The Great Revenue Divergence

Alexander Lee* Jack Paine†

May 16, 2022

Forthcoming in *International Organization*

Abstract

This article describes and explains a previously overlooked empirical pattern in state revenue collection. As late as 1913, central governments in the West collected similar levels of per-capita revenue as the rest of the world, despite ruling richer societies and experiencing a long history of fiscal innovation. Western revenue levels permanently diverged only in the following half century. We identify the twentieth–century great revenue divergence by constructing a new panel dataset of central government revenue with broad spatial and temporal coverage. To explain the pattern, we argue that sustainably high levels of revenue extraction require societal demand for an activist state and an existing supply of effective bureaucratic institutions. Neither factor in isolation is sufficient. We formalize this insight in a game-theoretic model. The government can choose among low-effort, legibility-intensive, and crony-favoring strategies for raising revenues. Empirically, our theory accounts for low revenue intake in periods of low demand (nineteenth-century West) or low bureaucratic capacity (twentieth–century former colonies), and for eventual revenue spikes in the West.

Keywords: Government revenues, Fiscal capacity, State capacity, War, Bureaucracy

*Associate Professor of Political Science, University of Rochester, Harkness Hall, Rochester, NY 14627. Email: alexander.mark.lee@rochester.edu.

†Associate Professor of Political Science, University of Rochester, Harkness Hall, Rochester, NY 14627. Email: jackpaine@rochester.edu.

The gap in the quality of government services in Western and non-Western countries is founded on disparities in taxation. Between 2010–19, central governments in Western European states and offshoots extracted, on average, 43% of their country's annual GDP in government revenues, compared to 27% in non-European countries.[1] Disparities in per–capita revenue intake are even greater, given much higher income levels in the West. When and why did these gaps in revenue collection emerge? Understanding this question is critical because tax collection and fiscal capacity are strongly associated with economic development, political order, and governance quality more broadly.[2]

We demonstrate that major discrepancies in state revenue collection are a recent phenomenon. On the eve of World War I, South American countries and some export-oriented colonies collected similar levels of per-capita revenue as in the West despite lower GDP; and often exceeded Western revenue collection when accounting for income differences. Many Asian and African countries and colonies lagged Europe, but these differences were small by twentieth-century standards. However, over the following half century, per–capita revenue intake skyrocketed in Western countries, compared to more modest increases elsewhere. The newfound revenue gap between Western and non-Western countries persisted, and in fact widened, even after former European colonies gained independence. In sum, the *great revenue divergence* occurred in the twentieth century.

We identify this previously overlooked pattern by constructing a new panel dataset of central government revenue. We combined data on central government revenue from Mitchell with historical exchange rates, gold prices, and population.[3] The main contribution of our dataset is its spatial and temporal breadth: at least one year for 18 Western countries (including 15 with at least one data point in the nineteenth century) and 76 non-Western countries (42 in the nineteenth century). This contrasts with existing government revenue datasets that have coverage before the twentieth cen-

[1] Calculated by authors using data from the International Monetary Fund 2017.
[2] Levi 1989; Besley and Persson 2014; Dincecco 2017.
[3] Mitchell 1998.

tury only (or mainly) for European countries,[4] or only the late twentieth century.[5] By combining depth and breadth, our dataset is uniquely suitable for analyzing comparative historical trends in government revenues.[6] We supplement our measure of revenues per capita with data on taxes/GDP from Andersson and Brambor.[7] We demonstrate a qualitatively similar pattern of revenue divergence when accounting for income differentials, albeit on a truncated sample.

Existing theories of revenue extraction, taken in isolation, cannot explain the great revenue divergence. Some scholars analyze *fiscal demand*. These theories emphasize how some states have greater needs than others to extract revenues, often because of participation in external wars. By contrast, *fiscal supply* explanations focus on the bureaucratic institutions used to gather information about the population. Bureaucracies enable states to accurately assess tax burdens and to efficiently extract revenues by making production legible to the state, which is alternatively referred to as high fiscal capacity.

However, existing bellicose and state-legibility explanations cannot answer two key questions about the twentieth–century revenue divergence. First, why did it occur so late? Existing accounts date large and permanent discrepancies in revenue collection to the nineteenth century or earlier. In the late eighteenth century, England and France each collected higher revenue per capita than major non-Western empires.[8] Over the next century, Western states improved their fiscal capacity by collecting voluminous information about their populations,[9] enacting modern fiscal

[4] Beramendi, Dincecco and Rogers 2019; Dincecco 2011; Karaman and Pamuk 2010; Scheve and Stasavage 2016.

[5] International Monetary Fund 2017; Queralt 2019.

[6] Although other scholars have also constructed datasets using Mitchell 1998, later we explain why our approach to making data points comparable across countries yields a much broader sample.

[7] Andersson and Brambor 2019.

[8] Karaman and Pamuk 2010, 623; Rosenthal and Wong 2011, 175; Hoffman 2015, 51; Dincecco 2017, 69.

[9] Brambor et al. 2020.

devices such as income taxes,[10] and undergoing modern industrial development.[11] By contrast, much of the rest of the world was under, or had recently gained independence from, Western colonial rule. Yet in the early twentieth century, the West was not clearly distinguished in its revenue intake. Thus, a permanently large revenue divergence occurred well after these discrepancies in latent fiscal capacity and economic development had emerged.

Second, in the twentieth century, why did non-Western countries continue to fall behind even after gaining independence? Leading existing explanations focus on how non-European countries during this period either fought too few wars, or only limited international wars funded by debt and civil wars.[12] Yet many newly independent states exhibited high demand for public expenditures. Anti-colonial activists believed that jurisdictional sovereignty would engender higher levels of public spending by aligning the government's incentives with its citizens rather than with European bondholders and civil servants.[13] Anti-colonial movements sought to use government to provide greater services for citizens. Furthermore, international competition was high in some parts of the post-colonial world (Middle East, South Asia, and East Asia), and most colonies experienced mass franchise expansion shortly before gaining independence. These pressures created additional demand for public expenditures.

To unravel the puzzle of the great revenue divergence, we develop a formal-theoretic framework to explain why high levels of revenue intake require the conjunction of high demand for an activist government and high supply of bureaucratic, or fiscal, capacity. A state with high fiscal capacity has a *latent advantage* at raising revenues. However, absent demand from politically relevant actors for fiscal expenditures, the governing elite will keep taxes low and underutilize the state's fiscal potential. Conversely, a state with weak bureaucratic capacity cannot immediately remedy this shortcoming because levels of societal legibility are highly persistent, at least in the short term. Faced with a pressing demand for revenues, a low-capacity state turns to *crony-favoring*

[10] Mares and Queralt 2015.

[11] Pomeranz 2009.

[12] Herbst 2000; Centeno 2002; Besley and Persson 2011; Queralt 2019.

[13] Furnivall 2014.

extraction: gaining revenues in return for granting a favorable market position to cronies. Although less lucrative than an effective income tax, low-capacity states lack the bureaucratic infrastructure to efficiently administer a legibility-intensive tax. Only states with high fiscal demand and high fiscal supply can monitor production effectively enough to extract high revenues from *legibility-intensive* revenue sources, and have the political will to tax heavily.

We also explain how political actors can bolster bureaucratic capacity over time. By choosing legibility-intensive extraction in the present, the government can bolster societal legibility in the future via learning-by-doing effects. Although bellicose factors can encourage states to invest in future fiscal capacity, other factors matter as well. Contrary to bellicose theories, anticipation of high demand in the future does not necessarily engender "common value states" (as discussed by Besley and Persson) that refrain from predating their economy.[14] Instead, crony-favoring extraction remains the best strategy for raising revenues if the stock of bureaucratic capacity will remain low regardless of the state's actions. This is true when the initial stock of bureaucratic capacity is sufficiently low or the potential for bureaucratic growth is low. Conversely, states for which either of these conditions are more favorable can gain from establishing legibility-intensive extraction. If fiscal demand is low in the present, the government incurs up-front costs to implementing extractive taxes, and initially underutilizes its fiscal capacity. However, by investing in fiscal capacity, the government positions itself to collect high levels of revenue in the future—if demand increases. Furthermore, the net costs of pivoting to legibility-intensive taxation are lower from the perspective of the governing elite when customs taxes entail high deadweight losses and income taxes create more favorable distributional effects for them.

In sum, our main theoretical implication is that large revenue intake requires the conjunction of high fiscal supply and high fiscal demand. Over time, divergence in revenue intake occurs if demand grows. High-capacity states distinguish themselves in revenue collection only when fiscal demand is high.

[14] Besley and Persson 2011.

Evidence from the nineteenth and twentieth centuries supports these expectations. By the nineteenth century, Western countries had amassed important advantages in latent fiscal capacity relative to other countries, in part because of prior warfare.[15] Western states expanded their bureaucratic capacity throughout the century because of ongoing industrialization, the spread of national identity and public education,[16] and elite coalitions that wanted to shift the distributional burden of taxes.[17] But the relative lack of intra-European wars between 1816 and 1913—a period that also predated modern welfare states—limited the demand for revenues. Consequently, Western states underutilized their growing fiscal capacity. Primary product exporters in South America as well as some colonial dependencies could generate similar levels of revenue simply by collecting customs taxes. Furthermore, some non-Western empires partially caught up because threats from the West created high fiscal demand. Thus, on the eve of World War I, there was a small or non-existent gap between the West and various groups of non-Western countries.

The two World Wars and Great Depression changed this calculus for Western states, who restructured their economies to fight total war. These stimuli unleashed permanently higher demand for social spending because of ensuing franchise expansion and the creation of welfare states. Prior investments in fiscal capacity enabled Western states to raise historically unprecedented levels of revenues, in particular through legibility-intensive sources such as income and value-added taxes.

By contrast, most European colonies suffered from low fiscal capacity. Colonial governments promoted primary product exports or collected low-yield direct taxes locally.[18] If these revenue sources were insufficient on their own to balance the budget, colonial governments could take advantage of permissive conditions for international borrowing.[19] However, the bureaucratic infrastructure was wholly inadequate for meeting heightened demand after countries gained indepen-

[15] Brewer 1990.
[16] Zhang and Lee 2020.
[17] Mares and Queralt 2015; Saylor and Wheeler 2017.
[18] Frankema and Booth 2019.
[19] Queralt 2019.

dence. Many post-colonial states with high fiscal demand turned toward crony-favoring strategies. Prior underinvestment in bureaucratic capacity prevented these states from effectively collecting legibility-intensive taxes.

The main non-Western exceptions were East Asian states, in particular Japan. Intense geopolitical pressure combined with a history of bureaucratic government enabled large increases in legibility-intensive taxation, similar to the West.

1 The Great Revenue Divergence: Trends Over Time

After introducing our new data, this section provides descriptive evidence of a great revenue divergence between Western Europe (and Japan) and the rest of the world starting around 1914. We then contrast the late onset of this revenue divergence with the earlier economic gap that had emerged between the West and the rest of the world.

1.1 Introducing the Revenue Data

Our main measure is central government revenue per capita in gold grams, which we constructed using the following steps.

1. We use data on *central government revenues* from Mitchell.[20] We translated fiscal years into calendar years to measure each country's annual revenues in thousands of local-currency units (although in some cases revenue is listed in U.S. dollars).

2. We use population data from Mitchell to calculate *revenue per capita*.[21] Exact population estimates are typically available only in census years, and we estimated population in non-census years by linearly interpolating between censuses (although we drop observations for

[20] Mitchell 1998.
[21] Mitchell 1998.

which no census occurred within two decades). For this reason, we cannot estimate revenue per capita before the date of the first census, even when earlier revenue data are available.[22]

3. We converted all currency measures to their equivalents in British pounds to generate a *common scale* for revenue levels. This required constructing a new time series of historical exchange rates into pounds.[23] We were unable to perform this step for country-years with non-convertible currencies, and thus we drop such observations even if revenue and population data are available.[24]

4. Finally, we converted revenue per capita in British pounds into *gold grams*.[25] Although unnecessary for cross-national comparisons, this step reduces problems in data visualization stemming from the volatility of the pound. It also makes our revenue series directly comparable to existing historical datasets that measure revenue in gold grams.

Our revenue variable advances existing quantitative data on state revenue in both geographical and

[22] An exception is that, for Africa, we incorporated Frankema and Jerven's (2014) data for 1850–1960; otherwise, almost all these observations would drop from the sample. We also incorporated additional population data for Russia that we discuss in Appendix A.7.

[23] We converted local currency units into British pounds. Correlates of War (COW) trade data (Barbieri, Keshk and Pollins, 2008) provide the main source for historical exchange rates. COW does not include data before 1870 or from colonies (although most colonies used the mother country's currency), and we supplement their data using Denzel (2010) and Officer (2016). Because COW data uses market quotes, it exhibits frequent short gaps for smaller countries. To reduce this problem, we interpolated rates in cases in which the data coverage gap was less than five years and the difference in rates on either side of the gap did not exceed 5%.

[24] Although we included some currencies with fixed exchange rates, we excluded currencies for which published exchange rates bore no relation to market supply and demand, or the exchange rate exhibited sharp year-to-year fluctuations. In many cases, this meant excluding periods of instability when a country's link to either gold or the dollar changed.

[25] Prices for gold ounces from Officer 2016.

chronological coverage. The amount of data available is extensive, extending back to the early nineteenth century in Western Europe and the late nineteenth century in most of the rest of the world. Specifically, the revenue data include at least one year for 18 Western countries and 76 non-Western countries. Fifteen Western countries have at least one data point in the nineteenth century, as do 42 non-Western countries. Relative world currency prices have fluctuated violently since the Bretton Woods system ended. For this reason, we analyze data only through 1969. Appendix Figure A.1 plots revenues over time for each territory in the dataset.

We are not the first to use the Mitchell revenue data for historical analysis.[26] However, our approach to weighting the data points enables us to incorporate more information than in existing studies. For example, Besley and Persson compute an unweighted average over time for eighteen rich countries.[27] Consequently, they do not calculate revenue collection for poorer countries in the nineteenth or early twentieth centuries, nor make time-series cross-section comparisons across a broad country and time sample. Others use Mitchell data from the nineteenth century, but only for European countries.[28] Yet others use Mitchell and other sources to construct a sample that is expansive globally after 1945, but confined to Western Europe, Japan, and the Southern Cone in the nineteenth century.[29]

Despite clear advantages of expansive country and time coverage, our approach to measuring state revenue intake also has drawbacks. Although we follow existing work on historical revenue collection by expressing revenue in gold or silver,[30] research on contemporary fiscal extraction typically examines government revenue as a percentage of GDP. Thus, comparing trends in per–capita revenue extraction does not rule out the possibility that differences in revenues mostly reflect changes in societal income. We address this concern in two ways. First, we analyze patterns for taxes as a

[26] Mitchell 1998.
[27] Besley and Persson 2014.
[28] Mann 1993, 358-401; Karaman and Pamuk 2010, 2013.
[29] Cagé and Gadenne 2018; Beramendi, Dincecco and Rogers 2019.
[30] Dincecco 2011, 2017; Karaman and Pamuk 2010, 2013.

fraction of GDP using data from Andersson and Brambor,[31] albeit at the cost of a restricted non-Western sample. Second, we compare the timing of revenue and income divergence to show that large increases in revenue intake among Western countries lagged large income gains by at least a half century.

