Supplement to

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

Sebastian Ottinger
(CERGE-EI)

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 Core Dataset and Coverage

Dataset

Our main dataset is based on a list of reigns for 13 states – all those covered by Woods (1913). We also adhere to the time period covered by Woods (1913) for each state. States can enter the dataset by splitting from former, larger states. This was the case for Sweden when it split from the Kalmar Union with Denmark in 1523.[1] Likewise, states can exit our dataset when they are taken over by other states. In our dataset, Castile became part of Aragon in 1504 (referred to as "Spain" in Woods, 1913), and Scotland became a part of England in 1625.[2] Woods (1913) provides tables on pages 305-403, listing – for each reign – the time period, the name of the ruler (or the type of reign if no ruler was in power, e.g., for interregna), 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 "+."[3] For the few cases where more than one ruler appears for the same reign,

[1]Sweden existed as an independent state before the Kalmar Union was formed in 1397, but Woods (1913) did not code this earlier period, noting that "[h]istorical material for a knowledge of this period is unfortunately very limited" (p.123) and that rulers from this period "were not members of royal families, so their study does not properly come within the scope of this work" (p. 124).

[2]In 1603, James VI of Scotland succeeded his heir-less cousin, the English queen Elizabeth I, and became king of the Union of England and Scotland as James I. In Woods' coding convention, Scotland "ends" with the death of James VI (James I of England) in 1625, as he had been the last monarch of an independent Scotland. We adhere to this convention in our baseline analysis and show that our results do not change if we instead drop all reigns potentially affected by this or similar unions between states (see Appendix B.4).

[3]In cases where Woods expressed a doubt by, say, "+ or $\pm$," we use the average (in this example, 0.5). In a robustness check in Appendix B.2, we recode all these cases conservatively so as to work against our baseline findings.

we focus on the ruler whose coding works against our baseline results.[4]

Sample Coverage

Table A.1 provides detail on the sample size. In total, 366 reigns are recorded by Woods (1913), starting with Hugh Capet's reign of France (987-996), and ending with Paul of Russia, who reigned from 1796 to 1801. For 353 of these, Woods was able to assess state performance. The others are very short reigns, or Woods could not make a definitive assessment due to 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 987, when Hugh Capet founded the Capetian dynasty), and the last state to enter is Sweden, after it split from Denmark in 1523 under Gustavus Vasa and became a separate political entity (see also Appendix footnote 1).

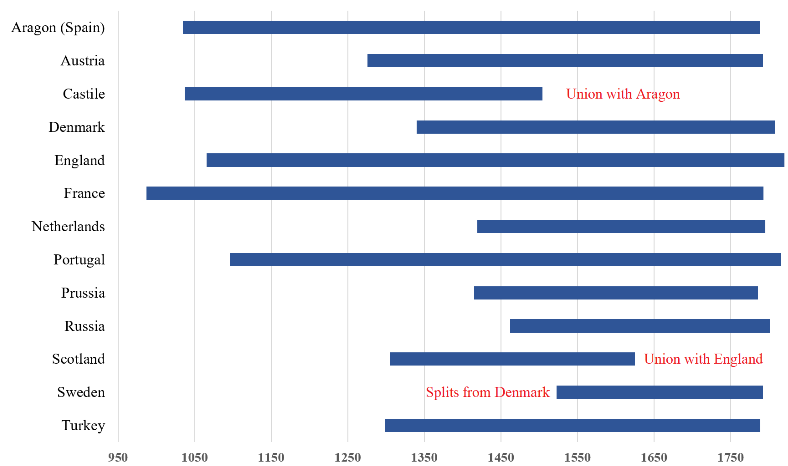

Figure A.1: Timeline of Sample Coverage: States in Sample

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

For 341 reigns, Woods assessed the ability of the ruler. He was unable to do so for instances where a reign was short or for episodes of Republican government in the Netherlands. We manually add ruler ability to three rulers for which Woods (1913) – based on his statements in the main text – presumably forgot to add assessments in the tabular form

[4]For instance, Austria was co-ruled by two regents from 1406 to 1411, Leopold and Ernest. Woods has a negative assessment of *State Performance* in this period and assesses Leopold mixed, but Ernest positively. We therefore code Ernest rather than Leopold. The one exception to this rule is the joint reign of Ferdinand and Isabella, who jointly ruled over Spain (Aragon) from 1479 to 1504. See Appendix A.5 for a detailed discussion of this case. In the cases where Woods only codes one of the contemporaneously reigning rulers, as is, for instance, the case for the reign of Albert and Otto in Austria from 1330-1359, we take the coding for Albert since none is available for Otto.

at the end of the book, where we take our data from.[5] In total, our main explanatory and outcome variable – *Ruler Ability* and *State Performance* assessed by Woods – are available for 339 reigns. Only 238 of these were reigns with an individual in power who is also listed in our genealogical data source, so that we have information on the coefficient of inbreeding. Our more 'objective' measures – territorial changes and the change in urban population within the state borders – are available for 320 and 316 reigns, respectively.[6]

Table A.1: Sample

Sample	Obs.
All reigns	366
Reigns with assessed *State Performance*	353
Reigns with assessed *Ruler Ability*	344
Reigns with information on border changes	320
Reigns with information on urban population	316
Both: *Ruler Ability* & *State Performance*	339
Both + individuals (gender assigned)	315
Both + coefficient of inbreeding (I)	238
Both: *Ruler Ability* & border changes	304
Both + coefficient of inbreeding (I)	209
Both: *Ruler Ability* & urban population	300
Both + coefficient of inbreeding (I)	208

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 Additional Variables

In addition to the variables described in Section 3 of the paper, we code other characteristics of rulers and reigns whenever this information is available. We collect this information from the English-language Wikipedia, but amend it whenever required by information from the corresponding national language Wikipedia. In addition, we refer to the Encyclopedia Britannica for verification and complementary information.

[5]For instance, Woods (1913) describes Ferdinand III, who ruled Castile from 1217 to 1252, as a "great military talent ... sincere friend of learning ... a real genius". Yet, the table records no assessment of ruler ability in his case. We manually assign a "1" in this case. Woods (1913) makes similarly unambiguous statements in the main text for Rudolph II of Austria (ruled 1576-1611, coded as "-1") and for Christian V, who ruled Denmark 1670-99 (coded as "0", as Woods remarks that he and his predecessor, Frederick III, whom he in turn coded as "0", "were not distinctive"(p.118).

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

Characteristics of Reigns

Dummy: Reign by Regents. About one-fifth of our sample (65 reigns) are instances of regents ruling. In most cases, these occurred because of so-called "minorities" of the heir to the throne: While the heir could not perform the duties due to young age, close relatives or other influential persons at the court took over as regents. Often, these regencies were divided among more than one person. We follow Woods in specifying these as separate reigns and identifying the most important among the regents whenever possible. For more than half of the regencies, either a specific regent cannot be identified, or the regent does not have any known relationship links in our genealogical data, so that these observations do not affect our IV results, since the degree of inbreeding is unknown.[7] For every reign, we identify whether the monarch in power was a regent. We then code a dummy indicating whether the title of the reign/name of the ruler contains the string "minority," "regency," or "regent," and manually verify this. For instance, while Alfonso V of Portugal (born in 1432) ruled de jure from 1438 to 1481, starting at the age of six, Woods (1913) splits this reign into three, one from 1438 to 1439 where both his mother, Leonor, and his uncle, Peter, ruled. Subsequently, Peter became the sole regent of Alfonso V's "minority" (second reign in Woods), until the latter became of age in 1449, when his actual reign started (third reign in Woods). The first two of these three reigns are classified as rule by regents.

Length of Reign. We calculate the length of each reign by subtracting each reign's start year from the end year. Both are listed in Woods and validated by us. We opt for Woods' choice in a few ambiguous cases, so as to ensure that his assessments concern the same period. We standardize the measure for comparability of coefficients with other variables.

Dummy: Regicide. For every reign in which an identifiable person is in power, we code a variable indicating regicide. This variable indicates whether the ruler died by execution after a trial or was murdered, and it is zero if the ruler exited office through natural death, an accident, or death in battle. Note that individuals executed or murdered *after* their reign ended are not coded as a regicide. In our empirical analysis we use the lagged variable for regicide, indicating whether a ruler's *predecessor* was murdered or executed.

[7]For this subset of regencies, Woods adopted the following convention: "When a regency is in non-royal hands ... I shall treat the period as if it were a period in which the monarchs were 'minus,' that is, absent or weak. If the conditions are favourable or 'plus' when the monarchical rule is absent, it counts so much against the influence of monarchs" (Woods, 1913, p. 13).

Personal Characteristics of Rulers

Dummy: Female. We assess the gender of the person in power based on the reported gender in their Wikipedia article. For reigns where no identifiable person was in power, e.g., for divided regencies for which no information on either person in power is provided, or for instances of Republican rule, this indicator is set to missing.

Dummy: Hereditary Succession. For every reign, we construct a variable indicating whether the person became the ruler due to hereditary succession. We code a dummy variable equal to one whenever a monarch is an offspring of the prior monarch in the state considered. We code cases when one sibling succeeds another as cases of hereditary succession. This was the case of Henry I of England, who came to power after his brother, William II, died childless.

Dummy: Tall. We identify individuals described as "tall" or "very tall" or taller than 1.79 meters (5'10") as tall. The information comes from online encyclopedias. Our choice of cutoff is inspired by Koepke and Baten (2005), who – estimating the height of the European population for our sample period, – state that "89.3% of the male observations fall into the range 164-178.9 cm." We assume a value of zero for all other individuals for whom no such descriptions are available. This reflects our assumption that if a ruler's body height had been noteworthy, our historical sources would have mentioned it.

Strong. We code physical appearance based on information available in online encyclopedias. The variable *Strong* takes on value "1" for those rulers described as "strong," "imposing," or along similar lines, and it takes the value "-1" for those described as "weak" or in similar terms. We assume a value of zero for all other individuals for whom no such descriptions are available.

Number of Children. We code this variable by summing up the number of all (legitimate and illegitimate) children mentioned in online encyclopedias and surviving beyond five. For instance, Charles VII of France (1422 - 1461) had 14 legitimate children, of which nine survived beyond the age of five. Louis (XI), his first-born, became the next King of France, ruling 1461 - 1483. Charles further had three (known) legitimate daughters, with one of his mistresses. We standardize the measure for comparability of coefficients with other variables.

Age at Ascension. We calculate this variable by subtracting the year of birth (collected from internet encyclopedias by us) from the first year of this persons' reign (as recorded by Woods). We standardize the measure for comparability of coefficients with other variables. We also code a dummy indicating whether a ruler had ascended to power relatively young, before the median age of 28 years.

Age at Death. We calculate this variable by subtracting the year of birth from the year of death (both collected from internet encyclopedias by us). We standardize the measure for comparability of coefficients with other variables.

Non-cognitive Ability. We draw on the adjectives provided by Woods (1913) to describe monarchs and code up a measure aimed at capturing non-cognitive abilities of monarchs. We assign values of "1" to individuals described as outgoing, emotionally stable, or using similar attributes, and a value of "-1" for those with negative non-cognitive abilities, such as neurotic or phlegmatic individuals. We assume a value of zero for all other individuals for whom no such descriptions are available. This reflects our assumption that if a ruler's non-cognitive traits had been noteworthy, our historical sources would have mentioned them. We standardize the measure for comparability of coefficients with other variables.

A.3 Summary Statistics

Table A.2 provides summary statistics for the main variable in our analysis and core sample, and Table A.3 does so for other variables used in the analysis describing characteristics of reigns and rulers.

Table A.2: Summary Statistics – Main Outcome and Explanatory Variables

A. Main Outcome Variables

	Mean	SD	Min	Max	N
State Performance	0.03	1.00	-1.30	1.04	339
$/Delta$ Log(Area)	0.07	0.41	-1.11	3.48	304
$/Delta$ Log(Urb. Pop.)	0.13	0.44	-1.91	2.56	300

B. Main Explanatory Variables

	Mean	SD	Min	Max	N
Ruler Ability	0.00	1.00	-1.14	1.14	339
Inbreeding	0.03	1.03	-0.80	5.54	238
Hidden Inbreeding	0.03	1.02	-0.97	4.42	238

Note: The table provides summary statistics for the main outcome and explanatory variables. These are for our core sample covered by Woods (1913). Ruler Ability, *State Performance*, *Inbreeding*, and *Hidden Inbreeding* are standardized.

Table A.3: Summary Statistics – Other Variables

A. Reign Characteristics						
	Mean	SD	Min	Max	N	Sources & Detail
Constrained Ruler	0.05	0.21	0.00	1.00	332	Section. 3.4
Length of reign (in years)	18.72	13.64	0.00	60.00	339	App. A.2
Dummy: Regency	0.19	0.39	0.00	1.00	339	App. A.2
Dummy: Regicide	0.11	0.32	0.00	1.00	203	App. A.2
Conflict: Dummy	0.88	0.33	0.00	1.00	339	App. C.4
Conflict: Share Years at War	0.56	0.36	0.00	1.00	339	App. C.4
B. Ruler Characteristics						
	Mean	SD	Min	Max	N	Sources & Detail
Dummy: Female	0.12	0.33	0.00	1.00	315	App. A.2
Hereditary succession (dummy)	0.77	0.42	0.00	1.00	293	App. A.2
Dummy: Tall	0.10	0.30	0.00	1.00	315	App. A.2
Strength	0.01	1.22	-2.78	2.68	315	App. A.2
Number of children	0.02	1.01	-1.25	3.79	291	App. A.2
Age at Ascension	-0.02	0.90	-2.25	3.63	308	App. A.2
Age at Death	-0.02	0.86	-2.32	1.86	308	App. A.2
Noncognitive Ability	-0.10	1.20	-1.23	1.53	315	App. A.2

Note: The table provides summary statistics for characteristics of rulers and reigns. These are for our core sample of reigns covered by Woods (1913). All non-dummy variables are standardized, except for the share of years at war.

A.4 Validation of Woods' State Performance and Ruler Ability Coding

This appendix checks and validates Woods' coding of State Performance and Ruler Ability. We are not the first to check Woods' coding. For example, Thorndike (1936) had numerous research assistants "grade" the morality and intellect of more than 300 members of the European nobility.[8] This data quality assessment resulted in correlations of the cognitive 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 confirms Woods' data (see Appendix A.4). To check Woods' (1913) coding of state performance and ruler ability, we asked our main research assistant to review the evidence in various encyclopedias and devise own assessments of ruler ability and state performance, using Woods' three-tier scale, but without using Woods as a direct source. As an additional check, we then asked other research assistants to assess a randomly selected subset of reigns. In what follows, we show that the assessments across our research assistants and Woods are highly comparable.

[8] 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 (who, like Woods, was a eugenicist) had five more research assistants independently do the same coding.

A.4.1 Validation of Woods by Main Research Assistant

The left panel of Figure A.2 provides a binned scatter plot of our main research assistants' assessment of ruler ability against that of Woods (1913). A clear assessment was possible based on online encyclopedias for 171 rulers. In 98 out of 171 assessed cases, our research assistant reached the same assessment as Woods did, while in 20 they reached the opposite assessment.[9] Among the remaining 53 cases for which our validation-coding deviates (but is not the opposite of Woods), 16 are explained by the fact that Woods assigned intermittent grades *between* -1,0, or 1. Our research assistant was not given this option and hence there cannot be exact agreement for those. This leaves 37 cases, among which 19 are instances where our research assistant assigned the monarch's ability a value 1, while Woods assigned 0; 7 cases for which our research assistant assigned a -1, while Woods assigned a 0; and 3 (8) cases for which Woods assigned a 1 (-1), while our research assistant assigned a 0. 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 assistant's assessment of *State Performance* with that of Woods (1913) ($\rho = 0.49$). Of the 233 reigns for which our research assistant was confident in making an assessment, in 123 they completely agreed with Woods' assessment. In 27 instances, they reached the opposite assessment, and in 83 instances, our research assistant and Woods disagree in their assessment of state performance, but not diametrically so. In 18 of these cases, Woods assigned intermittent values of state performance between the values of 0, 1, and -1, which were not an option for our research assistant.

