Online Appendix for *The Great Revenue Divergence*

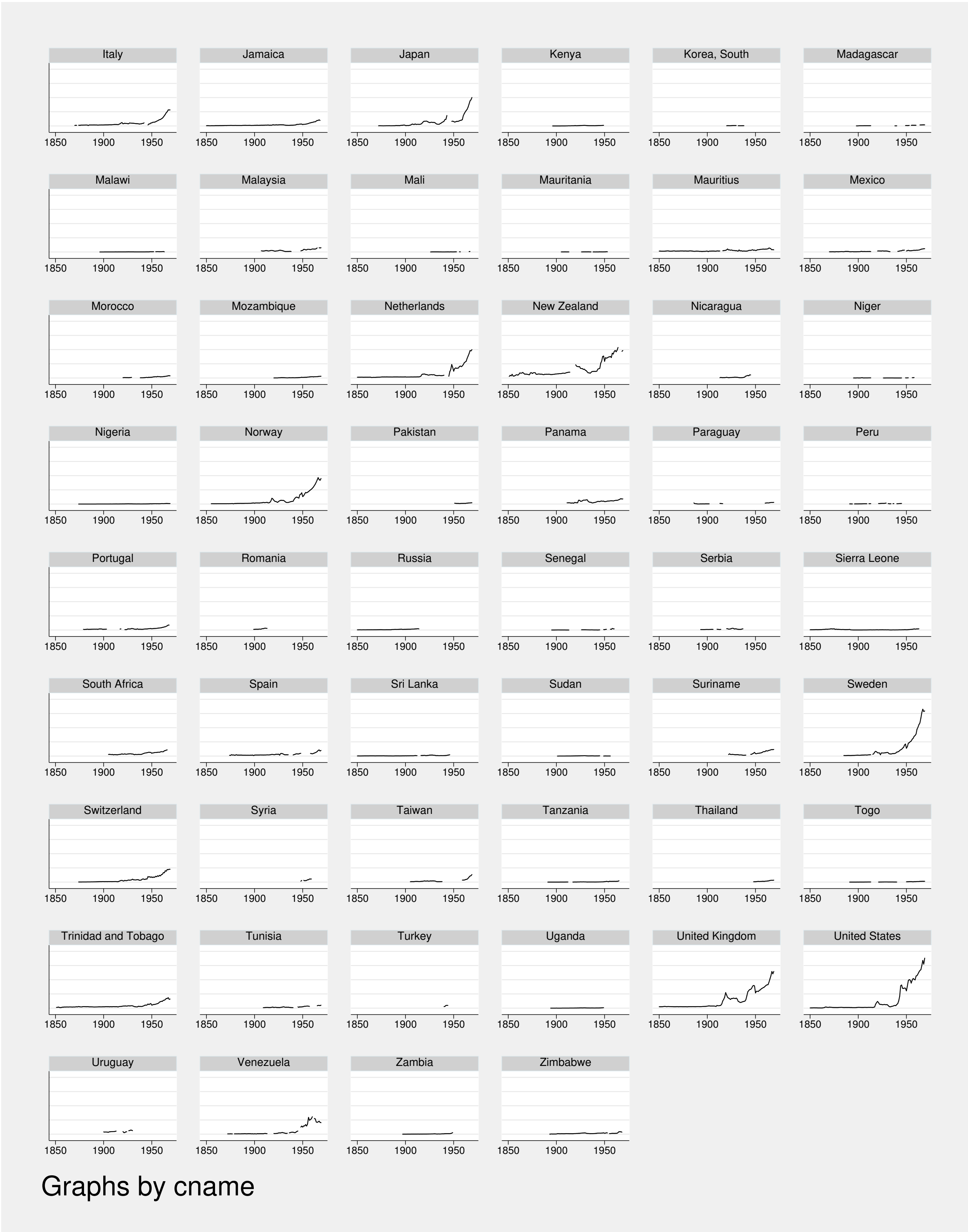
Contents

List of Tables

List of Figures

A Supporting Information on Data and Patterns

Figure A.1: Revenues Per Capita by Country, 1850–1969



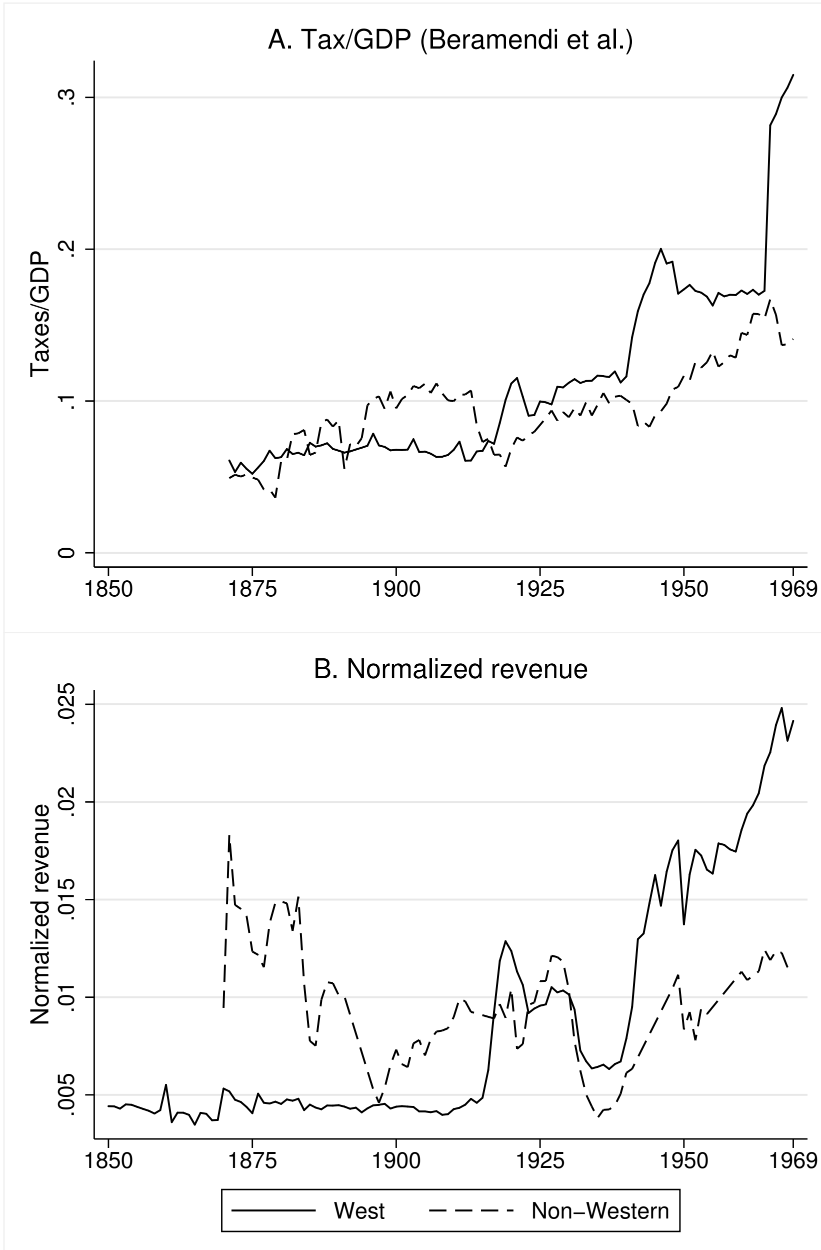
Notes. For each panel in Figure A.1, the lines show estimated central government revenue per capita in gold grams, converted at nominal exchange rates. Gaps indicates years of missing data. The range of the y-axis is 0 to 800.