1.2 Documenting the Great Revenue Divergence

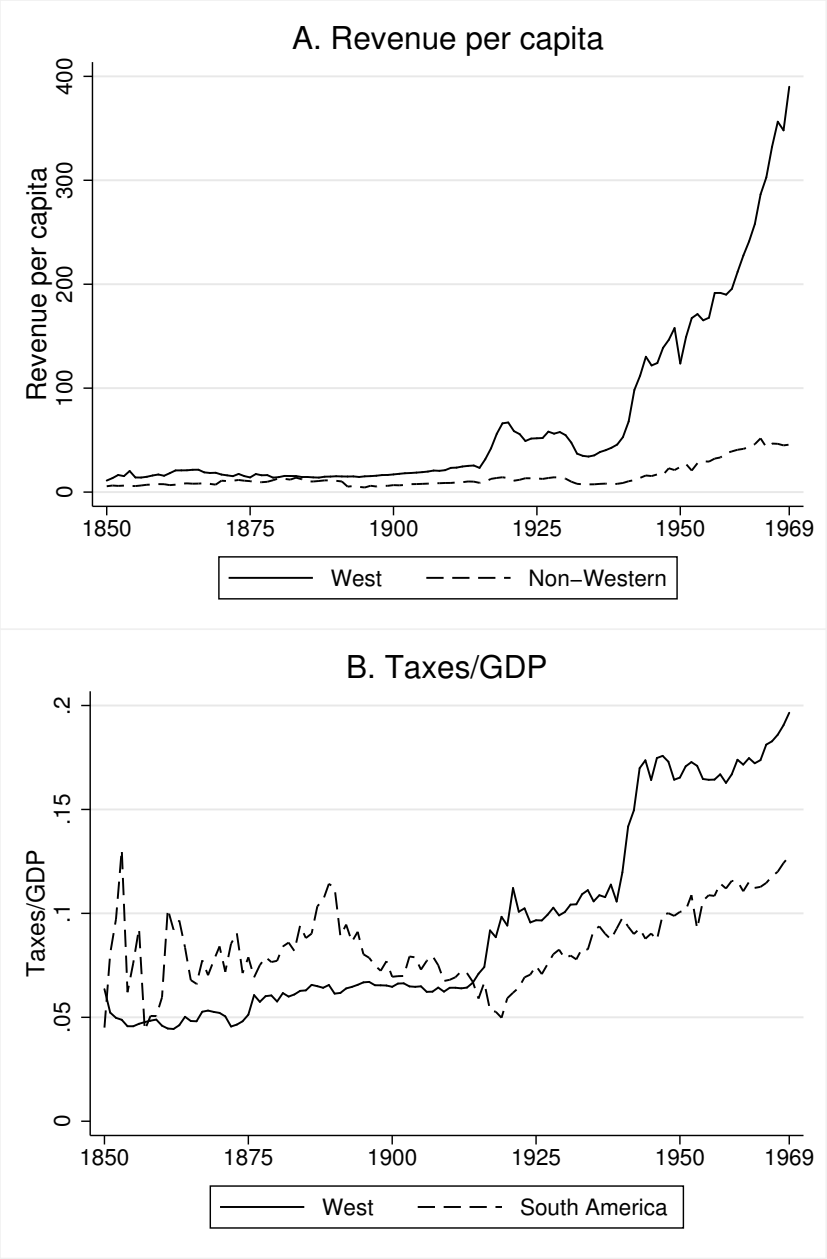
Figure 1 documents the great revenue divergence. Panel A presents our main measure of per-capita revenues. Panel B presents taxes as a fraction of GDP, albeit at the cost of a smaller sample in which non-Western countries are restricted to South America (plus Mexico).

Before World War I, the West did not dwarf other areas in revenue collection. In 1913, Chile and Uruguay each collected more revenue per capita than any country in Western Europe. Denmark collected less revenue than these two as well as Trinidad and Tobago, South Africa, Malaysia, Cuba, and Panama. The United States collected even less than that, and was slightly behind Brazil and slightly ahead of Jamaica. When compared to *all* non-Western countries, Western countries collected somewhat more revenue (2.5 times). However, this discrepancy is small by contemporary standards, and is mostly driven by meager revenues in many African colonies. Of the 49 non-Western countries in our dataset in 1913, 23 are in Africa, and Western countries collected 5.6 times more revenue than these colonies.

These patterns changed drastically after 1913. Between 1914 and 1969, per–capita revenue intake increased on average by 1,547% in Western countries. Sharp gains were not entirely confined to the West, as Japan experienced a sixteen-fold increase during this period. However, other non-Western countries failed to keep pace, and on average their revenues grew by 446%. Thus, gains outside the West were 71% smaller than those among Western countries. The patterns are largely similar among different subsets of non-Western countries: 359% increase in South America and the Caribbean, 428% in Africa, and 564% in Asia (Middle East, South Asia, Southeast Asia, East Asia

[31] Andersson and Brambor 2019.

Figure 1: The Great Revenue Divergence



Notes. The lines depict revenue intake averaged over Western and non-Western countries. In Panel A, the measure is central government revenue per capita in gold grams (converted at nominal exchange rates), as described above. In Panel B, the measure is taxes/GDP from Andersson and Brambor 2019.

excluding Japan).[32] By 1969, the ordering of countries in terms of per–capita revenue collection largely mirrors contemporary rankings, with nearly every country outside the West (except Japan)

[32]The countries in the sample in each region are not identical 1913 and 1969, and we verified that the magnitude of the increases were qualitatively similar when restricting the comparisons to

trailing every Western country. Overall, by this time, the average Western state collected 8.5 times more in per-capita revenue than the average non-Western state.

The divergence is also stark when assessing taxes as a fraction of GDP. As Panel B shows, South America outpaced the West on this measure throughout the nineteenth century. In 1913, Britain lagged Brazil, Uruguay, Chile, and Argentina. However, between 1914 and 1969, taxes rose from 6.4% to 19.7% of GDP in Western countries, a three-fold difference. The gains among South American countries were smaller, rising from 7.1% up to 12.7%. Overall, these gains were 41% smaller than those in the West.

The robustness of the main pattern to differences in GDP is unsurprising when we consider historical timing. When economic historians discuss a "great divergence," they mean the divergence in per–capita economic output between Western and non-Western countries.[33] Although scholars debate the timing and causes of this divergence, they agree it occurred no later than the mid-nineteenth century amid the spread of the Industrial Revolution across Europe. Figure 2 compares Western countries to non-Western countries on both revenues per capita and GDP per capita. Until World War I, Western countries typically had a larger advantage in GDP than they did in revenue collection. In the following decades, the revenue ratio increased more sharply than the GDP ratio.

1.3 Robustness Checks

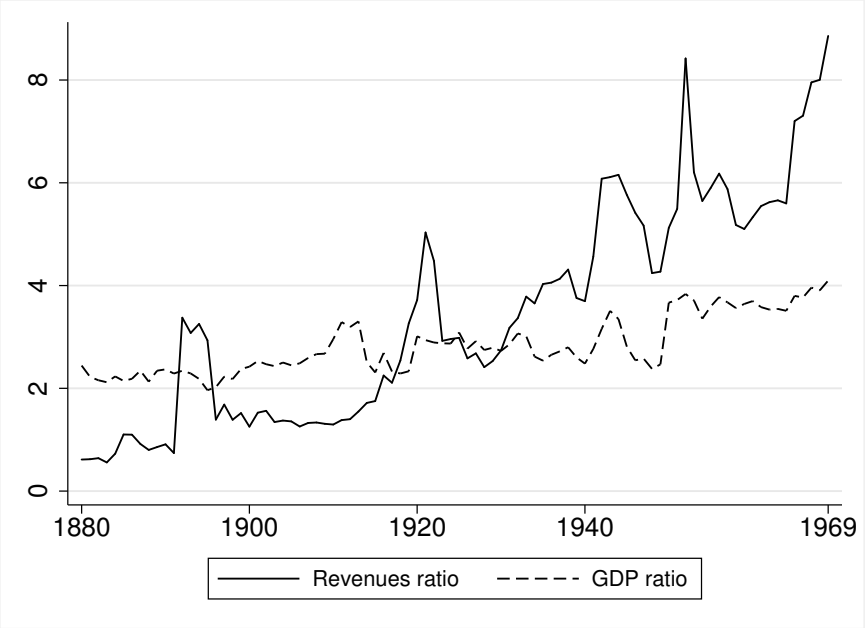
In the appendix, we analyze the robustness of our core pattern. In Figure 1, we use taxes/GDP from Andersson and Brambor.[34] Compared to alternatives, it (a) has lesser missingness relative to our core dataset (52% of country-years are missing) and (b) uses natural units. In Appendix A.1, we demonstrate qualitatively similar trends when using alternative datasets that account for differences

a constant basket of countries.

[33] Pomeranz 2009.

[34] Andersson and Brambor 2019.

Figure 2: Comparing Income and Revenue Divergence



Notes. Each line divides the average value of the outcome among Western countries by the average value among non-Western countries. The solid line depicts our core measure of revenues per capita, and the dashed line depicts GDP per capita in constant 2011 U.S. dollars from Bolt et al. 2018. In this figure, we include only country-years with both revenue and income data.

in GDP. We analyze taxes/GDP from Beramendi, Dincecco, and Rogers,[35] which is missing 69% of the country-years from our core dataset, and has 47% fewer observations for non-Western countries than Andersson and Brambor.[36] We also constructed a panel of normalized revenue data. Despite relatively better data coverage (missing 42% of country-years compared to our core sample), the units are non-natural because we divide *nominal* revenue intake in the local currency by *constant-U.S. dollar* GDP estimates from Bolt et al.'s update of Angus Maddison.[37]

One concern with our main measure of revenue per capita is that, by using nominal exchange rates, longitudinal changes in revenues may reflect changes in the foreign exchange market rather than changes in actual revenue. Appendix A.2 explains two ways in which our main measure guards against this concern. We also present intra-imperial comparisons, hence comparing territories that used the same currency or a highly stable peg. Nor can we directly account for differences in

[35] Beramendi, Dincecco and Rogers 2019.

[36] Andersson and Brambor 2019.

[37] Bolt et al. 2018.

purchasing power or directly measure tax intake. However, Appendixes A.3 and A.4 explain why these shortcomings are unlikely to influence the findings. We also created a separate series that expresses central government revenue per capita in silver, rather than gold, grams. Ultimately, the choice of precious metal does not qualitatively alter the main pattern. In Appendix A.5, we discuss why we chose gold rather than silver for our primary measure. Finally, in Appendix A.6, we estimate regression coefficients for the interaction of regional location and time period to express the core pattern from Figure 1 in more precise quantitative terms.

2 Existing Theories

Why did a large and permanent revenue divergence occur in the twentieth century, but not earlier? To answer this question, we build upon the rich existing literature on government revenues and state capacity. We categorize existing theories based on whether they focus on the *demand* for greater public spending, or the *supply* of bureaucratic institutions that facilitate revenue collection. Although both perspectives yield important insights, each is incomplete for explaining the twentieth–century great revenue divergence.

2.1 Fiscal Demand

Demand-based theories of taxation focus on factors that create stronger preferences for central government revenues. The most commonly studied demand factor in the literature is international warfare. Scholars broadly accept that external wars played an important role in facilitating modern European states.[38] Other authors make the converse argument that less intense geopolitical competition in many ex-colonies in Sub-Saharan Africa and Latin America has undermined their state-building efforts.[39]

[38] Tilly 1992; Brewer 1990.

[39] Herbst 2000; Centeno 2002.

Directly, preparation for and participation in an external war raises the state's need for revenue to pay and deploy soldiers for the conflict. Indirectly, these conditions may persist in a post-war ratchet effect. States need to service debt accumulated during the conflict, and wars can also spur permanent institutional changes. Mass-mobilization wars create political consensus for egalitarian taxation systems and franchise expansion.[40] These changes create pressure to sustain programs of social redistribution, which require high taxes to fund, that emerged during the war.[41] Besley and Persson formalize the war-demand logic.[42] The key choice in their model is government investment in future tax-collection capacity. A high valuation for public goods in "common-interest states" increases the value of future revenues, which boosts incentives for fiscal investments. External threats correspond with a high value of their public goods parameter.[43]

Despite highlighting some important elements, the bellicose perspective cannot explain the great revenue divergence on its own. Why did the Western revenue advantage remain large after its former colonies gained independence, which sparked high demand outside the West? For example, between 1940 and 1975, India fought in a world war under threat of invasion (during which it raised the largest volunteer army in world history), achieved independence alongside mass franchise expansion and an ascendent political elite strongly committed to social welfare measures, and engaged in three wars with Pakistan. Yet per–capita central government revenue intake was 67 times higher in Western Europe than India in 1969. Similar international pressures in the twentieth-century Middle East[44] and nineteenth-century South America[45] also failed to engender sustainably large revenue collection.

Some scholarship in the bellicose tradition incorporates additional factors that condition the effort of war on revenue collection. Examples include the presence of parliamentary institutions and

[40] Scheve and Stasavage 2016.
[41] Lindert 2004.
[42] Besley and Persson 2011.
[43] Besley and Persson 2011, 46-7, 58.
[44] Barnett 1992.
[45] Centeno 2002.

level of urbanization (Karaman and Pamuk, 2013), a pre-existing political union between the military and dominant social class (Centeno, 2002), and access to international debt markets (Queralt, 2019). The latter factor in particular likely provides a contributing factor to understanding the great revenue divergence. However, without incorporating bureaucratic capacity into bellicose theories, we cannot simultaneously explain why a large revenue divergence arose and became permanent in the twentieth century.

2.2 Fiscal Supply

Other scholars focus on the supply of fiscal institutions that facilitate revenue collection. The core element of fiscal, or bureaucratic, capacity is information about where citizens and other producers live and how much they produce. Standardized records enable bureaucrats to determine appropriate tax quotas and to sanction non-payers effectively, and make society "legible."[46] In low-legibility societies, citizens and other producers can exit by either physically migrating or engaging in informal economic activity beyond the state's reach. Throughout history, states have needed some bureaucratic capacity to collect taxes on land and to directly tax production. Modern income and value-added taxes are even more information-intensive.

Although the concept of bureaucratic capacity is inherently multi-faceted, recent research measures key components of states' information-collection abilities across broad comparative samples. Brambor et al. collected data on civil registration systems and state statistical offices dating back to the eighteenth century.[47] Data on births, deaths, and marriages is essentially a precondition for effective direct taxation because otherwise bureaucrats face difficulties to simply identifying the citizenry. Similarly, Lee and Zhang compiled data on the effectiveness of censuses in the twentieth century, which correlates strongly with public goods provision.[48]

[46] Scott 1998; Lee and Zhang 2017; Stasavage 2020.
[47] Brambor et al. 2020.
[48] Lee and Zhang 2017.

