Online Appendix

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

Sebastian Ottinger
(Northwestern and 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.[1] 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 "+." For the few cases where more than one ruler appears, we focus on the ruler whose coding works against our baseline results.[2]

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 1623 under Gustavus Vasa and became a separate political entity.

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. In total, our

[1]States enter the dataset when the ruler first mentioned by Woods appears. States can enter the dataset by splitting from former, larger states. This was the case for Sweden, which emerged as a separate state in 1623 when it split from Denmark. 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]For instance, for Ferdinand and Isabella, who jointly and successfully (as assessed by Woods) ruled over the Habsburg Empire from 1479 to 1504, we focus on Isabella, who had a higher inbreeding coefficient than her husband.

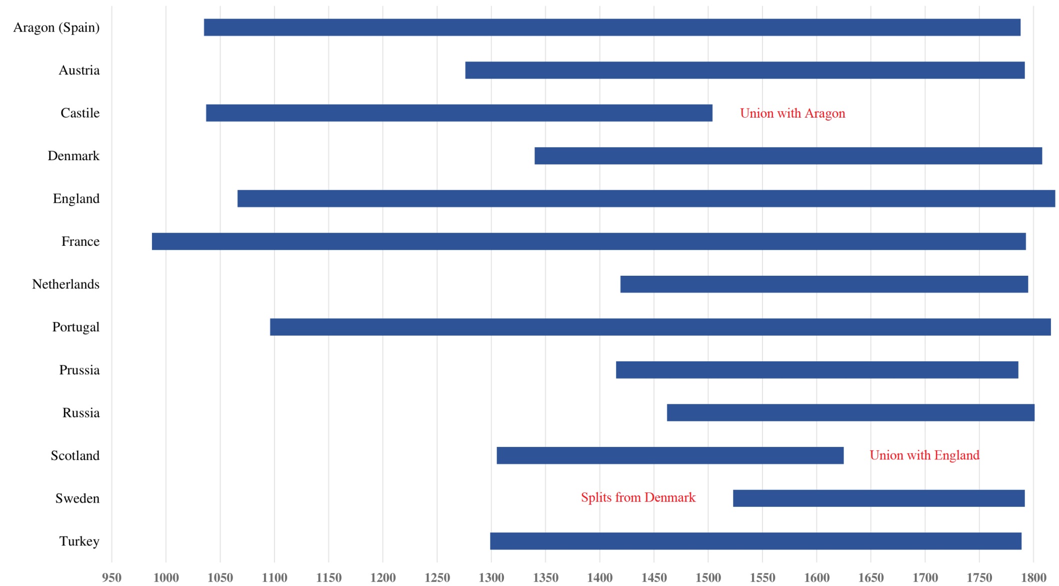

Figure A.1: Timeline of Sample Coverage: States in Sample

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

main explanatory and outcome variable – *Ruler Ability* and *State Performance* assessed by Woods – are available for 336 reigns. Only 235 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 'objective' measures of state performance based on territorial changes and the change in urban population within the state are available for 317 and 307 reigns, respectively.[3]

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.

Characteristics of Reigns

[3] 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.

Table A.1: Sample

Sample	Obs.
All reigns	366
Reigns with assessed *State Performance*	353
Reigns with assessed *Ruler Ability*	341
Reigns with information on border changes	317
Reigns with information on urban population	307
Both: *Ruler Ability* & *State Performance*	336
Both + individuals (gender assigned)	312
Both + coefficient of inbreeding (F)	235
Both: *Ruler Ability* & border changes	298
Both + coefficient of inbreeding (F)	203
Both: *Ruler Ability* & urban population	289
Both + coefficient of inbreeding (F)	198

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.

Dummy: Reign by Regents. For every reign, we identify whether the monarch in power was a regent. To this end, we 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. About a fifth of our sample are reigns by regents. 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).

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.

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: Educated as Heir. We assess from biographical accounts for every reign if the ruler had been educated to be the future heir or ruler. For instance, Albert II (who co-ruled Austria from 1330-1359) initially prepared for an ecclesiastical career and became a bishop early on. Only when his elder brother, Frederick, died did Albert become (together with his surviving brother) ruler of the Habsburg lands.[4] On the other hand, Albert's son, Rudolph (IV), was – as the oldest son – predetermined to become heir and likely educated as such. We presume that a ruler was educated as heir if s/he was the direct successor to the throne and there is no indication in the historical record suggesting that another career had been envisioned.

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.

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

[4]Albert co-ruled with his brother, the "obscure" Otto, which is why we focus on Albert instead when gathering additional variables for this reign.

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.

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.

A.3 Validation of Woods' State Performance and Ruler Ability Coding

To check Woods' (1906; 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

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	336
Δ Log(Area)	0.08	0.43	-1.11	3.48	294
Δ Log(Urb. Pop.)	0.16	0.47	-1.09	3.00	285

B. Main Explanatory Variables

	Mean	SD	Min	Max	N
Ruler Ability	0.00	1.00	-1.14	1.14	336
Inbreeding	0.02	0.99	-0.77	5.33	235
Hidden Inbreeding	0.03	0.99	-0.92	4.08	235

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*, and *Inbreeding* are standardized.