A.1 Accounting for Differences in Income

In Panel B of Figure 1, we depict patterns for taxes/GDP over time using data from Andersson and Brambor (2019). Here we show a similar pattern for two alternative measures that account for GDP. Figure A.2 contains two panels. Panel A uses taxes/GDP data from Beramendi et al. (2019). Compared to Andersson and Brambor (2019), they contain fewer South American countries (only Argentina, Brazil, Chile, Mexico, and Uruguay before 1920) and, as noted in the text, fewer country-years overall. For Panel B, we constructed a "normalized" revenue variable with non-natural units: we divide our data on nominal revenues by constant-dollar GDP estimates from Bolt et al.'s (2018) update of Angus Maddison. Unfortunately, because we do not have the GDP data in nominal local-currency denominations, we cannot directly combine their data with the Mitchell revenue data to calculate revenues as a fraction of GDP. Mitchell (1998) provides some data points for nominal GDP in the local currency, although these data are unavailable for almost every non-Western country before 1950. The skewed sample makes it difficult to assess the robustness of our core pattern using this source.

In both alternative datasets, Western and non-Western countries experience a reversal of fortunes. In 1913, taxes/GDP in the West was 48% lower than taxes/GDP in South American countries according to Beramendi et al.'s (2019) measure (6.1% vs. 10.7%). Similarly, normalized revenue was 46% lower in Western countries. By 1969, these advantages had flipped. Taxes/GDP was 2.2 times greater in Western countries (31.5% vs. 14.1%), and normalized revenue was 2.1 times greater. Thus, these replicate the core finding of a great revenue divergence in the twentieth century, despite a restricted sample in one dataset and non-natural units of analysis in the other.

Figure A.2: Alternative Measures: Accounting for Differences in Income



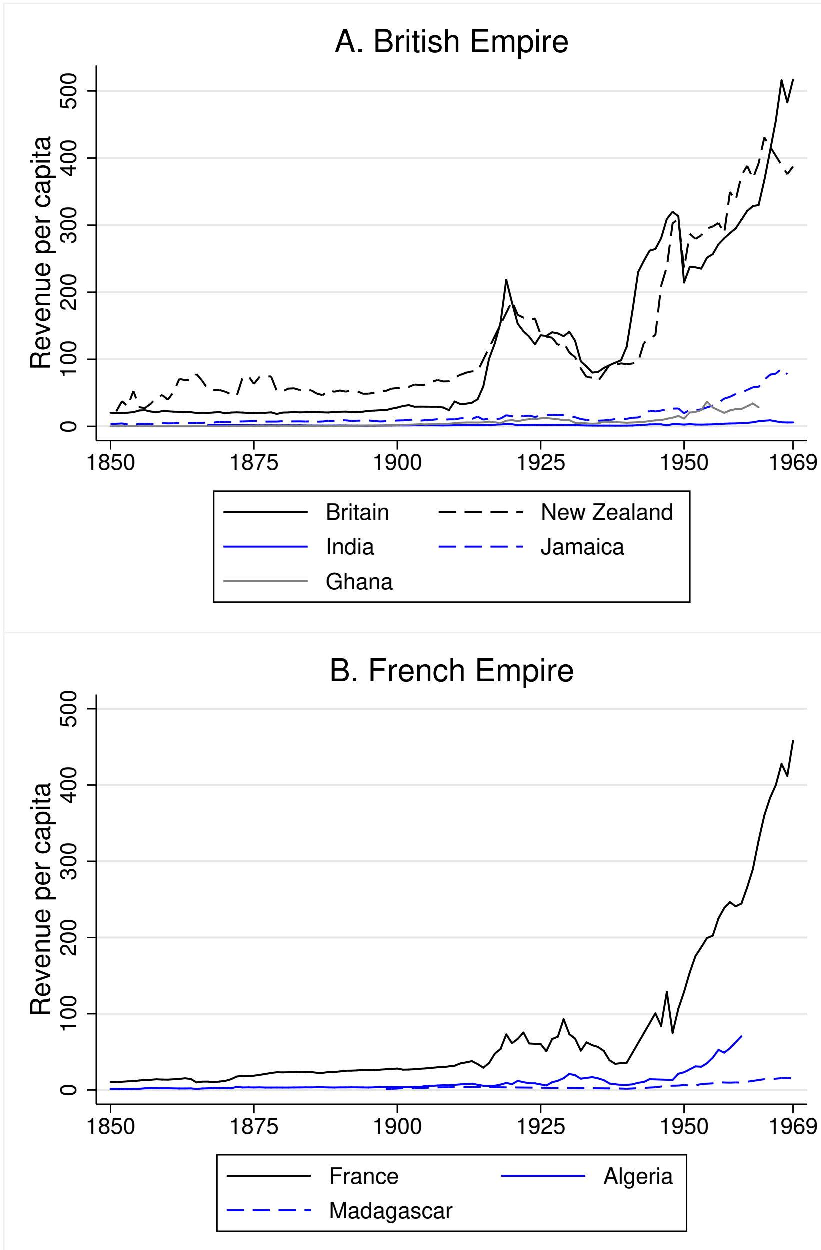
Notes. Taxes/GDP from Beramendi et al. (2019). Panel B: normalized revenue. The above text provides more details on each variable.

A.2 Exchange Rate Effects and Within-Empire Comparisons

One possible concern with our main measure of revenues is that, by using nominal exchange rates, longitudinal changes in revenue per capita may reflect changes in the foreign exchange market rather than changes in actual revenue. In the short term, the data exhibit many sharp short-term changes that clearly reflect currency revaluations. Two of our scope conditions address this concern: we examine only data through 1969, when the Gold Standard and Bretton Woods regimes stabilized exchange rates; and we excluded currencies for which the published exchange rate was grossly manipulated and not convertible (e.g., the Soviet ruble).

To provide more direct evidence that fluctuations in the foreign exchange market do not qualitatively alter the main pattern, here we present within-empire comparisons. Exchange rates remain constant over time within these samples, either because the colonies used the same currency as the mother currency or a highly stable peg. Figure A.3 examines the British Empire, and each of the five countries (except for India before 1899 and after 1947) used sterling or a currency pegged to sterling throughout the period. Although New Zealand and Britain had higher levels of revenue per capita than the other colonies in 1913, these differences were small by modern standards; and they grew immensely over time. For example, per capita revenue in Britain was 3.2 larger than in Jamaica in 1913, and by 1968 it was 6.2 times larger. Normalized revenue was two times larger *in Jamaica* in 1913 than in Britain. However, by 1968, the advantage had flipped: normalized revenue was 2.2 times larger in Britain than in Jamaica. Similar trends are apparent within the French Empire.

Figure A.3: Within-Empire Comparisons



Notes. The lines show estimated central government revenue per capita in gold grams, converted at nominal exchange rates.