Fiscal *capacity* differs from revenue *intake*. States can collect information about production and life events (birth, death, marriage) without using it for taxation. States can govern a literate population capable of filling out written tax forms, without requiring them to do so. In such scenarios, fiscal capacity is *latent* and ready to use when the ruling group wishes.

Despite adding another important piece, analyses of fiscal supply also offer incomplete explanations for the great revenue divergence. Why did a large and permanent divergence not occur earlier? For most of the nineteenth century, Western countries outpaced others in terms of collecting information about their citizens and educating their population. Britain imposed the world's first modern income tax during the Napoleonic Wars. However, it suspended the income tax after the wars, and did not match its 1810 per–capita revenue record until 1915. Britain's high fiscal capacity remained largely latent throughout the nineteenth century. Similarly, other Western European countries improved their tax bureaucracies during the nineteenth century but underutilized their potential until the twentieth.

2.3 Sources of Fiscal Capacity

Conceptually, it is useful to distinguish war-based pressures from bureaucratic capacity. Yet empirically, these factors are not completely independent of each other. A key idea in bellicose theories of European state building is that participation in wars encouraged states to improve their bureaucracies. We agree that war is one important contributor to improvements in bureaucratic capacity over the longer run. This helps to explain why latent revenue-raising capacity was higher in the West than most of the rest of the world following the Napoleonic Wars. However, bellicose theories of fiscal-capacity investments cannot answer key questions about how fiscal capacity can increase in the absence of warfare, or can fail to increase for countries within a competitive international system.

Our empirical examination begins in the early nineteenth century. At this time, historical participation in wars had likely contributed to a divergence in bureaucratic capacity between European

countries and much of the rest of the world. European history provides numerous examples of states enacting bureaucratic reforms to gain a coercive edge. For example, Britain introduced the Bank of England in 1694 during the Nine Years' War with France, which created a major financing advantage.[49] Similar pressures during the Napoleonic Wars propelled Britain's first income tax. Mann examines several great powers (Britain, France, Prussia, Austria) and argues that warfare stimulated bureaucratic reforms before the French Revolution.[50] These reforms introduced standards for hiring and promotion, and shifted toward salaried rather than office-owning state officials.

However, even for explaining historical levels of fiscal capacity in Europe, purely bellicose theories are incomplete. Battles occurred as often in China as they did in Europe between 1000 AD and 1800,[51] and other aspects of their regional state systems influenced why bureaucratic capacity grew in Europe but fell in China.[52] Nor was the effect of war uniform throughout European history. Wars often generated crippling debt and encouraged leaders to take irresponsible actions such as debasing the currency (e.g., Louis XIV in France), as opposed to promoting fiscal systems that could generate consistent tax revenues over the longer term. Even in cases like Britain where scholars largely agree that participation in wars contributed to bureaucratic development in the eighteenth century, they also expound specific scope conditions such as early centralization, island location, and the *lack* of participation in wars during the early part of the Military Revolution.[53]

Even more pertinent for our empirical analysis, bellicose theories cannot account for two facts about fiscal capacity development in the nineteenth and twentieth centuries. First, at the beginning

[49] Brewer 1990.

[50] Mann 1993, 444-78.

[51] Dincecco and Wang 2018.

[52] Hoffman 2015. See Huang and Kang 2021 for a complementary discussion of how Japan and Korea built bureaucratic capacity in the first millennium AD. Despite not typically facing a threat of Chinese invasion, mimicking China's institutions solidified the domestic power of the ruling coalitions in Japan and China.

[53] Brewer 1990.

of the nineteenth century, no Western state (with the possible exception of Britain) possessed a "modern" bureaucracy. The historically unprecedented revenue increases that began during World War I would not have been possible without intensive improvements in fiscal capacity throughout the nineteenth century. Yet Western European states fought few wars with each other between 1816 and 1913. Why did fiscal capacity grow during an extended period of low warfare?[54]

Second, many ex-colonies faced high fiscal demand following independence from European powers. In some cases, high demand stemmed from a competitive regional environment and bellicose pressures. Why did most of these states fail to develop strong bureaucracies?

3 Overview of Theoretical Premises

To unravel the puzzle of the great revenue divergence, we develop a theoretical framework that combines bellicose and state-legibility factors. A government chooses how to raise revenues, and citizens decide whether to comply with tax demands or exit the formal economy. The strategic interaction occurs over two periods. We first motivate the core premises of the theory. In the next section, we present and solve a game-theoretic model.

3.1 Structure of the Tax System

The government chooses among three options for structuring the tax system. The first option is to exert low fiscal effort, such as relying on existing infrastructure to collect *customs revenues*. Col-

[54]During this period, Brambor et al. 2020, 202 find that participation in warfare is uncorrelated with information-capacity levels; in fact, their coefficient estimate is negative. Goenaga, Sabate and Teorell (2021) show that warfare was not a major driver of fiscal expansion in European countries during the long nineteenth century. Mann (1993) argues that although wars propelled bureaucratic reform in the eighteenth century, this factor was unimportant in much of the nineteenth century.

lecting customs taxes requires relatively few agents at one or several major ports. These indirect taxes are easy to collect if the economy is already organized in a manner to facilitate international trade. This was true of Western states by the nineteenth century. In many colonies and ex-colonies, intervention by the colonizer restructured the economy to produce certain cash crops. Centeno notes the contrast between "administratively simple but inelastic customs taxes" and "more politically challenging, but potentially more lucrative, domestic sources of revenue," which require greater bureaucratic capacity to collect.[55]

Low fiscal effort carries two drawbacks. First, low fiscal effort may not meet societal demand for revenues. Second, despite requiring low administrative effort, narrowly based taxes such as customs taxes often entail higher deadweight loss than more broadly based taxes such as an income tax.[56] Customs taxes can also create adverse distributional consequences for the ruling elite relative to an income tax.[57] As we discuss later, this created an impetus to reform tax systems in nineteenth-century Europe, despite the absence of strong war pressures.

These drawbacks may propel governments to choose either of two high-effort strategies. On the one hand, they can target a subset of producers to offer economic privileges (e.g., state-run monopolies or crony-owned firms) in return for revenue. This strategy can be lucrative because the government concentrates economic gains among highly legible cronies. However, what we classify as *crony-favoring extraction* entails high effort because it requires significant state involvement in and restructuring of the economy. Such restructuring enables only highly legible citizens to produce valuable goods, or facilitates direct government control over valuable assets. The clearest examples of crony-favoring economic interventions occur when governments construct state-owned enterprises or otherwise favor monopolies in certain industries. This creates a symbiotic political relationship whereby the government easily accesses information about the firm's production, and favored firms gain economic advantages. Collectivized agriculture in the Soviet Union provides

[55] Centeno 2002, 104.
[56] Albon 1997; Irwin 2010.
[57] Mares and Queralt 2015.

an extreme example. More typical cases are ones like Egypt and India in which the government actively intervenes in the economy to create a "captive tax base."[58] As Chaudhry describes, "In cases where the government becomes the primary employer and producer and assumes the role of setting prices, its task is simplified to monitoring the activities of corporations and agencies that it owns and manages."[59]

On the other hand, the government can create widespread economic rights and attempt to tax this broader base. *Legibility-intensive extraction* requires information about the populace as a whole. Efficient collection of income taxes and value-added taxes requires detailed information about the identities and productivity of citizens and firms, as well as complex bureaucracies to collect and process this information.

3.2 Bureaucratic Capacity

One factor that influences the government's strategy for raising revenues in each period is bureaucratic (or fiscal) capacity, about which we make two assumptions. First, societal legibility sticky in the short run. Thus, the state inherits a stock of bureaucratic capacity in each period. Second, over time, states can take concerted actions to boost fiscal capacity. Hence, choosing legibility-intensive extraction in period 1 can bolster fiscal capacity in period 2.

First, the idea that bureaucratic capacity is persistent is widespread in the literature. Consequently, states facing a demand shock are rarely able to rapidly and dramatically improve societal legibility. Dincecco's discussion of the historical origins of state capacity in Europe dates back to the fall of the Carolingian Empire in the 800s and extends into the twentieth century.[60] Conversely, once created, bureaucracies tend to self-perpetuate even when fiscal demand is temporarily low. Some legibility reforms can persist without any spending at all. For example, introducing last names

[58] Waterbury 1993, 134.
[59] Chaudhry 1993, 252.
[60] Dincecco 2017. See also Stasavage 2020.

and addresses enables states to find and distinguish citizens.[61] This information facilitates revenue extraction, even if not used immediately.

Second, states can take concerted efforts to improve societal legibility over time. Although this can occur through various channels, we focus on learning-by-doing effects. Thus, choosing legibility-intensive extraction early on can bolster future bureaucratic capacity, similar to fiscal-capacity investments in Besley and Persson.[62]

4 A Formal Model of Revenue-Extraction Strategies

The interaction of bellicose and state-legibility factors affects revenue intake, as we show in our formal model. Only states with high fiscal demand and high fiscal supply can achieve large revenue intake. In the short term, countries with varying levels of bureaucratic capacity may nonetheless generate similar levels of revenues, either because fiscal demand is low or because no states are greatly advantaged at collecting legibility-intensive taxes. However, over time, (exogenous) increases in demand or (endogenous) increases in bureaucratic capacity can generate a divergence in revenue collection.

Various factors can push governments to take concerted actions to increase fiscal capacity over time, including (1) relatively high initial bureaucratic stock, (2) anticipation of high fiscal demand in the future, (3) high potential for bureaucratic growth, and (4) high deadweight loss from easy-to-collect taxes. We discussed the first factor above: past participation in war contributed to the European advantage in fiscal capacity in the early nineteenth century, although other factors also influenced Europe's initial stock of bureaucratic capacity. The second factor is the same stimulant that creates "common interest states" in Besley and Persson.[63] Yet we also depart from bellicose theories by highlighting alternative features that are not intrinsically tied to high fiscal demand.

[61] Scott 1998.
[62] Besley and Persson 2011.
[63] Besley and Persson 2011.

Below we operationalize and elaborate upon the third and fourth factors in nineteenth-century Europe: industrialization created high potential for bureaucratic growth, and the ruling elite in many countries shifted to income taxes with low rates to reduce deadweight loss from narrowly based taxes.

4.1 Setup

We model an interaction between a government and citizens, which unfolds over two periods. Time is denoted by $t \in \{1, 2\}$, and players value consumption equally across the two periods. The sequence of moves within each period is: (1) the government chooses how to structure the tax system, (2) the government proposes a specific tax rate to each citizen, (3) each citizen either complies with the tax and produces in the formal sector, or exits to the informal sector.

Society consists of a continuum of atomless citizens with mass N, which has a lower bound that strictly exceeds 1 (see below) and an upper bound of $\overline{N}$ (defined in Appendix C). The set of citizens is denoted as $\mathcal{N}$, and each citizen is indexed by i. In both periods, each citizen produces output worth Y_i, the value of which is determined by the government's actions (see below). Each citizen also has an exit option that yields consumption of a fraction $e_i \in (0, 1)$ of ones output.[64] The exit option is individual-specific and, in each period, is independently drawn for each citizen from a smooth density function $H(e_i)$ with positive support on $[0, 1]$.[65]

In each period, the government decides how to structure the tax system. It begins each period with an endowment of customs revenues worth $R^{\text{cus}} > 0$. One option for structuring the tax system is to exert low effort and rely solely on customs revenues. Alternatively, the government can choose either of two high-effort strategies to collect additional taxes.

First, under a *legibility-intensive* strategy, the government grants legal rights to participate in the

[64] We omit time subscripts on Y_i and e_i to reduce notational clutter.

[65] We write the associated pdf as h. One proof requires the additional assumption $h' \leq 0$, which, for example, the uniform distribution satisfies.

formal economy to all citizens. This choice yields output $Y_i = 1$ for each citizen, which is subject to taxation. Due to limitations in bureaucratic capacity, the government does not know the value of the exit option for each citizen. A fraction $l_t \in (0, 1)$ of citizens are *legible*, and the government perfectly knows the value of e_i for each legible citizen. The remaining fraction $1 - l_t$ of citizens are illegible. The government knows only the prior distribution of possible values of e_i for such citizens. In any period that the government chooses legibility-intensive extraction, it pays a fixed cost $F \in \left(0, \overline{F}\right)$, with an upper bound $\overline{F} > 0$ defined later (see footnotes 70 and 74). We interpret F as a relative cost. When political elites perceive low-effort taxes such as customs taxes as creating extreme economic inefficiencies or unfavorable distributional consequences for themselves, in effect, the cost F of shifting the tax base is lower.

Second, under a *crony-favoring* strategy, the government favors a subset of legible citizens, normalized to mass 1.[66] This could involve limiting economic production to specific cronies, or putting economic production directly under state ownership. Restructuring the economy to reduce competition enables each favored citizen to produce $Y_i = Y$, which is subject to taxation, but pushes any production by the mass $N - 1$ of non-cronies outside the reach of the state, and hence $Y_i = 0$. To make the tradeoffs non-trivial, we assume (a) the crony-favoring strategy bolsters the production of favored citizens relative to their production under legibility-intensive extraction and (b) crony-favoring extraction diminishes total output. Formally, $1 \leq Y < N$, which also forms a lower bound for N. Structuring the tax system to favor cronies also incurs a fixed cost. Despite not requiring a similar bureaucratic effort as the legibility-intensive strategy, subsidies paid to favored firms and the difficulty of displacing vested economic interests create costs for a government to actively intervene to distort market competition. To isolate the role of bureaucratic capacity in distinguishing the two high-effort strategies, we assume that the government pays the same fixed cost F in any period it chooses crony-favoring extraction.

[66] Favored citizens are randomly drawn, and therefore have the same distribution of the exit option, $H(e_i)$, as the full set of citizens.

The fraction of legible citizens, l_t, reflects bureaucratic capacity (equivalently, fiscal supply). We assume l_1 is an exogenous parameter. However, l_2 depends in part on the revenue-collection strategy in period 1. If the government chooses legibility-intensive extraction in period 1, then $l_2 = \min\{\Delta \cdot l_1, 1\}$. Higher values of $\Delta > 1$ indicate stronger learning-by-doing effects, and hence higher potential bureaucratic growth. By contrast, if the government chooses low effort or crony-favoring extraction in period 1, then $l_2 = l_1$.[67] The endogeneity of l_2 to choices in period 1 is the only way in which decisions in period 1 affect those in period 2.

If, in a particular period, the government chooses a high-effort strategy for structuring the tax system, it then proposes an individual-specific tax rate $\tau_i \in [0, 1]$ to each citizen. For every legible citizen, the value of e_i is revealed before the government chooses τ_i; and for every illegible citizen, this value is revealed immediately afterwards. Each citizen knows its individual-specific values of τ_i, e_i, and Y_i when moving. Each simultaneously responds to its tax proposal either by complying and consuming $(1 - \tau_i) \cdot Y_i$, or exiting and consuming $e_i \cdot Y_i$.

