we code detailed outcome variables for various economic and political aspects of each reign based on both Woods' text and information from encyclopedias (see Appendix D.2 for detail). We find that ruler ability had particularly strong effects on law and order, administrative efficiency, and diplomatic prestige of a state. Capable rulers also fostered economic performance (both agriculture and commerce) as well as the living conditions of their populace.

Third, in Appendix D.3 we examine how ruler ability affected war and conflict. We find that overall, states with capable rulers were significantly less likely to experience conflict. Distinguishing between domestic and international conflicts as outcome variables, we then show that this finding is entirely driven by the latter: Capable rulers were much less likely to engage in external conflicts (Table A.21). This suggests an interesting mechanism, given that capable rulers also expanded their states' territory and urban population (see Table 5): Capable rulers avoided conflicts

[52] In this context, we note that a rich literature connects birth order to individual and social capabilities, typically finding favorable effects for first-born children (c.f. Rohrer, Egloff, and Schmukle, 2015). Motivated by these differences, Oskarsson, Dawes, Lindgren, and Öhrvall (2021) show that firstborn sons are more likely to become politicians today. This could potentially also be a mechanism behind our results: if firstborn children were more capable, and if inbreeding led to more infant death, then the next-in-line successors of inbred royal parents may have been less capable because of a birth-order effect (instead of – or in addition to – a direct effect of inbreeding on intellectual capability). However, this is not the case. In unreported results (available upon request), we show that birth-order fixed effects do not affect our IV estimates.

overall, but especially so when these were risky, with potential territorial losses. Fourth, we decompose the change in urban population into an intensive component (growth or decline of urban population within a state's existing borders) and an extensive component (changes in urban population due to territorial gains or losses). We find that able rulers mostly expanded the (taxable) urban population via the extensive margin (Appendix D.4). In contrast, on average, capable rulers did not cause faster urban growth *within* their states' boundaries. This is compatible with historical facts across early modern Europe, where strong, capable rulers had an ambiguous effect on domestic city growth because they fostered economic prosperity on the one hand, but they also kept cities' ambitions to become independent in check, thereby curbing their potential to grow further (c.f. Angelucci, Meraglia, and Voigtländer, 2020). In sum, the evidence suggests that capable rulers fostered administrative efficiency, the rule of law, and economic prosperity within their realms, while choosing wisely which external conflicts to engage in – with the result that they managed to expand their territories into valuable, urbanized areas.

5 Constraints on Ruler Power

Were European states inevitably at the mercy of incapable, inbred rulers? The modern literature in political economy and management suggests that leaders matter particularly strongly when they act in institutionally unconstrained environments. Examining CEOs, Clark et al. (2014) show that "leaders matter most when ownership and governance structures correspond with a weak or ambiguous institutional logic." Similarly, at the national level in modern data, Jones and Olken (2005) find particularly strong changes in growth when autocratic leaders die, while Besley and Reynal-Querol (2017) document higher economic growth under hereditary leaders when constraints on them were weak. In our setting, all leaders were hereditary, but there were also important differences in the extent to which their actions were legally and de-facto constrained. In addition, in contrast to previous work, we observe ruler ability. We can thus examine whether institutional constraints mitigated the effects of ruler ability on state performance. We first describe a motivating example – monarchs in England only mattered before a strong parliament emerged – and then present our results.

5.1 Example: Constraints on England's Monarchs in the 17th Century

Consider the cross-country variation of our baseline OLS association documented in Figure 3. The coefficient for England was rather small, especially when compared to other Western European monarchies, such as France and Spain. In Figure A.5 in the appendix, we split England into two separate observations, one containing the reigns before the turbulent seventeenth century and one after 1600. In the seventeenth century, the Civil War and the Glorious Revolution led to increased constraints on the monarch in power (see Figure 6 and the discussion in Section 3.4). This change is also reflected in the relationship between ruler ability and state performance: We document a strong coefficient for England before 1600, which is very similar to other Western European states. After 1600, in contrast, the coefficient is close to zero.

5.2 Results: Constrained Monarchs Matter Less

To assess whether the ability of constrained European monarchs mattered less we estimate the following specification with interactions:

$$y_{r,s} = \beta_1 RulerAbility_{r,s} + \beta_2 Constrained_{r,s} + \beta_1 RulerAbility_{r,s} \times Constr_{r,s} + \delta_s + \varepsilon_{r,s} \quad (4)$$

where *Ruler Ability*$_{r,s}$ is the assessed capability of monarch of state s in reign r and *Constrained*$_{r,s}$ is a dummy variable indicating whether the ruler faced institutional constraints, based on the two variables described in Section 3.4 that indicate whether (prior to the start of a reign) constraints on the executive were substantial: the Polity-IV-based constraints on the executive and parliamentary activity. Regarding the outcome variables $y_{r,s}$, a valid concern is that Woods' (1906; 1913) assessment of ruler capability may have been affected by the extent to which their power was constrained. For this reason, we only use our two 'objective' measures of state performance as outcome variables, namely the change in territory and in urban population during each reign. Finally, δ_s denotes state fixed effects. In our IV results, we instrument for *Ruler Ability* with the coefficient of inbreeding $F_{r,s}$, and for the interaction term with *Constrained*$_{r,s} \times F_{r,s}$.[53]

Table 8 presents our results, beginning with our own reign-specific coding of constraints on rulers in columns 1-4. While we draw our conclusions from the IV results, we also report the OLS coefficient for both outcome variables, because these results i) provide a consistency check and b) draw on a larger sample of rulers, since our instrument – the coefficient of inbreeding – is not observed for all rulers. We find a sizable *negative* interaction term that is statistically significant in both IV specifications, and of very similar magnitude as the (positive) coefficient on *Ruler Ability*. In words, these results imply that the ability of rulers did not matter when they faced "substantial limitations on executive authority." The indicator for constrained rulers itself also has a statistically significant effect on territorial change and on urban growth. That is, constraints on executive power boost performance of the state also directly. If we were to take the estimated coefficients at face value, a one-std increase (about 1) in the ability of an *unconstrained* ruler would have a similar effect as introducing substantial institutional constraints on an average ruler (with *Ruler Ability*=0).[54]

Columns 5-8 repeat the analysis, using the activity of parliaments from Van Zanden et al.

[53]The exclusion restriction is that the interaction term *Constrained*$_{r,s} \times F_{r,s}$ affected changes in territory and urbanization only via the ruler ability – constraints channel. While it is possible to imagine violations of this condition, two features can help to address these: The variable *Constrained* in levels is included in both the first and second stage regressions, and the constraints on rulers are measured *before* the respective ruler came to power.

[54]Recall that we standardize *Ruler Ability* so it has mean zero and standard deviation 1.

(2012) to construct the dummy variable that indicates constrained rulers. While this indicator is much more coarse, it broadly confirms the results from the earlier columns: Throughout, capable rulers have positive effects (and incapable rulers, negative effects), and this finding is significantly weakened for rulers constrained by more active parliaments.

In sum, our results suggest that the capability of monarchs mattered less when and where their actions were constrained by institutions. In our setting, parliaments – and therefore the constraints on monarchs – became gradually stronger in North-Western Europe after the 16th century (Van Zanden et al., 2012). At the same time, the dynasties ruling Europe increasingly drew on an ever smaller pool of potentially suitable royal marriage partners. In turn, this increased the coefficient of inbreeding throughout, and particularly so in Northern Europe. One fascinating implication of our results is thus that the emergence of strong parliaments in North-Western Europe may have shielded these states from the negative effects of ever more inbred royal elites.

6 Conclusion

The importance of individual leaders for the course of history has been subject to continued debate since the times of Napoleon. The Emperor of the French also illustrates a central identification problem: rather than 'great men' shaping history, historical circumstances may give rise to 'great men,' who find their way into office even when born to a modest family on a far off Mediterranean island. In other words, it is hard to disentangle a causal effect of leaders on their country's performance from unobserved factors or even reverse causality. We explored the period that has been most prominently debated in this context: Europe between the 10th and 18th century.

This paper is the first to provide systematic causal evidence that more capable European rulers boosted outcomes for the states they governed. To identify these effects, we exploited the fact that European monarchs ascended to power by primogeniture, independent of their ability. In addition, ruler ability varied because of century-long inbreeding within dynasties. The detrimental effects of inbreeding were unknown until the 20th century; in fact, a popular belief among European dynasties was that kin marriage helped to preserve royal virtues. In addition, a significant part of consanguinity (the degree of genetic similarity) was 'hidden' in the history of kin marriage during previous generations. In combination, these features yield quasi-random variation in ruler ability, allowing us to identify its causal effect on state performance. We find sizeable coefficients, with capable leaders boosting their states' performance along multiple dimensions, including economic outcomes, administrative efficiency, urban growth, and territorial gains. The latter is particularly striking, given that capable rulers were less likely to engage in conflicts. In combination, these two observations suggest that able rulers chose wisely which conflicts to engage in, favoring those that promised territorial gains. Overall, our results imply that European rulers did 'make history,' with their actions shaping the European map during the period that laid the foundation for modern nation states.

We also showed that the effect of ruler ability on state performance was muted in states with strong institutional constraints on their monarchs. The most important institution exerting such constraints were parliaments, and these, in turn, were most active in North-Western Europe. At the same time, inbreeding of dynasties surged in North-Western Europe between the 15th and 18th century. Our results suggest that parliaments shielded Northern Europe's states from the adverse effects of inbreeding within their ruling dynasties.

References

Abramson, S. F. (2017). The Economic Origins of the Territorial State. *International Organization 71*(1), 97–130.

Acemoglu, D. and M. O. Jackson (2015). History, Expectations, and Leadership in the Evolution of Social Norms. *The Review of Economic Studies 82*(2), 423–456.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *American Economic Review 95*(3), 546–579.

Adams, R., M. Keloharju, and S. Knüpfer (2018). Are CEOs Born Leaders? Lessons from Traits of a Million Individuals. *Journal of Financial Economics 130*(2), 392–408.

Afzal, M. et al. (1993). Inbreeding Depression and Intelligence Quotient among North Indian Children. *Behavior Genetics 23*(4), 343–347.

Alston, E., L. J. Alston, and B. Mueller (2021). Leadership and Organizations. Technical report, National Bureau of Economic Research.

Alston, L. J. (2017). Beyond Institutions: Beliefs and Leadership. *The Journal of Economic History 77*(2), 353–372.

Alvarez, G., F. C. Ceballos, and C. Quinteiro (2009). The Role of Inbreeding in the Extinction of a European Royal Dynasty. *PLoS One 4*(4), e5174.

Andrews, I., J. H. Stock, and L. Sun (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics 11*(1), 727–753.

Angelucci, C., S. Meraglia, and N. Voigtländer (2020). How Merchant Towns Shaped Parliaments:From the Norman Conquest of England to the Great Reform Act.

Asdell, S. (1948). The Relation between Inbreeding and Intelligence. *Human Biology 20*(4), 171–181.

Assouad, L. (2020, September). Charismatic Leaders and Nation Building. working paper or preprint.

Ausenda, G. (1999). Kinship and Marriage Among the Visigoths. *The Visigoths from the Migration Period to the Seventh Century: An Ethnographic Perspective*, 129–90.

Bairoch, P., J. Batou, and P. Chèvre (1988). *La Population des villes Europeenees de 800 à 1850: Banque de Données et Analyse Sommaire des Résultats*. Geneva: Centre d'histoire economique Internationale de l'Université de Genève, Libraire Droz.

Bartlett, R. (2020). *Blood Royal: Dynastic Politics in Medieval Europe*. Cambridge University Press.

Baten, J., T. Keywood, and G. Wamser (2021). Territorial State Capacity and Elite Violence from the 6th to the 19th Century. *European Journal of Political Economy*, 102037.

Becker, S. O., A. Ferrara, E. Melander, and L. Pascali (2020). Wars, Taxation and Representation: Evidence from Five Centuries of German History. Technical report, CEPR Discussion Papers.

Becker, S. O. and H. K. Hvide (2013). Do Entrepreneurs Matter?

Benzell, S. G. and K. Cooke (2018). A Network of Thrones: Kinship and Conflict in Europe, 1495-1918. Technical report, Working Paper.

Berra, T. M., G. Alvarez, and F. C. Ceballos (2010). Was the Darwin/Wedgwood Dynasty Adversely Affected By Consanguinity? *BioScience 60*(5), 376–383.

Bertrand, M. and A. Schoar (2003). Managing with Style: The Effect of Managers on Firm Policies. *The Quarterly Journal of Economics 118*(4), 1169–1208.

Besley, T., J. G. Montalvo, and M. Reynal-Querol (2011). Do Educated Leaders Matter? *The Economic Journal 121*(554), F205–227.

Besley, T. and M. Reynal-Querol (2017). The Logic of Hereditary Rule: Theory and Evidence. *Journal of Economic Growth 22*(2), 123–144.

Bittles, A. H. (2012). *Consanguinity in Context*, Volume 63. Cambridge University Press.

Bloom, N. and J. Van Reenen (2007). Measuring and Explaining Management Practices across Firms and Countries. *The Quarterly Journal of Economics 122*(4), 1351–1408.

Bradbury, J. (1999). Philip Augustus and King John: Personality and History. *King John: New Interpretations*, 347–361.

Braudel, F. and S. Reynolds (1992). *Identity of France: People and Production*. The Identity of France. HarperPerennial.

Carlyle, T. (1840). *On Heroes, Hero-worship, & The Heroic in History*. Number 74.

Carr, R. (1991). Bourbon Spain, 1700–1808. By John Lynch.(A History of Spain.) Pp. xiv+ 40 incl. 7 plates, 3 maps and 8 tables. Oxford–Cambridge, Mass.: Basil lackwell, 1989.£ 35. 0 631 14576 1. *The Journal of Ecclesiastical History 42*(1), 131–133.

Ceballos, F. and G. Álvarez (2013). Royal Dynasties as Human Inbreeding Laboratories: The Habsburgs. *Heredity 111*(2), 114–121.

Charlesworth, D. and J. H. Willis (2009). The Genetics of Inbreeding Depression. *Nature reviews genetics 10*(11), 783–796.

Chattopadhyay, R. and E. Duflo (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica 72*(5), 1409–1443.

Clark, C. (2007). *Iron Kingdom: The Rise and Downfall of Prussia, 1600-1947*. Penguin UK.

Clark, J. R., C. Murphy, and S. J. Singer (2014). When Do Leaders Matter? Ownership, Governance and the Influence of CEOs on Firm Performance. *The Leadership Quarterly 25*(2), 358–372.

Dal Bó, E., P. Dal Bó, and J. Snyder (2009). Political Dynasties. *The Review of Economic Studies 76*(1), 115–142.

Devlin, B., M. Daniels, and K. Roeder (1997). The Heritability of IQ. *Nature 388*(6641), 468–471.

Dippel, C. and S. Heblich (2021). Leadership and Social Movements: The Forty-Eighters in the Civil War. (2), 1–35.

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

Do, Q.-A., S. Dray, E. Huillery, and J.-L. Keene (2020). Emperors without Scepters: Early Colonial Leaders' Personality and Civil Conflicts.