A.3 Price Effects

Even after we account for artificial exchange rates or short-term fluctuations in exchange rates, our comparisons do not capture differences in prices. Ideally, we would normalize currencies using a purchasing power index that measures state revenue at purchasing power parity. However, the rarity of reliable price data prior to the late twentieth century—let alone price data comparable across nations—implies that accounting for prices would severely constrict the sample and would make impossible many of the illuminating historical comparisons that we present. Cross-national purchasing power data are available only since 1950 (Feenstra et al., 2015), after the great revenue divergence we identify had already occurred.

However, differences in purchasing power are unlikely to explain our pattern for three reasons. First, the differences are still present when we measure state revenue as a percentage of GDP.

Second, differences in purchasing power in 1950 were modest compared to the differences in revenue that we observe. Although purchasing power in South Africa was 73% more than Britain in 1955, nominal per capita revenues were 441% higher in Britain than in South Africa. More broadly, there do not seem to be systematic differences in purchasing power across categories of countries. In 1950, average GDP purchasing power conversion factors were similar in Western Europe and East Asia compared to the rest of the world (0.102 versus 0.91).

Finally, the regression models in Appendix B with country fixed effects account for time-invariant cross-national differences in purchasing power. To confound the divergence trend, purchasing power would also have to diverge over time, with nominal revenue in the West increasing precipitously despite the real purchasing power of that revenue remaining static (at least relative to the non-Western world). The available data do not support this story, although we are limited to countries with PPP data in 1950 and afterwards in the Penn World Table dataset (Feenstra et al., 2015). Although purchasing power increased in the West relative to the rest of the world in this period, this increase was modest relative to differences in per capita nominal revenue increases. Between 1950 and 1968, the GDP conversion factor increased by 71% in the West compared to 20% elsewhere. However, during this period, revenue increased by 294% in the Western compared to 18% elsewhere.

A.4 Non-Tax Revenue

Conventional sources of tax revenue based on taxing output (head taxes, trade taxes, income taxes) do not provide the only possible source of government revenues. Governments may also benefit from natural resource production, foreign aid, and remittances from expatriates. A large literature documents the empirical importance of "rentier" revenue sources and examines their effects on political outcomes (Ross, 2012; Morrison, 2014; Menaldo, 2016). Alternatively, states can (at least in the short term) substitute for taxes by borrowing (Centeno, 2002; Queralt, 2019), which was a common strategy earlier in European history.

Although we not dispute the importance of non-tax revenues for many political outcomes, we not engage with them in depth here because they are unlikely to explain our core pattern. The West began to distinguish itself from the rest of the world in the early twentieth century because of its countries' superior ability to increase tax revenues (Scheve and Stasavage, 2016), not because of their superior exploitation of natural resources, which was not especially high in these countries. Nor do non-tax revenues convincingly explain relative stagnation in much of the non-Western world. There are certainly some cases, such as Nigeria and Sierra Leone, where natural resource abundance plausibly contributed to fiscally weak states. However, most countries outside the OECD that extract large revenue streams are also oil-rich (Ross, 2012), and therefore their abundance in natural resources biases *against* a great revenue divergence occurring. Nor can resource curse arguments explain why many *resource poor* countries have also failed to catch up to the West. Similarly, Western countries have had better (and cheaper) access to loans for a much longer period than other parts of the world (Stasavage, 2007).

A.5 Gold versus Silver

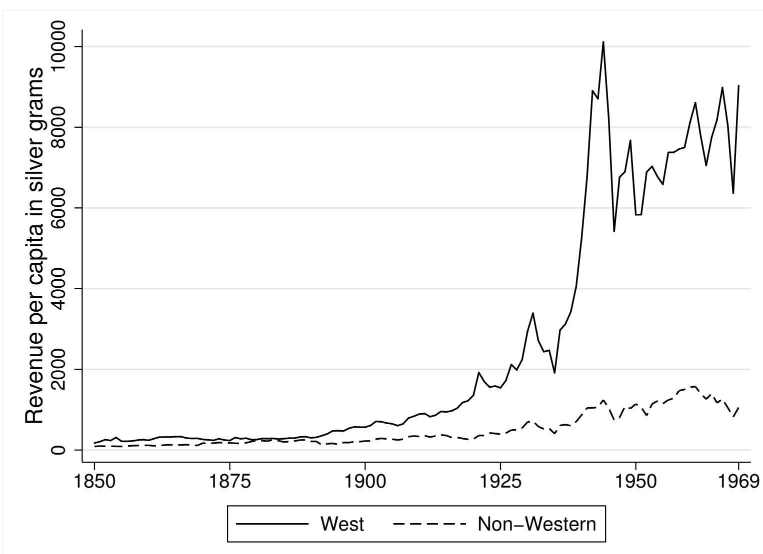
Before 1970, global currencies were typically fixed in relation to gold or silver (or both). Which precious metal is a more appropriate yardstick for value across time? This consideration is important because gold and silver prices do not move in unison. Silver prices fell by about 50% relative to gold between the 1870s and the 1890s, and then plunged even lower in the 1930s before recovering.

We use gold for our main measure of revenues for two reasons. First, for most of our time period, the majority of currencies were linked to gold rather than silver. Consequently, denominating in gold minimizes exchange rate volatility. Of the 69 countries with revenue data in 1900, 53 were on the gold or gold exchange standard, 36 had been on it for at least 20 years, and another seven would adopt the gold standard by 1907. The predominance of gold did not reflect intrinsic superiority, but instead that the core Western European nations adopted it and imposed it on their colonies and economic clients.

Second, existing evidence suggests that denominating in gold makes our measure more stable with respect to prices than denominating in silver. Bampinas and Panagiotidis (2015), for instance, find that in the two centuries following 1792, gold hedges inflation in the United States and Britain much better than does silver. Nations that retained the silver standard, in particular China, had higher inflation than elsewhere, which influenced debates over metallic standards in the West (Van der Eng, 1999).

However, as Figure A.4 shows, the choice of precious metals makes little substantive difference. We compiled an alternative revenue series in silver grams, which exhibits similar cross-national and temporal patterns as Figure 1.

Figure A.4: Revenue Trends in Silver



Notes. The lines show estimated central government revenue in silver grams (converted at nominal exchange rates). Silver-to-gold price ratios from Officer (2016).

A.6 Regression Evidence of the Great Revenue Divergence

Table A.1 estimates regression coefficients to substantiate the core pattern highlighted in Figure 1: Western countries diverged from other countries only after 1913. In Columns 1 and 2, the dependent variable is the logged version of our main revenue variable (revenue per capita in gold grams converted at nominal exchange rates). In Columns 3 and 4, the dependent variable is taxes/GDP from Andersson and Brambor (2019). These measures correspond to those used in the two panels of Figure 1. Every specification contains a lagged dependent variable, and we cluster standard errors by country. In the odd-numbered columns, we pool the data and regress revenues on an indicator for Western countries, an indicator for post-1913, and their interaction. In the even-numbered columns, we include only the interaction term and additionally include country and year fixed effects (perfect collinearity precludes including the lower-order terms in these specifications). The year fixed effects account for time-specific factors such as changes in the price of gold or international shocks, and the country fixed effects account for country-specific sources of heterogeneity that remain constant over time.

The regression estimates confirm the intuitions from Figure 1. In all specifications, the interaction term is positive and statistically significant. The marginal effect estimates for Columns 1 and 3 additionally show that Western countries raised significantly more revenue than other countries after 1913, but not before.