Table A.3: Summary Statistics – Other Variables

A. Reign Characteristics

	Mean	SD	Min	Max	N	Sources & Detail
Constrained Ruler	0.05	0.23	0.00	1.00	329	Section. 3.4
Length of reign (in years)	18.59	13.63	0.00	60.00	336	App. A.2
Regency	0.22	0.42	0.00	1.00	336	App. A.2
Dummy: Regicide	0.14	0.34	0.00	1.00	200	App. A.2
Conflict: Dummy	0.88	0.33	0.00	1.00	336	App. C.3
Conflict: Share Years at War	0.57	0.36	0.00	1.00	336	App. C.3

B. Ruler Characteristics

	Mean	SD	Min	Max	N	Sources & Detail
Dummy: Female	0.12	0.33	0.00	1.00	312	App. A.2
Dummy: Educated as Heir	0.78	0.42	0.00	1.00	141	App. A.2
Dummy: Hereditary succession	0.77	0.42	0.00	1.00	290	App. A.2
Dummy: Tall	0.10	0.30	0.00	1.00	312	App. A.2
Strength	0.01	1.20	-2.70	2.59	312	App. A.2
Number of children	0.09	1.01	-1.18	3.89	289	App. A.2
Age at Ascension	-0.04	0.90	-2.28	3.61	305	App. A.2
Age at Death	-0.02	0.87	-2.33	1.86	305	App. A.2
Noncognitive Ability	-0.09	1.17	-1.19	1.49	312	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.

ruler capability 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.

A.3.1 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 170 rulers. In 97 out of 170 assessed cases, our research assistant reached the same assessment as Woods did, while in 20 they reached the opposite assessment.[5] 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 234 reigns for which our research assistant was confident in making an assessment, in 124 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.3.2 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: Main RA – Woods: 0.53 (170); Main RA – RA1:

[5] 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.

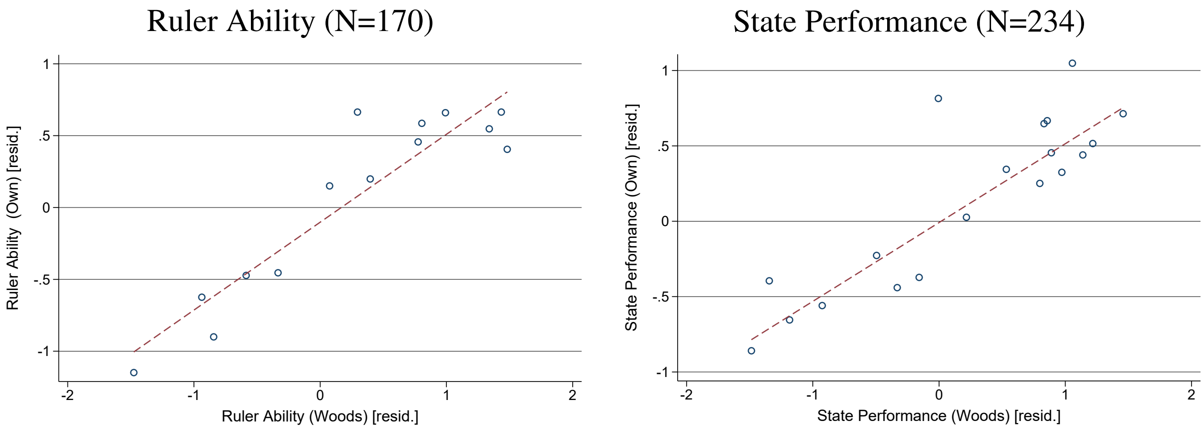

Figure A.2: Validation: Binscatters with state FE

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*.

0.53 (195); Main RA – RA2: 0.69 (30); RA1 – Woods: 0.50 (228); RA1 – RA2: 0.63 (32); RA2 – Woods: 0.44 (48). 2) for *State Performance*: Main RA – Woods: 0.50 (234); Main RA – RA1: 0.56 (253); Main RA – RA2: 0.63 (41); RA1 – Woods: 0.52 (237); 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.4 Territorial Changes: Details and Example

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.

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

A.5 Extended Sample: Coding

We extend Woods's original sample both in terms of time period and countries covered: We include Poland(-Lithuania) and Hungary, and we extend the coding until Word War I.[6] To do so – similar as in our validation in Appendix A.3 – we asked our main research assistant to assess the capability of rulers from all of the states covered by Woods who reigned after Napoleon until World War

[6] We add Poland (960-1795) and Hungary (1000-1526) because they are two large (and ultimately unsuccessful) states not covered by Woods. This addresses the concern that selection into our sample (by Woods) may have depended on the ultimate success of states, possibly affecting our findings. The fact that the inclusion of these two states does not change our coefficient of interest speaks (along with the state FE) against this concern.

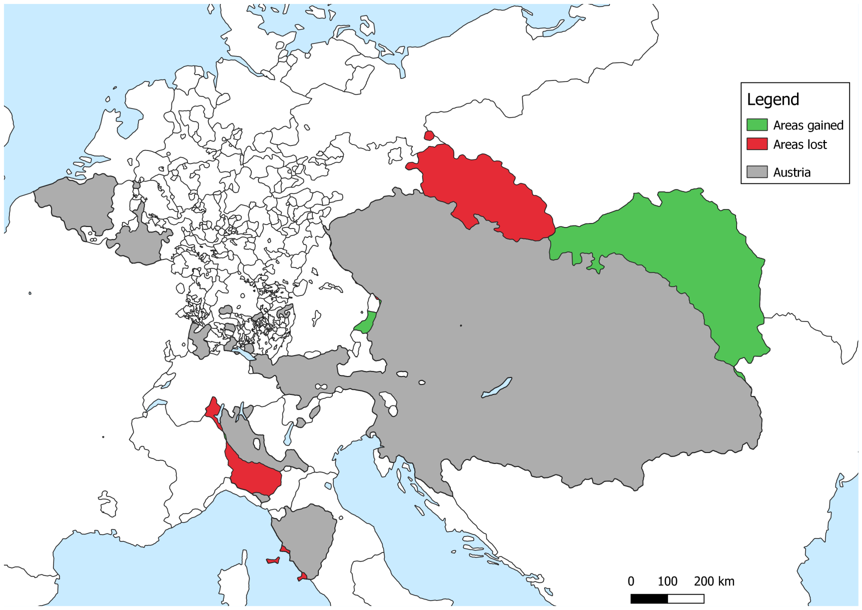

Figure A.3: Austria's Territorial Changes During the Reign of Maria Theresa

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.