Dube, O. and S. Harish (2020). Queens. *Journal of Political Economy 128*(7), 2579–2652.

Easterly, W. and S. Pennings (2020). Leader Value Added: Assessing the Growth Contribution of Individual Mational Leaders.

Engels, F. (1968). Letter to W. Borgius, January 25, 1894. *Karl Marx and Friedrich Engels, Selected Works*, 704–705.

Eslava, F. (2020). Conflict and Gender: Female Mayors in Colombia.

Fareed, M. and M. Afzal (2014a). Estimating the Inbreeding Depression on Cognitive Behavior: A Population Based Study of Child Cohort. *PLoS One 9*(10), e109585.

Fareed, M. and M. Afzal (2014b). Evidence of Inbreeding Depression on Height, Weight, and Body Mass Index: A Population-based Child Cohort Study. *American Journal of Human Biology 26*(6), 784–795.

Fareed, M., M. K. Ahmad, M. A. Anwar, and M. Afzal (2017). Impact of Consanguineous Marriages and Degrees of Inbreeding on Fertility, Child Mortality, Secondary Sex Ratio, Selection Intensity, and Genetic Load: A Cross-Sectional Study from Northern India. *Pediatric Research 81*(1), 18–26.

Ferreira, F. and J. Gyourko (2014). Does Gender Matter for Political Leadership? The Case of US Mayors. *Journal of Public Economics 112*, 24–39.

George, S. E. and D. Ponattu (2018). Like Father, Like Son? The Effect of Political Dynasties on Economic Development. Technical report, Working Paper.

Goody, J. (1983). *The Development of the Family and Marriage in Europe*. Cambridge University Press.

Grant, M. (1919). The Passing of the Great Race, or The Racial Basis of European History. 1916. *Preface: Henry Fairfield Osborn. New York: Charles Scribner's Sons*.

Gueniffey, P. (2020). *Napoleon and de Gaulle: Heroes and History*. Harvard University Press.

Hamamy, H., S. E. Antonarakis, L. L. Cavalli-Sforza, S. Temtamy, G. Romeo, L. P. Ten Kate, R. L. Bennett, A. Shaw, A. Megarbane, C. van Duijn, et al. (2011). Consanguineous Marriages, Pearls and Perils: Geneva International Consanguinity Workshop Report. *Genetics in Medicine 13*(9), 841–847.

Hamilton, E. J. (1938). Revisions in Economic History: VIII.–The Decline of Spain. *The Economic History Review 8*(2), 168–179.

Horowitz, M. C., A. C. Stam, and C. M. Ellis (2015). *Why Leaders Fight*. Cambridge University Press.

Jones, B. F. and B. A. Olken (2005). Do Leaders Matter? National Leadership and Growth since World War II. *The Quarterly Journal of Economics 120*(3), 835–864.

Judge, T. A., A. E. Colbert, and R. Ilies (2004). Intelligence and Leadership: A Quantitative Review and Test of Theoretical Propositions. *Journal of Applied Psychology 89*(3), 542.

Kennedy, P. (1988). *The Rise and Fall of the Great Powers*. New York: Random House.

Kennedy, P. (1989). The Rise and Fall of the Great Powers. *US: Random House*.

Kokkonen, A. and A. Sundell (2014). Delivering Stability—Primogeniture and Autocratic Survival in European Monarchies 1000-1800. *American Political Science Review*, 438–453.

Lewontin, R. C. (1970). Race and Intelligence. *Bulletin of the Atomic Scientists 26*(3), 2–8.

Logan, T. D. (2018). Do Black Politicians Matter? Technical report, National Bureau of Economic Research.

MacCulloch, D. (2018). *Thomas Cromwell: A Life*. Penguin UK.

Malinowski, B. (1927). Sex and Repression in Savage Society.

Malmendier, U. and G. Tate (2005). CEO Overconfidence and Corporate Investment. *The journal of finance 60*(6), 2661–2700.

March, J. G. and T. Weil (2009). *On Leadership*. John Wiley & Sons.

Marshall, M. G., K. Jaggers, and T. R. Gurr (2017). Polity IV Project: Political Regime Characteristics and Transitions, 1800-2016 Dataset Users' Manual. *available at https://www.systemicpeace.org/inscr/p4manualv2016.pdf*.

Marx, K. (1907). The eighteenth brumaire of louis bonaparte. *Charles H Kerr and Co.*.

McQuillan, R., N. Eklund, N. Pirastu, M. Kuningas, B. P. McEvoy, T. Esko, T. Corre, G. Davies, M. Kaakinen, L.-P. Lyytikäinen, et al. (2012). Evidence of Inbreeding Depression on Human Height. *PLoS Genet 8*(7), e1002655.

Mitchell, S. Z. (2013). *Mariana of Austria and Imperial Spain: Court, Dynastic, and International Politics in Seventeenth-Century Europe*. Ph. D. thesis, University of Miami.

Mobarak, A. M., T. Chaudhry, J. Brown, T. Zelenska, M. N. Khan, S. Chaudry, R. A. Wajid, A. H. Bittles, and S. Li (2019). Estimating the Health and Socioeconomic Effects of Cousin Marriage in South Asia. *Journal of biosocial science 51*(3), 418–435.

Montiel Olea, J. and C. Pflueger (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Morton, N. E. (1978). Effect of Inbreeding on IQ and Mental Retardation. *Proceedings of the National Academy of Sciences 75*(8), 3906–3908.

Neisser, U., G. Boodoo, T. J. Bouchard Jr, A. W. Boykin, N. Brody, S. J. Ceci, D. F. Halpern, J. C. Loehlin, R. Perloff, R. J. Sternberg, et al. (1996). Intelligence: Knowns and Unknowns. *American psychologist 51*(2), 77.

Nicholas, H. G. (2021). Winston Churchill.

Nunn, N. and N. Qian (2011). The Potato's Contribution to Population and Urbanization: Evidence from an Historical Experiment. *Quarterly Journal of Economics 126*(2), 593–650.

Oskarsson, S., C. T. Dawes, K.-O. Lindgren, and R. Öhrvall (2021). Big Brother Sees You, But Does He Rule You? The Relationship Between Birth Order and Political Candidacy. *The Journal of Politics 83*(3), 000–000.

Payne, S. G. (1973). A History of Spain and Portugal. Vol. 2. *Madison: U of Wisconsin P*.

Robert, A., B. Toupance, M. Tremblay, and E. Heyer (2009). Impact of Inbreeding on Fertility in a Preindustrial Population. *European Journal of Human Genetics 17*(5), 673–681.

Roberts, A. (2018). *Churchill: Walking with Destiny*. Penguin.

Röhl, J. C. (1996). *The Kaiser and his Court: Wilhelm II and the Government of Germany*. Cambridge University Press.

Rohrer, J. M., B. Egloff, and S. C. Schmukle (2015). Examining the Effects of Birth Order on Personality. *Proceedings of the National Academy of Sciences 112*(46), 14224–29.

Royuela-Rico, A. (2020). Inbreeding and Serious Mental Illness in the First Spanish Bourbons. *Revista de Neurologia 71*(2), 61–68.

Saini, A. (2019). *Superior: The Return of Race Science*. Beacon Press.

Scheidel, W. (1995). Incest Revisited: Three Notes on the Demography of Sibling Marriage in Roman

Egypt. *The Bulletin of the American Society of Papyrologists 32*(3/4), 143–155.

Schulz, J. F. (2016). The Churches' Bans on Consanguineous Marriages, Kin-networks and Democracy. Technical report, CeDEx Discussion Paper Series.

Simonton, D. K. (1983). Intergenerational Transfer of Individual Differences in Hereditary Monarchs: Genetic, Role-modeling, Cohort, or Sociocultural Effects? *Journal of personality and social psychology 44*(2), 354.

Simonton, D. K. (2006). Presidential IQ, Ppenness, Intellectual Brilliance, and Leadership: Estimates and Correlations for 42 US Chief Executives. *Political Psychology 27*(4), 511–526.

Spencer, H. (1896). *The Study of Sociology*. Appleton.

Stearns, P. N. and W. L. Langer (2001). *The Encyclopedia of World History*.

Thorndike, E. L. (1936). The Relation between Intellect and Morality in Rulers. *American Journal of Sociology 42*(3), 321–334.

Tolstoy, L. (2007). War and Peace (Richard Pevear & Larissa Volokhonsky, Trans.).

Tullock, G. (1987). Autocracy. *Dordrecht: Martinus Nijhoff*.

Van Den Berghe, P. L. and G. M. Mesher (1980). Royal Incest and Inclusive Fitness. *American Ethnologist 7*(2), 300–317.

Van Zanden, J. L., E. Buringh, and M. Bosker (2012). The Rise and Decline of European Parliaments, 1188–1789 1. *The Economic History Review 65*(3), 835–861.

Weber, M. (1921). Wirtschaft und Gesellschaft.

White, L. A. (1948). The Definition and Prohibition of Incest. *American Anthropologist 50*(3), 416–435.

William, J. (1880). Great Men, Great Thoughts, and the Environment. *The Atlantic Monthly 46*(276), 443.

Wolf, A. P. (2005). *Inbreeding, Incest, and the Incest Taboo: The State of Knowledge at the Turn of the Century*. Stanford University Press.

Woods, F. A. (1906). *Mental and Moral Heredity in Royalty: A Statistical Study in History and Psychology*. H. Holt.

Woods, F. A. (1913). *The Influence of Monarchs: Steps in a New Science of History*. Macmillan.

Wright, S. (1921). Systems of Mating. II. The Effects of Inbreeding on the Genetic Composition of a Population. *Genetics 6*(2), 124.

Xuetong, Y. (2019). *Leadership and the Rise of Great Powers*. Princeton University Press.

Yao, Y. and M. Zhang (2015). Subnational Leaders and Economic Growth: Evidence from Chinese Cities. *Journal of Economic Growth 20*(4), 405–436.

Young, M. W. (2004). *Malinowski: Odyssey of an Anthropologist, 1884-1920*. Yale University Press.

FIGURES

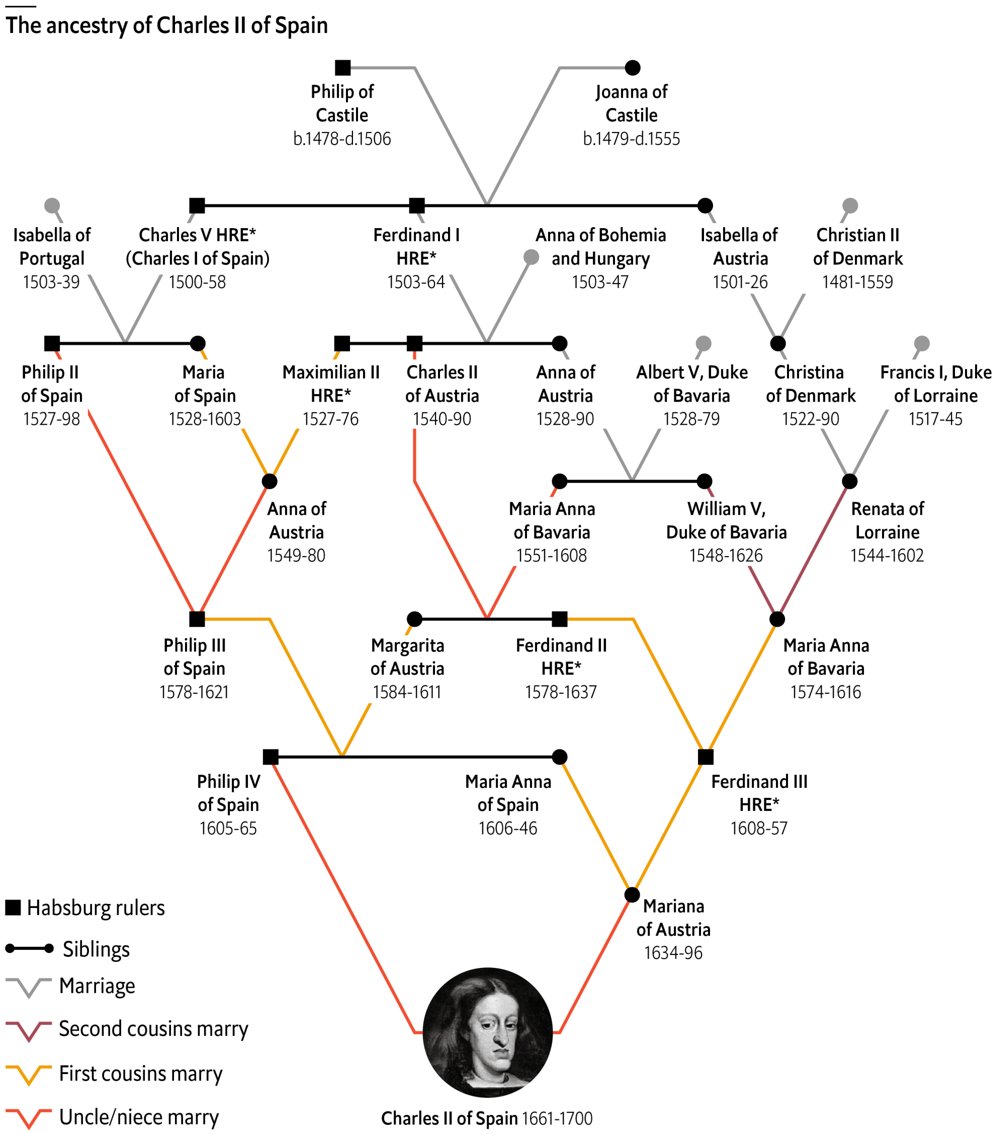

Figure 1: Pedigree of Carlos II. of Spain

Note: The figure shows the pedigree of Carlos II., King of Spain from 1665 to 1700. Note the intricate links to common ancestors of both his parents, stretching back over multiple generations. From The Economist's coverage of this paper on February 20th, 2021 © The Economist Newspaper Limited, London. All rights reserved.