Table A.1: The Great Revenue Divergence: Regression Evidence

DV:	Revenues p.c.	Revenues p.c.	Taxes/GDP	Taxes/GDP
	(1)	(2)	(3)	(4)
West*Post-1913	0.0552***	0.0742***	0.367***	0.631***
	(0.0111)	(0.0173)	(0.0802)	(0.164)
West	0.00285		-0.0354	
	(0.00674)		(0.0553)	
Post-1913	0.0222***		0.155*	
	(0.00542)		(0.0770)	
Country-years	5,878	5,878	2,874	2,874
Countries	94	94	28	28
R-squared	0.985	0.969	0.941	0.924
LDV	YES	YES	YES	YES
Country FE	NO	YES	NO	YES
Year FE	NO	YES	NO	YES
		Marginal effect estimates		
West | Pre-1913	0.00285		-0.0354	
	(0.00674)		(0.0553)	
West | Post-1913	0.0581***		0.331***	
	(0.00860)		(0.0696)	

Notes. Table A.1 summarizes a series of OLS regressions with country-clustered standard errors. *** p<0.01, ** p<0.05, * p<0.1.

A.7 Revenues in Major Non-Western Empires

In Panel D of Figure 5, we present revenue intake for select non-Western empires with available data in the nineteenth century. Here we address that sample in more detail and provide additional details about our estimates for Russia.

If we had data on more non-Western empires, we might observe a slightly larger gap between Western and other countries before World War I, but our core observation would remain qualitatively unchanged. The Ottoman empire and China collected less revenue than did Russia at the end of the eighteenth century (Karaman and Pamuk 2010, 623; Rosenthal and Wong 2011, 175; Hoffman 2015, 51; Dincecco 2017, 69), and scholars typically portray nineteenth-century reforms in these empires as considerably less successful than those in Japan (and, to a lesser extent, Russia). Karaman and Pamuk show that the gap between the Ottoman empire and the West remained large in the early twentieth century, as Britain collected over four times more revenue per capita. Income differentials are undoubtedly part of the story, although we lack the data to know definitively what percentage of this gap is explained by income. The first GDP point for Turkey is in 1950, when Britain's GDP per capita was roughly four times greater.

We present additional details on our estimates for Russia because, in the early twentieth century, our estimates differ somewhat from those in Karaman and Pamuk (2010). Table A.2 compares our revenue estimates using the five-decade averages presented in Karaman and Pamuk, plus an additional average for 1910–13 from our dataset. As the table shows, although our data are largely aligned in the nineteenth century, a discrepancy emerged in the twentieth century. For the decade 1900–09, our estimate is 64% higher, and we report a large increase over that figure by 1910–13.

Table A.2: Russian Revenue Data

Decade	*Karaman and Pamuk*	*Our data*
1780–89	1.7	no data
1820–29	2.5	2.1
1850–59	3.6	3.2
1880–89	6.5	6.1
1900–09	7.5	12.3
1910–13	no data	17.4

Notes: For both series, revenue amount is annual per capita revenue in gold grams, averaged over the time periods specified.

This discrepancy is unexpected because we both use Mitchell (1998) as the source data for revenues in the local currency as well as McEvedy and Jones (1978) for population data (until 1897, when the first census occurred and hence Mitchell's population data begins). Given the importance of Russia as a comparison point for non-Western empires, we make an exception to our general coding rule to not include data points before the first census (which occurred in 1897) and to not interpolate if there was more than twenty years in between censuses (the next one occurred under the Soviet Union in 1926). We believe this choice is justified in this case given Karaman and Pamuk's usage of the same population data.

Our revenue estimates differ from those in Karaman and Pamuk (2010) because of a technical consideration about currency conversion. We convert revenue amounts in the local currency into British pounds based on nominal exchange rates. That is, we measure how many pounds a country would receive if they exchanged all their annual revenue into pounds. In this case, the ruble was pegged to the franc, and thus we are in effect converting francs into pounds. We then use pound-to-gold exchange rates to express revenue in gold grams, although this is purely for convenience of interpretation (given greater volatility in the pound than in gold). By contrast, as they explain in their appendix, Karaman and Pamuk convert Russia's revenues in rubles into its value in silver *based on the silver content of the ruble*, before then converting this amount into gold based on the silver-to-gold exchange rate. Thus, they evaluate revenue intake based on the intrinsic value of the local currency (as measured in silver), rather than on the amount at which the local currency could be exchanged for pounds. Although the ruble might well have been overvalued given its low underlying silver content, we view our estimation procedure as more faithfully estimating the international market value of a given amount of revenue intake. This, in turn, yields a higher estimate for per-capita government revenues in Russia in the early twentieth century compared to existing studies.

A.8 Mobilization for Intra-European vs. Imperial Wars

One claim in the article is that fiscal demand tended to be low in Western countries in the nineteenth century. Yet if proxying fiscal demand by participation in wars (as we do in Appendix B), it is crucial to differentiate between type of war. Britain fought few years of intra-European wars, but many years of imperial wars (although not the only European power involved in imperial wars, Britain was the most frequent participant). Thus, it is crucial to assess whether different types of wars diverge in their costs and levels of mobilization. Among available data, Onorato et al.'s (2014)'s variable for the percentage of the population mobilized in the state military most directly captures mobilization for conflict. We use this as our dependent variable in Table A.3, in which we analyze data for Britain. In the first column, we regress mobilization on participation in intra-European wars (data from Onorato et al. 2014). Examining a long time sample (1689–1913), there is a statistically significant positive correlation between participation in war and mobilization of troops. In the average war year, Britain mobilized 1.49% of its population, compared to 0.73% in non-war years. In the second column, we restrict the temporal sample to 1816–1913. Although the correlation is based on a much smaller number of years with intra-European wars, the coefficient remains statistically significant. By contrast, when we analyze imperial wars over the same time frame in Column 3 (extra-state wars from Correlates of War; Sarkees and Wayman 2010), the magnitude of the coefficient estimate shrinks dramatically and is no longer statistically significant (p=0.424). On average, Britain mobilized 0.83% of its population in years of imperial wars compared to 0.76% of other years. In Column 4, we show that these conclusions are unchanged when including both intra-European wars and imperial wars as regressors.

Table A.3: Wars and Mobilization in the British Military

	DV: Fraction of population in state military			
	(1)	(2)	(3)	(4)
Intra-European war	0.00272***	0.00127***		0.00129***
	(0.000412)	(0.000351)		(0.000351)
Imperial wars			0.000160	0.000197
			(0.000199)	(0.000187)
Observations	209	98	98	98
R-squared	0.825	0.778	0.749	0.781
LDV?	YES	YES	YES	YES
Years	1689–1913	1816–1913	1816–1913	1816–1913

Notes. Table A.3 presents OLS regression estimates. *** p<0.01, ** p<0.05, * p<0.1.

B Regressions: Interacting Fiscal Demand and Supply

Our main theoretical implication about combining supply and demand yields a natural statistical test, for which we provide evidence here: the interaction of these variables should positively associate with revenue intake. This evidence complements the historical discussion in the article.

B.1 Data Setup

In Table B.1, the revenues variable is our main measure, central government revenues per capita in gold grams converted at nominal exchange rates, although we log it for the regressions. The core sample includes 94 countries and consists of all country-years up to 1969 with available revenue data (including colonies with data), although missing data on covariates reduces the number of observations in some specifications.