I – or until the last monarch available or ruling before the start of WWI. For example, the list of monarchs of France ends with Napoleon III, who ruled from 1852 to 1870. We assess the ruler ability and *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 to also obtain coefficients of inbreeding from http://roglo.eu/. In addition, adding Poland and Hungary to the dataset gives 63 additional reigns (during our baseline sample period), for 40 of which we also recover the coefficient of inbreeding.

A.6 Data and Background on Coefficient of Inbreeding

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

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

where $paths$ is each path through which an individual can derive identical alleles from a common ancestor of both parents, n is the number of individuals in the paths (excluding the individual itself), and $1 + FA$ is a correction factor for the inbreeding coefficient of the common ancestor in the path. The 0.5 component comes from the fact that each individual has a 50% chance to pass on a particular allele to an offspring.

Figure A.4 provides an illustrative example of the calculation of the inbreeding coefficient. A is the offspring of B and another individual. Let us, for now, assume that the parents of individual A are unrelated, so that we do not have to apply a correction factor for the common ancestor B. Lines signify blood relationship. If A were to mate with B, the offspring I would be inbred. To calculate the coefficient of inbreeding for I, we first note that only one common ancestor exists, B, and only one path (a path is an entire circle, starting and ending at I: here $I \rightarrow A \rightarrow B \rightarrow I$). In this path, there are two individuals who are not I: A and B. Hence, the only path has length $n = 2$ and the coefficient of inbreeding of individual I is $F(I) = 0.5^2 = 0.25$. Were B inbred as well, we would have to adjust for that "ancestral" degree of inbreeding (FA) of B. Consider for instance the case in which B itself would similarly be the offspring of parent and child. Then, FA would be 0.25. Hence, $F = 0.5^2 \cdot (1 + 0.25) = 0.3125$.

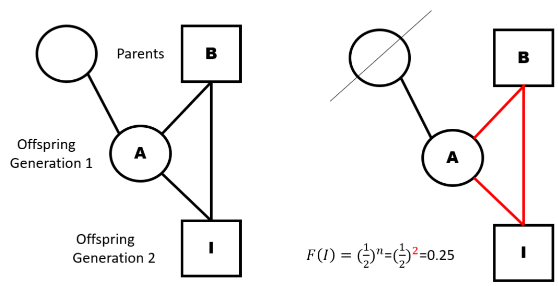

Figure A.4: Example Calculation of Inbreeding Coefficient

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

A.7 Details on Calculation of Hidden Component of Inbreeding

In our main analysis, the instrument is the coefficient of inbreeding, F. As is evident from the discussion in A.6 and the pedigree of Carlos II (Figure 1), high values of F need not necessarily imply very closely related parents. Instead, relationship links in temporal distance from the parents of an individual can build up over time, and account for a sizable share of the observed coefficient of inbreeding. In equation (A.1), this can be reflected in either one or more of the ancestors being inbred themselves (and thus showing up as FA in equation (A.1)), or by new paths opening up due to the often intricate relationship links between any of the person's ancestors (reflected as new *paths* in equation (A.1)). Consider Carlos II again. With $F = 25.36$, he is the monarch with the highest coefficient of inbreeding in our dataset. Yet, his parents were 'merely' uncle and niece, with most of the similarity in genes actually coming from a multitude of pathways through many common ancestors. The 'naive' coefficient of inbreeding of Carlos II, based on his parents being uncle-niece, would be $F = 12.5$, implying that more than half of the observed F of Carlos would require the knowledge of relationship links beyond that of his parents' generation. We calculate the 'hidden' component of F by subtracting the 'naive' coefficient of inbreeding, which is obtained from the closest relationship link between the rulers' parents:

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

where $F(naive)$ is 12.5 for monarchs whose parents were uncle and nieces (4 monarchs in total), and 6.25 for the (17) monarchs whose parents were (first) cousins. We obtain information on the blood relationship of rulers' parents from from *roglo.com*. For Carlos, we compute $F(hidden) = 12.86$.