The government's consumption in each period depends on revenues raised. All revenues are assumed to be spent on public goods. The government is rewarded for increasing revenue to get closer to the amount of expenditures demanded by society, denoted as R_t^{dem} for fiscal demand, and penalized for raising taxes such that total revenue exceeds the socially desired amount. The rationale here is that citizens want to fund desired public goods, but do not want to contribute taxes for undesired projects or private rents. Formally, total revenues R_t equal the customs endowment R^{cus} plus any additional taxes collected upon pursuing a high-effort extraction strategy. The government

[67] Assuming a deterministic relationship between revenue-extraction strategies and bureaucratic development reduces moving pieces, although the results would be qualitatively similar under alternative assumptions. For example, we could assume that bureaucratic capacity can atrophy over time under crony-favoring policies (e.g., India). We could also assume a small probability that states pursuing crony-favoring policies nonetheless experience a gain in bureaucratic capacity in period 2 (e.g., South Korea).

gains a marginal benefit of 1 from any (endogenously raised) taxes raised such that total revenues do not exceed R_t^{dem}, and a marginal benefit of -1 from any taxes that push total revenues above R_t^{dem}. Thus, if $R_t^{\text{dem}} < R^{\text{cus}}$, then the government gains negative utility from raising any taxes beyond the customs endowment. If instead $R_t^{\text{dem}} > R^{\text{cus}}$, then the government receives positive utility from any taxes raised up to $R_t^{\text{dem}} - R^{\text{cus}}$, and negative utility from raising any additional taxes. We assume $R_1^{\text{dem}} > 0$ is a fixed parameter and that Nature draws $R_2^{\text{dem}} > 0$ in between periods according to a distribution described later. To parallel our assumption about how endogenous tax revenue affects the government's utility, we assume the government's marginal benefit from any exogenous customs receipts is 1. However, the precise assumption here is immaterial because the government consumes customs revenues regardless of its strategic choices. Finally, in any period in which the government chooses a high-effort strategy for structuring the tax system, it pays the fixed cost F. This subtracts from its consumption.

4.2 Short-Term Revenue Intake: Analysis of Period 2

We solve backwards to characterize subgame perfect strategies, and all proofs are in Appendix C. In period 2, the government cares solely about short-term revenue intake. We first derive maximum possible revenues, and then we explain the government's optimal approach.

Each legible citizen will comply with a tax proposal that satisfies $\tau_i \leq 1 - e_i$. Thus, maximizing revenue requires the government to set the individual-specific tax rate to make each legible citizen indifferent between complying and exiting: $\tau_i^* = 1 - e_i$. This strategy induces every legible citizen to comply.[68] By contrast, a lack of discriminating information forces the government to set the same tax rate for each illegible citizen. The optimal rate balances two considerations. A higher tax rate yields a higher fraction of income, τ, from each illegible citizen who complies, but decreases the fraction of illegible citizens who comply. Only illegible citizens with low-valued exit options

[68] Any strategy profile in which citizens reject offers with positive probability when indifferent creates an open set problem for the government's offer, and hence cannot be an equilibrium.

(relative to the tax proposal) comply, $e_i \leq 1-\tau$. The revenue-maximizing tax proposal for illegible citizens solves:[69]

$$\hat{\tau} \equiv \arg\max_{\tau \in [0,1]} \int_0^{1-\tau} \tau \cdot dH(e_i). \tag{1}$$

Structuring the tax system to favor cronies yields a unit mass of favored citizens who each produce $Y_i = Y$ and are perfectly legible, and the remaining mass of $N-1$ citizens produce $Y_i = 0$. Legibility-intensive extraction enables all citizens to legally produce $Y_i = 1$, but only a fraction l_2 are legible. If the government makes the revenue-maximizing tax proposal to each citizen, expected revenues under each way to structure the tax system are:

$$R^{\text{crony}} = \underbrace{Y \cdot \int_0^1 (1-e_i) \cdot dH(e_i)}_{\text{All favored citizens are legible}}. \tag{2}$$

$$R^{\text{leg}}(l_2) = N \cdot \Bigg[\underbrace{l_2 \cdot \int_0^1 (1-e_i) \cdot dH(e_i)}_{\text{Legible citizens}} + \underbrace{(1-l_2) \cdot \int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)}_{\text{Illegible citizens}} \Bigg]. \tag{3}$$

Comparing these two terms shows that legibility-intensive revenue extraction yields higher revenues than crony-favoring policies if and only if bureaucratic capacity is high enough.

Lemma 1 (Maximum tax extraction)**.** *A unique threshold $\bar{l} \in (0,1)$ exists such that $R^{leg}(\bar{l}) = R^{crony}$. Given this threshold, we can express maximum tax revenues in period 2 as:*

$$R_2^{max} = \begin{cases} R^{crony} & \text{if } l_2 < \bar{l} \\ R^{leg}(l_2) & \text{if } l_2 \geq \bar{l}. \end{cases} \tag{4}$$

This result establishes the importance of fiscal supply. The next question is whether it is optimal for the government to maximize revenues, which depends on fiscal demand. If $R_2^{\text{dem}} < R^{\text{cus}}$, then the government can fund all desired expenditures while exerting low effort at tax collection. At the other extreme, if $R_2^{\text{dem}} > R^{\text{cus}} + R_2^{\text{max}}$, the government taxes maximally. Within these lower and

[69] In Appendix C, we prove that the maximizer is unique and strictly bounded between 0 and 1.

upper bounds, if fiscal demand is close to R^{cus}, then the fixed cost deters the government from pursuing high-effort extraction. If instead fiscal demand is close to $R^{\text{cus}} + R_2^{\max}$, then the government pays the fixed cost for either legibility-intensive or crony-favoring extraction, but intentionally sets taxes to collect less-than-maximum revenues. Proposition 1 presents a subgame perfect Nash equilibrium strategy profile.[70]

Proposition 1 (Optimal revenue extraction in period 2)**.**

- ***Low fiscal demand.*** *If $R_2^{dem} \leq R^{cus} + F$, then the government structures the tax system to exert low effort at tax collection, which yields $R_2 = R^{cus}$.*
- ***Intermediate fiscal demand.*** *If $R^{cus} + F < R_2^{dem} < R^{cus} + R_2^{max}$, then the government structures the tax system to exert high effort at tax collection. If $l_2 < \bar{l}$, the specific mode is crony-favoring; and legibility-intensive otherwise. The government sets $\{\tau_i\}_{i \in \mathcal{N}}$ to achieve total tax intake of $R_2^{dem} - R^{cus}$. This yields less-than-maximum revenues, $R_2 = R_2^{dem}$.*[71]
- ***High fiscal demand.*** *If $R_2^{dem} \geq R^{cus} + R_2^{max}$, then the government structures the tax system to exert high effort at tax collection. It sets $\tau_i = 1 - e_i$ for each legible citizen and $\tau_i = \hat{\tau}$ (see Equation 1) for each illegible citizen, which maximizes tax intake.*
  - ***Low fiscal supply.*** *If $l_2 < \bar{l}$, then the specific mode of high-effort extraction is crony-favoring and total revenues are $R_2 = R^{cus} + R^{crony}$.*
  - ***High fiscal supply.*** *If $l_2 \geq \bar{l}$, then the specific mode of high-effort extraction is legibility-intensive and total revenues are $R_2 = R^{cus} + R^{leg}(l_2)$.*[72]
- ***Citizens' responses.*** *Each citizen complies with any tax proposal satisfying $\tau_i \leq 1 - e_i$, and exits otherwise.*

[70] To eliminate strategically uninteresting cases, we restrict the upper bound on the fixed costs such that $\overline{F} < R^{\text{crony}}$. This ensures that the fixed costs are not high enough, on their own, to deter high-effort extraction.

[71] For reasons described in the appendix, a continuum of actions yield payoff-equivalent equilibria in this parameter range.

[72] Note that when fiscal demand R_2^{dem} is intermediate or high, the optimal choice for structuring the tax system does not depend on F. The government compares the expected utilities from legibility-intensive and crony-favoring extraction, and F cancels out because the fixed cost is identical for each.

Overall, existing bellicose arguments are correct that high fiscal demand stimulates governments to collect more revenues. However, by not incorporating the supply side of revenues, they cannot tell us the means by which governments will attempt to raise revenues nor how successful they will be. For states with low bureaucratic capacity that face a demand shock, distorting the economy yields greater revenue intake than attempting to tax a broader base. Yet in equilibrium, a government with high bureaucratic capacity that chooses legibility-intensive extraction will bring in more revenues than a low-capacity government that chooses crony-favoring extraction, even though both are acting optimally given their stock of bureaucratic capacity. To see why, recall the (substantively plausible) assumption that a broader-based economy yields a larger potential tax base, $Y < N$. This implies that revenues from crony-favoring extraction form a lower bound for maximum revenues. The government chooses legibility-intensive extraction only if this strategy yields higher revenues than this lower bound (see Lemma 1).

4.3 Investments in Fiscal Capacity: Analysis of Period 1

In period 1, the government cares not only about how the structure of the tax system affects contemporaneous revenues, but also revenue intake in period 2. The core implication is qualitatively unaltered: only governments with high fiscal demand and (the potential for) high fiscal supply choose legibility-intensive extraction. However, the threshold for "high" fiscal supply (i.e., bureaucratic capacity) is lowered because the shadow of the future heightens incentives to pursue legibility-intensive extraction. To make the analysis parallel with the preceding section, we characterize threshold values of initial bureaucratic capacity that determine optimal actions. We then take comparative statics on other variables that influence the government's calculus for investing in fiscal capacity.

If the starting level of bureaucratic capacity takes an extreme value, then the government's choice for how to structure the tax system in period 1 does not influence its choices in period 2. Suppose initial societal legibility is very high, $l_1 > \bar{l}$ (see Lemma 1 for this threshold). This guarantees

that legibility is high enough in period 2 that, regardless of the government's actions in period 1, legibility-intensive extraction yields more revenues than crony-favoring extraction. If instead initial legibility is very low, $l_1 < \underline{l} \equiv \frac{\bar{l}}{\Delta}$, then the converse implication is true. Even if the government gets the learning-by-doing boost to bureaucratic capacity in period 2, crony-favoring extraction would yield more revenue than legibility-intensive extraction.

Only if initial legibility is between these two thresholds does the shadow of the future yield new insights. To illuminate the substantively important insight, we focus on the following specific case within the range $l_1 \in \left(\underline{l}, \bar{l}\right)$. Fiscal demand is low in period 1, $R_1^{\text{dem}} < R^{\text{cus}}$. This ensures that the government will not seek additional revenues in period 1 beyond the customs endowment, even if it chooses to invest in fiscal capacity. This is the interesting case illuminated by the dynamic analysis because, in a single-shot game, the government would never choose legibility-intensive extraction if fiscal demand is low. We also assume that Nature draws fiscal demand for period 2 $\left(R_2^{\text{dem}}\right)$ from a Bernoulli distribution that takes value $R_{\text{high}}^{\text{dem}}$ with probability $p_{\text{high}} \in (0,1)$, and $R_{\text{low}}^{\text{dem}}$ with complementary probability. We set these values so that for a low draw of fiscal demand, the government chooses a low-effort tax structure and does not seek additional revenues in period 2 beyond the customs endowment; and for a high draw of fiscal demand, the government seeks maximum extraction. The formalization of these thresholds follows directly from Proposition 1: $R_{\text{low}}^{\text{dem}} < R^{\text{cus}} + F$ and $R_{\text{high}}^{\text{dem}} > R^{\text{cus}} + R^{\text{leg}}(1)$.[73]

If the government chooses legibility-intensive extraction in period 1, then its total expected consumption across the two periods is:

$$\underbrace{R^{\text{cus}} - F}_{\text{Period 1}} + \underbrace{R^{\text{cus}} + p_{\text{high}} \cdot \left[R^{\text{leg}}\left(\min\{\Delta \cdot l_1, 1\}\right) - F\right]}_{\text{Period 2}}. \tag{5}$$

If instead the government does not invest in bureaucratic development in period 1, then its expected

[73] For the lower bound of $R_{\text{high}}^{\text{dem}}$, note that $\max\{R^{\text{crony}}, R^{\text{leg}}(l_t)\}\big|_{l_t \in [0,1]} = R^{\text{leg}}(1)$.

utility is:

$$\underbrace{R^{\text{cus}}}_{\text{Period 1}} + \underbrace{R^{\text{cus}} + p_{\text{high}} \cdot \left(R^{\text{crony}} - F\right)}_{\text{Period 2}}. \tag{6}$$

Comparing these two terms uncovers the conditions under which the government invests in fiscal capacity for period 2: legibility-intensive extraction (after getting the learning-by-doing boost) must yield sufficiently more revenues than crony-favoring extraction. Equating the two preceding expressions enables us to define an implicit threshold $\hat{l}$ that determines whether the government makes this investment:[74]

$$R^{\text{leg}}\left(\Delta \cdot \hat{l}\right) = R^{\text{crony}} + \frac{F}{p_{\text{high}}}. \tag{7}$$

As in the period 2 analysis, the stock of bureaucratic capacity influences the government's optimal revenue-raising strategy. However, in period 1, the government may choose legibility-intensive extraction even if fiscal demand is low and crony-favoring extraction yields higher maximum revenues. The future gains created by bureaucratic growth change the government's calculus. Proposition 2 formalizes this intuition.[75]

> **Proposition 2** (Optimal fiscal capacity investments)**.** *Suppose $l_1 < \bar{l}$. Assuming the scope conditions for the special case described above, the government chooses legibility-intensive extraction if and only if $l_1 > \hat{l}$ (defined in Equation 7). This threshold satisfies $\hat{l} > \underline{l}$ for all parameter values, and $\hat{l} < \bar{l}$ for low-enough F.*

We now take comparative statics on factors that determine when the government invests in fiscal capacity in period 1. Given the preceding proposition, it is immediately apparent that higher levels of initial bureaucratic stock encourage fiscal capacity investments. We discussed the empirical application of this result above: various factors, including past participation in warfare, precipitated higher l_1 in Western than non-Western countries in the early nineteenth century. In Proposition 3, we analyze three additional factors that increase the range of parameter values in which the gov-

[74] To ensure that $\min\{\Delta \cdot \hat{l}, 1\}$ is interior, we need to place another restriction on the upper bound $\overline{F}$, which we discuss in the appendix.

[75] The proposition characterizes only the optimal investment decision. The thresholds that characterize the full set of equilibrium actions are intuitive given the preceding analysis of period 2.

ernment invests in fiscal capacity (i.e., decrease $\hat{l}$). One is a greater likelihood that fiscal demand is high in the future. This recovers a core result from Besley and Persson,[76] and highlights how the anticipation of future wars (or other sources of high fiscal demand) can propel fiscal capacity investments.