A.4.2 Validation by Additional Research Assistants

We asked two additional research assistants to assess state performance and ruler ability for randomly selected subsets of reigns. One research assistant (RA) was asked to review 343 reigns of the extended sample, while a second one reviewed 48 reigns of the core sample. For these reigns, the RAs coded both ruler ability and *State Performance*. We report the pairwise correlations for RA1 and RA2 with Woods' coding and with our main RA's coding (with the respective number of reigns in parentheses). 1) For ruler ability:

[9] One such example is 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 assistant. Nevertheless, for consistency, we keep Woods' original coding. This observation does not affect our IV results because the inbreeding coefficient for Peter III is not available.

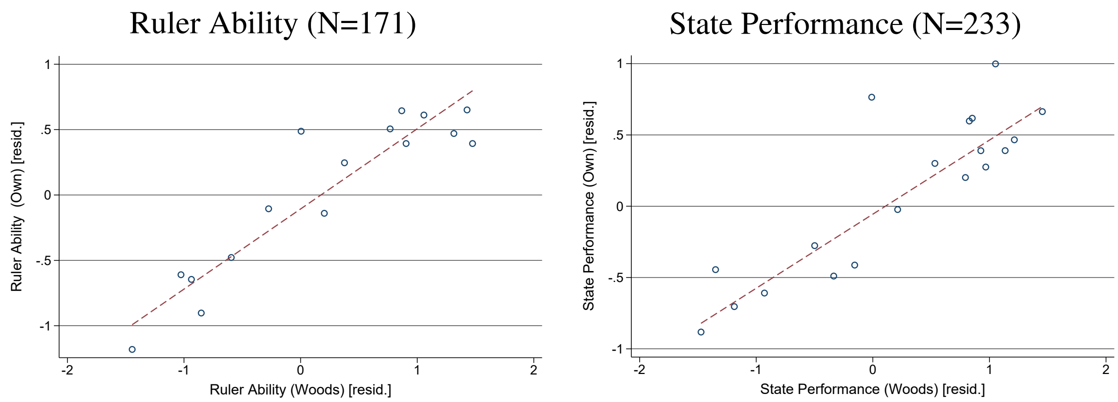

Figure A.2: Validation: Binscatters with state FE

Note: The figure shows our validation of the coding of ruler ability and *State Performance* by Woods (1913), controlling for state fixed effects (as in our baseline regressions). Each point in the binned scatter plot represents more than 10 underlying observations. The left figure is for ruler ability, the right figure is for *State Performance*.

Main RA – Woods: 0.53 (171); Main RA – RA1: 0.53 (194); Main RA – RA2: 0.69 (30); RA1 – Woods: 0.50 (227); RA1 – RA2: 0.63 (32); RA2 – Woods: 0.44 (48). 2) for *State Performance*: Main RA – Woods: 0.50 (233); Main RA – RA1: 0.55 (252); Main RA – RA2: 0.63 (41); RA1 – Woods: 0.52 (236); RA1 – RA2: 0.72 (32); RA2 – Woods: 0.60 (48). Overall, there is a high agreement on the assessment both between our research assistants as well as between each of them and Woods.

A.5 Territorial Changes: Details, Manual Adjustments, and Example

A.5.1 Methodology

The state borders provided by Abramson (2017) are available at five-year intervals, while reigns in Woods (1913) have year-specific start and end dates. We link the end year of each reign to the subsequent five-year observation, and start dates to the preceding five-year observation. We then compute the change in area during each reign as $\Delta logArea = \log(Area_{end}) - \log(Area_{start})$. Note that this methodology ensures that the first or last reign during which a state exists does not influence our results (because $\log(Area_{start})$ or $\log(Area_{end})$, respectively, are not defined for those cases with zero start or end area).

A.5.2 Manual Adjustments

The five-year frequency at which area is available is necessarily crude and might lead to wrongly assigning gains or losses of the subsequent or prior ruler to the current one. While this should not lead to systematic bias, we check all entries in our dataset and manually assign several of them a different start or end year. Furthermore, in some cases, we manually adjust the territory at the beginning or end of particular reigns. In total, 17 reigns are affected. We detail all these cases, the adjustment, and our motivation for doing so below. Table A.7 shows that just dropping these cases or using the unchanged values does not affect our core results.

France

- Louis XVI of France, who officially ruled from 1774-1793, was deposed during the French Revolution in 1792. Abramson's data ends in 1790, and since no meaningful area changes took place between 1790 and the deposition, we use the area from 1790 as the terminal area for Louis.

Sweden

- Woods begins to code the rulers of Sweden in 1523, before the state appears in Abramson in 1530. To avoid not having territory recorded at the beginning of the first reign in Sweden, we assign the earliest available year in Abramson as the start area.
- The last reign in Sweden recorded by Woods, that of Gustavus III, who ruled from 1771-1792, ends with his assassination in 1792. Abramson's data ends in 1790, and since no meaningful area changes occur between 1790 and the assassination, we use the area from 1790 as the terminal area for Gustavus.

England

- Stephen of Blois ruled England from 1135-1154. Our coding convention would assign the territory in 1155 at the end of his reign. This would give Stephen the territorial gains made by his successor Henry II in France. Therefore, we use 1150 as Stephen's terminal area.
- Elizabeth ruled England from 1558-1603. Her childlessness at death in 1603 triggered the union with Scotland under James. If we used 1605 as her terminal area, we would assign the territory increase to Elizabeth. To avoid that, we instead use 1600 as her terminal area.
- James II of England (ruled from 1685-1688) was deposed during the Glorious Revolution, and during his reign, England's territory did not change. Yet, if we were to assign the area in 1690 of Abramson as James II's terminal area, the territory of England would fall during his reign. Therefore, we instead use Abramson's area in 1685 as the terminal area for James II.

- William III of England ruled from 1689-1702. At the time of his death, the Netherlands was still under the control of England. If we were to assign the Abramson area in 1705 as England's territory at the end of William's reign, the Netherlands would already be lost for England. Hence, we instead use 1700 as the terminal area for William III.

Spain

- John II ruled Aragon from 1458-79. After his reign, Aragon was united with Castile to form Spain. If we use 1480 for his reign-end territory (the year implied by our standard procedure), we would assign the territorial increase to John II. To avoid this, we instead use Aragon's area in 1475 as the end value of the area for John's reign.
- Isabella ruled Castile from 1474-1479 and was the last ruler of independent Castile listed in Woods. Her marriage with Ferdinand united Castile with Aragon in 1479. In Abramson, no observation was available for Castile in 1480 since it was no longer independent. We assign the area of Spain in 1480, so Isabella receives credit for the union on Castile's side.[10] Without this adjustment, Isabella's reign would not be included in our core regression with territorial change as the outcome since the territory at the end of her reign would be missing. We show below that our results are robust to exluding this and similar cases of 'windfall' territorial gains due to unions.
- The rulers of unified Spain Ferdinand V (1506-1516) and Cardinal Ximenez (regent for Joanna 1516-17), would both get the 'windfall' area of Spain's union with Austria, which occurred in 1517. To avoid this, we instead use Spain's area in 1515 as the end value of the area for Ferdinand and Ximenez's reign.
- Charles V ruled Spain from 1517 to 1556. By our baseline procedure, he would have been assigned the area of Aragon in 1515, before its union with Austria. Since he inherited joint Austria-Spain, we set the initial size to the size of Spain in 1520 instead, which by then already includes Austria-Spain and avoids assigning the large territorial gain due to the union to Charles V.

[10]Note that Ferdinand's territory also increased, as his original kingdom Aragon became united Spain at the end of his reign. Thus, both Ferdinand and Isabella are credited for the territorial increase, which is motivated by the fact that both were involved in securing the dynastic union of Spain. None of our results depend on this coding choice (see Appendix B.4).

- Philip II ruled Spain from 1556-1598. By our baseline procedure, we would assign him the territory of Spain in 1555 as the starting value. This would wrongly give him the Austria-Spain empire at its largest extension, which occurred during Charles V's reign. Therefore, we use Spain in 1560 as the start area for Philip II.
- Philip III ruled Spain from 1598-1621. By our baseline procedure, he would receive the area in Abramson as of 1625 for his end year. However, this would credit him for the successful Dutch campaign that his successor, Philip IV, led. To avoid this, we assign Spain's area of 1620 as the end value for Philip III's reign.
- Carlos II ruled Aragon from 1679-1700. Spain signed the Treaties of Nijmegen in 1678 before Carlos II took over government affairs. This treaty returned areas of the Spanish Netherlands to Spain. To avoid assigning this territorial gain to Carlos II, we use the territory of Spain in 1680 (rather than 1675, as our baseline procedure would imply) as Carlos II's start area.

Habsburg Austria-Spain

- Maximilian I ruled Austria from 1493-1519. He pulled the strings to unite Austria with Spain. Abramson does not report Austria's area separately during the union with Spain. We assign the combined Austria-Spain Habsburg 1520 territory as the territory at the end of Maximilian I's reign.
- Charles V, who also ruled Spain (see above), ruled over Austria from 1519 to 1521. Since he merely inherited the union of Austria-Spain, we assign the combined Austria-Spain Habsburg 1520 territory as the territory at the start and end of his reign.
- Ferdinand I became ruler of Austria in the name of his brother, Charles V, and ruled Austria from 1521-1564. Since the combined Austria-Spain Habsburg was still in place when he became ruler of Austria, we assign the combined Austria-Spain Habsburg 1520 territory as his start territory.

A.5.3 Example

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

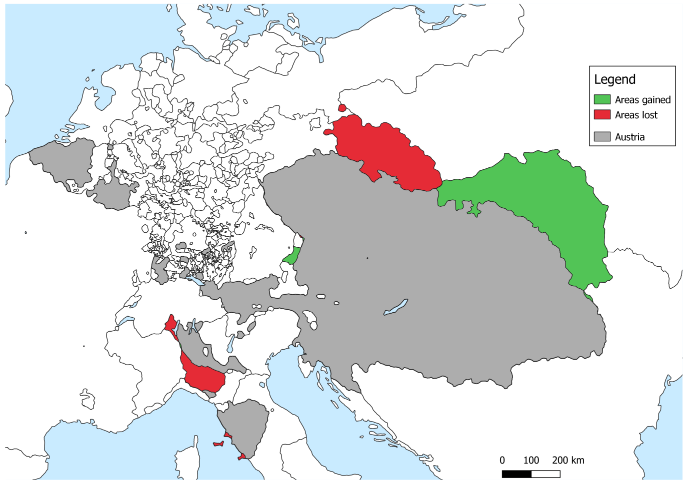

Figure A.3: Austria's Territorial Changes During the Reign of Maria Theresa
Note: The figure shows the change in land area under the control of the Austrian Habsburg from the beginning to the end of Queen Maria Theresa's reign from 1740 to 1780. The data on state borders is from Abramson (2017), and we calculate net gains of 7% during the reign of Maria Theresa.

A.6 Urban Population: Details

We use the amended Bairoch, Batou, and Chèvre data from Voigtländer and Voth (2013). These data are available at the century or half-century level. We perform the following adjustments. First, we interpolate yearly population estimates for each city, assuming linear growth rates. Second, we identify the actual coordinates of all cities. Then, we adjust the coordinates since the Abramson borders are slightly off unsystematically. We first identified the 1'391 ($\approx$ 63%) cities within 60 km of any of the Woods states. We had a research assistant manually infer the locations of the cities so that their location corresponded to borders as if those were correct. With this, we can identify the Abramson polity each city lay in at each five-year interval and their approximate population, and aggregate the total urban population of each polity at each five-year interval, to finally link it to the reign's end and start years as we do with territorial size. We link the end year of each reign to the subsequent five-year observation and start dates to the preceding five-year observation. We then compute the change in urban population during a reign as $\Delta Urb.Pop. = \log(Urb.Pop.end) - \log(Urb.Pop.start)$. For the exercise in section Appendix D.4, we also distinguish between urban growth in cities already part of the territory at the beginning of the reign and that of cities in territory lost or won.

We perform the comparable manual adjustments described in Appendix A.5 but for the urban population. For instance, John II ruled Aragon from 1458-79, just before its union with Castile. If we were to calculate urban population at the end of his reign using the year 1480 implied by your standard procedure, we would assign him the gains in urban population due to the union of Castile and Aragon. To avoid this, we instead use the urban population in the territory of Aragon in 1475 as his terminal value.

Finally, consider the example of Maria Theresa again. During her reign, the urban population of Habsburg Austria increased from about 1.998 million to 2.436 million, amounting to 19.8%. Most of this urban population gain stems from increases in areas that Maria Theresa started with and kept during her reign: the population in cities that belonged to Austria in both 1740 and 1780 boasted a population of 1.818 million in 1740 and 2.314 million in 1780.

A.7 Extended Sample: Coding

In addition to our validation of Woods' coding, we also provide robustness checks with an extended sample – both in terms of time period (until World War I) and states covered (adding Hungary, Poland, Bohemia, and Bavaria). We coded this extended sample using Woods' original sources, as well as modern encyclopedias.

Our extended sample adds Poland(-Lithuania), Hungary, Bohemia, and Bavaria until 1790 (i.e., the end of Woods' time period). For each of these states, we begin their coverage with the earliest-mentioned *hereditary* monarchs ruling the entirety of the state. These are Stephen I, who ruled Hungary as its first king from 1000 to 1038; Mieszko I, the first Christian ruler of Poland; Premislaus Ottokar I, the first hereditary king of Bohemia until his death in 1230; and Albert IV (reigned 1460 - 1508), who established primogeniture in Bavaria.[11] The four states add 73 reigns for which ruler ability and state performance could be assessed, and we are able to obtain the coefficient of inbreeding for 50 of these.

We also extend the time period for the states in Woods' original sample until 1914. To do so, we first compiled a list of rulers from all of the states covered by Woods who reigned after Napoleon until World War I – or until the last monarch available – based on Wikipedia, and then assessed the capability of rulers. For example, the list of monarchs of France ends with Napoleon III, who ruled from 1852 to 1870. We assess the ruler ability and

[11] We cover Poland from 960 to 1795, Hungary until Habsburg rule, i.e. 1526, because they are two large (and ultimately unsuccessful) states not covered by Woods. Furthermore, we collected data on Bohemia until the beginning of Habsburg rule as well (1526) and data on Bavaria until Napoleon (1777). Both of these states were similarly also ultimately unsuccessful and became part of larger polities. Bavaria became part of the German Empire in 1871. This addresses the concern that selection into our sample (by Woods) may have depended on the ultimate success of states. The fact that the inclusion of these four states does not change our coefficient of interest (see Appendix B.5) speaks (along with the state fixed effects) against this concern.

State Performance on the same three-point scale as Woods (1913), using Woods' original sources as well as modern encyclopedias. By extending the time period of the states coded by Woods, we add 38 reigns, for 29 of which we are able also to obtain coefficients of inbreeding from http://roglo.eu/.

B Additional Empirical Results

This appendix provides additional empirical results and robustness checks.

B.1 OLS Results Using Authors' Coding for Woods' Reigns

In Table A.4 we compare our baseline regressions using Woods' (1913) assessment and our own coding (as described in Appendix A.4). Column 1 repeats the baseline OLS regression (corresponding to Table 1, col 2 in the paper). Column 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 somewhat smaller, but still sizable and highly significant association.

Table A.4: 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	0.622***	0.419***	0.395***	0.462***
	(0.050)	(0.067)	(0.064)	(0.076)
State FE	✓	✓	✓	✓
R^2	0.41	0.22	0.23	0.26
Observations	339	226	175	158

Note: *State Performance* is a comprehensive measure that was originally coded by Woods (1913). Columns 2 and 4 use our 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 our assessment in cols 3 and 4. See Appendix A.4 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 OLS Results with Modified and Conservative Coding

In Table A.5 we exclude reigns for which Woods (1913) chose intermittent values of ruler ability or *State Performance*. In column 1 we exclude all reigns that do not indicate a clearly good ("+") or bad ("-") *State Performance*. Excluding these intermediate cases, the point estimate increases considerably as compared to our baseline regression in Table 1, col 2. 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 state performance.[12] Still, the coefficient remains sizable and significant.