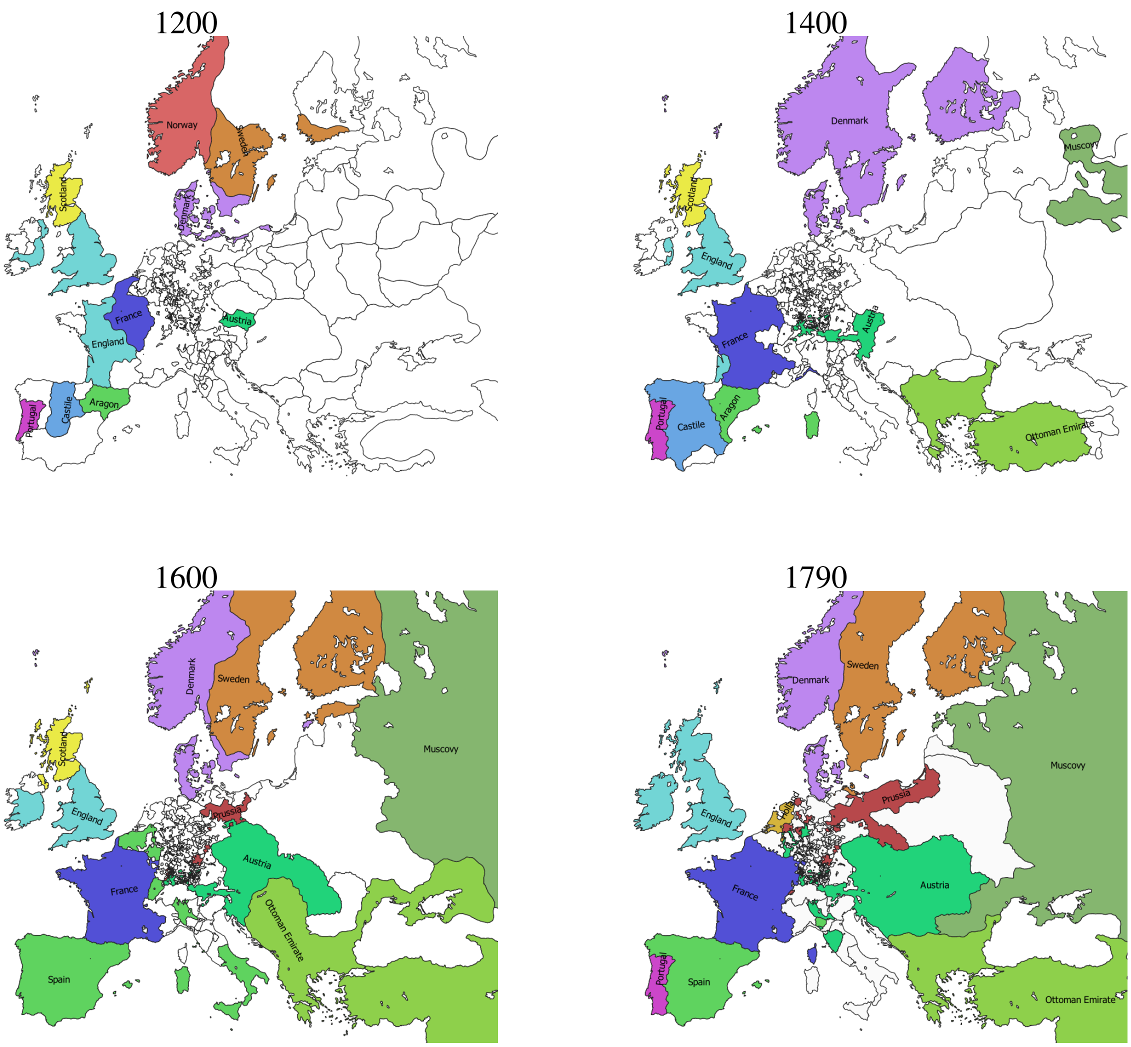

Figure 2: States in Sample

Note: The figure shows the boundaries of the states in our baseline sample at four points in time: 1200, 1400, 1600, and 1790. Data on state boundaries are from Abramson (2017). See Appendix A.1 for detail

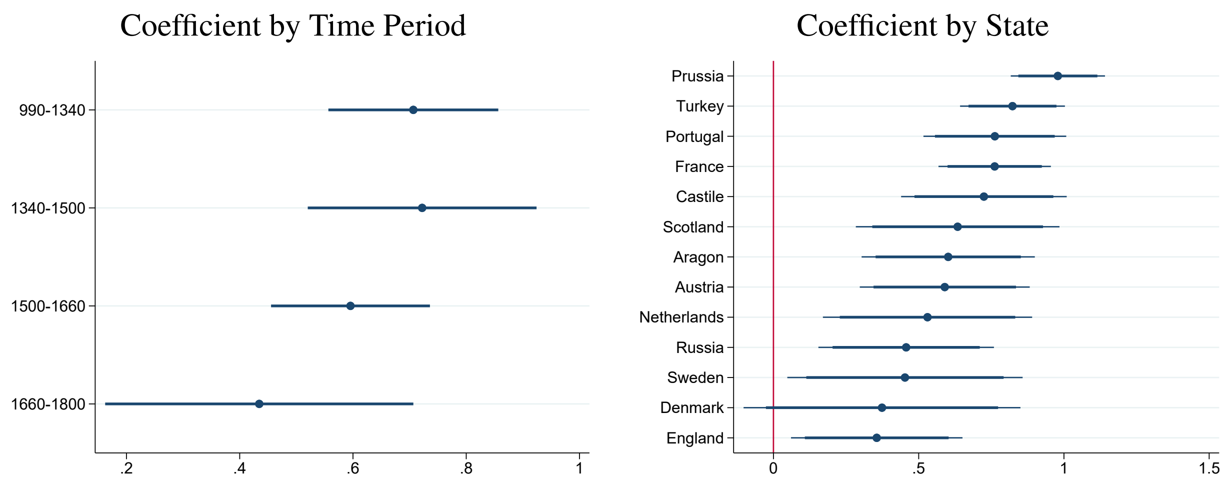

Figure 3: Association between Monarch Ability and State Performance by Period and State

Note: The figure shows coefficients of regressing ruler ability on state performance by time period (left panel) and by states in our baseline sample (right panel). Underlying each panel, we run a joint OLS estimation that includes state fixed effects. The corresponding regressions are reported in Appendix B.3. The figure also shows 90% confidence intervals (based on standard errors clustered at the state level).

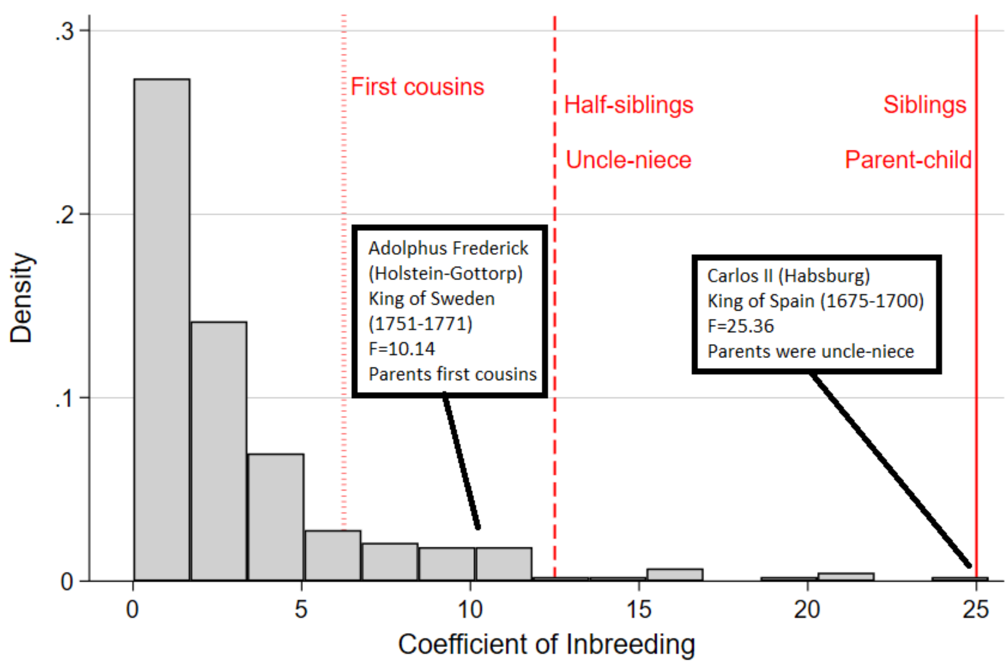

Figure 4: Histogram: Coefficient of inbreeding of Monarchs

Note: The figure shows the distribution of the coefficient of inbreeding (F) – the instrument for ruler ability in our analysis – for the 246 European Monarchs with available genealogical information in our baseline dataset. $F = 0$ indicates no relation among the parents of a monarch, $F = 50$ would theoretically result from self-fertilization.

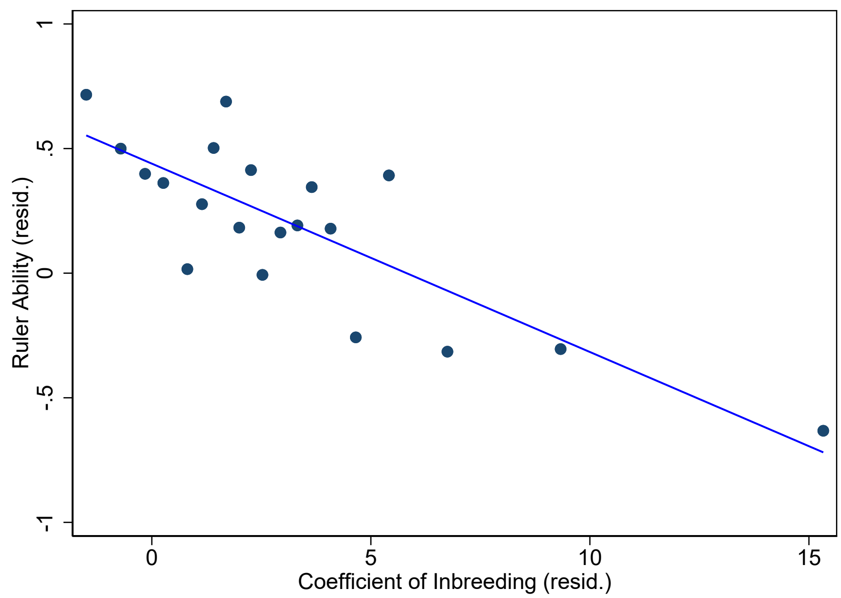

Figure 5: First Stage: Binscatter with Country Fixed Effects

Note: The figure shows a binned scatter plot for our first-stage regression of a ruler ability on the coefficient of inbreeding, controlling for state fixed effects. Each of the 20 bins in the graph corresponds to more than 10 individual rulers.

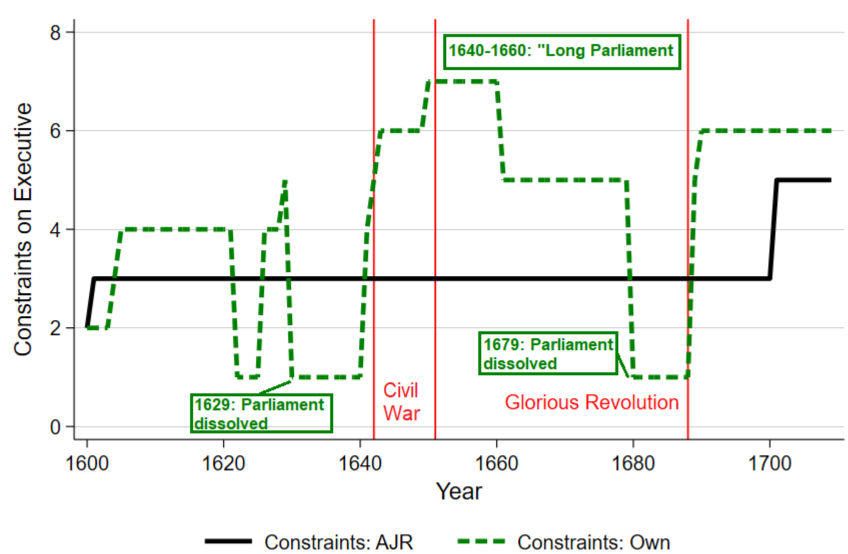

Figure 6: Constraints on Executive: Year-by-year, 17th Century England

Note: The figure shows changes in constraints on the executive for England in the 17th century, using the Polity IV score that ranges from 1-7. The black solid line depicts the century-level coding by Acemoglu et al. (2005), while the green dashed line shows our annual variable, which can then be mapped to individual reigns.

TABLES

Table 1: Monarchs and Performance of State – OLS Results

Dependent variable as indicated in table header

Dep. Var.	*State Performance*		$\Delta log(Area)$		$\Delta log(UrbPop)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Ruler Ability	0.616***	0.618***	0.117***	0.113***	0.110***	0.099***
	(0.052)	(0.050)	(0.032)	(0.035)	(0.028)	(0.027)
State FE		✓		✓		✓
R^2	0.38	0.41	0.07	0.11	0.06	0.10
Observations	336	336	298	298	289	289

Note: The table documents a strong relationship between ruler ability and our three measures of performance of the state at the reign level. *State Performance* in columns 1-2 is a comprehensive measure based on the coding by Woods (1913). The dependent variable in columns 3-4 is the change in a state's land area during a monarch's reign, and in columns 5-6, it is the change in total urban population. All regressions are run at the reign level. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 2: Robustness of OLS Results: Different Samples

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Notes:	Baseline	Only Dynasty Members	Exclude Regencies	Exclude Foreign Rule	Exclude Multi-Reign Rulers	All Restric-tions	Extended Sample until 1914	Extended Sample incl. PL & HU
Ruler Ability	0.618***	0.653***	0.687***	0.626***	0.610***	0.670***	0.532***	0.554***
	(0.050)	(0.058)	(0.070)	(0.052)	(0.052)	(0.071)	(0.044)	(0.042)
State FE	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.41	0.45	0.50	0.42	0.40	0.50	0.34	0.34
Observations	336	290	262	325	319	235	374	437

Note: The table documents the robustness of our baseline regression (col 2 in Table 1) to using different samples. *State Performance* is a comprehensive measure based on the coding by Woods (1913). See Section 4.2 for a detailed description of the sample restrictions for cols 2-6. In col 7 we extend the sample (based on our own coding) for all states included in Woods until 1914. In col 8 we extend the sample to further include Poland and Hungary, again until 1914 (see Appendix A.4 for detail). All regressions are run at the reign level. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 3: OLS Results – Heterogeneity by Ruler Characteristics

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)
Dummy for:	Female	Young Ascension	Designated Heir	Regicide (lagged)
Ruler Ability	0.613***	0.557***	0.578***	0.623***
	(0.065)	(0.094)	(0.186)	(0.062)
Dummy × Ruler Ability	0.021	0.111	0.015	-0.071
	(0.126)	(0.101)	(0.193)	(0.180)
Dummy	-0.207*	0.000	-0.233	-0.176
	(0.107)	(0.067)	(0.154)	(0.222)
State FE	✓	✓	✓	✓
R^2	0.41	0.40	0.45	0.41
Observations	311	305	141	191

Note: The table shows results of interacting the baseline regression with dummy variables. In column 1, this indicator is one if the ruler was a woman. In column 2, the interaction variable is a dummy for rulers ascending to the throne below the median age of 28 years. In column 3, it indicates rulers who were raised as designated heir, while in column 4 it indicates whether the prior ruler was murdered or executed after trial. The dependent variable, *State Performance*, is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Inbreeding and Monarch Ability – First-Stage Results

Dependent Variable: Ruler Ability

	(1)	(2)	(3)	(4)	(5)
Note:			F < 25	F < 12.5	Documented PG†
Coefficient of Inbreeding	-0.068***	-0.076***	-0.077***	-0.084***	-0.057***
	(0.012)	(0.012)	(0.014)	(0.017)	(0.012)
State FE		✓	✓	✓	✓
R^2	0.09	0.15	0.15	0.12	0.15
Observations	235	235	234	227	136

Note: The table shows results of our first-stage regressions of ruler ability on monarchs' coefficient of inbreeding. The coefficient of inbreeding measures the degree of similarity in the genes of offspring due to common ancestors, and thus the increased risk of genetic disorders resulting from the consanguinity of the monarch's parents. Column 3 excludes Carlos II of Spain, whose parents shared as many genes as offspring of siblings. Column 4 excludes all monarchs whose parents shared at least as many genes as offspring of half-siblings. All regressions are run at the reign level. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
† Subsample includes only documented cases where rulers ascended to power due to primogeniture.