The other comparative statics results in Proposition 3 reflect factors that are independent of war: greater potential for bureaucratic growth (higher Δ) and lower costs to implementing legibility-intensive extraction (lower F). Later we discuss why these factors propelled investments in fiscal capacity in nineteenth-century Europe. Industrialization and urbanization made it, in principle, easier to collect information about citizens. This increased the returns to developing a meritocratic bureaucracy, hence raising Δ. Additionally, elites became increasingly concerned about the dead-weight losses and adverse distributional consequences of customs taxes, relative to income taxes. A more costly status quo made elites more willing to tolerate the disruption created by reforming the tax system, hence lowering F.

> **Proposition 3** (Comparative statics on fiscal capacity investments)**.** *Each of the following changes in parameter values decrease* $\hat{l}$*:*
>
> - *Higher* p_{high}
> - *Higher* Δ
> - *Lower* F

4.4 Implications for Revenue Divergence

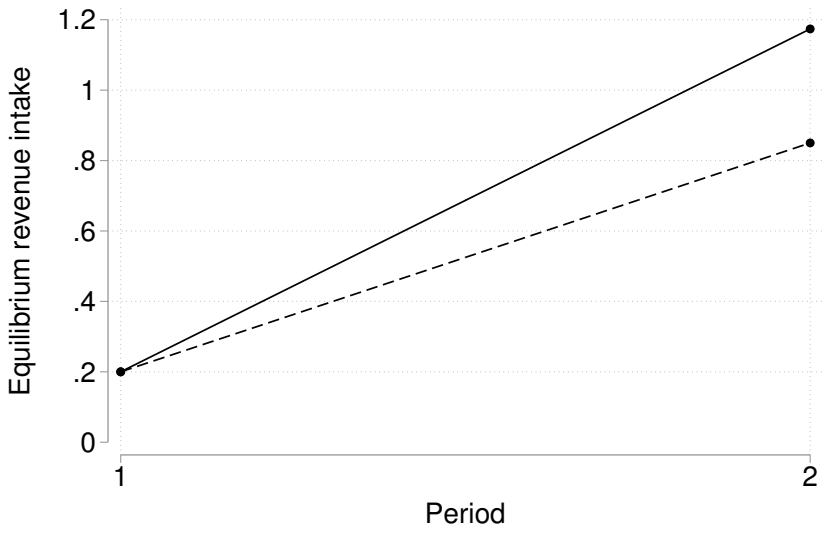
The model explains why increases in fiscal demand over time can create a revenue divergence between states that vary in bureaucratic capacity. High-capacity states distinguish themselves in revenue collection only when fiscal demand is high.[77]

[76] Besley and Persson 2011.

[77] Footnote 67 highlights that adding stochastic elements to the model would not qualitatively change the core intuitions.

In Figure 3, we depict hypothetical revenue trajectories for two countries that are identical except in their initial level of bureaucratic capacity, one with low $l_1 < \hat{l}$ and one with intermediate $l_1 \in (\hat{l}, \bar{l})$. We assume that fiscal demand is low for both in period 1. Consequently, neither government raises revenue beyond their endowed customs taxes. However, the government with intermediate initial bureaucratic capacity (solid line) nonetheless will sink the fixed cost to create future foundations for legibility-intensive extraction. By contrast, the low-legibility state (dashed line) has no incentive to invest in future fiscal capacity despite anticipating an identical probability of high fiscal demand in period 2. The learning-by-doing effects from investing in bureaucratic capacity are sufficiently small that this government would continue to respond to high fiscal demand in the future with crony-favoring extraction.

Figure 3: Hypothetical Revenue Trajectories



Notes. Parameters are $H \sim U(0,1)$, $N = 1.9$, $F = 0.05$, $R^{\text{cus}} = 0.2$, $Y = 1.3$, $p_{\text{high}} = 0.8$, and $\Delta = 3$. For these parameter values, the threshold values are $\underline{l} = 0.122$, $\hat{l} = 0.167$, and $\bar{l} = 0.368$. For the solid line, $l_1 = 0.35$. Any value $l_1 < \hat{l}$ yields the trajectory depicted by the dashed line.

If Nature draws high fiscal demand for both governments in period 2, then revenue divergence occurs. Although both extract maximally, the government with higher bureaucratic capacity chooses legibility-intensive extraction and gains higher revenue intake.[78] By contrast, the low-legibility

[78]For these parameter values, the higher-capacity state would not choose legibility-intensive

state chooses crony-favoring extraction. Existing models, such as Besley and Persson,[79] cannot account for this divergence. In their model, any state that anticipates high fiscal demand in the future is a "common value state" that will invest in fiscal capacity to capitalize on demand shocks. However, in our model, bureaucratic capacity conditions the effect of demand shocks. In the short term, low fiscal supply pushes high-demand states toward crony-favoring rather than legibility-intensive extraction. In the long term, states with poor prospects for bureaucratic growth will have lower incentives to invest in fiscal capacity even if they anticipate high fiscal demand in the future.

5 Empirical Evidence for Theoretical Implications

Our main theoretical implication is that large revenue intake requires the conjunction of high fiscal supply and high fiscal demand. Evidence from the nineteenth and twentieth centuries supports this expectation. Western countries enjoyed an advantage in bureaucratic capacity, which grew over time. However, fiscal demand was low until World War I. Numerous non-Western primary product exporters and some agrarian empires kept pace or caught up with the West. Later, high fiscal demand in the West propelled these states to capitalize on their latent advantages in fiscal capacity. Huge discrepancies in revenue collection emerged relative to non-Western countries. After independence, their colonies also experienced demand shocks. However, low supply prevented high levels of revenue collection, which explains why the large revenue divergence was permanent.

In Appendix B, we propose one way to operationalize fiscal demand and supply for a large-N sample. Using two-way fixed-effects models, we demonstrate that participation in war (which proxies for demand shocks) exhibits a positive and statistically significant association with revenues only in countries with an experienced civil registration system (which proxies for high bureaucratic capacity). Thus, the interaction effect is positive.

extraction in period 2 absent investment in bureaucratic capacity in period 1.

[79] Besley and Persson 2011.

5.1 Low Fiscal Demand in the Nineteenth-Century West

Between the conclusion of the Napoleonic Wars and the start of World War I, revenue intake was low in Western countries. Bellicose theories anticipate this pattern because infrequent intra-European warfare lowered fiscal demand. However, despite minimal pressures from war, fiscal capacity grew throughout the century. Our theory accounts for non-bellicose stimulants to bureaucratic reform.

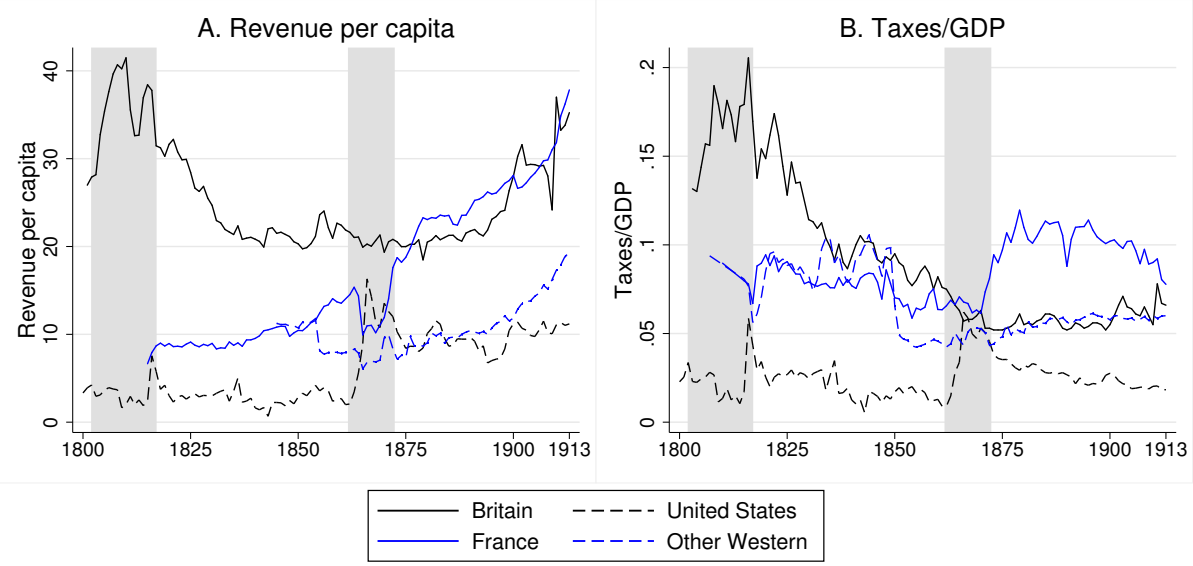
In Figure 4, we highlight low revenue intake in important states. Britain imposed the world's first modern income tax during the Napoleonic Wars,[80] a period we highlight in gray. However, per–capita revenue intake declined afterwards, and this decline is even more pronounced when accounting for Britain's strong economic growth. Even when Britain reimposed an income tax in 1842, the marginal rate began at 2.9%, and it remained low into the twentieth century. France imposed a new set of direct taxes starting with its Revolution, but did not implement an income tax until World War I. The United States experienced a brief spike in revenues when it imposed an income tax during its Civil War and Reconstruction (1862–72), which we also highlight in gray. However, for most of the nineteenth century, the U.S. government was a "state of courts and parties." In the 1870s, customs revenues from the Port of New York accounted for more than half of all federal revenues.[81] Customs revenues were sufficient to cover the small federal budget, and during the century they constituted on average 72.1% of U.S. revenues. The major outlier among Western countries was New Zealand, which we omit from the Other Western average in the figure to not obscure the main pattern. Throughout the nineteenth century, New Zealand consistently collected high levels of customs taxes relative to the small white population.

Low revenue intake reflected low fiscal demand. The long nineteenth century was more peaceful than the preceding or subsequent periods. Britain, for instance, participated in a major war against at least one other European power for 76 of the 150 years from 1665 to 1815, but in only three

[80] Dincecco 2017.

[81] Skowronek 1982, 24, 61.

Figure 4: Western Revenue Intake Pre-WWI



Notes. See note for Figure 1. The range of the scale for the y-axis in Panel A is one-tenth that of Figure 1.

years between 1816 to 1913. Wars of imperial conquest or within established colonies occurred more frequently (61 years for Britain). However, these conflicts were much less costly than intra-European conflicts. For example, the First Anglo-Burmese war and the Anglo-Zulu war each cost £5 million, and the First Anglo-Afghan War cost £14 million. By contrast, Britain's participation World War I cost £3.25 billion. The main reason for lower costs was that imperial wars required small commitments from metropolitan troops.[82] To assess this claim systematically, we analyzed whether war years correlate with a larger mobilization of domestic British troops.[83] When defining "wars" as intra-European conflicts, the correlation is positive and statistically significant. By contrast, when defining "wars" as imperial, the correlation is null.

Similarly, until the very end of this period, a limited franchise dampened domestic incentives for social provision and redistributive taxation. Britain did not provide old-age pensions until 1908, unemployment insurance until 1911, or universal secondary education until 1918. Britain's upper-

[82] Crowder 1971, 6-10.

[83] Data from Onorato, Scheve and Stasavage 2014. Details in Appendix A.8.

class political leadership would not personally benefit from heavy taxation, and they faced minimal demand from their middle-class electorate to boost expenditures. Across Western Europe, demand for welfare provisions was low throughout the nineteenth century.[84]

Given low fiscal demand, bellicose theories anticipate retrenchment in Britain and many continental powers. However, these theories cannot account for why, nonetheless, fiscal capacity tended to increase during this period. All ten countries that introduced civil registration systems for births and deaths before 1850 are in Western Europe or its offshoots.[85] Age heaping in U.S. censuses declined by 62% among native whites between 1850 and 1900, with even larger improvements among other racial groups.[86] These gains were essential for enabling Western states to collect unprecedented levels of revenues when fiscal demand spiked starting in 1914.

Our model helps to explain why Western states invested in fiscal capacity despite low fiscal demand (and hence they underutilized their fiscal capacity). One relevant factor was that European countries began the period with greater fiscal capacity than most non-Western countries, that is, l_1 was high. This was, in part, of a legacy of prior wars, as discussed above. Yet two other key conditions from the model that explain bureaucratic growth were independent of war.

First, every Western country experienced industrialization prior to World War I.[87] This factor increased the potential for bureaucratic growth, which corresponds with higher Δ in the model. Industrialization reshaped citizens in ways that made it easier for states to control and tax them. Residents of cities are easier to monitor than villages, and it is easier to impose income taxes on cash wages than on harvests consumed as subsistence. The basic literacy, numeracy, and awareness of time required by industrial firms created positive spillovers for states. Zhang and Lee provide quantitative evidence that literacy is strongly associated with state capacity, and industrialization and urbanization facilitated increases in literacy.[88]

[84] Lindert 2004.
[85] Brambor et al. 2020.
[86] A'Hearn, Baten and Crayen 2009, 793.
[87] Pomeranz 2009.
[88] Zhang and Lee 2020.

Second, European states inherited irrational and inefficient state structures from the eighteenth century. This, in effect, reduced the net costs of reforming the tax system, hence lowering the fixed cost F in the model to implementing legibility-intensive extraction. Reforming the bureaucracy enabled political elites to reduce economic deadweight losses. Even the early nineteenth–century English bureaucracy, often cited as a global model, was recruited through patronage networks and purchase. Many so-called bureaucrats had sinecures and were recompensed by fees rather than salaries.[89] Statesmen took actions throughout the nineteenth century to eliminate these abuses and to recruit bureaucrats by examination.[90]

Similarly, many governing elites considered the income tax to be more equitable than the easily collected tariffs and excise taxes it replaced, which also bolstered quasi-voluntary compliance.[91] Sir Robert Peel's speech proposing the reintroduction of the British income tax in 1842 stated that the resulting surplus would be used in "making of great improvements in the commercial tariff in England; in addition to these improvements to abate the duties on some great articles of consumption,"[92] and in fact a general reduction in tariffs occurred around that time. Previously, state officials viewed the income tax as appropriate only at times of serious fiscal crises, such as wars. However, reformers like Peel argued that an income tax with low marginal rates offered a welfare-improving means to pay for the ordinary costs of peacetime administration.[93] Across the continent, landed elites often pushed for the introduction of income taxes not to finance greater social expenditures, but instead to ensure that rising capitalist elites incurred a greater share of the total tax burden.[94]

[89] Brewer 1990.
[90] Mann 1993.
[91] Levi 1989.
[92] Young and Douglas 1964, 430.
[93] For similar arguments in France, see Bonney 2010, 96.
[94] Mares and Queralt 2015.