Table A.5: 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	0.771***	0.635***	0.772***	0.663***	0.503***
	(0.046)	(0.050)	(0.049)	(0.046)	(0.056)
State FE	✓	✓	✓	✓	✓
R^2	0.52	0.50	0.59	0.47	0.30
Observations	247	251	206	285	339

Note: This table documents the robustness of our baseline OLS regression. *State Performance* is a comprehensive measure based on the coding by Woods (1913). Column 1-3 use Woods' coding and exclude all reigns that are not rated as either clearly bad ("-") or clearly good ("+"). Column 4 excludes all reigns that are not rated as either clearly bad, clearly good 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 that are not either clearly bad (-1), clearly good (1) or mediocre (0), such that they work against the positive association of state performance and ruler ability. We recode 36 entries for ruler ability and 24 for *State Performance*.

B.3 Robustness of OLS and IV Results to Different Samples

This appendix section examines the robustness of our baseline result with respect to a variety of sample restrictions. Following the baseline sample in column 1, Table A.6 introduces numerous restrictions on our sample, reporting OLS, IV, and reduced-form results in panels A-C, respectively. 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).[13] The

[12] 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 ruler ability was coded as low (-1), and the performance of the state as middling between 0 and 1, we recode state performance in this case to 1.

[13] Interregna are periods between the rule of two monarchs when no monarch was present. There are five such interregna in our core data set (e.g., the interregnum in England from 1649 – 1659, comprising the rule of

coefficients on ruler ability (Panel A and B) and inbreeding (Panel C) are actually strengthened slightly and remain highly significant. Column 3 excludes all regencies, independent of whether the regent was a dynasty member or not. All coefficients remain stable. Note that the variation explained (R^2) also 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.[14] 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 state, or who ruled in more than one state contemporaneously. In both columns 4 and 5, the coefficient remains significant and comparable in size to the baseline. Finally, column 6 presents results for the sample of rulers that came to power due to documented hereditary succession (see Appendix A.1 for detail). We confirm our baseline results both in terms of magnitude and statistical significance.

B.4 Excluding Reigns Affected by Unions and Manual Adjustments

A potential concern with our baseline analysis is that, in some cases, rulers came to inherit large territories without any direct merit of their own. Consider, for instance, James, the king of Scotland, who also became king of England after his cousin, the English Queen Elizabeth I, died. Woods coded state performance for both of James' reigns, in Scotland and England (see Appendix footnote 2). However, our area-based outcome measures do not include an observation for Scotland during James' reign: Since his reign over Scotland (as king of united England-Scotland) ended after Scotland ceased to exist in the data on polities from Abramson (2017), his reign in Scotland is missing.[15] Nevertheless, our base-

Oliver Cromwell in 1653-58). During an interregnum, Woods simply coded "interregnum" instead of a ruler name, and he categorically assigned the value -1 for the ruler's ability, with the justification that the actual monarch who *should* have been in power must have been weak for an interregnum to occur. Since we tie our hands by strictly following Woods, the five interregna do enter our baseline OLS regressions because "ruler ability" is coded by Woods. However, the interregna do not enter our IV results because we follow Woods and do not code a coefficient of inbreeding when no ruler name is listed by Woods. Regarging regencies, recall that these are coded as separate reigns (see Section 3.1). Note that in column 2, we still include cases of rule by (dynastic) relatives of the designated heir until the heir assumed office. For example, Mariana was regent for her son Carlos II of Spain until he reached adulthood.

[14] Foreign rule refers to instances when monarchs of one state temporarily ruled over another state. For example, Philipp II of Spain ruled Spain and Portugal from 1580 to 1598. 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 state (column 5), we drop both observations, his reign in Spain and that in Portugal.

[15] See Appendix A.5 on our calculation of area changes and note that if the area at the end of a reign is missing, the log difference in areas is not defined.

Table A.6: Robustness of Baseline Result: Different Samples

Dep. Var.: *State Performance*

Notes:	(1) Baseline	(2) Only Dynasty Members	(3) Exclude Regencies	(4) Exclude Foreign Rule	(5) Exclude Multi-Reign Rulers	(6) Hereditary Succession
Panel A. OLS Regressions						
Ruler Ability	0.622***	0.656***	0.681***	0.630***	0.616***	0.655***
	(0.050)	(0.057)	(0.069)	(0.051)	(0.053)	(0.055)
R^2	0.41	0.45	0.49	0.43	0.41	0.48
Observations	339	294	274	328	314	225
Panel B. IV Regressions						
Ruler Ability	0.805***	0.842***	0.828***	0.774***	0.695***	0.863***
	(0.094)	(0.110)	(0.124)	(0.071)	(0.078)	(0.126)
First Stage Effect. F-Stat	49.7	48.1	77.0	41.5	30.6	31.3
Observations	238	235	218	227	216	194
Panel C. Reduced-Form Regressions						
Inbreeding	-0.253***	-0.263***	-0.264***	-0.246***	-0.212***	-0.238***
	(0.038)	(0.040)	(0.037)	(0.037)	(0.035)	(0.038)
R^2	0.11	0.11	0.12	0.10	0.08	0.11
Observations	238	235	218	227	216	194
Fixed Effects (Panels A-C)						
State FE	✓	✓	✓	✓	✓	✓

Note: The table documents the robustness of our baseline regression (col 2 in Table 1) to using different samples. The dependent variable *State Performance* is a (standardized) comprehensive measure based on the coding by Woods (1913). *Ruler Ability* is also standardized. See Appendix B.3 for a detailed description of the sample restrictions in cols 2-6. 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.

line coding shows a territorial expansion for England during James' reign, as the Scottish territory was added. Similar concerns are conceivable for other political secessions and unions in our data set, such as Aragon and Castile.

Some concerns related to the *timing* of unions and secessions were addressed in our baseline by our manual assessment as described in Appendix A.5. Here, we show the robustness of our main IV result to these adjustments. We first describe the eight reigns affected by unions and then discuss IV results, excluding these reigns and all manually adjusted reigns.

Description of reigns affected by unions

Secession of Sweden from Denmark. Sweden split from the Kalmar Union with Denmark in 1523, leading to a decrease in the area of Denmark while putting Sweden on the map.

In the robustness check below, we exclude the reign of Gustavus Vasa from 1523 to 1560, the first ruler of Sweden recorded in Woods (see Appendix footnote 1 for detail on earlier reigns in Sweden). We also drop the corresponding reign in Denmark of Frederick l who ruled from 1523 to 1533.

Union of England and Scotland Scotland became a part of England in 1625. James ruled as king of Scotland (from 1587 to 1625) and England (from 1603 to 1625). We exclude the reign of James in both states (as described above, log area change for Scotland under James is already missing in our baseline sample; so this effectively amounts to dropping one observation).

Union of Castile, Aragon, and Austria The marriage and reign of Ferdinand and Isabella led to the union of Castile and Aragon in 1479. Their daughter Joanna's marriage to Philipp, the son of Maximilian of Austria, led to the union of all of Spain and the Habsburg lands in 1519. We account for the first of these unions by excluding the last reign of Isabella in Castile (from 1474 - 1504, before its union with Aragon), and her co-reign (with her husband Ferdinand) in Aragon from 1479 to 1504. For the second union between the united Spain and Austria, we exclude the reign of Charles V in Spain (from 1517 to 1556) and his concurrent reign in Austria from 1519 to 1521.

Results

Table A.7 presents the robustness check of our IV results for the coding of area changes (Panel A) and the corresponding urban population changes (Panel B). Column 2 excludes the seven reigns described above that gained territories due to unions with other states. Column 3 excludes all manually adjusted reigns that we described in Appendix A.5.[16] In column 4, we do not make any manual adjustments and use the area changes resulting from simply linking Woods (1913) to Abramson (2017)'s area data with the procedure described in Appendix A.5.[17] Throughout columns 2-4, our core results are essentially unaffected by

[16]While we describe 17 adjustments in Appendix A.5, the number of observations in column 3 drops by only 16 relative to our baseline. This is because the short-term regent Cardinal Ximenes (1516-17) has no recorded inbreeding. Thus, even though we make an adjustment for his reign, he was never part of our core IV sample.

[17]There are seven fewer observations in the unadjusted area data, relative to our baseline. This is because without the adjustment, several observations have no terminal observations (Isabella in Castile) or start area (Sweden). Also, we add values to some Austrian rulers who would have no observed area during the Austro-Spanish Habsburg union because Abramson does not report the area for Austria during this period (see for example Maximilian I, reported under 'Spain' in Appendix A.5). These observations are also missing in the unadjusted data in column 4.

these robustness checks.

Table A.7: IV Results: Accounting for the Unions and Manual Adjustment

Dependent variable as indicated in panel header

	(1)	(2)	(3)	(4)
Sample:	Baseline	Drop Unions	Drop manually adj.	Unadjusted
		Panel A: $\Delta log(Area)$		
Ruler Ability	0.161***	0.164***	0.148***	0.160***
	(0.046)	(0.058)	(0.046)	(0.042)
State FE	✓	✓	✓	✓
First Stage Effect. F-Stat	46.2	52.2	35.4	42.4
Observations	209	202	193	202
		Panel B: $\Delta log(UrbPop)$		
Ruler Ability	0.136***	0.130**	0.136***	0.130***
	(0.048)	(0.052)	(0.046)	(0.038)
State FE	✓	✓	✓	✓
First Stage Effect. F-Stat	43.7	49.1	33.6	40.1
Observations	208	201	192	201

Note: The table shows the results of second-stage regressions, accounting for the unions and manual adjustments. Column 1 repeats our baseline. Column 2 drops all eight reigns affected by the unions and secessions in our data set as described in this section. Column 3 drops all manual adjustments we perform to our baseline and describe in Appendix A.5. Column 4 instead uses unadjusted outcomes. 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.

B.5 Results in the Extended Sample

Table A.8 presents the results for our main outcome variable *State Performance* in extended samples. The corresponding coding is described in Appendix A.7. Panel A shows OLS and IV estimates, and Panel B presents first-stage and reduced-form estimates. For comparison, columns 1 and 4 repeat the results from the baseline sample. Columns 2 and 5 use the sample of all states coded by Woods (1913), extending the coding until WWI. In columns 3 and 6, we add Poland(-Lithuania), Hungary, Bavaria, and Bohemia (all until 1790, i.e., in line with the Woods sample). Throughout, our results are robust to extending the sample both until WWI and by including the four additional states.

B.6 Time Trends

There are no trends in our three main outcome variables over time. Figure A.4 plots the trends in each outcome variable over time. We report the coefficients (and their 90% con-

Table A.8: Results in the Extended Sample

Dependent Variable: *State Performance*

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Baseline	+ until WWI	+ 4 states	Baseline	+ until WWI	+ 4 states
	Panel A.					
	OLS			IV		
Ruler Ability	0.622***	0.534***	0.620***	0.805***	0.908***	0.772***
	(0.050)	(0.044)	(0.043)	(0.094)	(0.092)	(0.103)
R^2	0.41	0.34	0.41			
First Stage Effect. F-Stat				49.7	40.3	51.8
Observations	339	377	412	238	267	288
	Panel B.					
	First Stage			Reduced Form		
Inbreeding	-0.314***	-0.244***	-0.321***	-0.314***	-0.244***	-0.321***
	(0.045)	(0.038)	(0.045)	(0.045)	(0.038)	(0.045)
R^2	0.15	0.11	0.16	0.15	0.11	0.16
Observations	238	267	288	238	267	288
Fixed Effects (Panels A-B)						
State FE	✓	✓	✓	✓	✓	✓

Note: This table shows the results for our baseline sample based on the coding by Woods (1913) in cols 1 and 4, which covers 13 European states until 1790. Cols 2 and 5 extend this sample until World War I, and cols 3 and 6 add Poland, Hungary, Bavaria, and Bohemia to the sample until 1790. All regressions are run at the reign level. The dependent variable is standardized. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

fidence intervals) of regressing the respective variable on century fixed effects. Notice that while the territory of most states in our sample expands throughout the sample period, there are no trends for our area and population outcomes because we use *changes* in these variables over each reign. Thus, the absence of trends merely means that the growth in area and population was relatively stable over time.

Next, in Figure A.5, we examine trends in our instrument and in the main explanatory variable. The trends over time in these variables are compatible with our argument: As inbreeding built up over time, ruler ability decreased.

Because ruler ability and inbreeding show trends, we check to what extent these trends may drive the relationship between these variables (i.e., our first-stage regression). Panel A of Table A.9 shows that the first-stage coefficient remains quantitatively similar and statistically highly significant when we introduce time-trend controls: century fixed effects (for the century during which the majority of each reign occurred) in column 2; linear time trends (captured by including the first year of each reign) in column 3; quadratic time trends (also based on the first year of reign) in column 4; and even when we interact these time trends with state fixed effects, thus controlling for state-specific linear and quadratic time trends in columns 5 and 6, respectively. The remaining panels B-D in Table A.9

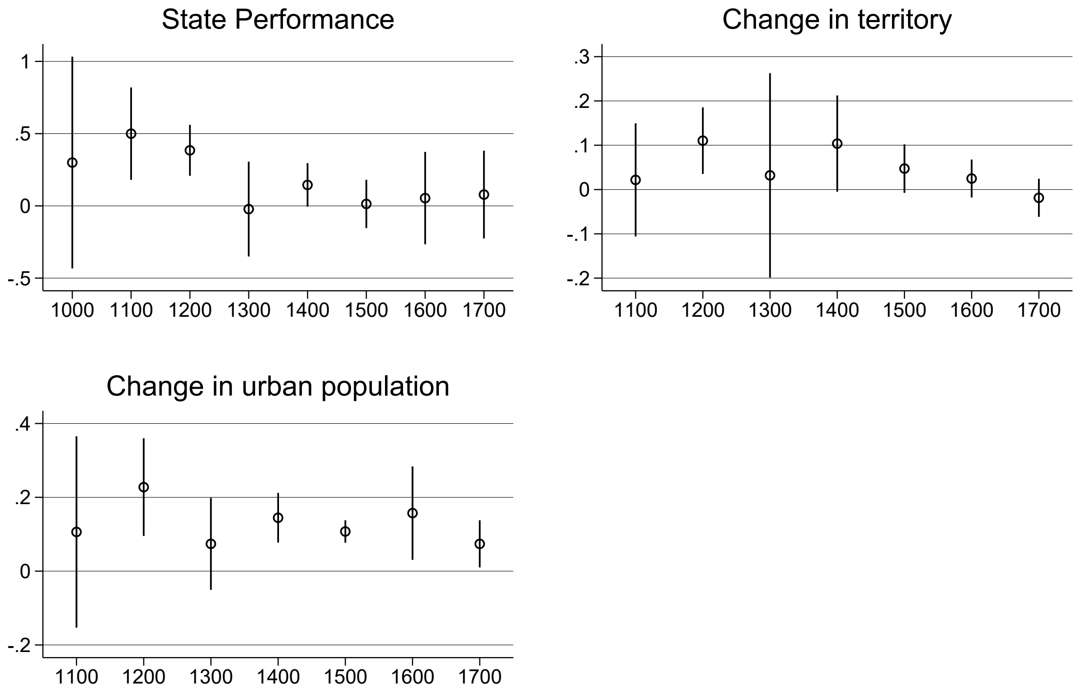

Figure A.4: Time Trends in Main Outcome Variables

Note: The figure shows that there are no time trends in our three main outcome variables. Dots represent the century-specific average of each variable, and whiskers are 90% confidence intervals.