Table 5: Monarchs and Country Performance – IV and Reduced-Form Results

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	*State Performance*		$\Delta log(Area)$		$\Delta log(UrbanPop.)$	
Note:		F < 12.5		F < 12.5		F < 12.5
			A. Second Stage Regressions			
Ruler Ability	0.794***	0.645***	0.176***	0.311**	0.191***	0.301**
	(0.100)	(0.185)	(0.051)	(0.123)	(0.049)	(0.121)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	45.8	26.7	42.4	22.9	37.8	23.1
Observations	235	227	203	196	198	191
			B. Reduced-Form Regressions			
Coefficient of Inbreeding	-0.061***	-0.054***	-0.012***	-0.025**	-0.013***	-0.023**
	(0.011)	(0.013)	(0.004)	(0.010)	(0.003)	(0.010)
State FE	✓	✓	✓	✓	✓	✓
R^2	0.11	0.08	0.10	0.11	0.07	0.08
Observations	235	227	203	196	198	191

Note: The table shows the results of second-stage and reduced-form regressions for our three outcome variables. The instrument, the coefficient of inbreeding, measures the increased risk of genetic disorders resulting from the consanguinity of the monarch's parents. Columns 1-2 use the assessment of political and economic conditions during each monarch's reign from Woods (1913) as measure of state performance, column 3-4 use the change in land area during each monarch's reign, calculated from Abramson (2017), and column 5-6 use the change in the urban population of the state during each monarchs reign. Column 2, 4, and 6 exclude all Monarchs whose parents share as many genes as offspring of half-siblings. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4. All regressions are run at the reign level and include state fixed effects. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Monarchs and Country Performance – ‘Hidden’ Inbreeding

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	*State Performance*		$\Delta log(Area)$		$\Delta log(UrbanPop.)$	
Note:		F < 12.5		F < 12.5		F < 12.5
A. Second Stage Regressions						
Ruler Ability	0.827***	0.807***	0.178**	0.259**	0.203***	0.276***
	(0.137)	(0.214)	(0.071)	(0.114)	(0.058)	(0.099)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	17.9	14.5	21.1	17.5	20.7	17.5
Observations	235	227	203	196	198	191
B. Reduced-Form Regressions						
‘Hidden’ Coefficient of Inbreeding	-0.097***	-0.090***	-0.020**	-0.028**	-0.023**	-0.030**
	(0.024)	(0.027)	(0.008)	(0.013)	(0.009)	(0.013)
State FE	✓	✓	✓	✓	✓	✓
R^2	0.12	0.10	0.10	0.11	0.07	0.08
Observations	235	227	203	196	198	191

Note: The table repeats all specifications from Table 5, but using only the ‘hidden’ component of inbreeding as an instrumental variable for ruler ability. The ‘hidden’ part of the overall coefficient of inbreeding was due to complex intermarriage patterns in the generations prior to a ruler’s parents. It could not be computed before the 20th century. See Appendix A.6 for details on the calculation of the ‘hidden’ measure of inbreeding. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 16.4. All regressions are run at the reign level and include state fixed effects. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Ruler-Pair Regressions

Dep. Var.: Δ_{ij} *State Performance*

	(1)	(2)	(3)	(4)	(5)
Note:					One-ruler match
			A. OLS		
Δ_{ij} Ruler Ability	0.631***	0.633***	0.613***	0.593***	0.544***
	(0.012)	(0.013)	(0.015)	(0.017)	(0.030)
R^2	0.44	0.46	0.75	0.81	0.87
Observations	5,512	5,512	5,512	5,491	1,755
			B. First Stage Regressions		
Δ_{ij} Coefficient of Inbreeding	-0.041***	-0.041***	-0.042***	-0.041***	-0.026***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
R^2	0.15	0.17	0.62	0.70	0.77
Observations	3,070	3,070	3,069	3,035	1,755
			C. Second Stage Regressions		
Δ_{ij} Ruler Ability	0.870***	0.848***	0.819***	0.667***	0.842***
	(0.093)	(0.094)	(0.072)	(0.071)	(0.121)
First Stage F-statistic	89.80	88.51	124.36	72.58	44.60
Observations	3,070	3,070	3,069	3,035	1,755
			D. Reduced Form Regressions		
Δ_{ij} Coefficient of Inbreeding	-0.035***	-0.035***	-0.035***	-0.028***	-0.036***
	(0.004)	(0.005)	(0.004)	(0.003)	(0.004)
R^2	0.13	0.15	0.62	0.71	0.81
Observations	3,070	3,070	3,069	3,035	1,755
State FEs	✓	✓	✓	✓	✓
State-pair FE		✓	✓	✓	✓
Ruler FE			✓	✓	✓
Country-pair × Century FE				✓	✓

Note: The table shows results from ruler-pair regressions. For each ruler, we compute the pair-wise difference in their ability, their coefficient of inbreeding, and the performance of their state relative to all concurrently ruling monarchs. Columns 1-4 include, for each ruler, all rulers of other states that overlapped for at least one year in their reign. Column 5 keeps for each ruler only the one ruler from all other states that he or she shared the largest temporal overlap in their reigns with. Section 4.5 provides further detail. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4. All regressions are run at the reign-pair level. Standard errors, clustered at the country-pair level, in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 8: The Role of Institutional Constraints on Ruler Power

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constraints coding:	—- Authors' Annual Polity IV Coding —-				– Century-Level Parliamentary Activity –			
Dep. Var.	$\Delta log(Area)$		$\Delta log(UrbanPop.)$		$\Delta log(Area)$		$\Delta log(UrbanPop.)$	
Estimation:	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Ruler Ability	0.104***	0.154***	0.096***	0.189***	0.095***	0.182***	0.099***	0.207***
	(0.034)	(0.042)	(0.026)	(0.051)	(0.030)	(0.056)	(0.030)	(0.054)
Constrained Ruler	0.108***	0.054*	0.300***	0.218***	-0.045	-0.113	-0.131***	-0.156
	(0.010)	(0.032)	(0.036)	(0.024)	(0.043)	(0.074)	(0.038)	(0.114)
Constrained Ruler × Ruler Ability	-0.109**	-0.106***	-0.041	-0.137**	-0.076	-0.183*	-0.108**	-0.401***
	(0.038)	(0.035)	(0.031)	(0.057)	(0.042)	(0.093)	(0.036)	(0.155)
State FE	✓	✓	✓	✓	✓	✓	✓	✓
First Stage F-Stat		15.4		13.4		15.4		13.6
R^2	0.10		0.10		0.09		0.09	
Observations	295	200	286	195	269	202	263	197

Note: The table shows that the effect of ruler ability on the performance of their states was muted when their executive power was constrained. In columns 1-4, the dummy *Constrained Ruler* indicates "substantial limitations" on ruler power, as by our own reign-level coding based on the Polity IV scale (see Section 5 and Appendix E for detail). In columns 5-8 we define a monarch to be constrained if the state was above the 90th percentile of the (century-level) measure of parliamentary activity from Van Zanden et al. (2012). All regressions are run at the reign level. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Online Appendix

History's Masters
The Effect of European Monarchs on State Performance

Sebastian Ottinger (UCLA)   Nico Voigtländer (UCLA, NBER, and CEPR)

A Data: Coverage, Validation, and Detail

This appendix provides background on the coverage of our dataset, the coding of variables, and summary statistics.

A.1 Detail: Data Set and Variables

Data Set

Our main data set is based on a list of reigns for 13 states. Woods (1913) provides tables on pages 305-403, listing – for each reign – the time period, the name of the ruler (or a description of the status when no monarch reigned, such as for interregna or Republican government in the Netherlands), an assessment of the rulers' ability, as well as the performance of the state during this reign. Ruler ability and state performance are coded categorically ranging from "-" to "+." For the few cases where more than one ruler appears, we focus on the ruler whose coding works against our baseline results.[1]

Sample Coverage

Table A.1 provides detail on the sample size. In total, 366 reigns are recorded by Woods (1913). For 353 of these, Woods was able to assess state performance. The others are either very short reigns or Woods was not able to make a definitive assessment based on scarce sources. Figure A.1 provides a timeline for all states in our main sample. The earliest state to enter our sample is France (in 990, when Hugh Capet founded the Capetian dynasty), and the last state to enter is Sweden, after it split from Denmark in 1623 under Gustavus Vasa to become a separate political entity.

For 341 reigns, Woods assessed the ability of the ruler. He was unable to do so for instances where rule was short or for episodes of Republican government in the Netherlands. Our alternative measures of state performance based on territorial changes and the change in urban population within the state are available for 317 and 307 reigns, respectively.[2] In total, both our main ex-

[1] For instance, for Ferdinand and Isabella, who jointly and successfully ruled over the Habsburg Empire from 1479 to 1504, we focus on Isabella, who had a higher coefficient of inbreeding than her husband.

[2] Most of the reduction in sample size is explained by the fact that the data in Abramson (2017) only ranges from

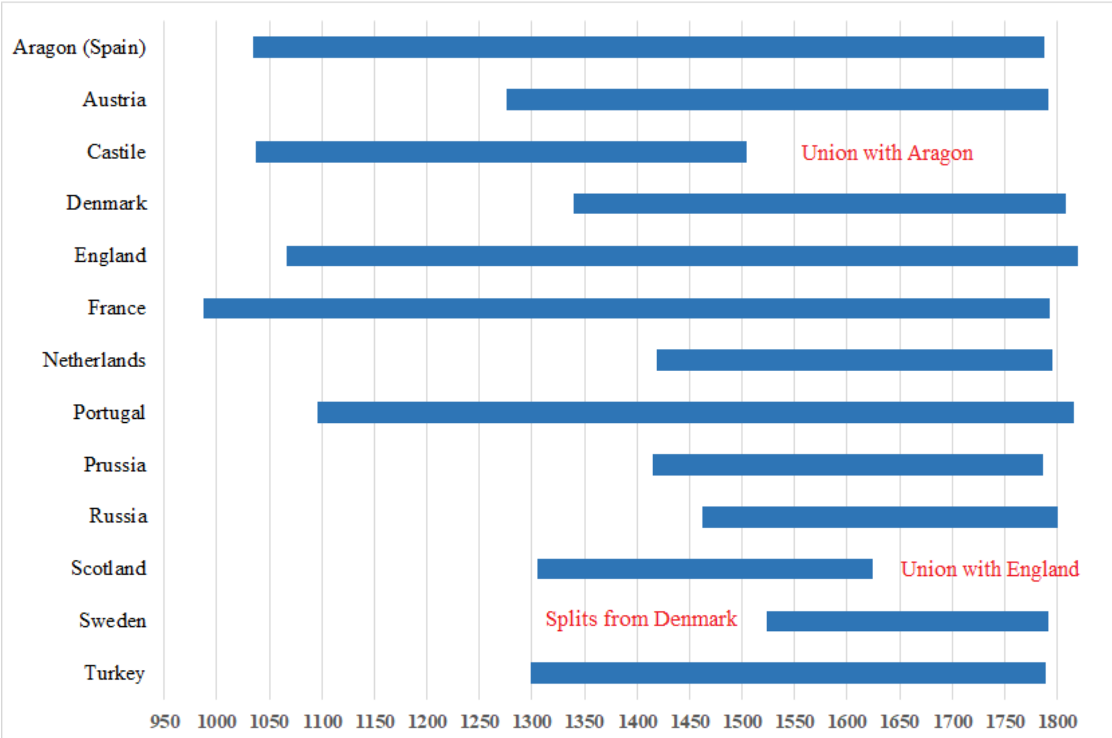

Figure A.1: Timeline of Sample Coverage: States in Sample

Note: The figure shows the states in our sample together with the time period over which they are covered.

planatory and outcome variable – *Ruler Ability* and *State Performance* assessed by Woods – are available for 336 reigns. Only for 235 of these were listed in our genealogical data source, so that we have information on the coefficient of inbreeding for those.

Additional Variables

In addition to the variables described in Section 3 of the paper, we code other characteristics of rulers whenever this information is available. Specifically, we collect for each monarch the birth and death year, sex, the dynasty the monarch belonged to, whether the monarch was raised as a designated heir or rather educated for a different role (say, for ecclesiastical life as a younger born royal offspring, or – as in the case of Catherine I. of Russia – for neither of such roles, but as a orphaned household servant), and whether the monarch ascended to the throne due to primogeniture as heir apparent (or, due to agnatic succession, by election of a council, or by starting a new dynasty). We collect this information from the English-language Wikipedia, but amend it whenever required by information from the corresponding national language Wikipedia.

1100 to 1790, so that we do not have areas at the beginning or the end of some reigns, or both. In a few other cases, Woods' list starts while the political entity is not yet de facto politically independent and therefore not covered by Abramson (2017), as for instance for the early years of the Netherlands.

Table A.1: Sample

Sample	Obs.
All reigns	366
Reigns with assessed *State Performance*	353
Reigns with assessed *Ruler Ability*	341
Reigns with information on border changes	317
Reigns with information on urban population	307
Both: *Ruler Ability* & *State Performance*	336
Both + individuals (gender assigned)	311
Both + coefficient of inbreeding (F)	235
Both: *Ruler Ability* & border changes	298
Both + coefficient of inbreeding (F)	203
Both: *Ruler Ability* & urban population	289
Both + coefficient of inbreeding (F)	203

Note: This table provides details on our baseline sample size for the three outcome variables (*State Performance*, territorial changes, and change in urban population during reigns) as well as the main explanatory variable *Ruler Ability* and our instrument – the coefficient of inbreeding of rulers.

A.2 Validation of Woods' State Performance and Ruler Ability Coding

To check Woods' (1906; 1913) coding of state performance and ruler ability, we asked research assistants to review the evidence in various encyclopedias and devise own assessments of ruler capability and state performance, using Woods' three-tier scale.

The left panel of Figure A.2 provides a binned scatter plot of our research assistants' assessment of monarch ability with that of Woods (1913). A clear assessment was possible based on online encyclopedias only for 169 rulers. In 96 out of 169 assessed cases, our research assistants reached the same assessment as Woods did, while in 20 they reached the opposite assessment. Those examples for instance include Peter III of Russia. He ruled for less than a year in 1762, and Woods characterized him as "[w]eak, dissolute, violent." However, this characterization has been reversed by historians since the time of Woods (c.f. Palmer, 2005)), and is reflected in the assessment of our research assistants. Of the remaining 53 cases, in 16 cases Woods assigned a grade between -1,0, or 1. Our research assistants were not given this option and hence there cannot be exact agreement for those. Of the remaining 37 cases, 19 are instances where our research assistant assigned the monarch's ability a value 1, while Woods assigned 0. These cases include James IV of England and Leopold I of Hapsburg. Overall, the correlation between our own and Woods' coding is $\rho = 0.52$.