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.3). 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	0.618***	0.422***	0.395***	0.461***
	(0.050)	(0.068)	(0.064)	(0.076)
State FE	✓	✓	✓	✓
R^2	0.41	0.22	0.24	0.25
Observations	336	224	176	159

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.3 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.[7] 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	0.768***	0.631***	0.770***	0.659***	0.498***
	(0.047)	(0.051)	(0.050)	(0.047)	(0.057)
State FE	✓	✓	✓	✓	✓
R^2	0.51	0.49	0.59	0.47	0.30
Observations	245	249	204	282	336

Note: This table documents 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 OLS and IV Results in the Extended Sample

Table A.6 presents our OLS and IV results for the extended sample for our main outcome variable *State Performance*. The corresponding coding is described in Appendix A.5. Columns 1-3

[7] 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.

show OLS estimates, columns 4-6 first-stage estimates, and columns 7-9 second-stage results. For comparison, columns 1, 4, and 7 repeat our results from the baseline sample. Columns 2, 5, and 8 use the sample of all states coded by Woods (1913), extended by our coding until WWI. The correlation between ruler ability and *State Performance* is slightly smaller in this extended sample, as is evident from column 2. In light of our results in Section 5 in the paper, this decrease in coefficient size may reflect the increase of executive constraints during this later time period. The first stage (col 5) is marginally weaker, while the second stage is somewhat stronger (col 8). A very similar picture emerges when we further add Poland(-Lithuania) and Hungary in columns 3, 6, and 9, respectively.

Table A.6: OLS and IV Results in the Extended Sample

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Specification:	OLS			First Stage			Second Stage		
Sample:	Woods	+ until WWI	+ PL, HU	Woods	+ until WWI	+ PL, HU	Woods	+ until WWI	+ PL, HU
Ruler Ability	0.618***	0.532***	0.554***				0.794***	0.903***	0.933***
	(0.050)	(0.044)	(0.042)				(0.148)	(0.098)	(0.096)
Coefficient of Inbreeding				-0.076***	-0.057***	-0.060***			
				(0.011)	(0.009)	(0.009)			
R^2	0.41	0.34	0.34	0.15	0.11	0.12	0.39	0.23	0.23
Observations	336	374	437	235	264	304	235	264	304

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

C 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 possible alternative mechanisms related to path dependence in state performance, strategic marriage (inside or outside of the kin network), as well as conflict.

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

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

Panel A of Table A.7 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. The one exception is the lagged change in urban population in column 6, which features a marginally significant negative coefficient. This may indicate that past expansions into urbanized territory render it more difficult to immediately add further such expansions in the next reign. Panel B A.7 shows that the 2SLS results are also robust to controlling for lagged state performance measures. Note that the IV coefficient in column 6 is no longer significant at common levels. However, in (unreported) regression results in which we include the 43 monarchs whose parents share no *known* relationship with each other (i.e., $F = 0$ – see footnote 32 in the paper), the p-value is 0.048.

Table A.7: 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***	-0.257***	-0.052***	-0.034***	-0.056***	-0.031
	(0.045)	(0.043)	(0.015)	(0.009)	(0.014)	(0.020)
L.State Performance		-0.029				
		(0.070)				
L.Δ Log(Area)				0.049		
				(0.057)		
L.Δ Log(Urb. Pop.)						-0.067*
						(0.035)
State FE	✓	✓	✓	✓	✓	✓
R^2	0.11	0.11	0.10	0.04	0.07	0.05
Observations	235	218	203	190	198	184
			B. Second Stage Regressions			
Ruler Ability	0.794***	0.806***	0.176***	0.115***	0.194***	0.105
	(0.100)	(0.147)	(0.051)	(0.035)	(0.051)	(0.068)
L.State Performance		0.034				
		(0.068)				
L.Δ Log(Area)				0.054		
				(0.045)		
L.Δ Log(Urb. Pop.)						-0.053
						(0.040)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	42.2	32.9	39.1	42.2	34.8	32.7
Observations	235	200	203	190	198	184

Note: The table shows the results of reduced-form and second-stage regressions, controlling for lags 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.

Table A.8 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.7 and A.8 suggest that neither of the conditions required for strategic kin marriage to affect our exclusion restriction are fulfilled.

Table A.8: Past State Performance as Confounder: First Stage

Dep. Var.: Ruler Ability

	(1)	(2)	(3)	(4)
Inbreeding	-0.314***	-0.326***	-0.284***	-0.285***
	(0.051)	(0.049)	(0.048)	(0.051)
L.State Performance		-0.065		
		(0.101)		
L.Δ Log(Area)			-0.061	
			(0.153)	
L.Δ Log(Urb. Pop.)				-0.095
				(0.157)
State FE	✓	✓	✓	✓
R^2	0.15	0.15	0.12	0.11
Observations	240	222	201	196

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.

C.2 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 $F = 0$ in the next generation.

Note that we exclude monarchs with (likely) completely unrelated parents from our baseline IV analysis (see footnote 32 in the paper). In Table A.9 below, 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.[8] 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 $F = 0$, especially for $\Delta log(Area)$. This is not the case: our results remain very similar

[8] 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 38, respectively.

to the baseline and statistically highly significant.

Table A.9: 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 $F=0$	Baseline	Include $F=0$	Baseline	Include $F=0$
Ruler Ability	0.794***	0.814***	0.176***	0.173***	0.194***	0.236***
	(0.100)	(0.099)	(0.051)	(0.051)	(0.051)	(0.053)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	42.18	38.35	39.07	41.76	34.80	36.51
Observations	235	278	203	243	198	236

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 $F = 0$. 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.

C.3 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 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.[9] 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* (albeit insignificant) association with conflict. Similarly, in our core sample, states

[9]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 first cousins with a coefficient of inbreeding of 11.45.

under more inbred rulers had a higher probability of conflict and a higher share of years under conflict.[10] 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.[11] We perform this analysis in Table A.10. 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.10 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").[12] 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 territorial changes in state performance, but also that of other aspects of state performance, beyond territorial changes.

C.4 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 de-

[10] We examine this further below, documenting in Table A.18 that higher ruler ability (due to lower inbreeding) reduced conflict.