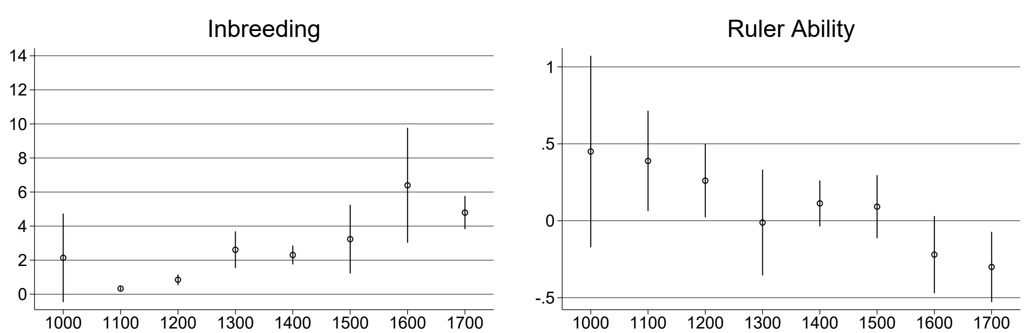

Figure A.5: Time Trends in Inbreeding and Ruler Ability

Note: The figure examines time trends our instrument (inbreeding) and in our main explanatory variable (ruler ability). Dots represent the century-specific average of each variable, and whiskers are 90% confidence intervals.

document that our OLS, IV, and reduced-form regressions for our main outcome variable *State Performance* are similarly unaffected by the inclusion of these additional controls.

B.7 Simulations with Randomly Assigned Inbreeding Coefficients

How “exceptional” are the reported coefficients for our three outcome variables? We implement a placebo exercise that randomly assigns coefficients of inbreeding from the actual inbreeding distributions to monarchs and then runs our baseline regressions, reporting reduced-form coefficients and significance levels.[18] In line with our baseline regressions, which include state fixed effects, we randomly assign inbreeding to monarchs from the state-specific inbreeding distribution. Thus, the inbreeding coefficient of Carlos II of Spain may be randomly assigned to Carlos III of Spain, but not to Henry V of England.[19] Figure A.6 shows the distributions of coefficients (left panel) and t-statistics (right panel) from 1,000 placebo regressions, for each of our three outcome variables. The solid red lines indicate our baseline reduced-form coefficients and t-statistics (i.e., with the correct assignment of inbreeding to rulers). The results shown in Figure A.6 show that our coefficients are indeed “exceptional.” For our broad outcome *State Performance*, none of the 1,000 iterations yields a coefficient that is either quantitatively larger (i.e., more negative) or statistically more significant than our baseline. For the other two outcome variables (area and population changes), only a very small percentage of all iterations yield a reduced-form coefficient that is more negative than our baseline: 2.0% for area changes and 10.7% for population changes. Regarding the t-statistics, our baseline results stand out even more clearly: The percentage of simulations that lead to smaller t-statistics than our baseline is zero for state performance, 0.5% for area changes, and 3.9% for population changes.

B.8 Possible SUTVA Violations: Adjusting for Areas Gained in Conflict

Using areas gained or lost as outcome variables introduces a “zero-sum” character that may lead to violations of SUTVA. In particular, suppose that state A, with a capable ruler, wins territory during a war with state B, which has an incapable ruler. This would lead to two observations where high ability is associated with territorial gains, and low ability, with territorial losses. Since these two observations are related, SUTVA would be violated.

[18] Note that this approach is not feasible for our IV analysis because randomly assigned inbreeding coefficients undermine the first stage. That is why we report this placebo for the reduced form.

[19] Results are very similar if we instead randomly assign inbreeding coefficients from *all* rulers (and all states) in our sample.

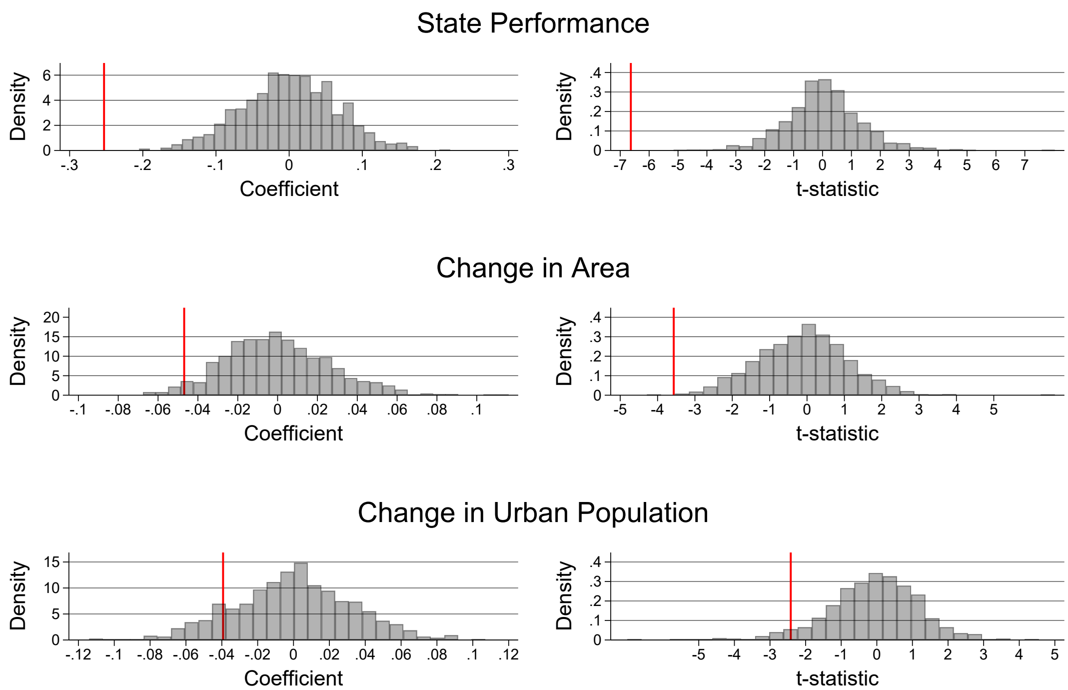

Figure A.6: Simulations with Randomly Assigned Inbreeding Coefficients

Note: The figure shows the results of 1,000 simulations that each randomly assign coefficients of inbreeding from the actual state-specific inbreeding distributions to monarchs and then run our reduced-form regressions for the three outcome variables (including state fixed effects). The solid red lines indicate our baseline reduced-form coefficients and t-statistics (i.e., with the correct assignment of inbreeding to rulers), as reported in Table 2, Panel A.

Table A.9: Accounting for Time Trends

Dep. Var.: Ruler ability (Panel A); *State Performance* (Panels B-D)

	(1)	(2)	(3)	(4)	(5)
Panel A. First Stage Regressions					
Inbreeding	-0.310***	-0.242***	-0.236***	-0.236***	-0.236***
	(0.047)	(0.060)	(0.052)	(0.052)	(0.060)
R^2	0.15	0.19	0.18	0.18	0.27
Observations	243	243	243	243	243
Panel B. OLS Regressions					
Ruler Ability	0.622***	0.623***	0.614***	0.620***	0.604***
	(0.050)	(0.040)	(0.040)	(0.039)	(0.050)
R^2	0.41	0.42	0.41	0.42	0.48
Observations	339	339	339	339	339
Panel C. IV Regressions					
Ruler Ability	0.805***	0.918***	0.863***	0.867***	0.806***
	(0.094)	(0.201)	(0.191)	(0.196)	(0.188)
First Stage Effect. F-Stat	49.7	16.8	23.0	23.0	16.5
Observations	238	238	238	238	238
Panel D. Reduced-Form Regressions					
Inbreeding	-0.253***	-0.219***	-0.204***	-0.205***	-0.192***
	(0.038)	(0.062)	(0.057)	(0.058)	(0.061)
R^2	0.11	0.13	0.12	0.12	0.25
Observations	238	238	238	238	238
Fixed Effects and Control Variables					
State FE	✓	✓	✓	✓	✓
Century FE		✓			
Linear Trend			✓	✓	✓
Quadratic Trend				✓	✓
Linear Trend × State FE					✓
Quadratic Trend × State FE					✓

Note: The table documents that our results are robust to the inclusion of century-fixed effects, linear and quadratic time trends, and their interactions with state-fixed effects. We use the century during which the majority of each reign occurred for the century-fixed effects and the start year of each reign for the time trends. The dependent variable *State Performance* is a (standardized) comprehensive measure based on the coding by Woods (1913). *Ruler Ability* is also standardized. All regressions are at the reign level. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01. 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.

In what follows, we address this concern for our outcome variable where it matters most obviously: territorial changes. We adjust for territories that may have been acquired by conflict from another state in our dataset in both our baseline regressions and in our ruler pair regressions. Below in Appendix C.4 also show that our broader outcome variable *State Performance* is robust to controlling for (or residualizing with respect to) area changes. We also note that overall, territorial changes were not zero-sum in our setting because most states were initially relatively small and expanded into territories outside of our dataset as the European map consolidated (see footnote 32 in the paper).

Adjusting for Areas Gained in Conflict: Baseline

We first implement a correction to our baseline (reign-level) regressions, excluding the component of territorial change that may have been acquired by conflict from another state. To do so, we identify from Brecke's Conflict Catalogue the participants on both sides of any conflict in our sample and period.[20] We combine the conflict data with the territory borders during each monarch's reign from Abramson (2017). This allows us to construct, for each *winner* of territories in our sample, an upper bound for the territory that this monarch may have won from other rulers in the sample. We then subtract this territory from the overall territorial gains of the winning monarch. This avoids double-counting, and the territorial changes of rulers who *lost* territory do not need to be adjusted.

To illustrate our procedure, consider Charles VII, who ruled France 1422-61 and expelled the English from France, ending the Hundred Years' War. According to Abramson's data, Charles VII oversaw a sizable territorial increase of France by around 81,000 square kilometers (23.4%). Next, we check how much of this area may have been taken from England. According to Brecke's data, Charles VII was engaged in conflict with England during his reign, specifically with Henry VI of England (1440-61). During Henry's reign, England lost about 57,000 km^2 of territory — the English possessions in France. We then adjust the gains by Charles VII of France to ignore the losses of England. That is, instead of gaining 81,000 km^2, by our adjustment, Charles VII only gained 24,000 (=81,000-57,000) square kilometers. We leave the territorial change for Henry VI of England unchanged: a

[20]The underlying data are from Peter Brecke's Conflict Catalogue, which compiles information from numerous sources and is available at https://brecke.inta.gatech.edu/research/conflict/. We use both the original Conflict Catalog data covering events from 1400 AD onward, and introduced in Brecke (1999), as well as the extension of it, going back until 900 AD for Europe.

loss of 57,000 km^2. By using this adjustment, the territorial change of England is counted only once (as a loss for England), while France receives only those territories that did not result from conflict with any other state in our dataset: the residual of 24,000 km^2.[21]

This example illustrates a simple case, as Charles VII is only recorded to have been involved in conflict with England. We generalize this adjustment for all reigns that gained territories. Formally, for each reign that resulted in territorial gains, we compute the maximum combined area that the corresponding ruler may have won from all conflict partners in our dataset. In particular, denote by $C_{s,r}$ the set of states with whom ruler r of state s had a conflict according to Brecke. Denote by $AreaLoss_j$ the area lost by conflict partner $j \in C_{s,r}$ during the time period of r's reign. We set $AreaLoss_j = 0$ if one of three conditions hold: i) If conflict partner j did not lose any land while his/her reign overlapped with r (that is, $AreaLoss_j$ is bounded below by zero so that area *gains* of conflict partner j do not imply any adjustments for ruler r); ii) if ruler r did not gain any land while his/her reign overlapped with j; iii) if ruler r did not gain any land during his/her *entire* reign. We then compute the loss of territories of all conflict partners $j \in C_{s,r}$ as $\widehat{AreaLoss}_{s,r} = \sum_{j \in C_{s,r}} AreaLoss_j$, reflecting an upper bound for the area that ruler r of state s may have conquered from all conflict partners. Finally, we subtract $\widehat{AreaLoss}_{s,r}$ from the total area gains of state s during ruler r's reign. If that gives a negative number, we set the area change to zero, i.e., we set the adjusted area at the end of r's reign equal to that of the beginning of r's reign. This ensures that our adjustment does not turn rulers who gained territories according to the raw data into losers of area. Overall, the adjustment affects 71 reigns in our data. For these reigns, the territorial change falls from 19.1% in the baseline to 9.0% in the adjusted version.

Table A.10 presents the results. For comparison, columns 1 and 2 repeat our baseline

[21]This example is convenient since the reign of Henry VI completely falls into that of Charles VII. Where this is not the case, we only use the reign of the paired ruler that overlaps with the ruler considered. Take as another example their successors Charles VIII of France (1491-98) and Henry VII of England (1485-1509). Brecke's data show France and England in conflict during Charles' reign. During the reign of Charles, the area of France grew by 5,878 km^2 (1.25%), while England's area decreased by 268 km^2 (0.18%). Yet, some of England's losses accrued in the time of Henry's reign that did not overlap with Charles. For this and similar cases, we calculate the change in area of the paired ruler during the overlap. Specifically, Henry's losses during the overlap of his reign with Charles amount to 113 km^2 (0.08%). It is this territory — rather than the 268 km^2– that was potentially won by Charles. Consequently, Charles VIII's maximum possible area gain from England during his overlap with Henry VII is 113 km^2. We subtract this from France's overall gain, so that the adjusted area gain for Charles is 5,765 (=5,878-113) square kilometers, or 1.23%. We proceed in the same fashion for all other rulers with partial overlap.

IV and reduced-form coefficients for area growth (see Table 2 in the paper). Columns 3 and 4 use the adjusted area growth, as described above. This adjustment increases the size of the IV coefficient slightly, and the estimate remains statistically highly significant. The reduced-form results (columns 4) are also very similar to our baseline.

Table A.10: Baseline Regressions: Adjusting for Areas Gained in Conflict

Dependent Variable: $\Delta log(Area)$ during reign

	(1)	(2)	(3)	(4)
Dep. Var:	Baseline		Adj. for areas gained in conflict	
Specification:	IV	RF	IV	RF
Ruler Ability	0.161***		0.182***	
	(0.046)		(0.067)	
Inbreeding		-0.047***		-0.053**
		(0.013)		(0.020)
State FE	✓	✓	✓	✓
First Stage Effect. F-Stat	46.2		46.2	
R^2	0.13	0.10	0.12	0.13
Observations	209	209	209	209

Note: The table shows that our results are not driven by double-counting areas gained from other states in conflict. Columns 1 and 3 report IV estimates, cols 2 and 4 reduced form (RF) estimates. In columns 3 and 4, we exclude area gains that may have been won from contemporaneous other rulers in the sample. See the description in Appendix B.8 for 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 approximately 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

Adjusting for Areas Gained in Conflict: Ruler-Pair Regressions

Second, we apply our adjustment to the context where the 'zero-sum aspect' matters most obviously: the ruler-pair regressions, where one ruler's territorial gain may have been the paired ruler's loss. For each ruler pair, we check whether they were involved in conflict and then adjust the gains of the ruler who won territory during his/her reign. The example highlighted above — Charles VII of France and Henry VI of England — represents a ruler pair (i.e., rulers whose reigns overlapped by at least 5 years). In the ruler pair regressions for territory, the outcome variable is the difference in the percentage change of each of their territories. We compute the adjusted difference by using the conflict-adjustment for the winner, as described above: The winner of the conflict, Charles VII, only gained 24,000 km^2 (instead of 81,000 without the adjustment), which implies an increase in the French area by 7.7% (as opposed to 23.4% without the adjustment). England, in turn, lost 57,000 km^2 of its area, or 32.1%. Thus, the dependent variable for the ruler pair Charles VII and

Henry VI is 7.7% - (-32.1%) = 39.8% (instead of 55.5% without this adjustment). For more complex conflicts with several powers, we consider a ruler pair in conflict if both are involved. We then apply the correction to the winner of territories in any pairwise comparison if the paired ruler lost any territory. If both states of a pair were net winners of territory during the respective reigns, we make no adjustment.

Table A.11 shows the results for our ruler-pair regressions with adjusted area changes. We use the same set of fixed effects as in our main ruler-pair regressions in Table 4 in the paper. In the most stringent set of fixed effects (even columns) we only compare rulers within the same state pair (e.g., Prussia-France) over time, on top of the fixed effects for each state, paired state, century, and state-century. Both the IV results (cols 1 and 2) and the reduced form (cols 3 and 4) are very similar to our baseline results for ruler pairs reported in Table 4 (cols 4 and 5 in Panels C and D) in the paper.