The right panel of Figure A.2 provides a binned scatter plot of our research assistants' assessment of country performance with that of Woods (1913) ($\rho = 0.49$). Of the 234 reigns for which our research assistants were confident in making an assessment, in 124 they completely agreed with Woods' assessment. In 27 instances, they reached the opposite assessment than Woods did; in 18 cases Woods assigned a state performance between the values of 0, 1, and -1. The remaining 83 instances are cases where our research assistants and Woods disagree in their assessment of state performance by a value of one, but not diametrically so.

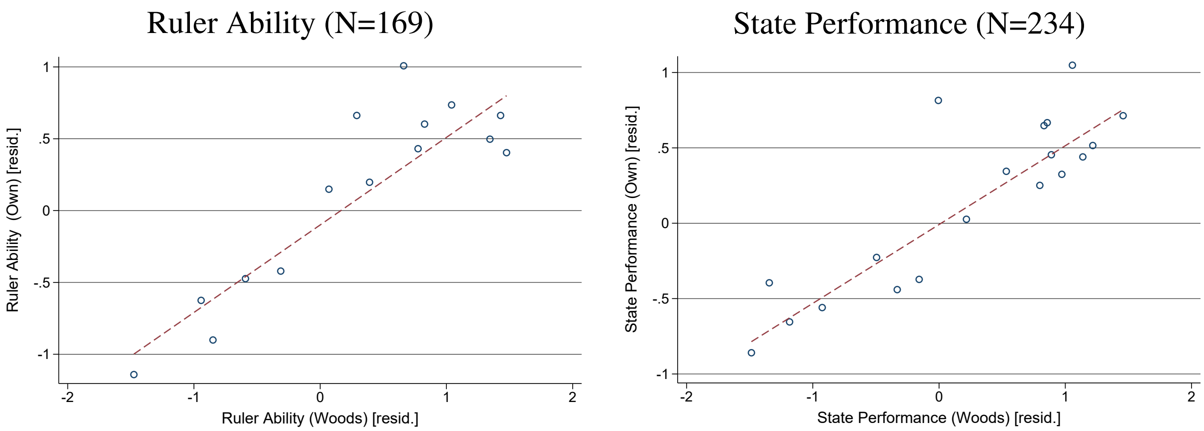

Figure A.2: Validation: Binscatters with country FE

Note: The figure shows our validation of the coding of ruler ability and state performance by Woods (1913). We code these variables during each reign possible from online encyclopedias and assess the association of our assessment with that of Woods. The left binned scatter shows residuals of state fixed effects of this association for ruler ability. The right binned scatter show this association for state performance.

A.3 Territorial Changes: Example

Figure A.3 provides an example for territorial change. It shows the change in land area of Habsburg Austria under Queen Maria Theresa (1740-1780). Austria lost territories in Silesia to Prussia, but it gained areas from Poland. Overall, Austria increased its area by 7%.

A.4 Extended Sample: Coding

In this Appendix, we extend Woods's original sample temporally until WWI and spatially to include Poland and Hungary. To do so – similar as in our validation in Appendix section A.2 – we asked a research assistant to assess the capability of rulers from all of the states covered by Woods reigning after Napoleon until World War I (or until the last monarch available or ruling until the start of World War I – for instance, the list of monarchs of France ends with Napoleon III, who ruled from 1852 to 1870).[3] In extending the time periods of the states coded by Woods, we assess

[3]In terms of procedure, we compiled a list of monarchs and asked research assistants to assess the capability of rulers and the performance of their countries on the same three-point scale as Woods (1913), using Woods' original sources as well as modern encyclopedias.

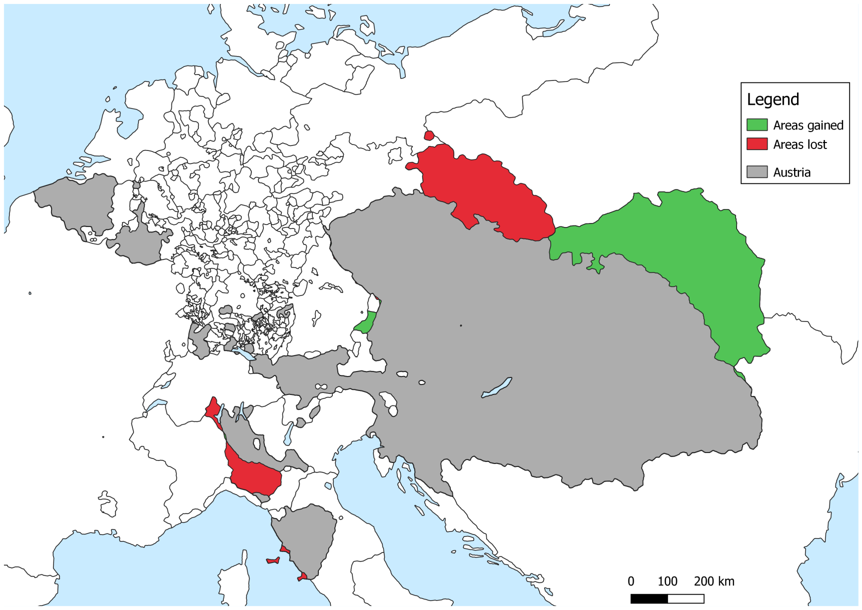

Figure A.3: Austria's Territorial Changes During the Reign of Maria Theresa

Note: The figure shows the change in land area under the control of the Austrian Habsburg from the beginning to the end of Queen Maria Theresa's reign from 1740 to 1780. The data on country borders is from Abramson (2017), and we calculate net gains of 7% during the reign of Maria Theresa.

ruler ability and state performance for 38 additional reigns, for 29 of which we are able to also obtain coefficients of inbreeding from roglo.com. In addition, we add Poland and Hungary to the dataset. Both reigns were not coded by Woods, but sufficient historical information is available for comprehensive coding. Adding Poland and Hungary (until World War I) adds 101 additional reigns, with 69 also having information on the coefficient of inbreeding.

A.5 Data and Background on Coefficient of Inbreeding

The coefficient of inbreeding measures of the degree of similarity in the genes of offspring due to common ancestors was developed by Wright (1921). It is the probability that both gene copies at any locus in an individual are identical by descent, i.e. from a common ancestor (Rédei, 2008), and is defined as follows:

$$F = \sum_{paths} (0.5)^n (1 + FA)$$

where F is the coefficient of inbreeding, $paths$ is each path through which an individual can derive identical alleles from a common ancestors of both parents, n is the number of individuals in the paths (excluding the individual itself), and $1 + FA$ is a correction factor for the inbreeding coefficient of the common ancestor in the path. The 0.5 component comes from the fact that each individual has 0.5 chance to pass on a particular allele to offspring.

Figure A.4 provides an illustrative example of the calculation of the inbreeding coefficient. A is the offspring of B and another individual. Let us assume that the parents of individual A are unrelated, so that we do not have to apply a correction factor for the common ancestor A. Lines signify blood relationship. If A were to mate with B, the offspring I would be inbred. To calculate the coefficient of inbreeding, we first note that only one common ancestor exists, B, and only one path. In this path, there are two individuals which are not I, A and B. Hence $F(I) = 0.5^2 = 0.25$. Were B inbred as well, we would have to adjust for that "ancestral" degree of inbreeding (FA).

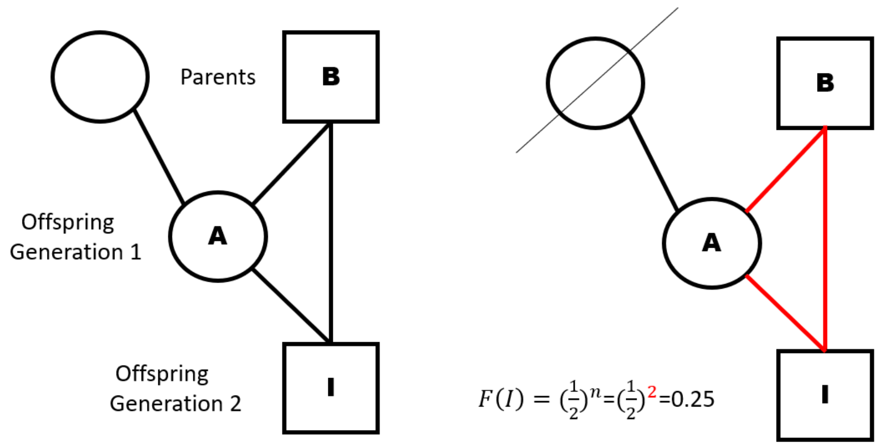

Figure A.4: Example Calculation of Inbreeding Coefficient

Note: The figures show the calculation of the coefficient of inbreeding for an parent-child offspring. Exactly one path through a common ancestor (B) of length $n = 2$ exists.

Rulers with Particularly High Coefficient of Inbreeding

Table A.2 shows the reigns with the highest coefficient of inbreeding in our sample.

A.6 Details on Calculation of Hidden Component of Inbreeding

In our main analysis, the instrument is the coefficient of inbreeding, F. As is evident from the discussion in A.5 and the pedigree of Carlos II (Figure 1), high values of F need not necessarily imply closely related parents. Instead, relationship links in temporal distance from the parents of an individual can build up over time, and account for a sizable share of the observed coefficient of inbreeding. Consider Carlos II again. With $F = 25.36$, he is the monarch with the highest coefficient of inbreeding in our data set. Yet, his parents were 'merely' uncle and niece, with most of the similarity in genes actually coming from a multitude of pathways through many distant common ancestor. The 'naive' coefficient of inbreeding of Carlos, based on his parents being uncle-niece, would be $F = 12.5$, implying that more than half of the observed F of Carlos would require knowledge of relationship links beyond that of his grandparents. We calculate a hidden

Table A.2: Rulers with the Highest Coefficient of Inbreeding

State	Ruler	Reign span	*Ruler Ability*	*State Performance*	Coefficient of inbreeding
Aragon	Charles ll	1679 - 1700	-1	-1	25.36
Aragon	Philip lll	1598 - 1621	-1	-1	21.26
Portugal	Philip lll of Spain	1598 - 1621	-1	-1	21.26
Portugal	Sebastian	1568 - 1578	-1	-1	20.25
Castile	Peter the Cruel	1350 - 1369	0	0	15.78
Aragon	Mary Anne (Regent for Charles II)	1665 - 1679	-1	-1	15.65
Austria	Leopold l	1657 - 1705	0	1	15.65
Austria	Ferdinand ll	1619 - 1637	0	-0.5	14.05
Austria	Ferdinand lll	1637 - 1657	0	0	12.07
Portugal	Philip ll of Spain	1580 - 1598	0.5	-1	11.45
Aragon	Philip ll	1556 - 1598	0.5	0.5	11.45
Aragon	Philip IV	1621 - 1665	-0.5	-1	11.32
France	Anne of Austria (Regent for Louis XIV)	1643 - 1661	0	0	11.32
Portugal	Philip IV of Spain	1621 - 1640	-1	-1	11.32
Sweden	Adolphus Frederick	1751 - 1771	-0.5	-1	10.4

Note: This table shows the 15 reigns in which the ruler had the highest coefficient of inbreeding (F) in our data set. Assessments of *State Performance* and *Ruler Ability* come from Woods (1913). The coefficient of inbreeding is calculated using http://roglo.eu/.

component of the coefficient of inbreeding by subtracting the coefficient of inbreeding implied by the closest relationship link between a rulers' parents indicated on *roglo.com*:

$$F(hidden) = F - F(naive)$$

where $F(naive)$ is 12.5 for monarchs whose parents were uncle and nieces (4 monarchs in total), and 6.25 for the (19) monarchs whose parents were (first) cousins. In the remainder we only use the hidden component as instrument for ruler ability. For Carlos, this would amount to $F(hidden) = 12.86$. Thereby, we isolate the component of the inbreeding coefficient that could be anticipated even without the advanced knowledge of calculating inbreeding coefficients and the intricate details of pedigrees.

B Additional Empirical Results

This appendix provides additional empirical results and robustness checks.

B.1 Baseline Results Using Author's Coding for Woods' Reigns

In Table A.3 we compare our baseline regressions using Woods' (1913) assessment and our own coding (as described in Appendix A.2). Column 1 repeats the baseline OLS regression (corresponding to Table 1, col 2 in the paper). Columns 2 uses our own coding of *State Performance*, combined with Woods' coding of ruler ability. Column 3 flips this specification, using Woods' coding of *State Performance* and our own coding of ruler ability. Finally, in column 4 we use our own assessments of both *State Performance* and ruler ability. For all checks in columns 2-4 we document a smaller but still sizable and highly significant association.

Table A.3: OLS Results Based on Woods' and Authors' Coding

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)
Coding of *State Performance*:	Woods	Own	Woods	Own
Coding of *Ruler Ability*:	Woods		Own	
Ruler Ability	0.618***	0.422***	0.393***	0.458***
	(0.050)	(0.068)	(0.063)	(0.056)
State FE	✓	✓	✓	✓
R^2	0.41	0.22	0.24	0.26
Observations	336	224	176	260

Note: *State Performance* is a comprehensive measure that was originally coded by Woods (1913). Columns 2 and 4 use the authors' own coding of state performance on the same scale. Similarly, the coding of ruler ability is based on Woods (1913) in cols 1 and 2, and based on the authors' assessment in cols 3 and 4. See Appendix A.2 for detail on the coding. All regressions are run at the reign level. Robust standard errors in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

B.2 Alternative Specifications, Conservative Coding, and Clustering

In this appendix section, we presents various robustness checks for different specification, for a conservative coding of ruler ability and state performance, and for different levels of clustering standard errors.

In Table A.4 we restrict attention to selected variable values as coded by Woods (1913). Specifically, in column 1 we exclude all reigns that do not indicate a clearly good or bad state performance. Excluding intermediate cases, the point estimate increases considerably. Column 2 focuses only on reigns of clearly capable or incapable rulers (i.e., a 1 or -1 coding), resulting in a point estimate that is very similar to the full sample. Column 3 restricts attention to cases where both ruler ability and state performance are required to be clearly good or clearly bad. In column 4, we exclude any reign where either variable takes the middling values of 0.5 or -0.5, and again find a very similar coefficient. For column 5, we recode all those middling values to work *against* a positive association between ruler ability and country performance.[4] Still, the coefficient remains sizable and significant.

Table A.4: Robustness: Different modifications of Woods' Coding

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)
Note:	"+" or "-" State	"+" or "-" Ruler	"+" or "-" Both	"+", "0", or "-" Both	Recoded Conservatively §
Ruler Ability	0.768***	0.631***	0.770***	0.659***	0.498***
	(0.047)	(0.051)	(0.050)	(0.047)	(0.057)
State FE	✓	✓	✓	✓	✓
R^2	0.51	0.49	0.59	0.47	0.30
Observations	245	249	204	282	336

Note: This table documents robustness of our baseline regression to the measurement of ruler ability and country performance. *State Performance* is a comprehensive measure based on the coding by Woods (1913). Column 1-4 use Woods' coding and exclude all reigns that are not rated as either clearly bad (-1)or clearly good (1). Column 4 excludes all reigns that are not rated as either clearly bad (-1), clearly good (1) or mediocre (0). All regressions are run at the reign level. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.
§ Recode all variables which are not either clearly bad (-1), clearly good (1) or mediocre (0), such that they work against the positive association of country performance and ruler ability. We recode 36 ruler abilities and 24 country performances.