We proxy fiscal demand with data from Correlates of War (Sarkees and Wayman, 2010) on participation in a major international war (at least 1,000 battle deaths). For fiscal supply, we use Brambor et al.'s (2020) data on the presence of a mandatory civil registration system for births, marriages, and deaths. Our main measure is the stock of years with such a system, although we also analyze an indicator for the presence of a civil registration system. We lag each measure by one year in the regressions, and we divide the stock variable by 100 (thus, effectively, the variable is hundreds of years with a registration system) to make the coefficient estimates more easily interpretable.

We also offer an important caveat about measurement. As highlighted in the qualitative discussion in the text, fiscal demand and fiscal supply are each multifaceted concepts that are difficult to operationalize with a single variable. For example, historically, participation in warfare has propelled fiscal demand, and it certainly played this role during the two world wars of the twentieth century. However, by this time, the scope of welfare provision by states (first in the West, and then elsewhere) had expanded such that *non*-participation in warfare does not necessarily indicate low demand. Similarly, the lack of permanent civil registration system is strongly indicative of low bureaucratic capacity, but a country that adopts one without other prerequisites (industrialization, a history of impartiality in bureaucratic recruitment) does not necessarily have high fiscal supply.

Thus, our measures offer reasonable ways to operationalize fiscal demand and supply for a large-N sample, although these concepts are inherently difficult to measure.

We estimate models with two-way fixed effects to eliminate sources of heterogeneity that are constant across countries or time. Of course, decisions to participate in war and to develop a civil registration system are likely driven to some extent by country-specific factors that vary over time. However, even if so, it is not clear that this source of confounding would bias the interaction term in a positive direction. In the text we cited evidence from Brambor et al. (2020, 202) that the development of civil registration systems was not, in general, driven by participation in war. It also seems unlikely that states can usually anticipate their war needs accurately and preventively ramp up fiscal capacity. For example, although World War I eventually yielded unprecedented revenue intake in Western countries, every participant was shocked by the scale of the war effort. These states waited several years after 1914 to impose high statutory rates on income taxes or introduce universal conscription (Scheve and Stasavage, 2016). However, given these unavoidable caveats about causal inference and data limitations, we regard these statistical associations as a plausibility probe for our theory rather than as conclusive evidence for a causal effect.

The statistical model is:

$$\begin{aligned}\text{ln(Revenue/pop.)}_{i,t} &= \beta_{lag}\text{ln(Revenue/pop.)}_{i,t-1} + \beta_{war}\text{War}_{i,t-1} + \beta_{reg}\text{Stock of reg. system}_{i,t-1} \\ &+ \beta_{inter}\text{War}_{i,t-1}\text{Stock of reg. system}_{i,t-1} + \beta_i + \beta_t + \epsilon_{i,t}. \qquad \text{(B.1)}\end{aligned}$$

We index countries by i and years by t. The main parameter of interest is β_{inter}, the coefficient estimate for the interaction term. Standard errors are clustered by country. In addition to the country and year fixed effects, every model also contains a lagged dependent variable. In unreported tests, we assessed the dependent variable for non-stationarity by running a series of Fisher-type unit-root tests based on augmented Dickey-Fuller tests. We calculated residuals from auxiliary regressions that include the country and year fixed effects, and these tests reject at the 1% significance level the null hypothesis that all panels contain unit roots.

B.2 Results

Main results. Table B.1 presents the main results. Column 1 contains the full sample of 5,878 country-years across 94 countries. This specification interacts war participation and the stock of years with a civil registration system, and the coefficient estimate is statistically significant. Column 2 adds covariates for logged population and whether the territory is independent, both lagged by one year, which minimally change the coefficient estimate. These covariates address two alternative explanations about country-specific time trends that may influence the coefficient estimates: demographic changes or comparing sovereign countries to colonized territories. Regarding the latter, in our baseline specification, we compare colonies with independent countries. This choice is appropriate because the ability to raise revenues matters, not where the revenues are spent. However, it is useful to show that such comparisons do not drive the results. Additionally, although colonized territories usually lacked a civil registration system, this is not imposed by definition in

Brambor et al.'s (2020) coding, as several colonized territories did indeed implement civil registration systems. This is consistent with our discussion in the article that European colonizers tended to not advance bureaucratic development. Below we elaborate upon measuring war participation for colonized territories.

In Columns 3 and 4, we consider an alternative version of the civil registration system variable. Our source data, Brambor et al. (2020), is missing for many countries in our sample. For the main version of the civil registration system variable, we code countries with missing data as never having a civil registration system. This is justified under the reasonable premise that countries for which Brambor et al. (2020) were unable to collect systematic information about their bureaucracy are unlikely to have a civil registration system. However, in Columns 3 and 4, we set the civil registration systems variable as missing for any countries not in Brambor et al.'s (2020) data. The sample decreases considerably to 3,176 country-years across 50 countries. Similar to the difference between Columns 1 and 2, Column 4 adds the two covariates to the specification from Column 3. The coefficient estimates for the interaction term are statistically significant in each specification, although slightly smaller in magnitude.

Scatterplot of cross section. Figure B.1 presents a scatterplot that corresponds with a cross-section of countries with revenue intake measured in 1969. The x-axis is cumulative years with participation in war between 1914 and 1969. We disaggregate countries by whether they had established an early registration system (specifically, before 1900; these countries are in black) or not (gray). We present separate regression lines for these two sets of countries. The line slopes steeply upward for countries with an early civil registration system, but is downward-sloping for other countries. Thus, the cross-sectional pattern recovers the positive interaction effect demonstrated in the panel regressions. Unsurprisingly, every country in the top right part of the scatterplot is Western European, Western offshoots, or Japan. An unreported regression specification shows that the coefficient for the interaction term is statistically significant. Note that the generally low participation of high-supply states in wars in the nineteenth century makes such a corresponding figure largely uninformative for this earlier period.

Figure B.1 highlights cases that support Schenoni's (2021) contention that the near absence of wars in South America in the twentieth century undermined fiscal-capacity building efforts in the region. As the figure shows, many of these countries developed civil registration systems early, but had relatively low revenue intake in 1969. Thus these cases differ in an important way from ones discussed in the article (such as India and Egypt) that had the opposite combination of high fiscal demand with low supply.

The scatterplot also highlights shortcomings of our demand measure, although in a direction that biases against finding a positive interactive effect. Several Western countries did not participate in World War I (or, for some cases, either world war). Yet there were clear spillover effects, as they experienced similar pressures as the participants given the threat of invasion and rising pressure for welfare spending. In Figure B.2, we highlight the spikes in revenue during World War I for the neutral states in Western Europe.

Robustness checks. In the remaining tables, we consider several additional robustness checks. In Table B.2, we consider two alternate measures (in each case, altering Columns 1 and 2 of Table B.1 with the following changes). In Columns 1 and 2 of Table B.2, we replace the stock of civil registration system years with an indicator for the presence of a civil registration system in the previous year. In Columns 3 and 4, we measure war participation differently for colonies. In the main measure, for the world-war belligerents, we code all their colonies as participants. This is the most appropriate coding decision given our theoretical interest, as the colonies supplied troops to the metropolitan country and there was greater impetus on not draining the metropolitan treasury for colonial expenses. However, in Columns 3 and 4, we do not code the colonies as participants in the world wars.