Table A.11: Ruler-Pair Regressions: Adjusting for Areas Gained in Conflict

Dep. var.: $\Delta log(Area)$ adjusted for conflict

	(1)	(2)	(3)	(4)
Specification:	IV Results		Reduced Form	
Δ_{ij} Ruler Ability	0.130**	0.164**		
	(0.055)	(0.073)		
Δ_{ij} Coefficient of Inbreeding			-0.033*	-0.036**
			(0.016)	(0.016)
First Stage Effect. F-Stat	82.50	38.96		
Observations	1,970	1,966	1,970	1,966
Fixed Effects				
State	✓	✓	✓	✓
Paired state	✓	✓	✓	✓
Century	✓	✓	✓	✓
State-pair		✓		✓
State × Century		✓		✓

Note: The table presents ruler-pair level results for area changes that are adjusted for conflict: For each ruler pair that had a conflict, we exclude the area gains that may have been won by one ruler from the other. More detail is provided in the text. 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 are multi-way clustered at the state, paired state, ruler, paired ruler, and ruler-pair level. * p<0.1, ** p<0.05, *** p<0.01.

B.9 Results with Annual GDP Data

In this appendix section, we use growth in GDP (both per capita and total) as an outcome variable for the six states in our sample with high-frequency GDP data before 1800. Those are England, France, Spain, the Netherlands, Portugal, and Sweden. Given that per-capita GDP varied relatively little before the Industrial Revolution, we view this exercise as mostly exploratory. However, this exercise provides a useful complement for those states whose borders changed relatively little over time (England, Spain, France, and Portugal).[22] We also use historical population data for the six states to compute the growth in total GDP. This accounts for the fact that in Malthusian economies, productivity increases translated predominantly into population growth. Importantly, the fact that GDP data are available annually allows us to include detailed time fixed effects. In what follows, we first describe the GDP data and then present annual regression results based on these data.

Historical GDP Data and Short-Run Variation

We use data on GDP per capita from Bolt and van Zanden (2020), who compile the estimates of several underlying papers.[23] These works generally employ two approaches to arrive at yearly GDP estimates. For states and periods when sufficient data exist (as is the case for England and Holland), source data on output in the different sectors is combined with price indices to construct yearly GDP estimates. When such data for an "output-based" approach is not available, a "demand-side" approach is used, for which annual wages form the backbone, along with much additional data at a lower frequency and various assumptions. Fouquet and Broadberry (2015) and de Jong and Palma (2018) provide more details on these approaches. Generally, estimates based on both approaches yield reasonable short-term variation in economic performance. For instance, Broadberry, Campbell, Klein, Overton, Leeuwen, et al. (2022) use these data to identify business cycles in England and find that business cycles "typically averaged 3 to 4 years from peak to peak" and that the short-term variation corresponds well with the historical evidence on "extreme weather, war, commercial dislocation or problems of money supply," and crucially, the Black Death. Meaningful short-term variation is also evident in the annual reconstruction of other European states' GDP data, as exploited by Broadberry and Lennard (2023). We impute yearly values of GDP p.c. for Spain, which is only available at ten-year intervals, and we match

[22] We are grateful to Steve Broadberry for suggesting this exercise.

[23] Specifically, Broadberry, Campbell, Klein, Overton, and Van Leeuwen (2015) for Britain, Ridolfi (2017) for France, Alvarez-Nogal and De La Escosura (2013) for Spain, Van Zanden and Van Leeuwen (2012) for Holland, Palma and Reis (2019) for Portugal, and Schön and Krantz (2015) as well as Krantz (2017) for Sweden.

the data on GDP p.c. of Spain to Aragon.[24] For the Netherlands in our dataset, we match GDP p.c. data for Holland, which is the largest province of the Netherlands.[25] For England, GDP p.c. data is available until 1700 and then onward for the United Kingdom. We match the data for both periods to England in our dataset and include a post-1700 dummy for England.

The data sources mentioned above only provide per-capita GDP. In order to construct estimates of aggregate country-wide GDP, we need annual population data. We build on population data from McEvedy and Jones (1978), including the corrections from Fenske (2013).[26] The population data is available at century intervals from 1000 to 1600 CE and at 50-year intervals from 1600 until 1800, except for Sweden, where it is available at century intervals from 1500 to 1800. We linearly interpolate between these intervals, assuming a constant annual population growth rate, with the exception of the period 1300-1400, which saw the Black Death. To account for the massive population shock in 1348-50, we implement the following correction: We assume that each country's population grew from 1301 to 1349 with the growth rate of the prior century (1200-1300) and assume a drop in population levels from 1349 to 1350 by 30%. Finally, we assume that each country's population grew at a constant rate from 1350 to 1400 onwards, consistent with the level implied by our prior assumptions in 1350 and estimate for 1400 in the underlying data.

Figure A.7 shows the per-capita GDP data before 1800 for the six states with high-frequency data. For England, the structural break of the Industrial Revolution around 1700 is clearly visible. To account for this, all our regressions include a post-1700 fixed effect for England (in addition to the various other fixed effects discussed below). For the remaining states, there are no clear long-term trends.[27] Importantly, the data does exhibit short-run

[24] Alternatively, matching Spain to both Aragon and Castile or only including Aragon from 1479 onward (once Aragon and Castile united) gives very similar coefficients with the same statistical significance.

[25] This probably overestimates GDP p.c. for the Netherlands, as Holland was the "most urbanized, dynamic and richest province of the Dutch Republic" Van Zanden and Van Leeuwen (2012). Nevertheless, over the long period of our analysis, and given that we use state and state-century fixed effects, the approximate use of Holland for the Netherlands is reasonable: Van Zanden and Van Leeuwen (2012) highlight that "[g]rowth in Holland was probably faster between 1500 and 1650, but after 1650 the 'periphery' of the country started to catch up, and it was probably more dynamic in the late 17th and 18th centuries, as a result of which the overall performance between 1500 and 1800 may have been very similar..." (Van Zanden and Van Leeuwen, 2012, footnote 9). Excluding the Netherlands from the analysis below yields comparable results with somewhat smaller but still highly significant estimates (available upon request).

[26] This data is available at https://warwick.ac.uk/fac/soc/economics/staff/jefenske/data/.

[27] There are some episodes of growth, such as the Netherlands in the 16th century. We account for these below by including century × state fixed effects.

variation, which is crucial for our identification of ruler effects.

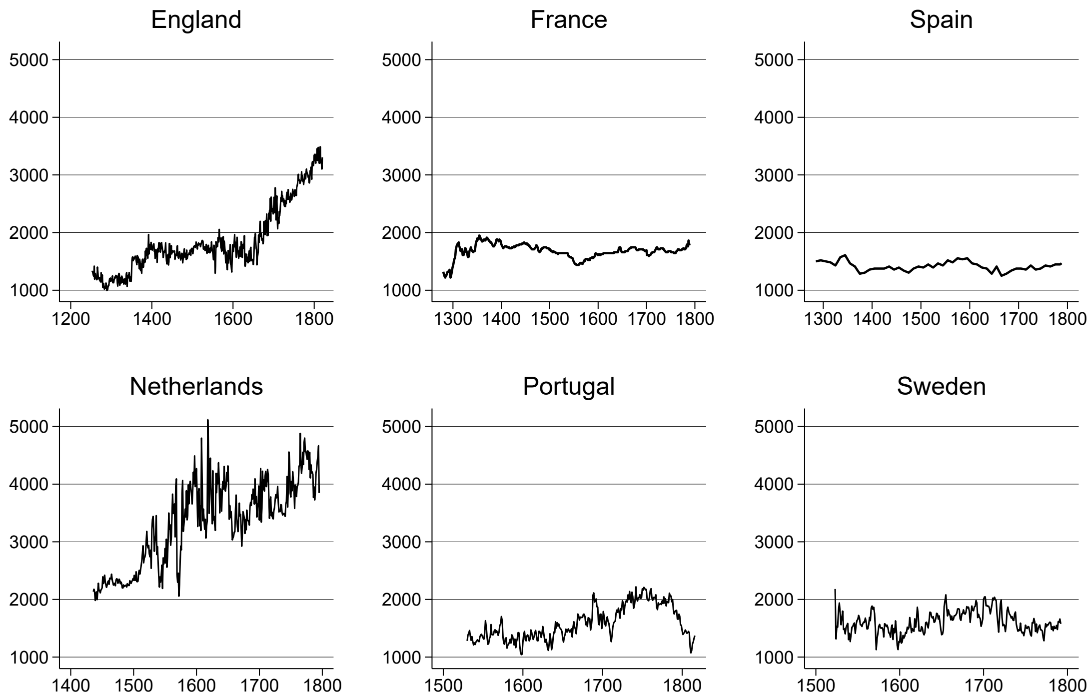

Figure A.7: GDP per Capita for States with High-Frequency Data before 1800

Note: The figure shows the time series of GDP per capita (in 2011 international dollars) for the six states in out sample with available high-frequency data before 1800. See the text for data sources and description.

Regression Results with GDP Data

Our annual GDP regressions take the following form:

$$GDPgrowth_{s,t} = \beta RA_{s,t} + \delta_s + \gamma_{[t]} + \varepsilon_{s,t} \ , \tag{A.1}$$

where the outcome variable is measured as 100 times the linear growth rate in GDP (per capita) and RA denotes ruler ability. Thus the coefficients of interest, β, can be interpreted as a percentage point change in growth due to a one-standard deviation change in ruler ability (or inbreeding).[28] In our IV results, we instrument the ability of the ruler of state s in year t, $RA_{s,t}$, with the corresponding inbreeding coefficient $Inbreeding_{s,t}$. Since each of these two variables takes on the same value for all years t during which a ruler r was in

[28]Recall that we standardize the explanatory variables. For GDP growth, we compute the first growth rate of each ruler between the year when s/he came to power and the following year. For example, if a ruler took over in 1508, the first growth rate is for 1508-09. Using instead the previous year, or using log instead of linear growth rates, gives very similar results.

power of state s, we cluster standard errors at the level of states and rulers. We also include fixed effects for each state (δ_s) and for different time periods – from centuries to individual years, denoted by $\gamma_{[t]}$. In addition, all regressions include a dummy for England after 1700, as motivated above.

Table A.12 presents our results. Columns 1-3 show the regressions for annual per-capita GDP growth, and columns 4-6, for annual growth in total GDP. The IV results in Panel A, column 1 (including century fixed effects), indicate that a one-standard deviation (std) increase in ruler ability led to a statistically highly significant increase in GDP per-capita growth by 0.21 percent per year. This effect is also economically meaningful, given the average of annual growth rate of 0.22 percent for the six countries over our sample period. In column 2, we add year fixed effects. Importantly, if anything, this makes our results even stronger, yielding a coefficient of 0.33 that continues to be statistically significant at the 1% level. In column 3 we add state $\times$ century fixed effects, capturing long-run trends. Note that this is a conservative specification, as such trends are not strongly pronounced in the GDP data (see Figure A.7), with the exception of England after 1700, which we already accounted for. The IV coefficient in this specification is even larger and remains highly significant, although it needs to be interpreted with caution, as the effective first stage F-statistic falls just below 10.

Columns 4-6 in Table A.12 repeat these regressions for total (instead of per-capita) GDP growth. This allows for the Malthusian mechanism, whereby higher productivity translates into population growth, while per-capita incomes may stagnate. Coherent with this mechanism, the IV coefficients are somewhat larger in magnitude: A one-std increase in ruler ability raises total GDP growth rates by about 0.35-0.5 percent per year, relative to a mean of 0.45 percent.

Finally, Panel B in Table A.12 presents the corresponding reduced-form results. These imply that an increase in ruler inbreeding by one std reduced per-capita GDP growth by about 0.1 percent, and total GDP growth by about 0.15 percent.

B.10 Heterogeneity of Results

This section examines whether the effect of ruler ability on state performance varied over time, or across ruler- (or reign-) specific characteristics.

Non-parametric results. To what extent do our results depend on where we 'draw the line'

Table A.12: Annual GDP Regressions

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	*Annual GDP p.c. Growth*			*Annual GDP Growth*		
	Panel A. IV Regressions					
Ruler Ability	0.210***	0.326***	0.561***	0.366***	0.477***	0.776**
	(0.063)	(0.082)	(0.215)	(0.095)	(0.107)	(0.365)
Mean Dep. Var.	0.220	0.220	0.220	0.447	0.447	0.447
First Stage F-Stat	31.6	21.4	8.3	31.6	21.4	8.3
Observations	2220	2220	2220	2220	2220	2220
	Panel B. Reduced-Form Regressions					
Inbreeding	-0.072**	-0.114**	-0.159*	-0.126**	-0.167***	-0.221**
	(0.025)	(0.034)	(0.064)	(0.033)	(0.037)	(0.078)
R^2	0.003	0.196	0.198	0.006	0.220	0.224
Observations	2220	2173	2173	2220	2173	2173
Fixed Effects (Panels A-B)						
State	✓	✓	✓	✓	✓	✓
England post-1700	✓	✓	✓	✓	✓	✓
Century	✓			✓		
Year		✓	✓		✓	✓
Century x State			✓			✓

Note: The table shows regressions with GDP and GDP per capita at the yearly level as outcome variables. Data on GDP per capita comes from the Maddison Project and is available at the yearly level for five states in our core sample (England, France, the Netherlands, Portugal and Sweden). For Aragon, we interpolate the 10-year GDP estimates. We construct total GDP using population data from McEvedy and Jones (1978), including the corrections from Fenske (2013), as described in the text. The table reports the results of IV regressions (Panel A) and reduced-form regressions (Panel B) at the year-state level, where the main explanatory variable is the ability of the ruler in charge during each year, instrumented for with his or her inbreeding coefficient. 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% (15%) relative bias is 16.4 (9.0). Standard errors, clustered at the level of states and rulers, in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

between high versus low inbreeding? We define the treatment variable as inbreeding $I > X$, where X represents the respective deciles of the inbreeding distribution.[29] Consider, for example, the 8th decile: All rulers with inbreeding above 4.87 are considered as 'treated,' (with the average inbreeding of 'treated' rulers above the 8th decile being 9.59). We then run a separate regression for each of the 10 variants of the treatment variable. The left panel of Figure A.8 shows the first stage, and the right panel, the reduced form results for our broad *State Performance* outcome variable. We obtain significantly negative coefficients for both specifications from the 7th decile onwards, corresponding to an inbreeding range above 3.49 (with the average inbreeding of 'treated' rulers above the 7th decile being 7.74). Thus, this highly flexible specification yields results that are in line with the biology and medical literatures – which emphasize serious negative impacts when inbreeding is above the level of first cousins (e.g., Fareed and Afzal, 2014).

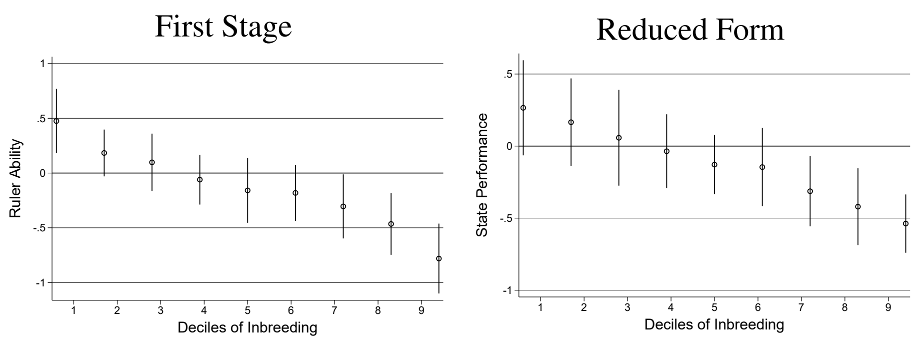

Figure A.8: Moving the Threshold for "High" vs. "Low" Inbreeding

Note: The figure shows the results of non-parametric estimates of our first-stage regression (left panel) and the reduced-form with *State Performance* as outcome variable (right panel). In both specifications, the significantly negative effects of inbreeding are obtained for rulers from the 7th decile of inbreeding onwards. The figure also shows 90% confidence intervals (based on standard errors clustered at the state level).