5.2 Primary Product Exporters and Empires in the Nineteenth Century

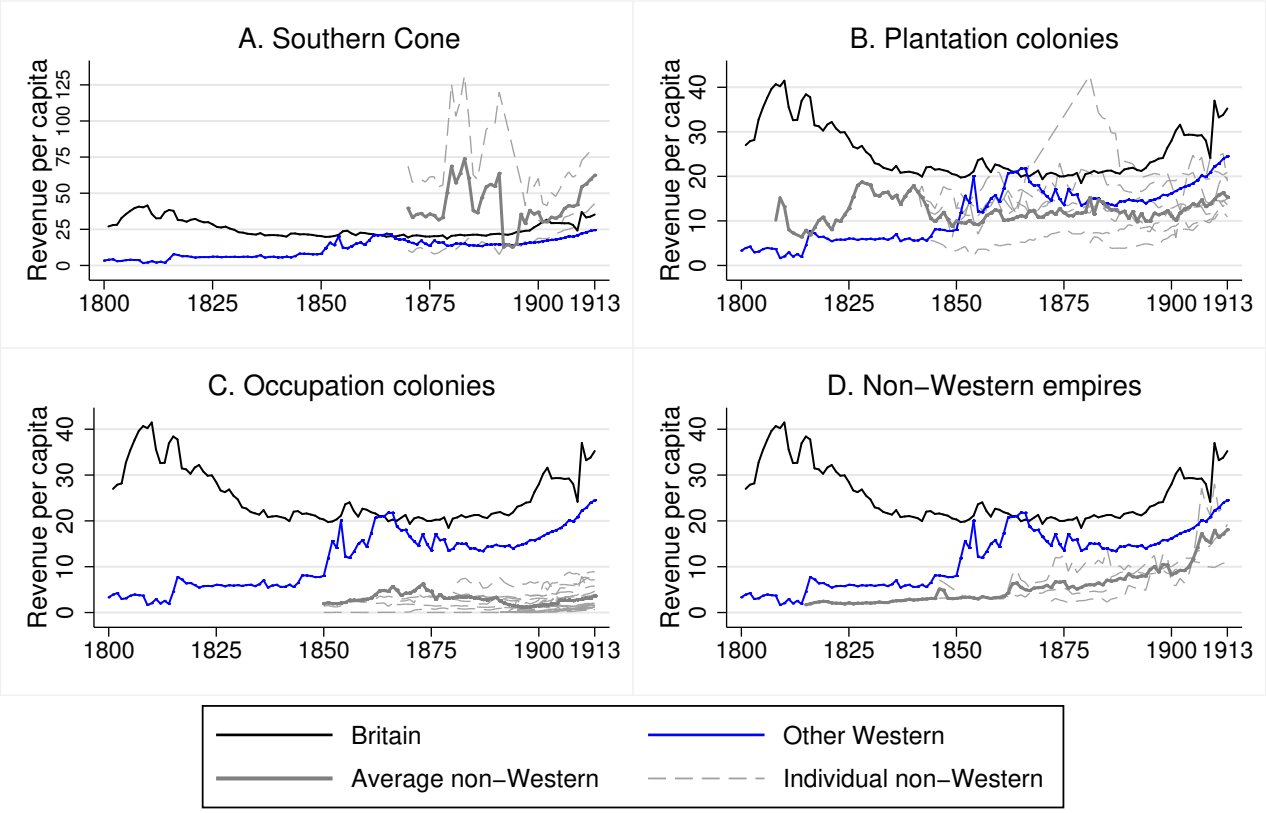
Customs revenues in primary product exporters. Even with low demand for public expenditures in the West, we might still expect these states to collect more revenue than states then (or recently) under Western colonial rule. European colonial rule was typically based on predatory extraction, which perpetuated low fiscal supply. Colonizers shaped the fiscal systems of dependencies to reflect metropolitan objectives. Most colonies made extensive use of either coercive labor institutions or local intermediaries, both of which tended to reduce the central colonial government's cash receipts. Spanish administrators in the Americas plundered their colonies for gold and silver, often using indigenous forced labor for mining and other production purposes. Elsewhere, Europeans forcibly imported millions of Africans to work as slaves on plantations throughout the West Indies and other areas where the climate permitted the production of sugar and other valued commodities. African colonies were characterized by high labor coercion and financial decentralization.[95] Britain collected head, hut, and other direct taxes in Africa through Native Authorities acting on the state's behalf.[96] Throughout Africa, colonists co-opted local indigenous institutions and aimed simply to collect enough taxes to balance the budget. Similarly, in the majority of areas in colonial India, colonial officials delegated the collection of the land revenue tax to princes or large landlords, and the government was usually unable to raise the rates they assessed.[97]

Despite not boosting legibility, European colonizers structured local economies to facilitate primary product exports. This enabled some dependencies to keep pace with the West without high fiscal effort. Europe's dominance in this period was based on superior military technology, scientific innovations and economic development, and epidemiological advantages, rather than high levels of taxation. In Figure 5, we compare Britain (black line) and average revenue intake in other Western countries (blue line) to four baskets of non-Western countries between 1800–1913. We plot individual non-Western countries in dashed gray, and their average in a thick, solid gray

[95] Van Waijenburg 2018.
[96] Gardner 2012.
[97] Lee 2019.

line.

Figure 5: Comparative Revenue Intake Pre-WWI



Notes. The lines show estimated central government revenue per capita in gold grams, converted at nominal exchange rates. The set of non-Western countries are as follows. Panel A: Argentina, Chile, Uruguay. Panel B: Barbados, Cuba, Fiji, Guyana, Jamaica, Malaysia, Mauritius, Trinidad and Tobago. Panel C: Algeria, Benin, Cameroon, Cyprus, Ghana, Guinea, India, Indonesia, Ivory Coast, Madagascar, Malawi, Niger, Nigeria, Senegal, Sri Lanka, Tanzania, Togo, Uganda, Zambia, Zimbabwe. Panel D: Egypt, Japan, Russia.

Countries in the Southern Cone of South America gained independence in the first half of the nineteenth century. On average, their revenue intake was relatively high. Between 1900–13, the Southern Cone countries collected 33% more in revenue per capita than Britain, and more than twice the amount of revenue of other Western countries, on average. Measuring taxes as a percentage of GDP reveals similar discrepancies. Whereas Britain and the average of all other Western countries each raised 6.4% of GDP in taxes, the corresponding figure for Southern Cone countries

was 9.2%. Revenues were particularly high in Chile, which reflected a boom in nitrate mining.[98] Between 1900–13, customs taxes constituted, on average, 71.3% of Chile's total revenues. In addition to the ease for primary product exporters to collect customs taxes from a handful of ports, high demand also contributed to revenue extraction in Chile. Victory in the War of the Pacific (1879–83) cemented the influence of domestic coalitions that favored an expansive, extractive state.[99]

In Panels B and C, we plot revenues from territories that were, at the time, subjected to colonial occupation. We distinguish between two types of colonies: plantation colonies in which a high fraction of the population was forced migrants engaged in production of cash crops on plantations, and colonies of occupation with largely indigenous populations. Plantation colonies, with more direct rule and high levels of trade, collected somewhat less revenue than European countries: from 1900–13, 55% less than Britain, and 31% less than other Western European countries.[100] However, these gaps are strikingly small compared to modern discrepancies or when considering the vastly superior bureaucratic institutions in the West. Furthermore, when normalizing by GDP, the advantage flips. Plantation colonies collected 72% more than Britain, and more than two times the average of other Western European states.[101]

Western countries were clearly distinguished from occupation colonies in revenue intake, even before World War I. Between 1900–13, Britain collected nearly thirteen times more in revenue per capita than occupation colonies, and other Western countries collected eight times more. Yet once again, these magnitudes were small by modern standards, and differences in GDP account for most of the discrepancy. When normalizing by income, Britain collected only two times more in revenue than occupation colonies, and the rest of the West collected only 65% more than occupation

[98] Brown 1963.

[99] Schenoni 2021, 418-19.

[100] These differences in per–capita revenue collection, as well as those for occupation colonies (see below), are similar in magnitude to Frankema's (2010) estimates, who compiled his revenue data for the British empire from colonial Blue Books.

[101] We lack taxes/GDP data for these observations; see Appendix A.1 for a discussion of our normalized revenue variable.

colonies.[102]

One possible concern is that comparing sovereign and non-sovereign polities yields misleading conclusions. Specifically, perhaps colonizers exploited their colonies to fund expenditures at home, which would enable them to keep domestic taxes low. However, this alternative explanation is unlikely to explain away the patterns presented here. It cannot explain why independent states in the Southern Cone extracted large amounts of revenue, nor why occupation colonies in Africa and Asia extracted minimal revenue. Research by economic historians shows that in the largest empires (Britain and France), colonial subsidies and defense expenditures exceeded in magnitude any revenue intake, which departed from the goal of financial self-sufficiency in the colonies. Analyzing Britain in the half century preceding World War I, Davis and Huttenback argue that the empire is better characterized as "a redistribution of income within the United Kingdom than as a transfer from the empire to the mother country."[103] Although many European investors benefited from colonial rule, this was possible because of the security environment funded by metropolitan taxpayers. Only in the small empires with one or several profitable colonies (Dutch, Belgian, Portuguese) did the empire contribute a significant net inflow to the metropole, mirroring patterns from imperial Spain in earlier centuries.[104] These authors also stress that "colonial revenues were first and foremost needed to secure *internal* order . . . [and] to pay the salaries of government officials who administered the government departments."[105]

Reforms in non-Western empires. In the final panel in Figure 5, we compare the West to major

[102]Despite broad coverage of revenue data, we lack GDP data for most occupation colonies (only India, Indonesia, and Sri Lanka). However, these colonies are not outliers in revenue collection. Between 1900–13, these three colonies collected, on average, 33% more in per-capita revenues than the entire group of occupation colonies. This suggests that differences in GDP between Western countries and occupation colonies explain most of the gap in per–capita revenue intake.

[103]Davis and Huttenback 1982, 119.

[104]Frankema and Booth 2019, 6-8.

[105]Frankema and Booth 2019, 5; emphasis in original.

non-Western empires. Many scholars highlight a large gap in revenue intake between the West (in particular Britain) and major non-Western empires at the end of the eighteenth century.[106] Despite this early mini-divergence, by the beginning of the twentieth century, the gap had narrowed between these empires and the West. We attribute this pattern to high fiscal demand, which stimulated either legibility-intensive extraction (Japan) or crony-favoring extraction (Russia and Egypt).

We have data for three major non-Western states before World War I: Egypt, Japan, and Russia.[107] Like several other empires (China, Ethiopia, Ottoman, Siam), these states engaged in defensive modernization programs to resist Western encroachment. Their ruling elites perceived high demand for centralized revenues, even in years that these states did not actively participate in war. Reforms in Japan followed two centuries of isolation and decentralized rule under the Tokugawa Shogunate, when demand for public expenditures was low. Japan enjoyed a long history of domain-level taxation and a professional state service,[108] which facilitated the implementation of a civil registration system in 1874. Consequently, Japan caught up to the West in per–capita revenue intake by 1913, and may have raised more when accounting for differences in GDP.[109]

Russia and Egypt highlight how crony-favoring extraction can yield comparable revenue intake to states with superior bureaucracies but that face low demand. In our dataset, Russia converged toward Western revenue intake during the nineteenth century. Our first data point is for 1815, when revenue collection in Britain was 22.6 times higher than in Russia, and in France was 3.9 times higher. This is consistent with an early revenue divergence shown by other scholars. In fact, this

[106] Karaman and Pamuk 2010, 623; Rosenthal and Wong 2011, 175; Hoffman 2015, 51; Dincecco 2017, 69.

[107] In Appendix A.7, we discuss this sample of non-Western empires and the revenue data in more detail.

[108] Sng and Moriguchi 2014.

[109] The datasets we use exhibit discrepancies on the latter point. Although Japan raised less on Andersson and Brambor's (2019) taxes/GDP measure, it raised more on Beramendi, Dincecco and Rogers's (2019) taxes/GDP measure and our normalized revenue measure.

gap between the West and Russia at the conclusion of the Napoleonic wars is even larger than the discrepancies listed by Dincecco in the 1780s,[110] which were 6.6 and 3 for Britain and France, respectively. However, the gap narrowed considerably by the onset of World War I. In response to defeat in the Crimean War, the Russian state initiated a drive to industrialize and build railroads. To finance this drive, the Russian state engaged in various crony-favoring methods to raise revenue. In 1902, state monopolies and state domains accounted for 56% of revenues, compared to only 7% for direct taxes. The liquor monopoly (established in 1895) itself constituted 25% of total revenues.[111] Between 1900–13, Britain collected only 83% more revenue per capita than Russia, and other Western states only 27% more. Although we lack GDP data for Russia during this period, it is likely that this relatively small gap is entirely explained by income differences.

In Egypt, Muhammad Ali unleashed an ambitious program to reform the military and economy. He engaged in bureaucratic reforms, but the state administration remained highly personalized. Instead, consistent with a crony-favoring strategy, he ordered the cultivation of numerous cash crops (in particular cotton) and established monopolies to buy them at low prices from peasants and then sell them on the world market for a profit.[112] In the 1870s, Western countries collected 62% more in revenue per capita than Egypt, and Egypt collected slightly more in normalized revenue.

5.3 Surging Revenues in the Twentieth-Century West

Starting with World War I, Western governments experienced permanently high demand for revenues. This change in conjunction with their bureaucratic foundations for effective legibility-intensive extraction yielded sustainably large revenue intake, as we demonstrated in Figure 1.

The two world wars required unprecedented mobilization of troops, reorganization and manage-

[110] Dincecco 2017, 69.

[111] Calculated by the authors from *The Statesmans Yearbook* for 1904.

[112] Ralston 1990, 84, 91.

ment of the economy to supply the war effort, and financing needs. European states overhauled their tax systems,[113] and experienced pressure to expand the franchise and provide citizens with a broad array of social welfare benefits to reward their sacrifices,[114] even in countries that did not directly participate in the wars.[115] The Great Depression as well as geopolitical competition during the Cold War also stimulated demand for activist states.