Finally, in Table B.3, we switch the dependent variable to taxes/GDP from Andersson and Brambor (2019); see Panel B of Figure 1. The overall number of countries and country-years drops precipitously, and the only non-Western countries are in South America (plus Mexico and Japan). Other than changing the dependent variable (and the lagged dependent variable), the models are identical to Columns 1 and 2 from Table B.1, and the four columns from Table B.2. The regressions are identical when using the alternative version of the registration system variable from Columns 3 and 4 of Table B.1 because no countries in this truncated sample are missing data on civil registration systems; hence we omit these duplicate specifications. In five of the six specifications, the interaction term is statistically significant (p=0.135 in Column 3). These findings, combined with the visual evidence from Panel B of Figure 1 and the later timing of revenue divergence relative to income divergence (see Figure 2), provide evidence against an alternative hypothesis that the great revenue divergence simply tracks changes in GDP over time.

B.3 Tables and Figures

Table B.1: Interacting War Participation with Civil Registration Systems

	DV: Log revenues p.c.			
	(1)	(2)	(3)	(4)
War*Stock of reg. system	0.0654***	0.0676***		
	(0.0207)	(0.0209)		
War*Stock of reg. system (alt.)			0.0427*	0.0424*
			(0.0244)	(0.0243)
War	-0.0110	-0.0114	0.0111	0.0119
	(0.0186)	(0.0186)	(0.0225)	(0.0226)
Stock of reg. system	0.0354	0.0289		
	(0.0283)	(0.0285)		
Stock of reg. system (alt.)			0.182***	0.236***
			(0.0250)	(0.0387)
Population		-0.0292**		-0.0221
		(0.0124)		(0.0184)
Independent		0.000875		-0.0272
		(0.0199)		(0.0227)
Country-years	5,878	5,878	3,265	3,265
Countries	94	94	50	50
R-squared	0.969	0.969	0.972	0.972
LDV	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes. Table B.1 presents OLS regression estimates with country-clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure B.1: Cross-Section of Interactive Effect

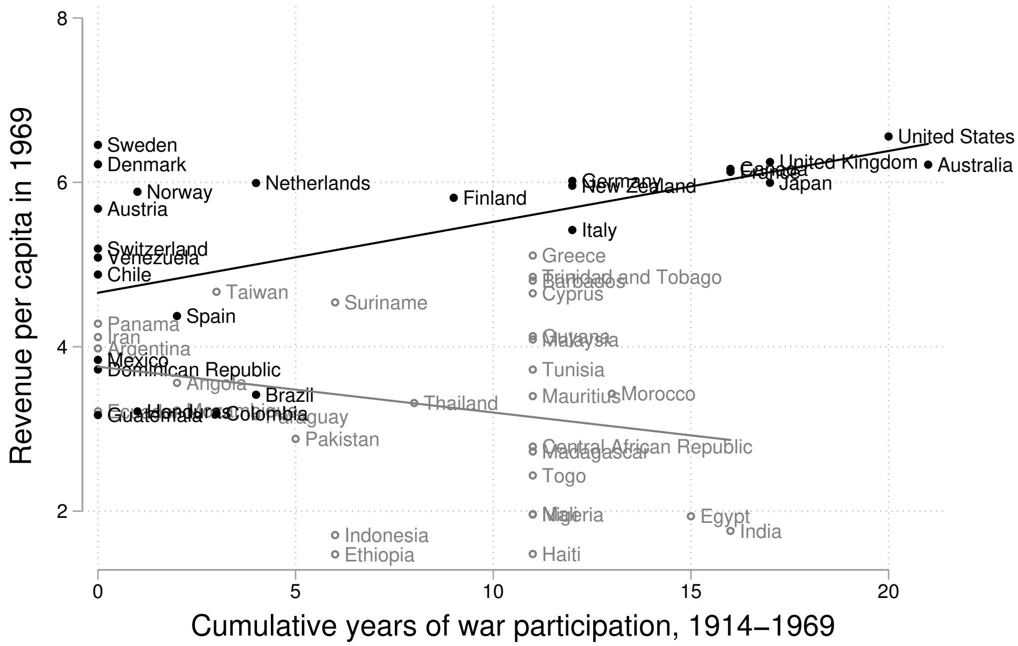
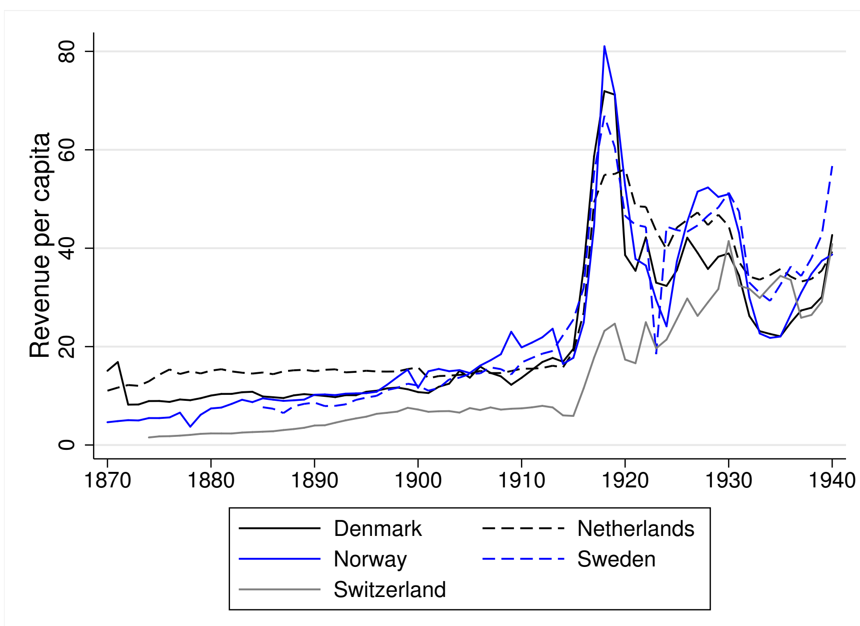


Figure B.2: Revenue Trends in WWI Neutrals



Notes. The lines show estimated central government revenue per capita in gold grams, converted at nominal exchange rates.

Table B.2: Alternative Measures

	DV: Log revenues p.c.			
	(1)	(2)	(3)	(4)
War*Reg. system indicator	0.0854***	0.0880***		
	(0.0291)	(0.0290)		
War (alt.)*Stock of reg. system			0.0502**	0.0513**
			(0.0249)	(0.0258)
War	-0.0292	-0.0296		
	(0.0201)	(0.0198)		
War (alt.)			0.0172	0.0181
			(0.0240)	(0.0250)
Reg. system indicator	-0.0680***	-0.0724***		
	(0.0140)	(0.0145)		
Stock of reg. system			0.0378	0.0316
			(0.0287)	(0.0289)
Population		-0.0326**		-0.0289**
		(0.0139)		(0.0127)
Independent		0.0112		0.000280
		(0.0166)		(0.0205)
Country-years	5,878	5,878	5,878	5,878
Countries	94	94	94	94
R-squared	0.969	0.969	0.969	0.969
LDV	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Notes. Table B.2 presents OLS regression estimates with country-clustered standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table B.3: Taxes/GDP