Heterogeneity over time. Did the relationship between monarchs' ability and state performance change over time? Figure A.9 depicts the OLS (left panel) and IV (right panel) coefficients on ruler ability for different time periods (we split the sample into four periods of about 150 years length each). We find statistically highly significant coefficients throughout. After 1650, the coefficient size decreases relative to the period 1500-1650. This period coincides with the rise of parliaments in Western Europe (van Zanden, Buringh, and Bosker, 2012). Thus, the timing in Figure A.9 is consistent with our results in

[29]For example, for "X = first decile," all rulers with inbreeding above 0.13 are defined as "inbred" (i.e., "treated"); the 5th decile has a value of 1.96. The inbreeding of first-degree cousins (value of 6.25) falls into the 8th decile, and 12.5 is in the 9th decile.

Section 5, where we show that stronger constraints on the executive diminished the importance of ruler ability for their states' performance.

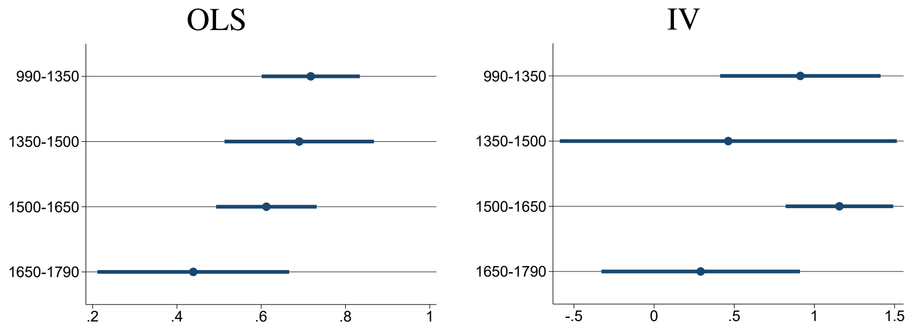

Figure A.9: Relationship between Ruler Ability and State Performance by Period

Note: The figure shows the coefficients on ruler ability in a regression with *State Performance* as the dependent variable, for OLS (left panel) and IV (right panel) by time period. Underlying each panel, we run a joint estimation that includes state fixed effects. The figure also shows 90% confidence intervals (based on standard errors clustered at the state level).

Heterogeneity by ruler- and reign-characteristics. In Table A.13, we include interaction terms between ruler ability and several characteristics. We collect the variables used in this section from encyclopedias and biographies, as explained in detail in Appendix A.1. We present IV results in Panel A, as well as reduced-form results in Panel B.[30] Columns 1-3 examine ruler-specific variables that were arguably unrelated to a given monarch's ability, but may nevertheless have interacted with his/her effectiveness as a ruler. For column 1, we define a dummy indicating whether a monarch was female, which was the case for 28 of the 238 reigns in our IV sample. The interaction term is quantitatively small, negative, and statistically insignificant. We thus cannot reject the null that the relationship between ruler ability and state performance was similar for male and female rulers, although it was possibly somewhat attenuated for the latter. 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). The interaction term is quantitatively somewhat larger and statistically marginally significant in the IV specification. If taken at face value, this result implies that ruler ability had a larger impact for rulers who came to power at a younger age, which makes intuitive sense. Note however, that there is no such effect in the reduced-form results. In column 3, we interact with a dummy indicating that the prior ruler was executed

[30]Note that in the first stage, we have to use the interaction of the ruler/reign characteristics with inbreeding. This gives rise to concerns that the interaction may be endogenous. The reduced-form results are not subject to these concerns.

after trial or murdered (Kokkonen and Sundell, 2014). We find no statistically significant difference in our IV results for those reigns. This speaks against the possibility that our results may be 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.

Columns 4 and 5 focus on the circumstances under which the reign started. The heterogeneity is with regard to whether in the prior reign, the state became a member of an alliance (col 4) or a target of one (5). We use data on alliances from Levy and Thompson (2005). These data were coded and made available by Benzell and Cooke (2021, Appendix A.6). The dataset covers 257 alliances between "great powers" from 1495 onwards, and it provides information on the year when an alliance was formed, its member states, as well as the states targeted by the alliance. Most states in our core sample are covered.[31] We create a "Target of Alliance" dummy that takes on value one if a ruler was the target of an alliance by other states for at least one year of his/her reign. In the same way, we create a "Member of Alliance" dummy. Both of the corresponding interaction terms with ruler ability are small and statistically insignificant. Thus, being a member or a target of an alliance did not alter the effect of ruler ability on state performance.

Overall, there is little evidence for heterogenous effects. These results suggest that inbred rulers had negative consequences for their states because they were inbred (and thus less capable), and not because of particular characteristics of themselves or their reigns. Finally, note that of the dummies for ruler or reign characteristic in Table A.13, none is itself statistically significant. In particular, this suggests that alliances (cols 4 and 5) did not systematically affect state performance.

B.11 Alternative Measure for Constraints on Rulers

We confirm the robustness of our results with an alternative measure of constraints on the executive. In the results shown below, we draw on data on parliamentary activity by van Zanden et al. (2012) to proxy for constraints on the monarch. This measure counts the number of years in a given century during which a state's parliament met at least once. The

[31]Castile and Aragon had already merged to united Spain by the time the alliance dataset begins. Alliance data for Scotland are not available. All other states in our sample are fully covered for the period from 1495 onwards.

Table A.13: IV and Reduced Form Heterogeneity

Dep. Var.: State Performance

	(1)	(2)	(3)	(4)	(5)
Dummy for:	Female Ruler	Young Ascension	Regicide of prior ruler	Previous ruler: Member of Alliance	Previous ruler: Target of Alliance
Panel A. IV Regressions					
Ruler Ability	0.835***	0.608***	0.709***	0.828***	0.867***
	(0.111)	(0.165)	(0.188)	(0.151)	(0.162)
Ruler Ability × Dummy	-0.259	0.407*	0.643	0.051	-0.109
	(0.281)	(0.237)	(1.113)	(0.368)	(0.642)
Dummy	-0.152	0.035	-0.509	0.107	0.106
	(0.120)	(0.105)	(0.699)	(0.280)	(0.241)
SW F-Stat (Ruler Ability)	159.1	102.9	6.3	34.4	29.2
SW F-Stat (Dummy × Ruler Abil.)	114.0	90.8	46.8	29.0	44.1
Observations	238	238	138	229	229
Panel B. Reduced-Form Regressions					
Inbreeding	-0.250***	-0.265***	-0.216**	-0.305***	-0.287***
	(0.028)	(0.075)	(0.070)	(0.049)	(0.044)
Inbreeding × Dummy	-0.025	0.013	1.143	0.060	0.066
	(0.158)	(0.098)	(1.101)	(0.136)	(0.162)
Dummy	-0.197	0.130	0.505	0.219	0.103
	(0.174)	(0.103)	(0.557)	(0.301)	(0.217)
State FE	✓	✓	✓	✓	✓
R^2	0.11	0.11	0.15	0.12	0.11
Observations	238	238	138	229	229
Fixed Effects (Panels A-B)					
State FE	✓	✓	✓	✓	✓

Note: The table shows the IV and reduced form results of interacting our baseline reduced form regression with different characteristics of rulers (columns 1 and 2) and prior reigns (columns 3-5). In column 1, the interaction variable is a dummy that takes on value one if the ruler was a woman; in column 2, for rulers ascending to the throne below the median age of 28 years. In column 3, the interaction variable indicates whether the prior ruler was murdered or executed after trial. Columns 4 and 5, respectively are dummies indicating whether, during the preceding reign, the state was either a member (column 4) or target (column 5) of an alliance. Data on alliances was compiled by Levy and Thompson (2005) and is provided by Benzell and Cooke (2021, Appendix A.6), which we use as our source. 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.

"Parliament Meetings" variable is then the share of years during which parliament met.[32] Table A.14 below presents the results for this variable and its interaction with ruler ability. We report the same six specifications as in Table 5 in the paper. The results are comparable in size and significance to those reported in Table 5 in the paper. For example, the IV results in column 2 in Table A.14 imply that a Parliamentary meeting frequency of 0.5 'neutralizes' the effect of ruler ability (as the interaction term is twice the size of the coefficient on Ruler Ability. This meeting frequency corresponds roughly to the 75th percentile. Alternatively, using a dummy indicator of parliament meetings at the 90th percentile (corresponding to a frequency of 0.79) also results in a statistically highly significant IV interaction coefficient, where the negative interaction term has a very similar size as the coefficient on ruler ability – i.e., a result that replicates our baseline estimates in Table 5 in the paper.

[32]The van Zanden et al. (2012) data covers the period from the 12th to the 18th century. For all countries except Turkey (which is not part of our IV sample in any case, as it is not covered by our source roglo.eu), we can link the parliamentary activity variable to our dataset. 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. Consistent with our baseline coding, we use the average value of this variable from five to ten years before the start of each reign. The downside is that the van Zanden et al. data are only available at the century frequency. Thus, the "five to ten year before the reign" effectively only makes a difference for reigns that started during the first decade of a century (as these are shifted back to the previous century's parliamentary activity).

Table A.14: The Role of Institutional Constraints on Ruler Power: Alternative Measure

Dependent variable as indicated in table header

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.		$\Delta log(Area)$			$\Delta log(UrbanPop.)$	
Estimation:	OLS	IV	RF	OLS	IV	RF
Ruler Ability	0.154***	0.238***		0.144***	0.236***	
	(0.035)	(0.062)		(0.028)	(0.062)	
Parliament Meetings	-0.116	-0.160***	-0.174**	0.021	-0.141*	-0.173
	(0.065)	(0.060)	(0.062)	(0.066)	(0.074)	(0.128)
Parliament Meetings × Ruler Ability	-0.236***	-0.428***		-0.154***	-0.542**	
	(0.058)	(0.163)		(0.048)	(0.275)	
Inbreeding			-0.069***			-0.068**
			(0.016)			(0.023)
Parliament Meetings × Inbreeding			0.139**			0.178*
			(0.045)			(0.098)
State FE	✓	✓	✓	✓	✓	✓
SW F-Stat (Ruler Ability)		52.9			47.3	
SW F-Stat (Constr. × Ruler Abil.)		31.2			28.5	
R^2	0.12		0.11	0.09		0.05
Observations	275	208	208	274	207	207

Note: The table shows that the effect of ruler ability on the performance of their states was muted when their executive power was constrained. The variable *Parliament Meetings Ruler* indicates the number of years in which parliament met at least once in the century in which the period five to ten years before a reign falls. 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.

C Potential Threats to the Exclusion Restriction: Additional Results

In this appendix, we provide additional results that complement our discussion of potential concerns with our identification strategy. We address a variety of potential threats to the exclusion restriction related to path dependence in state performance, strategic marriage (inside or outside of the kin network), as well as conflict.

C.1 Did the European Nobility Know about the Risks of Inbreeding?

Did the medieval and early modern European nobility know about the health risks of inbreeding? Our own reading of the historical literature found no evidence of such knowledge.[33] In addition, we have consulted academic experts on that question. In January 2023, we reached out to six historians who have worked extensively on European monarchies in the medieval and early modern period. These scholars published many of the salient books on the topic. Publishers include: Cambridge University Press (multiple books), Princeton University Press, Cornell University Press, University of Pennsylvania Press, Routledge, among many others. We asked: "Are you aware of any scholarly work we might have missed stating whether European monarchs (or nobility, or even the general population) were aware of the threats of inbreeding to their offspring's health/ability? If so, did they consider these negative effects when making marriage decisions?" None of the six scholars had come across any evidence that the European nobility was aware of negative health consequences of inbreeding. Below, we summarize some of the scholars' replies to our inquiry. The full correspondence is available upon request.[34]

- Scholar 1's response: "*I have never seen any indication of any medieval person being*

[33]After our period of analysis, there is some evidence that monarchs *suspected* that inbreeding might be related to physical shortcomings. One documented instance is Queen Victoria of England (1837-1901), who speculated that the hemophilia in her eighth child, Prince Leopold, may have been due to a lack of "some fresh blood...[f]or that constant fair hair and blue eyes makes the blood so lymphatic." (cited in Shaw, 2002, p. 69). However, immediately after this quote, Shaw points out that "It is unlikely that at the time of writing this letter the Queen had any idea precisely what was wrong with her family's blood." In fact, Shaw then suggests that the Queen may have formed her opinion based on the wide-spread legend of the "curse of the Coburgs," from whom she directly descended. The legend alleged that a jealous monk had cursed the Coburgs with the disease.

[34]We also asked these scholars about a statement that had been pointed out by a referee as indicating possible knowledge about the negative effects of inbreeding: In her popular book *A Distant Mirror*, Barbara Tuchman (1978, p.47) wrote: "Rulers likewise paid no attention whatever – with predictable results – to the prohibition of consanguinity in marriage, whose risks were well understood and forbidden by the Church within the fourth degree." Tuchman provides no evidence for her claim that "risks were well understood." The experts we consulted conjectured that this must have been pure speculation by Tuchman, and many of them explicitly stated that "Barbara Tuchman cannot be considered a reliable medieval historian" (in this exact or similar wording).

concerned about inbreeding from a biological perspective. Given that inbreeding was (and is) a way to lock in desirable traits in domestic animals, it would not have seemed a problem biologically."

- Scholar 2's response: "*I can confirm what you suggest. I think that Barbara Tuchman [see appendix footnote 34] was simply making an assumption based on the Church's prohibition (often flouted) of consanguineous marriages.*"

- Scholar 3's response: "*I have not come across any evidence that the nobles were aware of 'the threats of inbreeding to their offspring's health'. In fact, rather the contrary, since they were always trying to arrange cousin marriages in the face of ecclesiastical opposition. Clerical rhetoric criticizes such marriages as 'incestuous', with no implication of the physical problems of inbreeding but a strong suggestion of pollution and taboo.*"

- Scholar 4's response: "*The European nobility indeed was aware of the fact that they were closely related with each other. In their pleas for papal dispensation they state that fact, but I have not found a hint that they were aware of the biological consequences.*"

- Scholar 5's response: [Translation from German:] "*I asked myself the same questions, but in my research into the marriages of the Hohenzollerns I did not come across any sources that provide information about the level of knowledge of the actors. Inbreeding was not as common among the Hohenzollerns as it is among other dynasties. What I have found are sources that testify to concerns about the health and fertility of brides, as well as (in the 19th century) sources that show that people were concerned about the heredity of certain diseases.*"

- Scholar 6's response: "*The rationale for the [intermarriage] ban [by the church] was not however genetic. It was almost sociological. The idea was that intermarriage was a waste of a social bond. Kinship was a social bond already, so, if one could not marry kin, the social cohesion would be extended and enhanced. (...) I know of no discussion of genetic factors*"

C.2 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.

Panel A of Table A.15 documents that past state performance (for all three outcome variables) does *not* predict current state performance in our reduced-form regression: The coefficients in odd columns (baseline) and even columns (controlling for lags of state performance) are very similar. Panel B A.15 shows that the IV results are also robust to controlling for lagged state performance measures.[35]

Table A.16 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 (if anything, the small negative coefficients for the lags of all three outcome variables suggest the opposite). In sum, the results from Tables A.15 and A.16 suggest that neither of the conditions required for strategic kin marriage to affect our exclusion restriction are fulfilled.

C.3 Strategic Marriage Outside the Kin Network

Alternatively, rulers may have strategically married *outside* of their dynasty network when this implied future territorial expansion. This would mechanically increase state performance in the following period by enlarging the territory. Such a mechanism could result in a link between inbreeding and state performance, as a marriage between completely unrelated individuals would result in a coefficient of inbreeding of $I = 0$ in the next generation.

Note that we exclude monarchs with (likely) completely unrelated parents from our baseline IV analysis (see footnote 29 in the paper). In Table A.17, we include the 43 rulers whose parents (likely) had no relationship. For comparison, odd columns report our baseline IV results, while even columns add these 43 rulers.[36] If strategic marriage outside the kin network with unrelated individuals played a significant role, we would expect larger coefficients when including the 43 rulers with $I = 0$, especially for $\Delta log(Area)$. This is not

[35] We opt for one lag in the outcomes since this is what both the Akaike and Bayesian information criteria (AIC and BIC, respectively) suggest as the optimal number of lags for most outcomes and states. These criteria, developed in the time series literature, trade off the increased explanatory power of the models with more lags with a penalty for estimating more parameters, and the optimal number of lags comes from the model minimizing the criteria. However, we find that including up to four lags (the maximum optimal number of lags selected for any of the countries and outcomes separately by the AIC or BIC) of each outcome yields highly comparable results (available upon request).