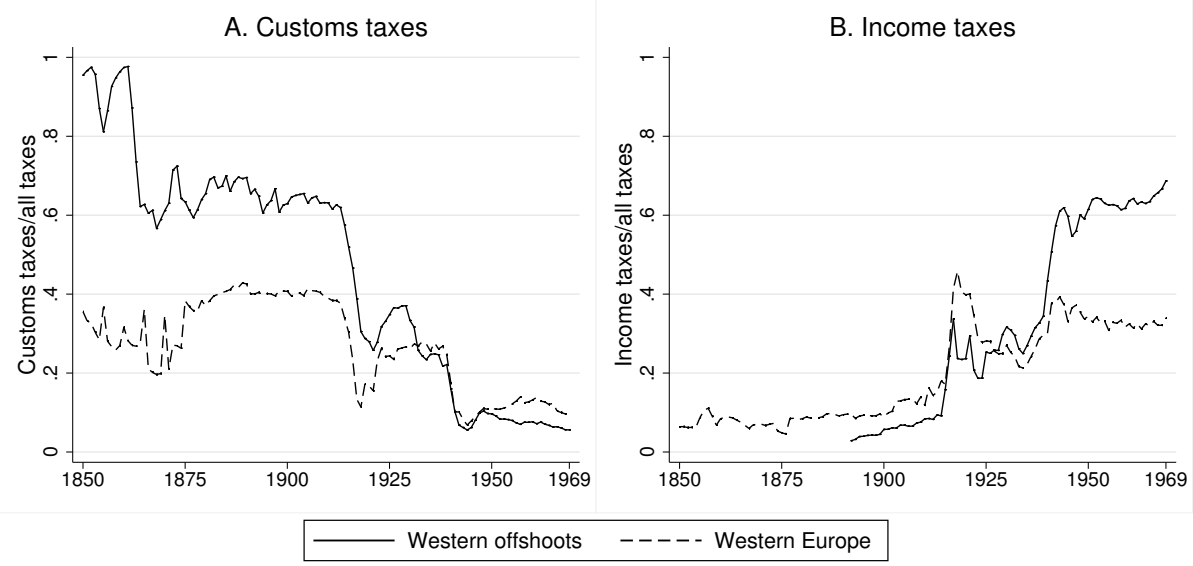
The prior legacy of high bureaucratic capacity enabled Western states to capitalize on these demand shocks. Improvements in fiscal capacity during the nineteenth century, which we described earlier, were a precondition for unprecedented increases in revenue intake. Legibility-intensive extraction—in particular, income taxes—replaced narrowly based taxes as the primary revenue source in Western states. Heavily reliant on bureaucratic competency and societal legibility, income taxes represented a major technological breakthrough in taxation capacity. Mares and Queralt praise the "unprecedented revenue generating capacity" of "the most advanced fiscal instrument to date."[116] This is also true of advanced consumption taxes, such as the value-added tax that became common in Western Europe.[117] Each tax requires high social legibility to collect efficiently. Income taxes are hard to evade if citizens receive monetary income by check or transfer, and value-added taxes are difficult to evade if firms routinely provide and receive invoices for sales and purchases.

In Figure 6, we present the fraction of revenue deriving from either customs or income taxes for Western states. Customs taxes once constituted the main source of revenues in Western offshoots and were also sizable in Western Europe. However, by the second half of the twentieth century, they were largely unimportant. In 1969, customs taxes comprised 6% of revenues in Western offshoots and 10% in Western Europe. By this time, income taxes were the main source of revenues for Western offshoots (69%). Income taxes were less important in Western Europe (34%)

[113] Scheve and Stasavage 2016.
[114] Lindert 2004.
[115] See Appendix Figure B.2 for evidence on the non-belligerents in WWI.
[116] Mares and Queralt 2015, 1975.
[117] Steinmo 1996.

because, instead, these countries relied more heavily on advanced consumption taxes. In 1969, all direct taxes plus advanced consumption taxes constituted, on average, 54% of revenues in Western European countries.[118]

Figure 6: Sources of Western Revenues: Customs and Income Taxes



Notes. Data from Andersson and Brambor 2019.

5.4 Low Fiscal Capacity in Former Western Colonies

The revenue gap between the West and the rest of the world increased exponentially after World War II, when most of the colonized world gained independence, as we showed in Figure 1. Insufficient bureaucratic reforms during the colonial period offer a more compelling explanation than bellicose theories for the general inability of non-Western states after gaining independence to converge toward Western revenue intake. Despite generally high demand in the post-colonial world, inadequate bureaucratic capacity hindered revenue intake.

Under colonial rule, the predominant strategies of taxing cash-crop exports and relying on local

[118]Computed by authors by summing the relevant categories from Mitchell 1998.

intermediaries did not require advanced bureaucracies. For example, for direct taxation, British administrators in Africa largely relied on head or hut taxes because they were "[u]nable to collect information on individual taxpayers and their incomes." They varied the rate of taxation by district based on the crude assumption that Africans "living in areas close to the railways or opportunities for wage labour could afford to pay a higher rate than those living in more remote regions."[119] Easy-revenue sources often satisfied the limited needs of colonial states before World War II, at least relative to the costs of constructing more intensive systems.

Independent countries in South America in the nineteenth century also had an alternative to building intensive bureaucracies, even when they competed in wars. Low global interest rates enabled states to pay for wars with debt, which they often renegotiated after the wars, rather than to develop intensive forms of domestic tax collection.[120] The availability of international capital reduced the elasticity problems inherent in low-legibility taxation. Instead, states could mortgage future customs and mineral revenues to gain short-term financing.

When fiscal demand rose after independence in the twentieth century, neglected bureaucratic reforms during the critical juncture of colonial rule became problematic. Most former colonies lacked a civil registration system at independence, which we use as a proxy for bureaucratic development in Appendix B, although most early-independence South American countries established a civil registration system in the late nineteenth century. At independence, India had 46 times as much census-age misreporting as the United States.[121] Given low supply, we anticipate that heightened fiscal demand after gaining independence should not discernibly boost revenue collection. Lee and Paine provide quantitative support for this contention by demonstrating null differences in countries' revenue intake before and after independence.[122]

In many post-colonial countries, low legibility persisted long after independence. Many lack ex-

[119] Gardner 2012, 116.
[120] Centeno 2002; Queralt 2019.
[121] Lee and Zhang 2017.
[122] Lee and Paine 2019.

tensive written or electronic records to monitor activity, or banking intermediaries that reduce the need for government agents to meet in person to collect taxes. In some African and Asian countries, customs revenues became *more* important in the mid-twentieth century. Governments gained freedom to set tariff rates, and older land or labor taxes declined in importance or were abolished by post-independence governments intent on reforming the colonial system. Bates explains how many African rulers after independence used funds from agricultural marketing boards (which serve the ostensible purpose of stabilizing prices for and revenues from primary products) to raise revenues by exploiting farmers.[123] Even when non-Western states have tried to impose modern legibility-intensive taxes, a lack of bureaucratic capacity has often impeded collection. In 1969, the average non-Western country collected 28% of its revenues from income taxes, and 20% of its revenues from customs taxes.[124] One exception was South Africa (51% of revenues from income taxes), which was highly effective at raising taxes within the white community.[125]

Egypt and India provide striking contrast cases for bellicose theories. Despite frequent participation in international warfare, these states collect low levels of revenue. A state needs high bureaucratic capacity to make its society legible. When this condition fails, rulers turn to crony-favoring extraction. Egypt and India each faced high demand for revenue given their participation in prolonged international rivalries (with Israel and Pakistan, respectively) that on several occasions flared into war. Yet both developed large and inefficient public sectors, as opposed to cultivating more sustainable sources of revenues. "The SOE sector does represent a captive tax base, and even as the SOEs run at a loss and seek financing abroad, they still generate a predictable source of taxes and compulsory payments to various fiscal agencies."[126] Egypt's attempt to implement a broad land reform in the 1950s and 1960s, which would have cut out large landowners as inter-

[123] Bates 1981.

[124] Data for eight South American countries (and Mexico) from Andersson and Brambor 2019. Other countries are authors' calculations from Mitchell 1998: Egypt, India, Indonesia, Iran, Pakistan, Philippines, South Africa, South Korea, Turkey.

[125] Lieberman 2003.

[126] Waterbury 1993, 134.

mediaries in the tax-collection process, failed due to basic problems of bureaucratic information about land titles and related issues.[127] In India, the proportion of revenue collected through direct taxes fell during the twentieth century, from 28% in 1900 to 15% in 2000 and with a low of 6.5% in 1987.[128] Tax-avoidance rates remained high and the government chose to raise import duties and nationalize large sectors of the economy. In 1969, each country collected less revenue per capita than the average non-Western country, and a low share of their revenues came from income taxes (15% in Egypt and 17% in India).

The main exceptions to the general pattern of fiscal weakness in the non-Western world are the "developmental states" of East Asia. Our theory anticipates these exceptions, which combined high supply and demand. Japan, Taiwan, and South Korea all had long traditions of professionalized bureaucracies, in fact, longer than those in the West. These countries experienced high demand for revenue to fund participation in World War II, their subsequent recovery, Cold War rivalries (including the Korean War), and ambitious programs of infrastructural development and public service provision. In 1969, Japan ranked ninth globally in per–capita revenue collection, and exceeded the Western average. Japan collected 62% of its revenue from income taxes, which was in line with Western offshoots. As late as 1964, Taiwan collected less per-capita revenue than the average non-Western country, but by 1969 nearly twice as much. South Korea collected 33% of its revenues from income taxes in 1969, in line with the Western European average.[129] Taiwan and South Korea further converged to Western patterns in subsequent decades.

[127] Migdal 1988.

[128] See *Statistical abstract relating to British India from 1894–95 to 1903–04*, Table 45; and *Handbook of Statistics on Indian Economy 2018–19*, Table 96.

[129] We lack internationally comparable data on per-capita revenue in South Korea because its currency was not convertible.

6 Conclusion

During the twentieth century, a historically unprecedented divergence in revenue intake occurred between Western countries and the rest of the world. Vast and permanent differences in revenue collection emerged much later than existing theories would expect. We explain the cross-sectional and longitudinal trends by distinguishing existing explanations based on bellicose and state-legibility factors. Whereas most existing research examines these factors in isolation, we provide a theory of how demand shocks can cause governments to engage in either legibility-intensive or crony-favoring extraction. We show that the optimal choice depends on extant bureaucratic capacity as well as the ability to boost societal legibility in the future. We then provide empirical evidence from the nineteenth and twentieth centuries to establish that the conjunction of high fiscal demand and high fiscal supply produced sustainably high revenue increases funded by income and value-added taxes.

Our new theory and dataset enable us to push beyond particular regions, specific time periods, and individual types of taxes. We build on existing theories and empirical findings to facilitate a broad comparative analysis of transformations in revenue intake over the past two centuries. Our framework centers around the importance of bureaucratic development and states' information-gathering capabilities. Wars undoubtedly contributed to improved fiscal capacity and revenue collection in some European cases in both the early modern period and the twentieth century. However, this effect was not constant over either time or space. In many European states, fiscal capacity grew but remained latent for much of the nineteenth century. Conversely, a bellicose environment typically did not help non-Western countries to improve fiscal capacity. The main exceptions were East Asian countries, such as Japan, that experienced not only high demand, but also a prior history of bureaucratic development.

Our perspective also highlights insufficient bureaucratic development as central to understanding low taxation in the non-Western world, as opposed to states enduring too few or the "wrong" kinds of wars. In the nineteenth century, states with largely illegible societies but valuable primary

products—which required minimal bureaucratic capacity to generate revenues—could keep pace with the West. However, once demand picked up across the globe, states with low fiscal capacity were heavily restricted in their ability to raise modern sources of revenues such as income taxes. Where favorable bureaucratic preconditions were absent, revenue extraction remained modest regardless of the frequency or types of wars.

References

A'Hearn, Brian, Jörg Baten and Dorothee Crayen. 2009. "Quantifying quantitative literacy: Age heaping and the history of human capital." *The Journal of Economic History* 69(3):783–808.

Albon, Robert. 1997. "The Efficiency of State Taxes." *Australian Economic Review* 30(3):273–287.

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

Barbieri, Katherine, Omar Keshk and Brian Pollins. 2008. "Correlates of War Trade Dataset." *Codebook Version* 2.

Barnett, Michael N. 1992. *Confronting the Costs of War in Egypt and Israel*. Princeton University Press.

Bates, Robert H. 1981. *Markets and States in Tropical Africa*. University of California Press.

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

Besley, Timothy and Torsten Persson. 2011. *Pillars of Prosperity*. Princeton University Press.

Besley, Timothy and Torsten Persson. 2014. "Why Do Developing Countries Tax So Little?" *Journal of Economic Perspectives* 28(4):99–120.

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.

Bonney, Richard. 2010. The Apogee and Fall of the French Rentier Regime 1801–1914. In *Paying for the Liberal State*, ed. José Luís Cardoso and Pedro Lains. Cambridge University Press pp. 81–102.

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.

Brewer, John. 1990. *The Sinews of Power*. Harvard University Press.

Brown, Joseph R. 1963. "Nitrate Crises, Combinations, and the Chilean Government in the Nitrate Age." *Hispanic American Historical Review* 43(2):230–246.

Cagé, Julia and Lucie Gadenne. 2018. "Tax Revenues and the Fiscal Cost of Trade Liberalization, 1792–2006." *Explorations in Economic History* 70:1–24.

Centeno, Miguel Angel. 2002. *Blood and Debt*. Penn State Press.

Chaudhry, Kiren Aziz. 1993. "The Myths of the Market and the Common History of Late Developers." *Politics & Society* 21(3):245–274.

Crowder, Michael. 1971. *West African Resistance*. Africana Publishing Corporation.

Davis, Lance E and Robert A Huttenback. 1982. "The Political Economy of British Imperialism: Measures of Benefits and Support." *Journal of Economic History* 42:119–130.

Denzel, Markus A. 2010. *Handbook of World Exchange Rates, 1590-1914*. Ashgate Publishing, Ltd.

Dincecco, Mark. 2011. *Political Transformations and Public Finances*. Cambridge University Press.

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

Dincecco, Mark and Yuhua Wang. 2018. "Violent Conflict and Political Development over the Long Run: China versus Europe." *Annual Review of Political Science* 21:341–358.

Frankema, Ewout. 2010. "Raising Revenue in the British Empire, 1870–1940." *Journal of Global History* 5(3):447–477.

Frankema, Ewout and Anne Booth. 2019. *Fiscal Capacity and the Colonial State in Asia and Africa, C. 1850-1960*. Cambridge University Press.

Frankema, Ewout and Morten Jerven. 2014. "Writing History Backwards or Sideways: Towards a Consensus on African Population, 1850–2010." *Economic History Review* 67(4):907–931.

Furnivall, J.S. 2014. *Colonial Policy and Practice*. Cambridge University Press.

Gardner, Leigh. 2012. *Taxing Colonial Africa*. Oxford University Press.

Goenaga, Agustín, Oriol Sabate and Jan Teorell. 2021. "The State Does Not Live by Warfare Alone: War and Revenue in the Long Nineteenth Century." *Working Paper* .

Herbst, Jeffrey. 2000. *States and Power in Africa*. Princeton University Press.

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

Huang, Chin-Hao and David C Kang. 2021. "State Formation in Korea and Japan, 400–800 CE: Emulation and Learning, Not Bellicist Competition." *International Organization* pp. 1–31.

International Monetary Fund. 2017. "IMF Data." Available at http://www.imf.org/en/Data.

Irwin, Douglas A. 2010. "Trade Restrictiveness and Deadweight Losses from US Tariffs." *American Economic Journal: Economic Policy* 2(3):111–33.

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

Karaman, K. Kivanc and Şevket Pamuk. 2013. "Different Paths to the Modern State in Europe." *American Political Science Review* 107(3):603–626.

Lee, Alexander. 2019. "Land, State Capacity, and Colonialism: Evidence From India." *Comparative Political Studies* 52(3):412–444.

Lee, Alexander and Jack Paine. 2019. "What Were the Consequences of Decolonization?" *International Studies Quarterly* 63(2):406–416.

Lee, Melissa M. and Nan Zhang. 2017. "Legibility and the Informational Foundations of State Capacity." *Journal of Politics* 79(1):118–132.