Table A.5 presents further robustness checks with a focus on alternative specifications. Column 1 uses as outcome variable a dummy indicating that state performance was good, instead of using a continuous variable ranging from bad ("-1") to good ("1"). Column 2 retains this dummy outcome variable and furthermore uses dummies for each possible value of the independent variable, ruler ability, instead of a continuous version thereof. The coefficients are as one would expect. Incapable rulers are negatively associated with good state performance, while capable rulers show a positive coefficient. The middling values of ruler ability are imprecisely estimated, while the rulers that were not clearly good or bad ("0") are the omitted base level. Column 3 does justice to the categorical nature of the *Ruler Ability* variable, by estimating an ordered probit regression. As in column 2, the individual coefficients are sensible and statistically highly significant for the 'good' and 'bad' ruler categories (still using neutral ruler ability ("0") as excluded category).

[4]To do so, we reassign all the middling values of 0.5 or -0.5, where Woods was unsure to either of the closest value of 0,1, or -1. For this we consider the other variable and recode the variable to work against a positive association between both. For instance, if the ruler was coded as having low ability (-1), and the performance of the state as middling between 0 and 1, we recode state performance in this case to 1.

Table A.5: Robustness: Specification

Dependent Variable as Indicated in Table Header

Dep. Var.:	Dummy Performance †		*State Performance*
	(1)	(2)	(3)
Estimation:	OLS	OLS	Ordered Probit
Ruler Ability	0.520***		
	(0.067)		
Ruler Quality = -1		-0.147**	-0.945***
		(0.063)	(0.159)
Ruler Quality = -0.5		-0.010	-0.535*
		(0.154)	(0.276)
Ruler Quality = 0.5		0.101	0.011
		(0.150)	(0.285)
Ruler Quality = 1		0.488***	1.111***
		(0.082)	(0.182)
State FE	✓	✓	✓
R^2	0.29	0.33	
Observations	336	336	336

Note: This table documents robustness of our baseline regression to using dummy variables and probit estimation. *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Clustered standard errors (state) in parentheses. * p<0.1, ** p<0.05, *** p<0.01. † Dummy variables that takes on value 1 if Woods coded performance of state as "good," and that takes on value zero otherwise.

Table A.6 includes further fixed effects to our estimation and provides robustness to clustering of standard errors at alternative levels. In column 1 we cluster standard errors at the state level as in our baseline specification. Given the small number of clusters, we also employ the small-cluster bootstrapping technique by Roodman, Nielsen, MacKinnon, and Webb (2019). This goes back to Cameron, Gelbach, and Miller (Cameron et al.), who recommend to use cluster bootstrap-based procedures to provide asymptotic refinement to the standard test of significance (which is based on an asymptotic approximation). We report the wild bootstrap p-values in brackets below the standard errors, confirming the high level of statistical significance. In column 2, we cluster at the country, dynasty, and century level, effectively reducing variation to monarchs hailing from a dynasty. Again, size and significance of the main coefficient of interest is barely affected. Lastly, in column 3 we include fixed effects at all these levels, which further increases the size of the main coefficient.

Table A.6: Robustness: Clustering and FE

Dep. Var.: *State Performance*

	(1)	(2)	(3)
Standard errors:	Cluster: State	Cluster: State, Dynasty, Century	Cluster: State, Dynasty, Century
Ruler Ability	0.618***	0.653***	0.684***
	(0.050)	(0.048)	(0.074)
	[<0.0001]	[0.0019]	[0.0017]
State FE	✓	✓	✓
Century FE			✓
Dynasty FE			✓
R^2	0.41	0.45	0.60
Observations	336	290	290

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Wild bootstrapped p-value in brackets, following Roodman et al. (2019) to account for the relatively small number of clusters. * p<0.1, ** p<0.05, *** p<0.01.

B.3 Baseline Results by State and Time Period

Table A.7 shows how our baseline results varies by state. We interact state fixed effects with ruler ability and show these coefficients in Column 1. Column 2 further splits England into two separate entities, one before 1600 and one after 1600, and depicts coefficients for both. Only for England before 1600 do we find a sizable and significant association between ruler ability and state performance. This is visualized in Figure A.5. Columns 3 and 4 in Table A.7 probe the robustness of these coefficients by including century and dynasty fixed effects. Their inclusion renders the coefficients of Russia and Holland smaller and insignificant, but overall, our results are robust.

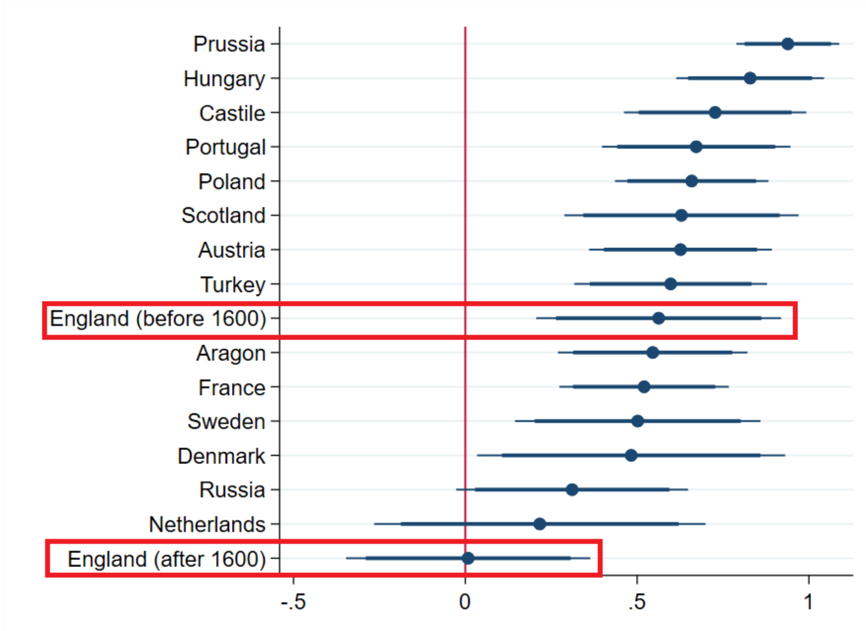

Figure A.5: Baseline OLS Results by State - England Before and After 1600

Note: The figure visualizes the estimation from column 2 in Table A.7, showing coefficients of each country and 90% confidence intervals. England is split into two separate observations, one including all reigns before 1600, and a second one including all those after 1600.

Table A.8 shows how our baseline result varies over time. Column 1 shows the coefficients of interactions of ruler ability and a dummy indicating whether the majority of a reign was before 1500, and one indicating that the majority of the reign lay in the years after 1500. Column 2 instead shows coefficients of interactions of ruler ability with a dummy indicating that the reign started in a specific time period. Lastly, column 3 shows the coefficients of interactions of ruler ability with an indicator for each century. This indicator take on value 1 whenever the majority of a reign lay in a specific century. Throughout, we document slightly smaller associations between ruler ability and state performance in later years.

B.4 OLS and IV Results in the Extended Sample

Table A.9 presents our OLS and IV results for the extended sample for our main outcome variable *State Performance*. The corresponding coding is described in Appendix A.4. Columns 1-3 show OLS estimates, columns 4 to 6 first-stage estimates, and columns 7-9 second-stage results. For comparison, columns 1, 4, and 7 repeat our results from the baseline sample. Columns 2, 5, and 8 use the sample of all states coded by Woods (1913), extended by our coding until the last monarch available or ruling until the start of World War I in 1914. The correlation between ruler ability and state performance is slightly smaller in this extended sample, as is evident from column 2. In light

Table A.7: Baseline By Country

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)
Aragon	0.603***	0.603***	0.456***	0.442***
	(0.000)	(0.000)	(0.048)	(0.054)
Austria	0.505***	0.505***	0.511***	0.525***
	(0.000)	(0.000)	(0.000)	(0.040)
Castile	0.726***	0.726***	0.789***	0.765***
	(0.000)	(0.000)	(0.027)	(0.041)
Denmark	0.374***	0.374***	0.242***	0.247***
	(0.000)	(0.000)	(0.000)	(0.049)
England	0.356***		0.559***	0.564***
	(0.000)		(0.043)	(0.060)
England (before 1600)		0.542***		
		(0.000)		
England (after 1600)		0.022***		
		(0.000)		
France	0.763***	0.763***	0.765***	0.772***
	(0.000)	(0.000)	(0.003)	(0.051)
Netherlands	0.551***	0.551***	0.548***	0.570***
	(0.000)	(0.000)	(0.081)	(0.127)
Portugal	0.763***	0.763***	0.698***	0.713***
	(0.000)	(0.000)	(0.014)	(0.014)
Prussia	0.981***	0.981***	0.981***	1.001***
	(0.000)	(0.000)	(0.000)	(0.029)
Russia	0.458***	0.458***	0.202***	0.199***
	(0.000)	(0.000)	(0.000)	(0.014)
Scotland	0.635***	0.635***	0.638***	0.627***
	(0.000)	(0.000)	(0.032)	(0.054)
Sweden	0.454***	0.454***	0.964***	0.934***
	(0.000)	(0.000)	(0.000)	(0.058)
Turkey	0.824***	0.824***	0.864***	0.920***
	(0.000)	(0.000)	(0.000)	(0.050)
State FE	✓	✓	✓	✓
Century FE				✓
Dynasty FE			✓	✓
R^2	0.43	0.44	0.57	0.58
Observations	336	336	336	336

Note: This tables documents the relationship between ruler ability and state performance by state. *State Performance* is a comprehensive measure based on the coding by Woods (1913). In column 1 we interact the baseline regression with a dummy for each state in the sample. Column 2 splits England into two observations, one for all reigns before 1600 and one for all those after 1600. All regressions are run at the reign level. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.8: Baseline By Time Period

Dep. Var.: *State Performance*

	(1)	(2)	(3)
pre1500	0.699***		
	(0.062)		
post1500	0.541***		
	(0.062)		
990-1340		0.710***	
		(0.071)	
1340-1500		0.694***	
		(0.101)	
1500-1660		0.606***	
		(0.066)	
1660-1800		0.441***	
		(0.130)	
10th century			0.930***
			(0.129)
11th century			0.635***
			(0.139)
12th century			0.581**
			(0.241)
13th century			0.868***
			(0.075)
14th century			0.606***
			(0.144)
15th century			0.590***
			(0.092)
16th century			0.489***
			(0.154)
17th century			0.524***
			(0.151)
State FE	✓	✓	✓
R^2	0.41	0.42	0.42
Observations	336	336	336

Note: This tables documents the relationship between ruler ability and state performance by broad time period. *State Performance* is a comprehensive measure based on the coding by Woods (1913). In column 1 we interact the baseline regression with a dummies indicating whether the reign began before or ater 1500. Column 2 shows coefficients of interactions with broader time periods, and column three shows the coefficient by century. All regressions are run at the reign level. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

of our results in Section 5 in the paper, this decrease in coefficient size might reflect the increase of executive constraints during this later time period. The first stage (col 5) is marginally weaker, while the second stage is actually somewhat stronger (col 8). A very similar picture emerges when we also add Poland(-Lithuania) and Hungary in columns 3, 6, and 9, respectively.

Table A.9: OLS and IV Results in the Extended Sample

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Specification:	———- OLS ———-			———- First Stage ———-			Second Stage		
Sample:	Woods	+ until WWI	+ PL, HU	Woods	+ until WWI	+ PL, HU	Woods	+ until WWI	+ PL, HU
Ruler Ability	0.618***	0.532***	0.554***				0.778***	0.885***	0.915***
	(0.050)	(0.044)	(0.042)				(0.149)	(0.101)	(0.098)
Coefficient of Inbreeding				-0.075***	-0.055***	-0.058***			
				(0.011)	(0.010)	(0.010)			
R^2	0.41	0.34	0.34	0.15	0.12	0.13	0.36	0.22	0.22
Observations	336	374	437	248	279	325	243	272	312

Note: This table shows OLS, first-stage and second-stage results for our baseline sample based on the coding by Woods (1913) in columns 1, 4, and 7, and when extending this sample until World War I (columns 2, 5, and 8), as well further including Poland and Hungary (columns 4, 6, and 9). *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

B.5 Ruler-Pair Regressions for Territorial Changes and Urban Population

Table A.10 presents the ruler-pair regressions explained in Section 4.5 in the paper, using the change in territory (cols 1 and 2) and in urban population (cols 3 and 4) as outcome variables. The table shows OLS results in Panel A, followed by IV results (including first stage and reduced-form) in Panels C-D.

Table A.10: Ruler-Pair Regressions: Territory and Urbanization

Dependent Variable as Indicated in Table Header

	(1)	(2)	(3)	(4)
Dep. Var. :	Δ_{ij} Change in Area		Δ_{ij} Change in Urban Pop.	
A. OLS				
Δ_{ij} Ruler Ability	0.233***	0.189***	0.208***	0.175***
	(0.020)	(0.023)	(0.018)	(0.022)
R^2	0.15	0.71	0.11	0.68
Observations	4,670	4,653	4,298	4,273
B. First Stage Regressions				
Δ_{ij} Coefficient of Inbreeding	-0.040***	-0.028***	-0.040***	-0.028***
	(0.005)	(0.005)	(0.005)	(0.005)
R^2	0.15	0.73	0.15	0.73
Observations	2,366	2,342	2,366	2,342
C. Second Stage Regressions				
Δ_{ij} Ruler Ability	0.203**	0.509***	0.250**	0.416**
	(0.102)	(0.170)	(0.096)	(0.202)
First Stage F-statistic	71.97	34.71	64.27	28.71
Observations	2,366	2,342	2,202	2,172
D. Reduced Form Regressions				
Δ_{ij} Coefficient of Inbreeding	-0.008*	-0.014***	-0.010***	-0.011***
	(0.004)	(0.004)	(0.004)	(0.004)
R^2	0.14	0.70	0.09	0.70
Observations	2,366	2,342	2,202	2,172
State FEs	✓	✓	✓	✓
Additional FE‡		✓		✓

Note: This table shows results from ruler-pair regressions for the outcome variables territorial change and change in urban population. For each ruler, we compute the pair-wise difference in their ability, their coefficient of inbreeding, and the respective outcome variables, to all concurrently ruling monarchs. We compare each ruler to all rulers of other states that overlapped for at least a year in their reign. All regressions are run at the reign-pair level. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4. Standard errors, clustered at the country-pair level, in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

‡ Additional fixed effects include state-pair FE, ruler FE, and state-pair × century FE.

C Potential Threats to the Exclusion Restriction: Additional Results

In this appendix, we provide additional results that complement our discussion potential concerns with our identification strategy. We begin by introducing the unanticipated ('hidden') component of inbreeding in our IV analysis. We then address a variety of possible alternative mechanisms related to path-dependence in state performance, strategic marriage (inside or outside of the kin network), or conflict.