	DV: Taxes/GDP					
	(1)	(2)	(3)	(4)	(5)	(6)
War*Stock of reg. system	0.429*	0.494**				
	(0.216)	(0.218)				
War*Reg. system indicator			0.336	0.426*		
			(0.218)	(0.249)		
War (alt.)*Stock of reg. system					0.437*	0.498**
					(0.221)	(0.224)
War	0.111	0.0370	0.159	0.0782		
	(0.181)	(0.200)	(0.114)	(0.150)		
War (alt.)					0.0994	0.0255
					(0.184)	(0.203)
Stock of reg. system	0.592***	1.258***			0.605***	1.255***
	(0.155)	(0.305)			(0.155)	(0.308)
Reg. system indicator			-0.451***	-0.427***		
			(0.109)	(0.126)		
Population		-0.333*		-0.266		-0.324*
		(0.167)		(0.158)		(0.168)
Independent		0.341**		0.257*		0.339**
		(0.128)		(0.135)		(0.128)
Country-years	2,874	2,780	2,874	2,780	2,874	2,780
Countries	28	28	28	28	28	28
R-squared	0.924	0.924	0.924	0.924	0.924	0.924
LDV	YES	YES	YES	YES	YES	YES
Country FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Notes. Table B.3 presents OLS regression estimates with country-clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

C Proofs for Formal Model

As a preliminary result, we prove that the revenue-maximizing tax rate is unique and strictly bounded between 0 and 1. This term is $\hat{\tau} \equiv \arg\max_{\tau \in [0,1]} \int_0^{1-\tau} \tau dH(e_i)$. Taking the first-order condition yields the implicit characterization $-\hat{\tau}h(1-\hat{\tau}) + \int_0^{1-\hat{\tau}} dH(e_i) = 0$. Applying the intermediate value theorem demonstrates that at least one $\hat{\tau} \in (0,1)$ exists satisfying Equation 1:

- For the lower bound at $\tau = 0$, we have $\int_0^1 dH(e_i) = 1 > 0$.
- For the upper bound at $\tau = 1$, we have $-h(0) < 0$ because there is positive mass at $e_i = 0$.
- The expression is continuous in τ.

Finally, to establish that $\hat{\tau}$ is a unique maximizer, we demonstrate that the second-order condition is strictly negative:

$$\tau h'(1-\tau) - h(1-\tau) - \int_0^{1-\tau} h(1-\tau) de_i < 0.$$

The last two terms are strictly negative for all $\tau \in [0,1]$, and the weak negativity of the first term follows from this and $h' \leq 0$.

Proof of Lemma 1. Applying the intermediate value theorem establishes that at least one $\overline{L} \in (0,1)$ exists satisfying $R^{\text{leg}}(\overline{L}) = R^{\text{crony}}$:

- For the lower bound, we need $R^{\text{leg}}(0) < R^{\text{crony}}$. This holds for $N < \overline{N} \equiv Y \frac{\int_0^1 (1-e_i) dH(e_i)}{\int_0^{1-\hat{\tau}} \hat{\tau} dH(e_i)}$. This threshold defines the upper bound for N expressed in the model setup. Note that the numerator of $\overline{N}$ is strictly positive, and the strict positivity of its denominator follows from the result above that establishes $\hat{\tau} \in (0,1)$.
- For the upper bound, we need $R^{\text{leg}}(1) > R^{\text{crony}}$. This follows directly from assuming $Y < N$.
- $R^{\text{leg}}(L_2)$ is continuous in L_2.

The unique threshold follows from establishing a strictly increasing relationship in L_2.

$$\frac{d}{dL_2}\Big[R^{\text{leg}}(L_2) - R^{\text{crony}}\Big] = N\Big[\int_0^1 (1-e_i) dH(e_i) - \int_0^{1-\hat{\tau}} \hat{\tau} dH(e_i)\Big]$$
$$= N\Big[\int_0^{1-\hat{\tau}} (1-\hat{\tau}-e_i) dH(e_i) + \int_{1-\hat{\tau}}^1 (1-e_i) dH(e_i)\Big].$$

The bounds of the first integral in the last line assume $e_i < 1-\hat{\tau}$, which makes the overall expression strictly positive.

In words, the last line states that the government raises more revenues from legible than illegible citizens. Among citizens with low-valued exit options ($e_i < 1-\hat{\tau}$) expressed in the first integral, the government can extract $1-e_i$ from each legible citizen because of sufficient information to hold them down to their reservation value, compared to the lower flat rate $\hat{\tau}$ for each illegible citizen. Among citizens with high-valued exit options ($e_i > 1-\hat{\tau}$) expressed in the second integral, the government still collects $1 - e_i$ from each legible citizen, but nothing from illegible citizens. They exit rather than pay a tax that, relative to their exit option, is too high to induce compliance. ■

Given assumptions stated in the setup and the preceding result, we have bounds $N \in \left(Y,\ Y\frac{\int_0^1 (1-e_i)dH(e_i)}{\int_0^{1-\hat{\tau}} \hat{\tau} dH(e_i)}\right)$. To prove this set is non-empty for all parameter values, it suffices to demonstrate $\int_0^{1-\hat{\tau}} \hat{\tau} dH(e_i) < \int_0^1 (1-e_i)dH(e_i)$. This easily rearranges to a true statement: $\int_0^{1-\hat{\tau}} (1-\hat{\tau}-e_i)dH(e_i) + \int_{1-\hat{\tau}}^1 (1-e_i)dH(e_i) > 0$. The end of the preceding proof provides intuition for why this expression is strictly positive.

Proof of Proposition 1.

- There are two cases to consider in the *low fiscal demand* region. If $R_2^{\text{dem}} < R^{\text{cus}}$, then the government gains negative utility from levying any taxes in excess of its endowment R^{cus}, and therefore chooses the low-effort tax system. If $R^{\text{cus}} < R_2^{\text{dem}} \leq R^{\text{cus}} + F$, then the government would (absent costs) choose to raise more revenue than R^{cus}, but only slightly more: $R_2^{\text{dem}} - R^{\text{cus}}$. Yet the fixed cost F to pursuing either high-effort extraction strategy exceeds the extra revenues that the government wants to raise, that is, $R_2^{\text{dem}} - R^{\text{cus}} < F$. Consequently, the costs outweigh the benefits of high-effort extraction, and the government chooses the low-effort tax system.
- In the *intermediate fiscal demand* region, the government's desire for additional revenues implies that the benefit of a high-effort tax system exceeds the cost, that is, $R_2^{\text{dem}} - R^{\text{cus}} > F$. However, fiscal demand is not high enough for the government to desire maximum tax revenues, that is, $R_2^{\text{dem}} < R^{\text{cus}} + R_2^{\text{max}}$. Consequently, the government chooses whichever high-effort tax system enables higher maximum revenues (see Lemma 1). It sets $\{\tau_i\}_{i\in\mathcal{N}}$ to achieve total tax intake of $R_2^{\text{dem}} - R^{\text{cus}}$, which yields less-than-maximum revenues of $R_2 = R_2^{\text{dem}}$.

  This parameter range permits multiple equilibria for two reasons, although all equilibria are payoff equivalent. First, a continuum of choices of $\{\tau_i\}_{i\in\mathcal{N}}$ yield the desired amount of revenues. Second, the government is indifferent between the two high-effort tax systems if two conditions are met: R_2^{dem} is above but close to $R^{\text{cus}} + F$, and $R^{\text{leg}}(L_2) > F$. In this case, either high-effort strategy enables the government to raise the desired amount of revenue. In such cases, we focus on the equilibrium in which the government chooses the high-effort strategy that, if used to its full potential, would yield more revenues.