[36] In columns 4 and 6, not all 43 rulers have corresponding observations for the outcome variable. Thus, the number of observations increases only by 40 and 39, respectively.

Table A.15: Past State Performance as a Confounder: IV and RF Results

Dependent variable as indicated in table header

Dep. Var.	*State Performance*		$\Delta log(Area)$		$\Delta log(UrbPop)$	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Reduced-Form Regressions						
Inbreeding	-0.253***	-0.256***	-0.047***	-0.034***	-0.039**	-0.030*
	(0.038)	(0.034)	(0.013)	(0.009)	(0.016)	(0.016)
L.State Performance		-0.027				
		(0.066)				
L.Δ Log(Area)				-0.036		
				(0.051)		
L.Δ Log(Urb. Pop.)						0.013
						(0.081)
State FE	✓	✓	✓	✓	✓	✓
R^2	0.11	0.11	0.10	0.03	0.04	0.03
Observations	238	221	209	196	208	195
B. Second Stage Regressions						
Ruler Ability	0.805***	0.817***	0.161***	0.114***	0.136***	0.100**
	(0.094)	(0.139)	(0.046)	(0.031)	(0.048)	(0.050)
L.State Performance		0.035				
		(0.067)				
L.Δ Log(Area)				-0.024		
				(0.048)		
L.Δ Log(Urb. Pop.)						0.030
						(0.076)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	49.7	38.1	46.2	50.9	43.7	49.1
Observations	238	203	209	196	208	195

Note: The table shows the results of reduced-form and second-stage regressions, controlling for one lag of the outcome variable. 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. Lags vary in length depending on the length of the previous reign. All regressions are run at the reign level and include state fixed effects. Standard errors clustered at the state level in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

the case: our results remain very similar to the baseline and statistically highly significant.

C.4 Accounting for Conflict

Conflicts could pose a threat to our exclusion restriction if they were systematically associated with both inbreeding and state performance. Overall, there is mixed evidence on the relationship between relatedness and conflict: Spolaore and Wacziarg (2016) find evidence of a positive link between relatedness and the probability of conflict at the level of *societies*. Benzell and Cooke (2021) explore the relationship between family networks of *alive* ruler pairs and conflict between their states. Their core results use random deaths in kinship networks to instrument for ties between monarchs, documenting that stronger alive kinship ties *reduce* conflict between states (and vice-versa, random deaths of connecting network

Table A.16: Past State Performance as Confounder: First Stage

Dep. Var.: Ruler Ability

	(1)	(2)	(3)	(4)
Inbreeding	-0.310***	-0.320***	-0.284***	-0.287***
	(0.047)	(0.045)	(0.044)	(0.045)
L.State Performance		-0.063		
		(0.096)		
L.Δ Log(Area)			-0.118	
			(0.129)	
L.Δ Log(Urb. Pop.)				-0.170
				(0.159)
State FE	✓	✓	✓	✓
R^2	0.15	0.15	0.12	0.12
Observations	243	225	204	203

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.17: Strategic Marriage outside of Kin Network: IV results

Dependent variable as indicated in table header

Dep. Var.	*State Performance*		$\Delta log(Area)$		$\Delta log(UrbPop)$	
	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Baseline	Include $I = 0$	Baseline	Include $I = 0$	Baseline	Include $I = 0$
Ruler Ability	0.805***	0.824***	0.161***	0.159***	0.136***	0.164***
	(0.094)	(0.093)	(0.046)	(0.045)	(0.048)	(0.053)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	49.7	45.2	46.2	49.1	43.7	46.2
Observations	238	281	209	249	208	247

Note: The table documents that our results are robust when we include rulers with unknown coefficient of inbreeding, who were most likely unrelated, so that we assign them a value of $I = 0$. The table shows results from instrumental variable regressions, in which ruler ability is instrumented with *Inbreeding*. We report 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. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

members raise the probability of interstate conflict). However, these kinship linkages at the ruler-*pair* level do not necessarily translate into higher levels of inbreeding of the individual rulers.[37] In fact, when using relationship by blood between pairs of rulers rather than shared (alive) kinship ties, Benzell and Cooke (2021, Appendix D.1) find a *positive*

[37] For instance, consider two brothers – offspring of genetically unrelated parents – who rule two different states. They would have strong blood (and kinship) ties, while both having an inbreeding coefficient of zero. One such example is Henry III, who ruled Castile from 1400 to 1406, and his brother Ferdinand I of Portugal (reign 1412-1416). Both have an inbreeding coefficient of 1.30 (less than half that of second cousins), while their kinship ties were extremely close. Furthermore, cousins (also with relatively close kinship ties) can have *different* inbreeding coefficients. Consider the case of Philipp II, ruling Spain from 1556 to 1598, and his cousin Maximilian II, concurrently ruling Austria (1564-1576). The latter had a comparatively low coefficient of inbreeding (1.38), while the former was almost as inbred as offspring of uncle and niece with a coefficient of inbreeding of 11.45.

(albeit insignificant) association with conflict. Similarly, in our core sample, states under more inbred rulers had a higher probability of conflict and a higher share of years under conflict.[38] Here, we explore the extent to which conflict may confound our main results on state performance and territorial changes. Before presenting the empirical results we note that the most likely implication of inbreeding being positively associated with conflict is that inbreeding adds more variability in state performance and territorial changes due to more frequent wars. That is, incapable (inbred) rulers would fight more often – and our results suggest that they would lose more often. While this would add identifying variation in our data, it would not necessarily confound our results.

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.[39] We perform this analysis in Table A.18. 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 conflict years (column 3). In addition, the coefficients of the two conflict variables themselves are quantitatively small and statistically insignificant.

We can also address a possible role of conflicts in our measurement of state performance. Our main measure *State Performance* is a composite measure, including territorial changes as one of many assessed features (others being administrative reform, economic performance, etc). In column 4 of Table A.18 we use as outcome the residual of a regression of *State Performance* on the percentage change in territory under the control of a monarch during their reign from Abramson (2017). Column 5 instead residualizes *State Performance* with a categorical variable of territorial expansion (“1”) or decline (“-1”).[40] In columns 4 and 5 the coefficient size is slightly reduced. Yet, the effect of ruler ability retains statistical significance and remains sizable. This underlines the importance of ter-

[38] We examine this further below, documenting in Table A.24 that higher ruler ability (due to lower inbreeding) reduced conflict.

[39] See footnote 20 in Appendix B.8 for the source of the underlying data by Peter Brecke. Based on this data, covering conflicts in Europe from 900 AD onward, 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.

[40] See Appendix D.2 for detail.

Table A.18: IV Results Accounting for Conflict

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)
Note on Dep. Var.:	—— Baseline ——			Resid. wrt % territorial changes †	Res. wrt territorial changes ‡
Ruler Ability	0.805***	0.780***	0.748***	0.742***	0.426**
	(0.094)	(0.087)	(0.095)	(0.093)	(0.203)
Conflict: Dummy		-0.155			
		(0.191)			
Conflict: Share Years at War			-0.191		
			(0.171)		
State FE	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	49.7	41.5	29.2	53.8	23.8
Observations	238	238	238	205	118

Note: The table presents different specifications that control for a possible role of conflict. The table shows results from instrumental variable regressions, in which ruler ability is instrumented with *Inbreeding*. We report 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. Standard errors clustered at the state level. * 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).

ritorial changes in state performance, but also that of other aspects of state performance, beyond territorial changes.

C.5 Order within Dynasties: Founder and Descendant Effects

George and Ponattu (2018) show that dynastic politics can generate a "reversal of fortune" development pattern, whereby 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 inbreeding 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 third of the Spanish Bourbons. Yet, he also hails from the Bourbon dynasty ruling France. He is the eighth 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 in Table A.19 repeats our baseline IV result. Column 2 restricts the sample 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 states 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 remain quantitatively similar and statistically significant.

Table A.19: 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	0.805***	0.842***	0.888***	0.575**
	(0.094)	(0.110)	(0.120)	(0.260)
State FE	✓	✓	✓	✓
Order in Dynasty FE			✓	
Order in International Dynasty FE				✓
First Stage Effect. F-Stat	49.7	48.1	44.0	7.8
Observations	238	235	235	235

Note: "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 third of the Spanish Bourbons. Yet, he also hails from the Bourbon dynasty ruling France. He is the eighth of all Bourbons, ordered by the year in which his reign began. The table shows results from instrumental variable regressions, in which ruler ability is instrumented with *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% (15%) relative bias is 16.4 (9.0). Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

C.6 Selection among Offspring

One concern with our identification strategy revolves around the selection of less inbred and more capable individuals among the prior rulers' offspring. Fathers may have 'rid themselves' of incapable offspring, or offspring who were more affected by their parents' consanguineous relationships may have died at young age, thus leaving more capable surviving successors. Note that both these mechanisms would work against our first stage (provided that all potential heirs to the throne were offspring of the same marriage): 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. In column 2 of Table A.20 we show that our results are very similar to the baseline IV result (reported in column 1) when reducing the sample to those monarchs who were the first-born sons and thus the most commonly legally mandated heirs to the throne. More precisely, we focus on monarchs of whom we know that either they were the first-born offspring (irrespective of gender), or, if the first-born's gender was female, were the first-born male offspring. While this implies a reduction in sample size by about 50% (due to numerous cases where information on gender by birth order is not available), the second stage coefficient is remarkably stable and highly significant.[41]

What if monarchs remarried and selected an heir among the offspring from the less consanguineous marriage? We address this in column 3 of Table A.20, by restricting the sample to those monarchs whose parents either had only one marriage or, if they had more than one marriage, had no offspring from any other marriage. Again, we find similar results. In sum, selection among offspring from the same or any other marriage is not a concern for the validity of our IV results.

Table A.20: IV Result: Selection Among Offspring

Dep. Var.: *State Performance*

	(1)	(2)	(3)
Sample:	All	Firstborn Sons	No Competing Claims
Ruler Ability	0.805***	0.777***	0.729***
	(0.094)	(0.225)	(0.180)
State FE	✓	✓	✓
First Stage Effect. F-Stat	49.7	24.1	15.1
Observations	238	122	147

Note: The table shows results from instrumental variable regressions in which ruler ability is instrumented with *Inbreeding*. The first column repeats our baseline IV results. Column 2 restricts attention to firstborn sons (for cases where this information is unambiguously available), and column 3 to those monarchs whose parents had only one marriage or no offspring from any other marriage. 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% (15%) relative bias is 16.4 (9.0). 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 further examine potential mechanisms that connect ruler ability to state performance, with inbreeding as the underlying source of variation in ruler ability. We first ask which characteristics of rulers may have been affected by inbreeding and how these, in turn, affected state performance. We show that the effect likely works through the cognitive (as opposed to physical or non-cognitive) abilities of monarchs. Second, we examine which dimensions of state-level outcomes were affected by inbred rulers. We show that ruler ability affected both political and economic components of state performance.

[41] 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 et al. (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 cognitive capability). However, this is not the case: As described above, our results also hold in the subsample of firstborn sons.

D.1 Inbreeding and Ruler Characteristics

To what extent did the (reduced-form) effect of inbreeding on state performance run through other ruler characteristics that may also have been affected by inbreeding? In what follows, we code numerous ruler characteristics and then examine whether they are i) predicted by inbreeding and ii) whether controlling for them changes our main results. Panel A in Table A.21 shows the results of regressing various characteristics of rulers on the (standardized) inbreeding coefficient I. For a straightforward interpretation, we also standardize all outcome variables, except for the dummy variables. The construction and sources of all the variables is described in detail in Appendix A.1.

Columns 1 in Table A.21 repeats our baseline first-stage regression, regressing (cognitive) ruler ability on inbreeding. The "Ruler Ability" measure from Woods (1913) explicitly aimed at capturing cognitive abilities.[42] Column 2 shows that more inbred rulers also had lower non-cognitive ability. While the literature on the effects of inbreeding typically highlights negative effects on *cognitive* ability, the work on non-cognitive ability such as emotional stability is much more sparse because these are hard to (objectively) measure. However, there is evidence that non-cognitive ability affects leadership skills (Adams, Keloharju, and Knüpfer, 2018). We thus investigate below whether non-cognitive ability may be an alternative mechanism by which inbreeding affected state performance.

Columns 3 and 4 in Panel A of Table A.21 show that inbreeding does not have a sizable effect on the height or physical strength of rulers: the coefficients are small and statistically insignificant. Column 5 finds a marginally significant negative effect of inbreeding on the number of children of a monarch. Next, columns 6-8 show small and statistically insignificant coefficients on age at death, length of reign, and age at ascension to the throne. Finally, in column 9 we document that inbred rulers were significantly less likely to ascend to power after regicide of the prior monarch (murdered in office or executed after a trial). However, the estimated coefficient is small in magnitude. One explanation for the negative effect is that often, less inbred members of new dynasties came to power after the deposition of former monarchs.

[42] Woods (1913, p.5) focused on what he called "mental" (i.e., cognitive) ability as opposed to "morals" (closely related to non-cognitive ability). However, the adjectives used by Woods to describe rulers can be used to also infer non-cognitive traits. We use these to code an indicator of non-cognitive ability, as described in Appendix A.2. In his 1906 publication, Woods coded *both* the mental (cognitive) and moral (non-cognitive) ability for a subset of 104 dynastic rulers. Our results from columns 1 and 2 in Table A.21 hold in this smaller sample, using Woods' (1906) own coding of both variables (available upon request).

Table A.21: Mechanism: Inbreeding and Ruler Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variable X=	Ruler Ability Cognitive	Non-cogn.	Tall[‡]	Strong[§]	Number of Children	Age at Death	Length of Reign	Age at Ascension	Regicide of prev. ruler[†]
				Panel A. Inbreeding & Ruler Characteristics – Dep. var.: X					
Inbreeding	-0.314***	-0.268***	0.007	-0.035	-0.115*	0.004	0.058	-0.042	-0.025**
	(0.045)	(0.069)	(0.020)	(0.135)	(0.057)	(0.066)	(0.050)	(0.040)	(0.009)
R^2	0.15	0.14	0.15	0.06	0.06	0.08	0.06	0.05	0.04
Observations	238	238	238	238	227	238	238	238	138
				Panel B. IV Regressions Controlling for X – Dep. var.: State Performance					
Ruler Ability (Cognitive)	0.805***	0.785***	0.803***	0.808***	0.719***	0.805***	0.833***	0.816***	0.692***
	(0.094)	(0.136)	(0.093)	(0.090)	(0.101)	(0.092)	(0.082)	(0.098)	(0.202)
X		0.023	-0.063	-0.030	0.078	0.031	0.156***	-0.085	-0.186
		(0.057)	(0.157)	(0.039)	(0.092)	(0.121)	(0.037)	(0.077)	(0.224)
First Stage Effect. F-Stat	49.7	12.4	48.4	55.7	44.6	51.7	53.4	50.2	41.1
Observations	238	238	238	238	227	238	238	238	138
				Panel C. Reduced Form Regressions Controlling for X – Dep. var.: State Performance					
Inbreeding	-0.059*	-0.189**	-0.253***	-0.251***	-0.219***	-0.254***	-0.269***	-0.256***	-0.212**
	(0.032)	(0.065)	(0.039)	(0.039)	(0.045)	(0.034)	(0.033)	(0.039)	(0.069)
X	0.616***	0.236***	0.048	0.037	0.227**	0.198**	0.279***	-0.074	-0.010
	(0.059)	(0.070)	(0.197)	(0.056)	(0.078)	(0.088)	(0.027)	(0.046)	(0.235)
R^2	0.42	0.19	0.11	0.11	0.17	0.14	0.18	0.11	0.14
Observations	238	238	238	238	227	238	238	238	138
				Panel D. Sobel-Goodman Mediation[(a)]					
Proportion of total effect mediated	0.765	0.250	-0.001	0.005	0.107	-.003	-0.064	-0.012	-0.02
State FE (Panels A.-C.)	✓	✓	✓	✓	✓	✓	✓	✓	✓

Note: The table documents that most of the effect of inbreeding on State Performance operates through (cognitive) Ruler Ability. This variable is our baseline measure throughout the paper, where we refer to it simply as Ruler Ability. The dependent variables in panel A are indicated in the table header. These variables are then used as control variables in panels B and C, where the dependent variable is *State Performance*. All variables *X* variables, except for the dummy variables 'Tall,' 'Regicide,' and 'Regency' are standardized, and so is inbreeding in all columns. 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% (15%) relative bias is 16.4 (9.0). Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

[‡] 'Tall' is a dummy indicating that ruler was described as 'tall,' 'very tall,' or had a recorded height of over 179 cm.