C.1 Past State Performance and (Strategic) Kin Marriage

It would constitute a threat to our exclusion restriction if royals married kin when state performance was low, leading to a higher coefficient of inbreeding in the following generation, *and* if past low state performance reduced performance during the reign of their offspring.

As we document in Table A.11, past state performance does not predict current state performance in our reduced-form regression. Even more, we can account for dynamics of state performance in our analysis leaving our main results unaffected. Column 1 repeats our baseline reduced form.[5] Lags of state performance do not affect current state performance, as is evident from column 2.[6] Column 3 shows that the inclusion of such lags does not affect the coefficient on the current ruler's coefficient of inbreeding. Columns 4 includes the coefficients of inbreeding of earlier rulers, and column 5 further includes lags of state performance. From this, it appears that past state performance does not predict current state performance beyond the effect of the coefficient of inbreeding.

Table A.12, organized in a comparable manner, documents that past state performance further does not predict ruler ability in our first stage. Past bad state performance does not lead to significantly worse rulers. Hence, neither of the conditions required for strategic kin marriage to affect our exclusion restriction appear to be fulfilled. Therefore, including lags of state performance and lags of the coefficient of inbreeding, does also not affect our IV estimates, evident from Table A.13.

[5] Each regression uses the largest available sample for estimation.

[6] Note that "time periods" in this setting refer to reigns, which naturally vary in length.

Table A.11: Past State Performance as Confounder: Reduced Form

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)
Coefficient of Inbreeding	-0.061***		-0.057***	-0.072***	-0.066**
	(0.011)		(0.013)	(0.020)	(0.027)
L.Coefficient of Inbreeding				0.021	0.018
				(0.038)	(0.041)
L2.Coefficient of Inbreeding				-0.011	-0.011
				(0.015)	(0.016)
L.State Performance		0.034	-0.022		0.026
		(0.067)	(0.070)		(0.087)
L2.State Performance		0.003	0.013		-0.024
		(0.058)	(0.079)		(0.091)
State FE	✓	✓	✓	✓	✓
R^2	0.11	0.05	0.11	0.15	0.16
Observations	235	288	200	145	134

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Lag varies in length depending on ruler lifetime. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.12: Past State Performance as Confounder: First Stage

Dep. Var.: Ruler Ability

	(1)	(2)	(3)	(4)	(5)
Coefficient of Inbreeding	-0.076***		-0.071***	-0.088***	-0.087***
	(0.012)		(0.012)	(0.020)	(0.022)
L.Coefficient of Inbreeding				0.025	0.021
				(0.035)	(0.039)
L2.Coefficient of Inbreeding				-0.026	-0.027
				(0.017)	(0.018)
L.State Performance		-0.094	-0.070		-0.053
		(0.085)	(0.094)		(0.082)
L2.State Performance		0.086	0.057		-0.044
		(0.061)	(0.061)		(0.073)
State FE	✓	✓	✓	✓	✓
R^2	0.15	0.07	0.16	0.21	0.22
Observations	235	288	200	145	134

Note: All regressions are run at the reign level. Lag varies in length depending on ruler lifetime. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

Table A.13: Past State Performance as Confounder: Second Stage

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)
Ruler Ability	0.794***	0.806***	0.806***	0.813***	0.752***
	(0.100)	(0.147)	(0.147)	(0.155)	(0.224)
L.Coefficient of Inbreeding				0.001	0.002
				(0.020)	(0.022)
L2.Coefficient of Inbreeding				0.010	0.010
				(0.017)	(0.016)
L.State Performance		0.034	0.034		0.066
		(0.068)	(0.068)		(0.088)
L2.State Performance		-0.032	-0.032		0.009
		(0.083)	(0.083)		(0.085)
State FE	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	45.8	35.7	35.7	21.9	16.8
Observations	235	200	200	145	134

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4 in cols 1-4 and 15.4 in col 5. All regressions are run at the reign level. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

C.2 Strategic Marriage Outside of Kin Network

Alternatively, rulers might strategically marry *outside* of their dynasty network when they anticipate future territorial expansion. Marrying outside a dynasty network also potentially mechanically increases state performance in the following period by enlarging territory due to the strategic marriage. Such a mechanism could result in a link between inbreeding and country performance, as a marriage between completely unrelated individuals would give a coefficient of inbreeding of $F = 0$ in the next generation.

Note that we actually exclude monarchs with (likely) completely unrelated parents from our baseline IV analysis. For rulers without (known) family relations, our source *roglo.com* does not provide F. Yet, this does not imply that those are necessarily zero. In Column 2 of Table A.14 below, we include the 43 rulers whose parents (likely) had no relationship: our results are stronger compared to the baseline results excluding these 43 rulers, presented in Column 1, but not solely driven by these.

C.3 Wars and Conflict Among Dynasties

Benzell and Cooke (2018) find that more "blood links" between two rulers increases the probability

Table A.14: Strategic marriage outside of kin network: IV results

Dependent Variable: *State Performance*

	(1)	(2)
Sample	Baseline	Include $F = 0$
Ruler Ability	0.794***	0.814***
	(0.100)	(0.099)
State FE	✓	✓
R^2	0.39	0.41
Observations	235	278

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 23.1. All regressions are run at the reign level. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

of conflict.[7] Thus, conflict could pose a threat to our exclusion restriction. However, conflict by itself would only affect our outcome variables if wars systematically resulted in either territorial gains or losses. On average, of course, these should cancel out to a zero mean. With this in mind, success at wars arguably depends on ruler ability – which in turn leads straight back to our argument. In fact, the most likely implication of the Benzell and Cooke (2018) finding is that inbreeding adds more variability in state performance due to more frequent wars. To be concrete, suppose that more related rulers fight more often (the OLS finding in Benzell and Cooke). Also, suppose that our inbred rulers have on average more "blood links," so that they fight more often.[8] Then we would get that incapable (inbred) rulers have to fight more often – and our results suggest that they would lose more often. Thus, we would have more identifying variation in our data, but this would be ultimately driven by our mechanism of ruler ability.

Nevertheless, we also empirically address the concern that conflict may affect our results. To do so, we code a dummy for whether a ruler was involved in a conflict during his or her tenure, and include this in both stages of our IV regressions. In addition to the dummy for any conflict during a reign, we also compute the share of conflict years during each reign.[9] We perform this analysis

[7]Spolaore and Wacziarg (2016) find evidence that at the level of societies, closer relatedness also increases the probability of conflict between those.

[8]This is the more likely mapping from Benzell and Cooke's ruler-pair setting to our individual-ruler setting.

[9]The data comes from David Brecke's Conflict Catalogue (available from https://brecke.inta.gatech.edu/research/

in Table A.15. Column 1 presents our baseline results; column 2 shows that the IV coefficient barely changes when we control for conflict. This is also the case when we control for the share of years during each monarch's reign in which the state was involved in a conflict or war (column 3). In addition, the two conflict variables are themselves quantitatively small and statistically insignificant.

Table A.15: IV Results Controlling for Conflict

Dep. Var.: *State Performance*

	(1)	(2)	(3)
Ruler Ability	0.794***	0.769***	0.731***
	(0.100)	(0.093)	(0.101)
Conflict: Dummy		-0.156	
		(0.192)	
Conflict: Share Years at War			-0.201
			(0.173)
State FE	✓	✓	✓
First Stage F-statistic	42.18	35.53	25.20
Observations	235	235	235

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Table shows results from instrumental variable regressions, in which ruler ability is instrumented with the coefficient of inbreeding. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 23.1. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

We can also address a possible role of conflicts directly in our measurement of state performance. Our main measure of state performance is a composite measure, including territorial changes as one of many assessed features (others being administrative reform, economic performance, etc). This directly sidesteps the potential confounding effects of warfare. In column 1 of Table A.16 we show our baseline second stage results. In column 2 we use as outcome the residuals of a regression of the percentage change in territory under the control of a monarch during their reign from (Abramson, 2017) on our composite measure of state performance. Column 3 instead uses our measure of state performance residualized with a categorical variable of territorial expansion ("1") or decline ("-1") assessed by our research assistant.[10] In both column 2 and 3 the coefficient size is only marginally reduced; the effect of ruler ability retains statistical significance

conflict/) and starts in 900 AD. We first identify whether a state participated in any conflict (in Europe) within a given year. Then, we calculate the share of years of each reign in which a state participated in a conflict.

[10]See Appendix D.2 for detail.

and remains sizable. This underlines the importance of other aspects of state performance, beyond territorial changes.

Table A.16: IV Results using State Performance Excluding Territorial Gains

	Dep. Var.: *State Performance*		
	(1)	(2)	(3)
Note on Dep. Var.:	Baseline	Resid. wrt % territorial changes †	Res. wrt territorial changes ‡
Ruler Ability	0.794***	0.731***	0.559***
	(0.147)	(0.179)	(0.150)
State FE	✓	✓	✓
First Stage F-stat	42.15	44.50	42.15
Observations	234	200	234

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. The table shows results from IV regressions in which ruler ability is instrumented with the coefficient of inbreeding. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4. Clustered standard errors (at the state level) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

† Column 2 residualizes the dependent variable by the percentage change in area during a monarch's reign based on the borders from Abramson (2017).

‡ Column 3 residualizes the dependent variable by our own indicator of territorial change during each reign, where 1 (0,-1) indicate territorial growth (stagnation, decline).

C.4 Order within Dynasties: Founder and Descendant Effects

George and Ponattu (2018) show that dynastic politics generates a "reversal of fortune" development pattern, where places develop faster in the short run (due to "founder effects" where bequest motives increase the relevant time horizon), but are poorer in the long run, as descendant effects outweigh founder effects (i.e., intergenerationally transmitted political capital renders descendants less politically accountable). One could presume that incest was worst at the end of dynasties – at the same time when the "reversal" effect would also be strongest.

To address this concern, we code a categorical variable for the order of rulers within dynasties. For example, Carlos III is the 3rd of the Spanish Bourbons. Yet, he also hails from the Bourbon dynasty ruling France. He is the 8th of all Bourbons, ordered by the year in which his reign began. We account flexibly for the potential importance of dynasty and founder effects by including fixed effects for the order of monarchs within their dynasties. Column 1 of Table A.17 repeats our baseline IV result. Column 2 restricts attention to rulers with information on their dynasty. Column 3 includes fixed effects for all rulers of the same order within their dynasty, treating rulers hailing from the same dynasty across countries as part of different dynasties. Column 4 instead includes fixed effects that treat such rulers as hailing from the same international dynasty. In both cases, our estimates are sizable and significant. While "reversal of fortunes" development patterns resulting from founder and descendant effects are potentially capturing some the effect of ruler ability on country performance running through inbreeding, the latter is operating distinctively from these.

Table A.17: IV Regressions Accounting for Monarch's Order in Dynasty

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)
Sample	Baseline	—— Known Order in Dynasty ——		
Ruler Ability	0.794***	0.845***	0.712***	0.616**
	(0.100)	(0.116)	(0.236)	(0.257)
State FE	✓	✓	✓	✓
Order in Dynasty FE			✓	
Order in International Dynasty FE				✓
First Stage Effect. F-Stat	45.8	44.1	8.8	9.5
Observations	235	231	231	231

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. "Order in Dynasty" is the order of a monarch in their dynasty in the same state, and "Order in International Dynasty" is the order of a monarch in their dynasty, considering that certain dynasties ruled in more than one states. For example, Carlos III is the 3rd of the Spanish Bourbons. Yet, he also hails from the Bourbon dynasty ruling France. He is the 8th of all Bourbons, ordered by the year in which his reign began. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

D Details on Potential Mechanisms

In this Appendix we provide additional analyses of the potential mechanisms underlying our main finding that ruler ability had a causal effect on the performance of states. We first show that this effect likely stems from the intellectual (as opposed to physical) abilities of monarchs. Then we show that ruler ability affected both political and economic components of state performance. We then show that capable rulers were less likely to participate in conflicts, but that they nevertheless tended to increase the urban population of their territory by successful conquest. This suggests that capable rulers engaged in conflicts when these benefited their country, but avoided them otherwise.

D.1 Physical or Intellectual Ability?

Inbreeding has negative consequence for both intellectual and physical abilities. Theoretically, our results could therefore be driven by i) the (potentially anticipated) early deaths of monarchs and their lack of reproductive success due to inbreeding, or by ii) their lack of intellectual capabilities that rendered them ineffective leaders (Alvarez, Ceballos, and Quinteiro, 2009). To distinguish between these possibilities, we control for the longevity of monarchs and their number of children in our IV regressions.[11] Column 1 of Table A.18 shows our baseline IV regression comparison.

[11] We obtain this information from online encyclopedias.

In columns 2 and 3, respectively, we control for age at death of the monarchs and the number of offspring. Column 4 includes both variables simultaneously. Our main coefficient of interest, that of ruler ability, is unaffected by the inclusion of these controls. This renders it unlikely that physical features related to inbreeding driven our results. The more likely mechanism is thus the one advanced in the main body of the text: Inbred monarchs were incapable leaders because of the consequences of inbreeding for their *intellectual* abilities to effectively reign their states, and not because of inbreeding's consequences for the physical abilities to achieve longevity and produce heirs.

Table A.18: IV Results Controlling for Longevity and Number of Offspring

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)
Ruler Ability	0.794***	0.804***	0.711***	0.728***
	(0.100)	(0.091)	(0.103)	(0.096)
Age at death		0.006		0.005
		(0.006)		(0.005)
Number of children			0.020	0.013
			(0.023)	(0.019)
State FE	✓	✓	✓	✓
First-State F-statistic	42.2	45.9	44.5	46.0
R^2	0.39	0.39	0.44	0.44
Observations	235	235	225	225

Note: *State Performance* is a comprehensive measure based on the coding by Woods (1913). All regressions are run at the reign level. Table shows results from instrumental variable regressions, in which ruler ability is instrumented with the coefficient of inbreeding. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

D.2 Which Aspects of State Performance Mattered?

Our main outcome variable, *State Performance* as assessed by Woods (1913), is a composite measure. In particular, Woods covered various economic and political aspects of reigns: "finances, army, navy, commerce, agriculture, manufacture, public building, territorial changes, condition of law and order, general condition of the people as a whole, growth and decline of political liberty, and the diplomatic position of the nation, or its prestige when viewed internationally," (Woods, 1913, p. 10). While our main interest focuses on the composite assessment of state performance, we further assess the various components in order to examine which specific aspects drive our result. We asked a research assistant to read through the full text of Woods (1913), assessing each of the components. Then, we validated and extended this coding using information available in online encyclopedias. In total, we assess 14 components, which we roughly group into political aspects and economic aspects of reigns. Here we provide a brief list of each of these, and some questions that display what aspects are covered by these measures.