[11] The underlying data are from David Brecke's Conflict Catalogue (available from https://brecke.inta.gatech.edu/research/conflict/) and start in 900 AD. We first identify whether a state participated in any conflict (in Europe) within a given year. Then, we calculate the share of years of each reign in which a state participated in a conflict.

[12] See Appendix D.3 for detail.

Table A.10: 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	0.794***	0.769***	0.731***	0.734***	0.430**
	(0.100)	(0.093)	(0.101)	(0.109)	(0.203)
Conflict: Dummy		-0.156			
		(0.192)			
Conflict: Share Years at War			-0.201		
			(0.173)		
State FE	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	42.18	35.53	25.18	44.51	25.07
Observations	235	235	235	199	117

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).

scendants 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 3rd of the Spanish Bourbons. Yet, he also hails from the Bourbon dynasty ruling France. He is the 8th of all Bourbons, ordered by the year in which his reign began. We account flexibly for the potential importance of dynasty and founder effects by including fixed effects for the order of monarchs within their dynasties. Column 1 in Table A.11 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.

C.5 IV Results for Hereditary Succession Sample

Table A.12 presents second-stage 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

Table A.11: IV Regressions Accounting for Monarch's Order in Dynasty

Dep. Var.: *State Performance*

	(1)	(2)	(3)	(4)
Sample	Baseline	—— Known Order in Dynasty ——		
Ruler Ability	0.794***	0.845***	0.893***	0.619**
	(0.100)	(0.116)	(0.121)	(0.258)
State FE	✓	✓	✓	✓
Order in Dynasty FE			✓	
Order in International Dynasty FE				✓
First Stage Effect. F-Stat	42.2	40.6	47.1	8.8
Observations	235	231	231	231

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 3rd of the Spanish Bourbons. Yet, he also hails from the Bourbon dynasty ruling France. He is the 8th of all Bourbons, ordered by the year in which his reign began. The table 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 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

in terms of magnitude and statistical significance.

Table A.12: IV Results – Hereditary Succession Sample

Dependent variable as indicated in table header

	(1)	(2)	(3)
Dep. Var. :	*State Performance*	$\Delta log(Area)$	$\Delta log(UrbPop)$
Ruler Ability	0.848***	0.155***	0.148**
	(0.139)	(0.053)	(0.073)
State FE	✓	✓	✓
First Stage Effect. F-Stat	26.4	16.6	15.7
Observations	191	164	162

Note: This table shows the results of second-stage regressions for our three outcome variables for the sample of reigns where the monarch ascended to the throne by hereditary succession. The table reports the first-stage effective F-statistic from the Montiel Olea and Pflueger (2013) robust weak instrument test; the corresponding critical value for max. 10% relative bias is 16.4. All regressions are run at the reign level and include state fixed effects. Standard errors clustered at the state level in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

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.13 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.

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.13, 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.13: IV Result: Selection Among Offspring

Dep. Var.: *State Performance*

	(1)	(2)	(3)
Sample:	All	Firstborn Sons	No Competing Claims
Ruler Ability	0.794***	0.738***	0.642***
	(0.100)	(0.270)	(0.183)
State FE	✓	✓	✓
First Stage Effect. F-Stat	42.18	19.28	12.77
Observations	235	120	145

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% relative bias is approximately 16.4. Standard errors clustered at the state level. * p<0.1, ** p<0.05, *** p<0.01.

D Details on Potential Mechanisms

In this appendix we 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 were affected by inbreeding and how these, in turn, affected state performance. We show that the effect likely stems from 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.

D.1 Inbreeding and Ruler Characteristics

Table A.14 shows the results of regressing various characteristics rulers on *Inbreeding* (which is standardized). 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. Except for non-cognitive ability, we document small associations throughout.

Columns 1 and 2 in Table A.14 show that inbreeding does not have a sizable effect on the height or physical strength of rulers: the coefficients are small and statistically insignificant. Columns 3-6 show similarly small statistically insignificant coefficients for the number of children of a monarch, his or her age at death, length of reign, and age at ascension to the throne. In column 7 we document that rulers who ascend to power after regicide of the prior monarch (murdered in office or executed after a trial) are significantly less inbred. This is likely explained by the fact that new dynasties came to power after the deposition of former monarchs. However, again the estimated coefficient is small in magnitude. Column 8 shows that regents were not more or less inbred than other rulers. This makes sense since regents were often close relatives while the heir was still a minor. Finally, column 9 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.

D.2 Controlling for Physical Ruler Attributes and Non-Cognitive Ability

We have shown that inbreeding had negative consequences for both cognitive and non-cognitive abilities of early modern rulers (see Table 4 in the paper and Table A.14, col 9). 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 (Alvarez, Ceballos, and Quinteiro, 2009; Álvarez et al., 2019). In what follows, we control for these characteristics directly in our IV regressions for *State Performance*. Column 1 of Table A.15 shows our baseline IV regression for comparison. In columns 2 and 3, we control for whether a monarch was tall or strong. In columns

Table A.14: Results: Correlates of Inbreeding

Dependent variable as indicated in table header

	(1) Tall‡	(2) Strong§	(3) Number of Children	(4) Age at Death	(5) Length of Reign	(6) Age at Ascension	(7) Regicide of prev. ruler†	(8) Regency	(9) Non-cognitive Ability
Inbreeding	0.008	-0.034	-0.098	0.023	0.076	-0.037	-0.028**	-0.008	-0.277***
	(0.021)	(0.137)	(0.060)	(0.069)	(0.055)	(0.039)	(0.011)	(0.020)	(0.075)
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
R^2	0.13	0.05	0.04	0.08	0.06	0.04	0.06	0.07	0.12
Observations	256	256	243	256	256	256	149	256	256

Note: The table shows results from regressions of ruler characteristics on *Inbreeding*. All dependent variables, except for the dummy variables "Tall", "Regicide", and Regency" are standardized, and so in inbreeding in all columns. 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 physically "weak" (-1), "strong" (1), or neither (0).
† Dummy indicating whether the *prior* ruler was murdered or executed after a trial.