[§] 'Strong' indicates whether the monarch was described as 'strong' (1), physically 'weak' (-1), or neither (0).

[†] Dummy indicating whether the *prior* ruler was murdered or executed after a trial.

[(a)] The Sobel-Goodman mediation test computes the proportion of the total effect of inbreeding on *State Performance* that is mediated by variable *X* in the corresponding column.

Next, we examine whether these ruler characteristics (in particular, those that were also associated with inbreeding), may drive our results. We have found that inbreeding had negative consequences for both cognitive and non-cognitive abilities of early modern rulers. In principle, both could influence the effectiveness of rulers as heads of states. In addition – even though we did not find a strong relationship between inbreeding and physical features of rulers, these could nevertheless have played a role. For example, due to inbred rulers' *anticipated* early death, their lack of reproductive success, or lack of physical strength (Álvarez, Ceballos, and Quinteiro, 2009; Álvarez et al., 2019). In Panel B of Table A.21, we control for these characteristics directly in our IV regressions for *State Performance*. Column 1 shows our baseline IV regression for comparison. In columns 2-9 we control for the various ruler characteristics (denoted by X and listed in the table header). Our main coefficient of interest, that of ruler ability, is unaffected by the inclusion of any of these controls. Notably, this is true even in column 2, where we control for non-cognitive ability: The inclusion of this variable barely changes our main coefficient of interest on (cognitive) ruler ability. This underlines the importance of cognitive ability.[43] Similarly, our IV result is also robust with respect to the inclusion of all other ruler characteristics (columns 3-9), and only the length of the reign is itself significantly related to *State Performance*.

Panel C in Table A.21 controls for the various ruler characteristics the corresponding reduced-form results. Note that in column 1, we thus control for cognitive ruler ability in a regression of *State Performance* on inbreeding. This is similar in spirit to a mediation analysis, which examines the extent to which the explanatory power of an instrument (inbreeding) is reduced when a potential mediating variable (here: cognitive ruler ability) is included. Indeed, the coefficient on inbreeding is reduced substantially, while cognitive ruler ability is highly significant and sizeable. We run Sobel-Goodman tests and report in Panel D the proportion of the total effect of inbreeding on *State Performance* that is mediated by each of the control variables X (listed in the header of Table A.21). Accord-

[43] We find further support for the predominant role of *cognitive* ruler ability by turning the result in column 2 on its head, that is, by using inbreeding to predict *non-cognitive* ability while controlling for cognitive ability in our IV regression. In this regression, non-cognitive ability has no predictive power in the IV, while the control variable, cognitive ability, has a sizable and significant coefficient. Further, note that differential measurement error in non-cognitive ability (coded by us based on adjectives in Woods, 1913) and cognitive ability (coded directly by Woods, 1913) cannot account for the small and insignificant coefficient on non-cognitive ability in column 2 in Panel B of Table A.21. Our results hold in a subsample of 106 rulers for whom Woods (1906) coded *both* cognitive and non-cognitive ability (available upon request; see appendix footnote 42 for detail on this coding).

ing to this exercise (which we interpret as exploratory rather than conclusive), 76.5% of the effect of inbreeding on *State Performance* is mediated by cognitive ruler ability.[44] In stark contrast, none of the other ruler characteristics significantly lowers the coefficient on inbreeding, and only few of these are themselves significantly associated with *State Performance*. The second-largest mediator, according to the Sobel-Goodman tests in Panel D, is non-cognitive ability, accounting for about one-quarter of the effect of inbreeding on *State Performance*. This is in line with the evidence that non-cognitive ability affects leadership skills (Adams et al., 2018) and with related work that emphasizes the importance of both cognitive and non-cognitive intellectual abilities (Lindqvist and Vestman, 2011; Heckman, Stixrud, and Urzua, 2006). None of the remaining variables in columns 2-8 (which reflect mostly physical attributes) appear to play any meaningful role as mediators.

Overall, the results in Table A.21 suggest that inbreeding rendered leaders ineffective predominantly due to its negative effect on cognitive capabilities and, to some degree, via non-cognitive capabilities. Physical characteristics of rulers do not appear to account for the strong effect of inbreeding on state performance.

D.2 Different Aspects of State Performance

Our main outcome variable, *State Performance* as assessed by Woods (1913), is a composite measure. 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). In what follows, we further assess the underlying components in order to examine which specific aspects drive our results. While Woods did not provide a coding of these specific aspects, his text descriptions often include passages with the corresponding information. 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. We briefly list of all these, along with some questions that display which aspects were covered by these measures.

[44] We obtain a very similar percentage for our other two outcomes, territorial changes and changes in urban population: 62.6% and 69%%, respectively. Results available upon request.

The **political aspects** of state performance cover the following domains: *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 well 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?

The **economic aspects** of state performance cover the following domains: *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 famines, 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. For those reigns for which Woods (1913) did not reach an assessment of *State Performance*, we similarly set our assessments of particular aspects to missing rather than zero.

We discuss results for political and economic aspects separately. Table A.22 shows results of our baseline IV regressions, where the dependent variables – instead of our composite measure *State Performance* – are our assessments of political aspects during each reign. As in our baseline analysis of *State Performance*, we standardize the dependent variables to mean zero and standard deviation one. In column 1 of Table A.22, 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. The measure here is a categorically assessed variable based on historical sources, while the one in our main analysis is a continuous variable constructed from actual data on polity border changes from Abramson (2017). In columns 2-8 we also document sizable effects of ruler ability on law and order in their states, public liberty, on finances, on the effectiveness of the administration, and the diplomatic prestige of the state. The remaining outcomes are also positively affected by ruler ability, but while the coefficients are sizeable, they are not statistically significant.

Table A.22: IV Results: Political Aspects 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	0.670***	0.525***	0.301*	0.564***	0.313	0.177	0.522**	0.401**
	(0.199)	(0.156)	(0.176)	(0.214)	(0.259)	(0.160)	(0.220)	(0.174)
State FE	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	49.4	49.4	49.4	49.4	49.4	49.4	49.4	49.4
Observations	238	238	238	238	238	238	238	238

Note: The table shows our IV results for different aspects of state performance. The first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test is 49.4; 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.

Next, we examine the effect of ruler ability on economic aspects of state performance. Table A.23 documents strong effects 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 small and not statistically significant.

Table A.23: IV Results: Economic Aspects 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	0.454**	0.671**	0.825***	0.114	0.062
	(0.179)	(0.261)	(0.242)	(0.224)	(0.161)
State FE	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	49.4	49.4	49.4	49.4	49.4
Observations	238	238	238	238	238

Note: All regressions are run at the reign level. The table shows results from instrumental variable regressions in which ruler ability is instrumented with *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 examine how ruler ability affected a state's conflicts (both internal and external). The data on conflict are from Peter Brecke's Conflict Catalogue and start in 900 AD (cf. footnote 20 in Appendix B.8). We identify whether a state participated in any conflict (in Europe) within a given year. We differentiate between overall conflicts, internal conflict, and external conflicts. 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. We generate two outcome variables for each sub-category: 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.[45]

Column 1 in Table A.24 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 states? To answer this question, columns 3 and 4 of Table A.24 use internal conflict as the outcome variable, and columns 5 and 6 use external conflict. The results show that our previous finding 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 2 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. Consistent with this interpretation, in column 7, we further find that forces under incapable rulers were systematically more likely to lose battle engagements (conditional on participating in any battle during a given reign).[46]

D.4 Decomposition of the Change in Urban Population

In this section, we decompose the effect of ruler 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.

Section 3.2 in the paper describes how we calculate the change in total urban population between the beginning and the end of each reign. Note that such changes can result from

[45]Note that these outcomes are available even for reigns for which state performance measures are unavailable. Therefore, we have slightly more observations in this table than earlier ones.

[46]Data on battles are from Miller and Bakar (2023) who in turn build on Jacques (2007). Crucially, they code for each battle which state fought on each side and which side won. When a battle falls into more than one reign (for instance, during a year when control transitioned from one rulers to the next), we assign each battle to both reigns.

Table A.24: IV Results: Conflict as Outcome

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

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dep. Var:	All Conflicts		Internal		External		Battles
	Dummy	Share	Dummy	Share	Dummy	Share	Share Lost
Ruler Ability	-0.293***	-0.158**	-0.049	-0.036	-0.300***	-0.112**	-0.139***
	(0.088)	(0.068)	(0.051)	(0.070)	(0.082)	(0.051)	(0.027)
State FE	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	43.9	43.9	43.9	43.9	43.9	43.9	44.4
Observations	243	243	243	243	243	243	187

Note: The table shows IV results documenting that ruler ability had a negative effect on states' participation in external conflicts, and that this finding is driven by less participation in external conflicts. We report 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.

either changes in the population of the cities that remained in the polity throughout the reign ("intensive margin"), or from changes in the urban population located in areas lost or gained during a reign ("extensive margin"). We now distinguish between the cities and their population that remained under a state's control during each reign, and those that were gained, or lost, during the reign of each monarch. For each category, we derive the corresponding growth rates (γ) in urban population for all cities that i) remained within the state ($\gamma_{intensive}$), ii) were added to the states due to expansions (γ_{gained}), and iii) were lost to other states (γ_{lost}). In logarithms, this yields a decomposition of percentage change in urban population between reigns $r-1$ and r into an intensive and extensive margin:

$$log(Pop_r^{Urb}) - log(Pop_{r-1}^{Urb}) = log(1+\gamma_{intensive}) + log(1+\gamma_{gained}) - log(1+\gamma_{lost}) \ ,$$

where the latter two components together reflect the extensive margin of urban growth.

Table A.25 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 2 in the paper that a one standard deviation increase in the

Table A.25: 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	0.105***	0.097***	0.136***	0.025***	0.024**	0.004	0.080***	0.074**	0.132**
	(0.017)	(0.029)	(0.048)	(0.007)	(0.009)	(0.015)	(0.016)	(0.025)	(0.053)
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat			43.7			43.7			43.7
R^2	0.09	0.08		0.13	0.12		0.06	0.06	
Observations	300	208	208	300	208	208	300	208	208

Note: The table decomposes the results on change in urban population into an intensive and extensive margin. 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.

ability of a monarch raises urban population by 14%. Interestingly, this effect is entirely due to the extensive margin. The IV estimate for the intensive margin is minuscule with a relatively small standard error, indicating a 'reliably estimated zero.'[47] In contrast, the IV coefficient for the extensive margin in column 9 is as large as the total effect in column 3. A possible 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, 2022, ch. 7). Note also that the research that exploits urbanization rates as historical proxy for economic progress (e.g., Acemoglu, Johnson, and Robinson, 2005; Nunn and Qian, 2011), typically examines long time horizons: 50-100 year periods. In contrast, the median reign in our dataset has a length of 14 years. Thus, there was limited time for long-term dynamics of urban populations to materialize within individual rulers' reigns. In this context, the extensive margin (conquering existing cities) was the more feasible way for rulers to expand their state's urban population within their lifetime: The strong results for the extensive margin imply that capable rulers extended their territories into valuable, urbanized areas (and were less likely to lose such areas during conflicts).

[47]The OLS estimate for the intensive margin in the full sample is statistically significant but quantitatively small. In addition, the OLS coefficients for the extensive margin (cols 7 and 8) are also significantly larger than those for the intensive margin.

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.

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.

Álvarez, G. et al. (2019). Inbreeding in the Last Ruling Dynasty of Portugal: The House of Braganza. *American Journal of Human Biology 31*(2), e23210.

Álvarez, 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.

Alvarez-Nogal, C. and L. P. De La Escosura (2013). The Rise and Fall of Spain (1270–1850). *The Economic History Review 66*(1), 1–37.

Angelucci, C., S. Meraglia, and N. Voigtländer (2022). How Merchant Towns Shaped Parliaments: From the Norman Conquest of England to the Great Reform Act. *American Economic Review 112*(10), 3441–87.

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.

Benzell, S. G. and K. Cooke (2021). A Network of Thrones: Kinship and Conflict in Europe, 1495–1918. *American Economic Journal: Applied Economics 13*(3), 102–33.

Bolt, J. and J. L. van Zanden (2020). Maddison Style Estimates of the Evolution of the World Economy. A New 2020 Update. *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands*.

Brecke, P. (1999). Violent Conflicts 1400 AD to the Present in Different Regions of the World. In *meeting of the Peace Science Society*, pp. 8–10.

Broadberry, S., B. Campbell, A. Klein, M. Overton, B. v. Leeuwen, et al. (2022). British Business Cycles, 1270-1870. Technical report, Competitive Advantage in the Global Economy (CAGE).

Broadberry, S., B. M. Campbell, A. Klein, M. Overton, and B. Van Leeuwen (2015). *British Economic Growth, 1270–1870*. Cambridge University Press.

Broadberry, S. and J. Lennard (2023). European Business Cycles and Economic Growth, 1300-2000.

de Jong, H. J. and N. Palma (2018). Historical National Accounting. *An Economist's Guide to Economic History*, 395–403.

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

Fenske, J. (2013). Does Land Abundance Explain African Institutions? *The Economic Journal 123*(573), 1363–1390.

Fouquet, R. and S. Broadberry (2015). Seven Centuries of European Economic Growth and Decline. *Journal of Economic Perspectives 29*(4), 227–244.

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

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor economics 24*(3), 411–482.

Jacques, T. (2007). Dictionary of Battles and Sieges (Vol. 2).

Koepke, N. and J. Baten (2005). The Biological Standard of Living in Europe During the last Two Millennia. *European Review of Economic History 9*(1), 61–95.

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

Krantz, O. (2017). Swedish GDP 1300-1560: A Tentative Estimate. Technical report, Lund Papers in Economic History No. 152.

Levy, J. S. and W. R. Thompson (2005). Hegemonic Threats and Great-power Balancing in Europe, 1495-1999. *Security Studies 14*(1), 1–33.

Lindqvist, E. and R. Vestman (2011). The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics 3*(1), 101–28.

McEvedy, C. and R. Jones (1978). *Atlas of World Population History, Facts on File*. New York.

Miller, C. and K. S. Bakar (2023). Conflict Events Worldwide since 1468BC: Introducing the Historical Conflict Event Dataset. *Journal of Conflict Resolution 67*(2-3), 522–554.

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

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. et al. (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.

Palma, N. and J. Reis (2019). From Convergence to Divergence: Portuguese Economic Growth, 1527–1850. *The Journal of Economic History 79*(2), 477–506.

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

Ridolfi, L. (2017). The French Economy in the Longue Durée: A Study on Real Wages, Working Days and Economic Performance from Louis IX to the Revolution (1250–1789). *European Review of Economic History 21*(4), 437–438.

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.

Schön, L. and O. Krantz (2015). New Swedish Historical National Accounts since the 16th century in Constant and Current Prices.

Shaw, K. (2002). *Royal Babylon: The Alarming History of European Royalty*. Crown.

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

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

Tuchman, B. (1978). *A Distant Mirror: The Calamitous 14th Century*. Alfred A. Knopf.

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

Van Zanden, J. L. and B. Van Leeuwen (2012). Persistent But Not Consistent: The Growth of National Income in Holland 1347–1807. *Explorations in economic history 49*(2), 119–130.

Voigtländer, N. and H.-J. Voth (2013). The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe. *Review of Economic Studies 80*(2), 774–811.

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

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