- Political aspects of state performance
  - Territorial changes: Did the territory of the state expand or shrink?
  - Law and order: Did the executive maintain and promote law and order in the state?
  - Public liberty: Was there persecution of minorities? Was there serfdom?
  - Finances: What was the state of treasury, royal finances, and public debt?
  - Army: How well-equipped, large, and successful was the army?
  - Navy: Did a navy exist? How was the naval force equipped?
  - Administration: Was the public administration effective, was it corrupt?
  - Diplomacy and prestige: Was the diplomacy of the state effectively implemented, was its diplomatic strategy successful? How was the state rated among other powers in Europe?
- Economic aspects of state performance
  - Living conditions of inhabitants: Did the welfare of the general populace change during a reign?
  - Infrastructure: Were roads, bridges, ports built or destroyed, or did they decay?
  - Commerce: Was there more commercial activity, trade, and growing prosperity? Or were restrictions on commerce and trade implemented?
  - Agriculture: Were there droughts, loss of farm land, or emigration of farmers?
  - Manufacture: Did the state produce and export more or less manufactures during the reign?

For all these aspects, we code negative developments as "-1" and positive ones as "1." Where we have neither information on positive nor negative developments, we presume no change and code zeros.

We discuss results for political and economic aspects separately. Table A.19 shows results of our baseline second stage regressions, where the dependent variables – instead of our composite measure *State Performance* – are our assessments of political aspects during each reign. As with in our baseline analysis, we standardize the dependent and explanatory variables to mean zero and standard deviation one. In column 1 of Table A.19, we again document a sizable effect of ruler ability on territorial change. Note, however, that this is a different measure than the one used in the main body of the paper. This measure is a categorically assessed variable based on historical sources, while the earlier one employed actual data on polity borders from Abramson (2017). We also document sizable effects of ruler ability on law and order in their states, on finances, on the effectiveness of the administration, and the diplomatic prestige of the state. The remaining

Table A.19: IV Results: Political Components of State Performance

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.:	Territorial Change	Law and Order	Public Liberty	Finances	Army	Navy	Adminis-tration	Diplomatic Prestige
Ruler Ability	0.708***	0.540***	0.324	0.618***	0.347	0.240	0.502**	0.437**
	(0.225)	(0.179)	(0.204)	(0.237)	(0.294)	(0.187)	(0.250)	(0.189)
State FE	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	45.86	45.86	45.86	45.86	45.86	45.86	45.86	45.86
Observations	240	240	240	240	240	240	240	240

Note: All regressions are run at the reign level. Table shows results from instrumental variable regressions, in which ruler ability is instrumented with the coefficient of inbreeding. The first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test is 45.9; the corresponding critical value for max. 10% relative bias is approximately 23.1. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

outcomes are also positively affected by ruler ability, but while the coefficients are sizeable, they are not statistically significant.

Next, we consider economic aspects of state performance and the causal effect of ruler ability on each of these. Table A.20 documents strong effects of ruler ability on the living conditions of a state's populace, on agriculture, and on its commerce. The remaining components also have positive signs, but the corresponding coefficients are not statistically significant.

Table A.20: IV Results: Economic Components of State Performance

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)
Dep. var.:	Living Conditions	Agri-culture	Commerce	Manu-factures	Infra-structure
Ruler Ability	0.444**	0.656**	0.990***	0.108	0.050
	(0.207)	(0.299)	(0.288)	(0.267)	(0.189)
State FE	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	45.86	45.86	45.86	45.86	45.86
Observations	240	240	240	240	240

Note: All regressions are run at the reign level. Table shows results from instrumental variable regressions, in which ruler ability is instrumented with the coefficient of inbreeding. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

D.3 Conflict as Outcome

In this section we show that incapable rulers participated in international conflict more often. The data on conflict are from David Brecke's Conflict Catalogue and start in 900 AD.[12] We identify whether a state participated in any conflict (in Europe) within a given year. Based on this information, we generate two outcome variables: a dummy for at least one conflict during a reign, and the share of years of each reign in which a state participated in a conflict.

Column 1 in Table A.21 shows that capable rulers were less likely to participate in any conflict, and their reigns also saw a smaller share of years of conflict (col 2). Is this because of less domestic unrest under capable monarchs or because capable monarchs were less likely to attack or get attacked by other monarchs (Dube and Harish, 2020)? To answer this question, we classify conflicts as internal if only one state is listed as a participant, and as a external (international) whenever more than one state is listed as participant. In columns 3 and 4 of Table A.21, we use internal conflict as our outcome variable, and in columns 5 and 6, external conflict. The results show that our previous result is driven by external conflicts: More capable leaders tended to participate in fewer conflicts involving other states, while there is no meaningful difference for internal conflicts. This is remarkable, given that more capable rulers also managed to expand their territory and urban population (see Table 5 in the paper). The most likely explanation for these findings is that – on average – capable rulers were better at selecting external wars that promised territorial expansions, while they avoided those that would likely have been costly.

Table A.21: IV Results: Conflict as Outcome

Dependent Variable: Conflict during a reign; detail in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var:	All Conflicts		Internal		External	
	Dummy	Share	Dummy	Share	Dummy	Share
Ruler Ability	-0.315***	-0.160**	-0.054	-0.034	-0.323***	-0.136**
	(0.090)	(0.071)	(0.055)	(0.073)	(0.084)	(0.056)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	40.90	40.90	40.90	40.90	40.90	40.90
Observations	240	240	240	240	240	240

Note: All regressions are run at the reign level. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

[12]The data are available at https://brecke.inta.gatech.edu/research/conflict/. Accessed in May 2021.

D.4 Decomposition of Change in Urban Population

In this Appendix, we decompose the effect of monarchs ability on the change in urban population during their reign into changes stemming from (i) the growth of cities always under control of the monarch during the entire reign, and (ii) the acquisition and loss of territory containing cities during the reign.

We start by imputing the yearly population for each of the cities in Bairoch (assuming a linear growth rate), and identify which polities these cities lay in at each 5-year intervals using the borders provided by Abramson (2017). For each reign, we then calculate the total urban population between the beginning and the end of each reign (we use urban population at the 5-year intervals at which the territory data is available). Note that such changes can result from either changes in the population of the cities that remained in the polity throughout the reign ("intensive"), or from changes in the urban population located in areas lost or gained during a reign ("extensive"). We identify the cities and their population that have always remainder under control, and those that were gained, or lost, during the reign of each monarch.

We decompose changes in total urban population into these separate components. Note that the urban population in the area controlled by a monarch at the beginning of his or her reign consists of (i) urban population in areas that will remain under the control of that monarch until the end of the reign, and (ii) the initial urban population in areas are lost during the reign:

$$Pop_t^{Urb} = Pop_t^{Urb,remain} + Pop_t^{Urb,lost}$$

where t indicates the beginning of a reign, and Pop^{Urb} stands for urban population. Similarly, urban population at the end of a reign can be decomposed into a first component which remained under control by the monarch, and a second component, comprising the urban population at the end of a reign in areas that were gained due to territorial expansion during the reign:

$$Pop_T^{Urb} = Pop_T^{Urb,remain} + Pop_T^{Urb,gained}$$

Therefore:

$$\frac{Pop_T^{Urb}}{Pop_t^{Urb}} = \frac{Pop_T^{Urb,remain} + Pop_T^{Urb,gained}}{Pop_t^{Urb,remain} + Pop_t^{Urb,lost}}$$

Let $\gamma_{gained} = \frac{Pop_T^{Urb,gained}}{Pop_T^{Urb,remain}}$ be the urban population in territories gained during the reign relative to the that in territories that remained under control during the entire reign. Similarly, denote by $\gamma_{lost} = \frac{Pop_t^{Urb,lost}}{Pop_t^{Urb,remain}}$ the fraction of urban population in the beginning of the reign in territories lost, relative to the population in areas kept. Then:

$$\frac{Pop_T^{Urb}}{Pop_t^{Urb}} = \frac{Pop_T^{Urb,remain}(1+\gamma_{gained})}{Pop_t^{Urb,remain}(1+\gamma_{lost})} = (1+\gamma_{intensive})\frac{1+\gamma_{gained}}{1+\gamma_{lost}}$$

where $1+\gamma_{intensive} = \frac{Pop_T^{Urb,remain}}{Pop_t^{Urb,remain}}$ and $\gamma_{intensive}$ is the rate of urban population growth in areas that remained under a monarchs control during the reign.

Applying logarithms, this yields a composition of percentage change in urban population into an intensive and extensive margin:

$$log(Pop_T^{Urb}) - log(Pop_t^{Urb}) = \underbrace{log(1+\gamma_{intensive})}_{\text{intensive: city growth in areas remaining under control}} + \underbrace{\underbrace{log(1+\gamma_{gained})}_{\text{acquisition of cities}} - \underbrace{log(1+\gamma_{lost})}_{\text{loss of cities}}}_{\text{extensive margin}}$$

Table A.22 shows the results for log changes in total urban population (cols 1-3) as well as for its intensive (cols 4-6) and extensive (cols 7-9) components. For each outcome, we first report the OLS results in the full sample, followed by OLS results in the "IV sample," (i.e., reigns for which we have information on the coefficient of inbreeding), followed by the IV results. Column 1 shows a sizeable correlation between ruler ability and the overall change in urban population. The result is very similar in the subsample in column 2. The IV result in column 3 shows a large and highly significant coefficient, replicating our baseline result from Table 5 in the paper that a one standard deviation increase in the ability of a monarch raises urban population by almost 20%. Interestingly, this is effect is entirely due to the extensive margin. The IV estimate for the intensive margin is minuscule with a relatively small standard error, indicated a 'reliably estimated zero.'[13] In contrast, the IV coefficient for the extensive margin in column 9 is as large as the total effect in column 3. A plausible explanation for these findings is that strong, capable rulers had an ambiguous effect on domestic city growth (i.e., on the intensive margin) because they fostered economic prosperity on the one hand, but they also kept cities' ambitions to become independent in check (c.f. Angelucci, Meraglia, and Voigtländer, 2020).

E Detail on Coding State-Year Level Constraints on Executive

Constraints on the Executive refer to legal and de-facto constraints limiting the actions of the executive branch of government. In a widely used measure, the Polity IV project provides a categorical variable measuring the relative strength of these constraints across countries from 1800 onward (Marshall, Jaggers, and Gurr, 2017). Acemoglu, Johnson, and Robinson (2005) code a similar

[13]The OLS estimate for the intensive margin full sample is statistically significant but so quantitatively small, and it loses its significance in the subsample in column 5). In addition, the OLS coefficients for the extensive margin (cols 7 and 8) are also significantly larger than those for the intensive margin.

Table A.22: Decomposition of Changes in Urban Population

Dependent Variable: Log change in urban population during reign, detail in table header

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. Var:	Total Change in Urb.			Intensive Change in Urb.			Extensive Change in Urb.		
Specification:	OLS	OLS	IV	OLS	OLS	IV	OLS	OLS	IV
Note:		IV sample			IV sample			IV sample	
Ruler Ability	0.099***	0.106**	0.191***	0.022**	0.019	0.004	0.078**	0.088**	0.187***
	(0.027)	(0.041)	(0.049)	(0.009)	(0.013)	(0.030)	(0.026)	(0.038)	(0.062)
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat			37.8			37.8			37.8
R^2	0.10	0.10		0.11	0.09		0.07	0.10	
Observations	289	198	198	289	198	198	289	198	198

Note: All regressions are run at the reign level. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is approximately 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

variable at the 100- and 50-year interval from 1000 CE until 1850. They base the measure on an encyclopedia of world history (Langer, 1972; Stearns and Langer, 2001). We follow their approach, but additionally identify the exact year when constraints on the executive (whereby we focus on the monarchs exclusively) changed. We code this measure for all states in our data set except Turkey, which is not covered by these sources. After 1800, we use the year-by-year measure of constraints on the executive from Marshall et al. (2017).

The categories of "constraints on the executive" range from 1 to 7, where 1 indicates unlimited authority of the monarch and 7 indicates "Executive Parity or Subordination" to other branches of government. We define an indicator of a monarch being constrained when constraints on the executive are above 5 – "Substantial Limitations on Executive Authority."

We list the categories below:

- 1: Unlimited Authority: There are no regular limitations on the executive's actions (as distinct from irregular limitations such as the threat or actuality of coups and assassinations.)
- 2: [Intermediate Category]
- 3: Slight to Moderate Limitation on Executive Authority: There are some real but limited restraints on the executive.
- 4: [Intermediate Category]
- 5: Substantial Limitations on Executive Authority: The executive has more effective authority than any accountability group but is subject to substantial constraints by them.
- 6: [Intermediate Category]

- 7: Executive Parity or Subordination: Accountability groups have effective authority equal to or greater than the executive in most areas of activity

References

Abramson, S. F. (2017). The Economic Origins of the Territorial State. *International Organization 71*(1), 97–130.

Acemoglu, D., S. Johnson, and J. A. Robinson (2005). The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *American Economic Review 95*(3), 546–579.

Alvarez, G., F. C. Ceballos, and C. Quinteiro (2009). The Role of Inbreeding in the Extinction of a European Royal Dynasty. *PLoS One 4*(4), e5174.

Angelucci, C., S. Meraglia, and N. Voigtländer (2020). How Merchant Towns Shaped Parliaments:From the Norman Conquest of England to the Great Reform Act.

Benzell, S. G. and K. Cooke (2018). A Network of Thrones: Kinship and Conflict in Europe, 1495-1918. Technical report, Working Paper.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. Bootstrap-Based Improvements for Inference with Clustered Errors. (90), 414–427.

Dube, O. and S. Harish (2020). Queens. *Journal of Political Economy 128*(7), 2579–2652.

George, S. E. and D. Ponattu (2018). Like Father, Like Son? The Effect of Political Dynasties on Economic Development. Technical report, Working Paper.

Langer, W. L. (1972). *An Encyclopedia of World History, 5th Ed.* Boston: Houghton Mifflin Co.

Marshall, M. G., K. Jaggers, and T. R. Gurr (2017). Polity IV Project: Political Regime Characteristics and Transitions, 1800-2016 Dataset Users' Manual. *available at https://www.systemicpeace.org/inscr/p4manualv2016.pdf*.

Montiel Olea, J. and C. Pflueger (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics 31*(3), 358–369.

Palmer, E. (2005). *Peter III.: der Prinz von Holstein*. Sutton.

Rédei, G. P. (2008). *Encyclopedia of Genetics, Genomics, Proteomics, and Informatics*. Springer Science & Business Media.

Roodman, D., M. Ø. Nielsen, J. G. MacKinnon, and M. D. Webb (2019). Fast and wild: Bootstrap inference in stata using boottest. *The Stata Journal 19*(1), 4–60.

Spolaore, E. and R. Wacziarg (2016). War and Relatedness. *Review of Economics and Statistics 98*(5), 925–939.

Stearns, P. N. and W. L. Langer (2001). *The Encyclopedia of World History*.

Woods, F. A. (1906). *Mental and Moral Heredity in Royalty: A Statistical Study in History and Psychology*. H. Holt.

Woods, F. A. (1913). *The Influence of Monarchs: Steps in a New Science of History*. Macmillan.

Wright, S. (1921). Systems of Mating. II. The Effects of Inbreeding on the Genetic Composition of a Population. *Genetics 6*(2), 124.