4 and 5 we directly control for the number of children and the age at death of the monarch (as before, both are standardized). Our main coefficient of interest, that of ruler ability, is unaffected by the inclusion of any of these controls. This renders it unlikely that physical features caused by inbreeding drive our IV results. Columns 6 and 7 directly control for the standardized length of a ruler's reign and the age of ascension to the throne. Again, the main coefficient of interest is barely affected. Controlling for whether the prior monarch was deposed through a regicide – in column 8 – lowers the coefficient of interest, but it remains sizable and highly significant. Column 9 controls for whether a regent was in power. While this variable itself is negatively associated with *State Performance*, its inclusion in the IV regressions does not change our coefficient of interest.

Finally, in column 10 we control for non-cognitive ruler ability. Recent work emphasizes the importance of both cognitive and non-cognitive intellectual abilities (Lindqvist and Vestman, 2011; Heckman, Stixrud, and Urzua, 2006). The "Ruler Ability" measure from Woods (1913) explicitly aimed at capturing cognitive abilities.[13] The inclusion of this variable in column 9 barely affects our main coefficient of interest – (cognitive) ruler ability. This underlines the importance of cognitive ability.

In sum, our results suggest that inbreeding rendered leaders ineffective due to its negative effect on cognitive capabilities, rather than physical or non-cognitive features.

[13] 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.

Table A.15: IV Results Controlling for Physical Features of Monarchs and Aspects of their Reign

Dep. Var.: *State Performance*

Controlling for X=	(1)	(2) Tall‡	(3) Strong§	(4) Number of Children	(5) Age at Death	(6) Length of Reign	(7) Age at Ascension	(8) Regicide of prev. ruler†	(9) Regency	(10) Non-cognitive Ability
Ruler Ability	0.794***	0.793***	0.798***	0.711***	0.794***	0.820***	0.805***	0.572***	0.804***	0.775***
	(0.100)	(0.099)	(0.097)	(0.103)	(0.098)	(0.088)	(0.104)	(0.138)	(0.107)	(0.143)
X		-0.061	-0.028	0.082	0.034	0.158***	-0.085	-0.144	-0.301**	0.024
		(0.157)	(0.040)	(0.093)	(0.121)	(0.037)	(0.076)	(0.220)	(0.147)	(0.060)
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	42.18	40.89	44.66	44.53	46.14	44.93	42.50	37.58	44.44	11.84
Observations	235	235	235	225	235	235	235	136	235	235

Note: The table controls for various ruler characteristics in our IV regressions. 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.
‡ "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 physically "weak" (-1) or "strong" (1).
† Dummy indicating whether the *prior* ruler was murdered or executed after a trial.

D.3 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.

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.16 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.16, 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.

Next, we examine the effect of ruler ability on economic aspects of state performance. Table A.17 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.

D.4 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 David Brecke's Conflict Catalogue and start in 900 AD.[14] 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

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

Table A.16: 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	0.678***	0.498***	0.323*	0.574***	0.320	0.211	0.478**	0.438**
	(0.208)	(0.160)	(0.185)	(0.220)	(0.271)	(0.163)	(0.226)	(0.178)
State FE	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	46.62	46.62	46.62	46.62	46.62	46.62	46.62	46.62
Observations	235	235	235	235	235	235	235	235

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 45.9; 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.

Table A.17: 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	0.408**	0.583**	0.868***	0.102	0.050
	(0.185)	(0.263)	(0.253)	(0.234)	(0.169)
State FE	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	46.62	46.62	46.62	46.62	46.62
Observations	235	235	235	235	235

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.

dummy for at least one conflict during a reign, and the share of years of each reign in which a state participated in a conflict.

Column 1 in Table A.18 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.18 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 5 in the paper). The most likely explanation for these findings is that – on average – capable rulers were better at selecting external wars that promised territorial expansions, while they avoided those that would likely have been costly.

Table A.18: IV Results: Conflict as Outcome

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

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var:	All Conflicts		Internal		External	
	Dummy	Share	Dummy	Share	Dummy	Share
Ruler Ability	-0.315***	-0.160**	-0.054	-0.034	-0.323***	-0.136**
	(0.090)	(0.071)	(0.055)	(0.073)	(0.084)	(0.056)
State FE	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat	37.65	37.65	37.65	37.65	37.65	37.65
Observations	240	240	240	240	240	240

Note: The table shows 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.

D.5 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 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.19 shows the results for log changes in total urban population (cols 1-3) as well as for its intensive (cols 4-6) and extensive (cols 7-9) components. For each outcome, we first report the OLS results in the full sample, followed by OLS results in the "IV sample," (i.e., reigns for which we have information on the coefficient of inbreeding), followed by the IV results. Column 1 shows a sizeable correlation between ruler ability and the overall change in urban population. The result is very similar in the subsample in column 2. The IV result in column 3 shows a large and highly significant coefficient, replicating our baseline result from Table 5 in the paper that a one standard deviation increase in the ability of a monarch raises urban population by almost 20%. Interestingly, this 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.'[15] 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, 2017, ch. 7). At the same time, 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).