  Finally, note that assuming $\overline{F} < R^{\text{crony}}$ implies that $F < R_2^{\text{max}}$ for all L_2. This ensures that the set of parameter values for the intermediate fiscal demand region is non-empty.
- In the *high fiscal demand* region, the government chooses whichever high-effort tax structure enables higher maximum revenues (see Lemma 1), and then sets $\{\tau_i\}_{i\in\mathcal{N}}$ to maximize revenues. ■

Before proving Proposition 2, we place another restriction on the upper bound $\overline{F}$ to ensure that $\min\{\Delta\hat{L}, 1\}$ is interior, as previewed in footnote 74. This is true if $R^{\text{leg}}(1) > R^{\text{crony}} + \frac{F}{p_{\text{high}}}$. This solves to $F < p_{\text{high}} \cdot (N - Y) \int_0^1 (1 - e_i) dH(e_i)$. Thus, combining this with the upper bound for F stated in footnote 70, we require:

$$\overline{F} = \min\Big\{R^{\text{crony}},\ \ p_{\text{high}} \cdot (N - Y) \int_0^1 (1 - e_i) dH(e_i)\Big\}.$$

Proof of Proposition 2. Need to show that $\hat{L}$ satisfies the bounds stated in the proposition and is unique.

- To establish the lower bound, we first present the following string of equalities and inequalities:
$$R^{\text{leg}}\big(\Delta \underline{L}\big) = R^{\text{leg}}\big(\overline{L}\big) = R^{\text{crony}} < R^{\text{crony}} + \frac{F}{p_{\text{high}}} = R^{\text{leg}}\big(\Delta \hat{L}\big).$$
The first equality follows from $\underline{L} \equiv \frac{\overline{L}}{\Delta}$. The second equality follows from Lemma 1. The inequality follows from $F > 0$ and $p_{\text{high}} < 1$. The third equality follows from Equation 7. We can remove the middle terms to state:
$$R^{\text{leg}}\big(\Delta \underline{L}\big) < R^{\text{leg}}\big(\Delta \hat{L}\big).$$
Because R^{leg} is a strictly increasing function, we have $\Delta \underline{L} < \Delta \hat{L}$, which simplifies to $\underline{L} < \hat{L}$.
- The upper bound follows immediately from the preceding proof for the lower bound and the fact that $\hat{L}$ strictly increases in F (see the proof of Proposition 3): $\hat{L} < \overline{L}$ for low-enough F.
- For uniqueness, see the first part of the proof for Proposition 3. ■

Proof of Proposition 3. Define $G(L_1) \equiv R^{\text{leg}}\big(\Delta L_1\big) - \Big(R^{\text{crony}} + \frac{F}{p_{\text{high}}}\Big)$. We first establish

$$\frac{\partial G(L_1)}{\partial L_1} = N\Delta \Bigg[\int_0^{1-\hat{\tau}} (1 - \hat{\tau} - e_i) dH(e_i) + \int_{1-\hat{\tau}}^1 (1 - e_i) dH(e_i)\Bigg] > 0.$$

See the proof of Lemma 1 for why this expression is positive.

Given this, applying the implicit function theorem to compute each derivative implies that the overall sign is opposite that from the constituent partial derivative.

$$\frac{\partial G(\hat{L})}{\partial p_{\text{high}}} = \frac{F}{p_{\text{high}}^2} > 0.$$

$$\frac{\partial G(\hat{L})}{\partial \Delta} = N\hat{L} \Bigg[\int_0^{1-\hat{\tau}} (1 - \hat{\tau} - e_i) dH(e_i) + \int_{1-\hat{\tau}}^1 (1 - e_i) dH(e_i)\Bigg] > 0.$$

$$\frac{\partial G(\hat{L})}{\partial F} = -\frac{1}{p_{\text{high}}} < 0.$$

■

References

Andersson, Per F. and Thomas Brambor. 2019. "Financing the State: Government Tax Revenue from 1800 to 2012. Version 2.0.". Lund: Lund University.

Bampinas, Georgios and Theodore Panagiotidis. 2015. "Are Gold and Silver a Hedge Against Inflation? A Two Century Perspective." *International Review of Financial Analysis* 41:267–276.

Beramendi, Pablo, Mark Dincecco, and Melissa Rogers. 2019. "Intra-Elite Competition and Long-Run Fiscal Development." *Journal of Politics* 81(1):49–65.

Bolt, Jutta, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden. 2018. "Rebasing 'Maddison'." Maddison Project Working paper 10. Available at https://www.rug.nl/ggdc/historicaldevelopment/maddison.

Brambor, Thomas, Agustín Goenaga, Johannes Lindvall, and Jan Teorell. 2020. "The Lay of the Land: Information Capacity and the Modern State." *Comparative Political Studies* 53(2):175–213.

Centeno, Miguel Angel. 2002. *Blood and Debt: War and the Nation-State in Latin America*. Penn State Press.

Dincecco, Mark. 2017. *State Capacity and Economic Development: Present and Past*. Cambridge University Press.

Feenstra, Robert C., Robert Inklaar Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10):3150–3182.

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

Karaman, K. Kivanç and Şevket Pamuk. 2010. "Ottoman State Finances in European Perspective, 1500–1914." *Journal of Economic History* 70(3):593–629.

McEvedy, Colin and Richard Jones. 1978. *Atlas of World Population History*. Penguin Books.

Menaldo, Victor. 2016. *The Institutions Curse: Natural Resources, Politics, and Development*. Cambridge University Press.

Mitchell, Brian. 1998. *International Historical Statistics: Europe 1750-1993*. Springer.

Morrison, Kevin M. 2014. *Nontaxation and Representation*. Cambridge University Press.

Officer, Lawrence H. 2016. "Exchange Rates Between the United States Dollar and Forty-one Currencies." MeasuringWorth.

Onorato, Massimiliano Gaetano, Kenneth Scheve, and David Stasavage. 2014. "Technology and the Era of the Mass Army." *Journal of Economic History* 74(2):449–481.

Queralt, Didac. 2019. "War, International Finance, and Fiscal Capacity in the Long Run." *International Organization* 73(4):713–753.

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

Ross, Michael L. 2012. *The Oil Curse: How Petroleum Wealth Shapes the Development of Nations*. Princeton University Press.

Sarkees, Meredith Reid and Frank Whelon Wayman. 2010. *Resort to War: 1816-2007*. CQ Press.

Schenoni, Luis L. 2021. "Bringing War Back in: Victory and State Formation in Latin America." *American Journal of Political Science* 65(2):405–421.

Scheve, Kenneth and David Stasavage. 2016. *Taxing the Rich: A History of Fiscal Fairness in the United States and Europe*. Princeton University Press.

Stasavage, David. 2007. "Cities, Constitutions, and Sovereign Borrowing in Europe, 1274–1785." *International Organization* 61(3):489–525.

Summers, Robert and Alan Heston. 1991. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988." *The Quarterly Journal of Economics* 106(2):327–368.

Van der Eng, Pierre. 1999. "The Silver Standard and Asia's Integration into the World Economy, 1850-1914." *Review of Asian and Pacific Studies* 18:59–85.