E Detail on Coding State-Year Level Constraints on Executive

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

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

Table A.19: 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	0.098***	0.105**	0.194***	0.022**	0.019	0.004	0.077**	0.086**	0.190***
	(0.027)	(0.041)	(0.051)	(0.009)	(0.013)	(0.030)	(0.026)	(0.038)	(0.063)
State FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
First Stage Effect. F-Stat			34.8			34.8			34.8
R^2	0.10	0.10	0.07	0.11	0.09	0.08	0.07	0.10	0.04
Observations	289	198	198	289	198	198	289	198	198

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.

at the 100- and 50-year interval from 1000 CE until 1850. They base the measure on an encyclopedia of world history (Langer, 1972; Stearns and Langer, 2001). We follow their approach, but additionally identify the exact year when constraints on the executive changed. We code this measure for all states in our dataset except Turkey, which is not covered by these sources.[16] The categories of "constraints on the executive" range from 1 to 7, where 1 indicates unlimited authority of the monarch and 7 indicates "Executive Parity or Subordination" to other branches of government. We define an indicator of a monarch being constrained when constraints on the executive are at least 5 – "Substantial Limitations on Executive Authority."

As another illustrative example, consider France in the 18th century. Until 1789, our secondary source (Stearns and Langer, 2001) describes several events that demonstrate the complete absence of constraints on the French monarch. For instance, in 1749, a "five percent tax on all incomes [was] introduced, but the clergy were exempted after they protested. The Parliament objected to their exclusion but the king overruled that objection" (Stearns and Langer, 2001, p.329). This period therefore receives a "constraints" coding of 1. The constraints on the monarch changed in 1789, when the "[t]he king suspended meetings of the [newly-assembled] Estates General for three days and closed the hall. Members met at a neighboring tennis court and took an oath not to separate until they had given the realm a constitution. June 23. The king ordered each estate to meet separately, but deputies refused" (Stearns and Langer, 2001, p.431). These and other developments

[16]Note that for the first three reigns in France recorded by Woods, as well as for the first reign in Castile, Portugal, Prussia, Russia, and Scotland, we do not have information on constraints in our sources and thus code these as 'missing.' Treating these reigns as unconstrained leaves our result unaffected (unreported).

indicate that – as the French Revolution started – some constraints on the executive emerged. Therefore, we code France as having "Slight to Moderate Limitation on Executive Authority" (a coding of 3) starting in 1789. In 1792, Stearns and Langer (2001, p.431) then indicates strong constraints on the executive: "Executive Parity or Subordination: The Accountability groups have effective authority equal to or greater than the executive in most areas of activity," as the National Convention suspended "all functions of the monarchy. The assembly voted to enact all legislation vetoed by the king, who was confined to the Temple." This yields a "constraints" coding of 7 after 1792. Note that the century-level measure of Acemoglu et al. (2005) retains a level of 1 throughout the 18th century.

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.

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

Angelucci, C., S. Meraglia, and N. Voigtländer (2017). How Merchant Towns Shaped Parliaments: From the Norman Conquest of England to the Great Reform Act. Working Paper 23606, National Bureau of Economic Research.

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.

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

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

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.

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.

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

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.

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

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

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

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

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

Stearns, P. N. and W. L. Langer (2001). *The Encyclopedia of World History: Ancient, Medieval, and Modern, chronologically arranged*. Houghton Mifflin Company.

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

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

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

Quick Guide to Identification, Mechanisms, and Measurement Concerns

Section 4 presents empirical evidence that supports a causal relationship between ruler ability and state performance. The tables below provide a brief overview of how we address possible concerns with our identification strategy and the measurement of central variables in our analysis.

Evidence in Support of a Ruler-Ability Mechanisms	
We argue in Section 4 that inbreeding reduced the cognitive ability of rulers, which in turn had negative effects on state performance under their reign. Here we summarize the evidence in support of this mechanism.	
Type of evidence: **H**istorical/**E**mpirical	Description of Mechanism and Evidence
E: Appendices D.3, D.4, and D.5	*Inbred monarchs were bad for state performance because they pursued ineffective policies.* By decomposing state performance into different components, we show that incapable monarchs reduced law and order, administrative capacity and the diplomatic position of their state; they also reduced commerce and agricultural activity, as well as general living conditions (Tables A.16 and A.17). We also show that less inbred (and thus capable) rulers participated in fewer international conflicts, yet they acquired more urbanized territories (Tables A.18 and A.19). This suggests that capable rulers chose more wisely which international conflicts to engage in.
H	*Inbred monarchs were bad for state performance because they selected incapable advisers.* This interpretation is in line with our argument. Note that selection of capable advisers is a core task of national leaders, and failure to do so reflects upon their own capability. For example, Charles II was not only incapable of governing personally, but also of selecting able ministers (Hamilton, 1938, p. 174).
E: Appendix D.2	*Does cognitive or non-cognitive ability drive our results*? Woods (1913) specifically aimed at assessing the cognitive ability of monarchs, and research in biology documents a strong negative effect of inbreeding on cognitive abilities. Yet, recent research also emphasizes non-cognitive intellectual ability as a determinant of leadership. In Appendix D.2, we document that controlling for rulers' non-cognitive abilities in our IV regressions does not affect our main coefficient of interest (i.e., cognitive ruler ability).

Identification Concerns *Our IV strategy exploits two features: i) Pre-determined hereditary appointment of rulers independent of their ability; ii) the fact that inbreeding led to variation in ruler ability. The exclusion restriction is that inbreeding affected state performance only through its impact on the ability of the hereditary ruler. This table discusses potential concerns with our identification strategy.*	
Type of evidence: **H**istorical/**E**mpirical	Description of Concern and Evidence
H: Section 2.3 **E**: Section 4.5	*Could monarchs have known and anticipated the negative effects of inbreeding and thus avoided them?* This is historically unlikely: A common belief among European monarchs was that consanguineous marriages were beneficial because they helped to keep a "pure" blood line. See footnote 44 for a (late) counterexample from the 19th century showing that even were skepticism about inbreeding was expressed, it resulted from superstition in idiosyncratic cases, rather than from systematic concerns. Note also that the actual degree of inbreeding was very difficult to anticipate, as it requires both the mathematical tools (only devised by Wright in 1921, cf. footnote 4) and detailed pedigrees hailing back generations. But even if monarchs implicitly approximated such a measure, they would have tended to apply it towards a preference for *more* inbreeding. In addition, the degree of a ruler's inbreeding was determined at his/her parents' generation. We can thus address remaining concerns by excluding any inbreeding from the parents' generation when we focus on 'hidden' inbreeding from previous generations (see Section 4.5).
E: Section C.1 and Section 4.4	*Strategic marriage within kin networks?* Did unsuccessful monarchs strategically choose more closely related marriage partners to fend off external threats? (a) Note that this decision must have been made in a ruler's parents' generation. Thus, to bias our results, this concern requires that state performance in the previous reign was correlated with subsequent state performance under the heir's reign. We directly address this concern in Appendix C.1, where we show that controlling for lags of state performance leaves our main results unaffected. (b) Our analysis in Section 4.4 further addresses concerns about strategic marriage. There, we only use the component of inbreeding "hidden" in the pedigree (i.e., excluding the part of inbreeding that resulted from the generation of a ruler's parents). We show that our results remain unaffected.

Identification Concerns (continued)	
Type of evidence: **H**istorical/**E**mpirical	Description of Concern and Evidence
E: Appendix C.2 and C.3	*Strategic marriage outside of kin networks?* Did able monarchs aim to strategically "marry out" in order to expand their realms? (a) Note that marrying completely unrelated partners would result in a zero inbreeding coefficient. In the baseline analysis we exclude these cases (because they have no known family relationships). Including these cases does not affect our results, as Appendix C.2 documents. (b) In Appendix C.3 we further exclude the territorial component from our measure of state performance (by residualizing Woods' comprehensive assessment of state performance with our measure of territorial change) and show that our main results hold.
H/E: Appendix C.3	*Conflict among related dynasties?* If related dynasties were more likely to fight wars with each other and were more inbred, this could pose a threat to our identification. (a) Note first that conflict a priori likely only increases the variance of state performance, and does not systematically lead to territorial acquisitions or losses. (Apart from the effect of ruler ability on the probability of winning, which leads back to our main argument). (b) We can further address concerns related to our outcome – state performance – resulting from (successful or unsuccessful) conflicts by excluding territorial acquisitions from our main outcome and by controlling for conflict in our regressions (Table A.10). Our main results are essentially unaffected, as shown in Appendix C.3.
H/E: Appendix C.4	*Founder effects?* Could our main result be driven by regression to the mean from early successful founders to incapable and mechanically more inbred later rulers within dynasties? We address this concern by controlling for the order of individual monarchs within their dynasties in Table A.11.

Ruling out Alternative Mechanisms

We argue in Section 4 that inbreeding reduced the cognitive ability of rulers, which in turn had negative effects on state performance under their reign. Here we summarize potential alternative mechanisms and evidence that speaks against them.

Type of evidence: **H**istorical/**E**mpirical	Description of Concern and Evidence
H: Tables A.14 and A.15	*Inbreeding reduced fertility, thus creating a threat to succession.* We find only a weak association between rulers' coefficient of inbreeding and the number of (surviving) children in Table A.14 in our sample. Also, controlling for the number of (surviving) offspring leaves our IV results unaffected (see Table A.15).
E: Tables A.14 and A.15	*Inbreeding reduces longevity and thus tenure, leading to a threat to succession, as well as to fewer opportunities for rulers to prove themselves.* Table A.14 shows that – at least in our sample – inbred rulers did not have significantly shorter tenure or longevity. Table A.15 controls for longevity in our IV specification, leaving our results unaffected.
E: Tables A.14 and A.15	*Inbreeding depressed body height and physical appearance, leading to lower perceived (rather than actual) capability.* Table A.14 documents that inbred rulers were not systematically less likely to be assessed as tall or as having a strong physical appearance. In Table A.15 we also control for height and physical strength, showing that our IV results are not affected.
H/E: Table A.13	*Did monarchs select more able successors among their (many) offspring and potential heirs?* For instance, monarchs may have systematically deposed of incapable older sons to clear the way for younger, more capable heirs. (a) If such a selection of more capable offspring had taken place among the offspring from the *same* marriage, it would merely reduce variation in our first-stage relationship, as all offspring from the same parents share an identical coefficient of inbreeding. (b) Selection of more capable offspring from *different* (and potentially less consanguineous) marriages could pose a concern. We address this in Table A.13 where we restrict attention to those monarchs who were the first-born offspring of their parents (or the second-born in case the first born was female). While this reduces our sample size, our main finding remains sizable and significant.