Levi, Margaret. 1989. *Of Rule and Revenue*. University of California Press.

Lieberman, Evan S. 2003. *Race and Regionalism in the Politics of Taxation in Brazil and South Africa*. Cambridge University Press.

Lindert, Peter H. 2004. *Growing Public*. Vol. 1 Cambridge University Press.

Mann, Michael. 1993. *The Sources of Social Power*. Vol. 2 Cambridge University Press.

Mares, Isabela and Didac Queralt. 2015. "The Non-Democratic Origins of Income Taxation." *Comparative Political Studies* 48(14):1974–2009.

Migdal, Joel S. 1988. *Strong Societies and Weak States*. Princeton University Press.

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

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.

Pomeranz, Kenneth. 2009. *The Great Divergence*. Princeton University Press.

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

Ralston, David B. 1990. *Importing the European Army*. University of Chicago Press.

Rosenthal, Jean-Laurent and Roy Bin Wong. 2011. *Before and Beyond Divergence*. Harvard University Press.

Saylor, Ryan and Nicholas C. Wheeler. 2017. "Paying for War and Building States." *World Politics* 69(2):366–408.

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*. Princeton University Press.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press.

Skowronek, Stephen. 1982. *Building a New American State*. Cambridge University Press.

Sng, Tuan-Hwee and Chiaki Moriguchi. 2014. "Asia's Little Divergence." *Journal of Economic Growth* 19(4):439–470.

Stasavage, David. 2020. *The Decline and Rise of Democracy: A Global History from Antiquity to Today*. Princeton University Press.

Steinmo, Sven. 1996. *Taxation and Democracy*. Yale University Press.

Tilly, Charles. 1992. *Coercion, Capital, and European States, AD 990-1992*. Blackwell Oxford.

Van Waijenburg, Marlous. 2018. "Financing the African Colonial State: The Revenue Imperative and Forced Labor." *Journal of Economic History* 78(1):40–80.

Waterbury, John. 1993. *Exposed to Innumerable Delusions*. Cambridge University Press.

Young, George M and David Charles Douglas. 1964. *English Historical Documents*. Eyre & Spottiswoode.

Zhang, Nan and Melissa M Lee. 2020. "Literacy and State–Society Interactions in Nineteenth-Century France." *American Journal of Political Science* 64(4):1001–1016.

Online Appendix

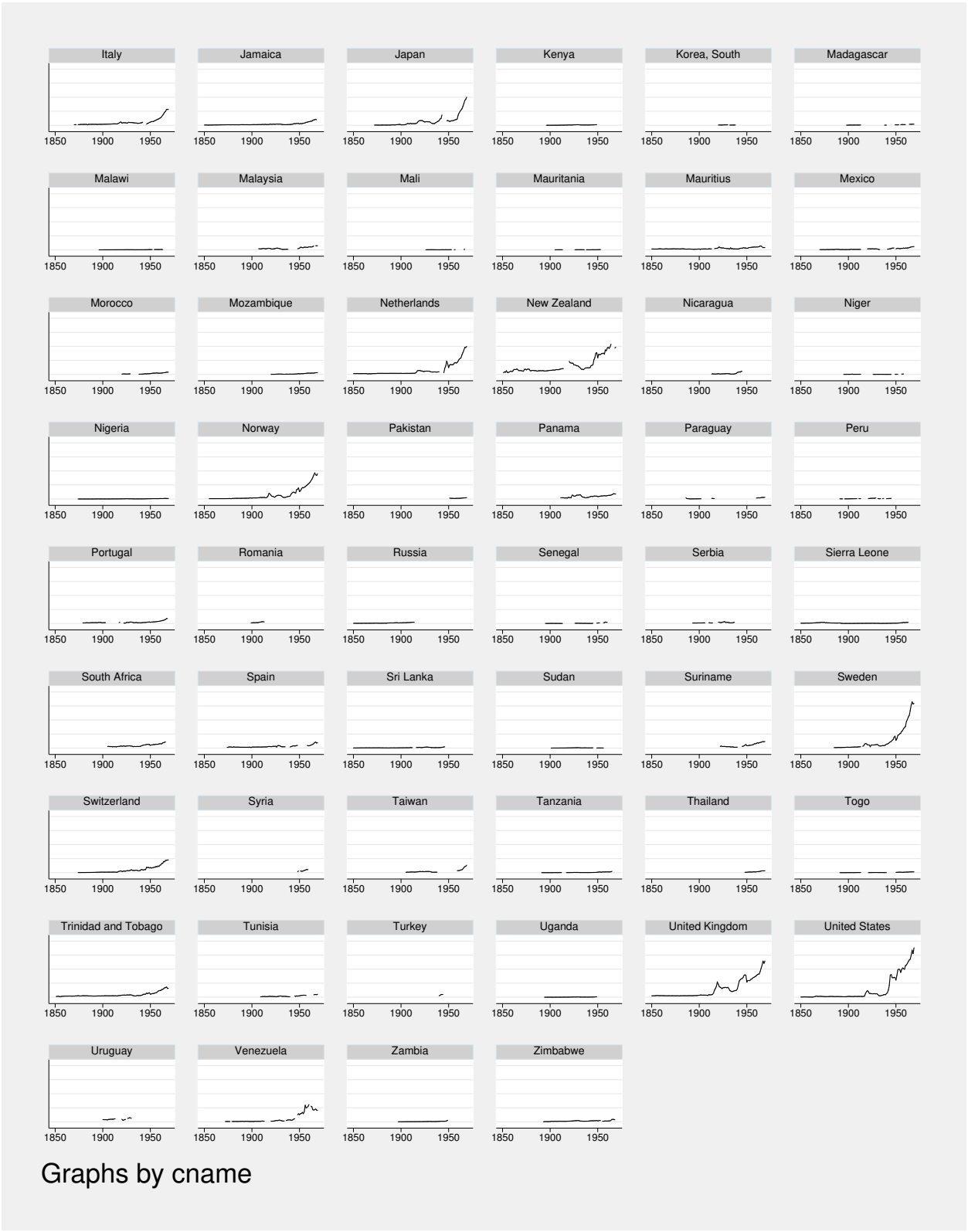
Contents

List of Tables

List of Figures

A Supporting Information on Data and Patterns

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



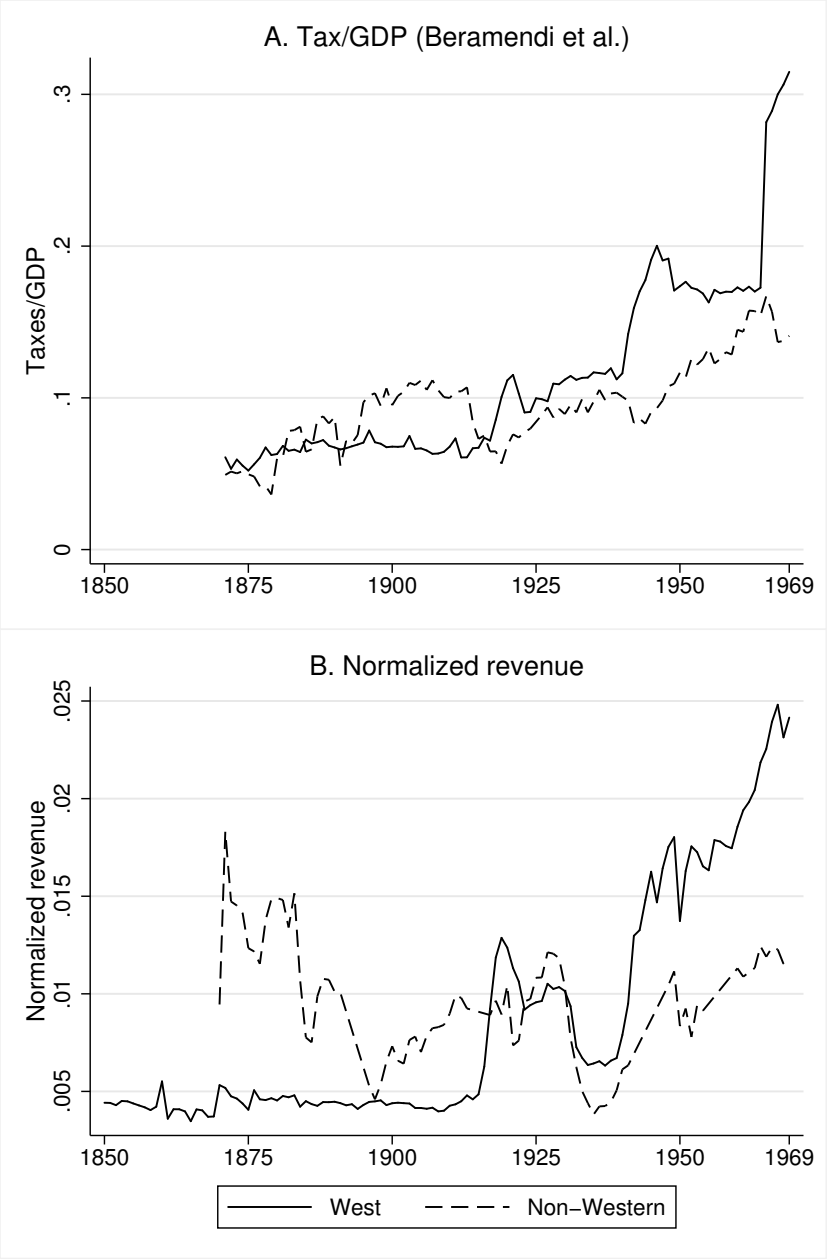
Notes. For each panel in Figure A.1, the lines show estimated central government revenue per capita in gold grams, converted at nominal exchange rates. Gaps indicates years of missing data. The range of the y-axis is 0 to 800.

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, Dincecco and Rogers (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, Dincecco and Rogers'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



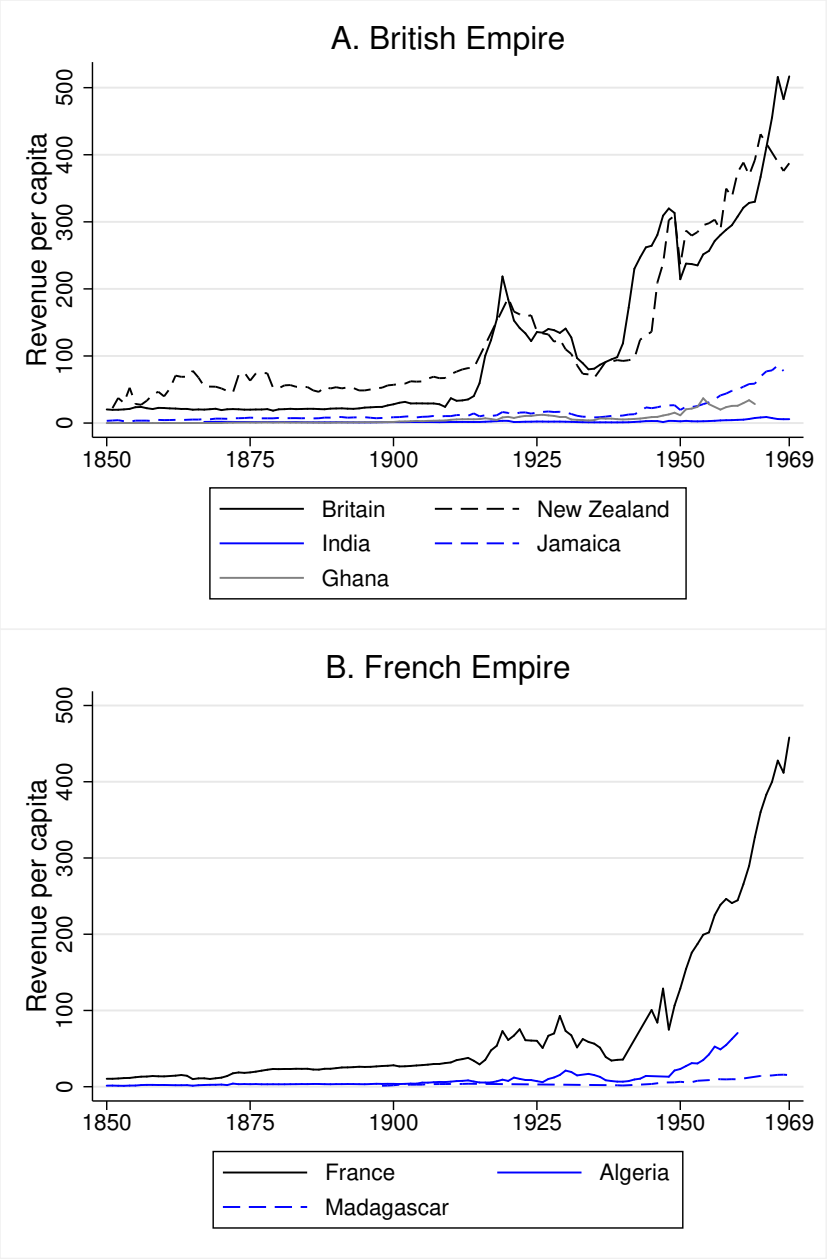
Notes. Taxes/GDP from Beramendi, Dincecco and Rogers (2019). Panel B: normalized revenue. The above text provides more details on each variable.

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 (Summers and Heston, 1991), 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 Tables dataset. 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

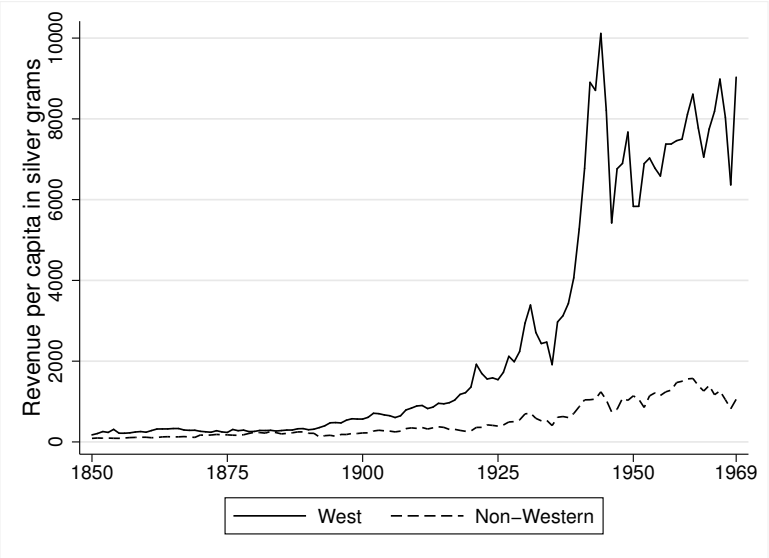
Before 1970, global currencies were typically fixed in relation to gold or silver (or both). Which precious metal is a more appropriate yardstick for value across time? This consideration is important because gold and silver prices do not move in unison. Silver prices fell by about 50% relative to gold 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, Scheve and Stasavage'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, Scheve and Stasavage 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:

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

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

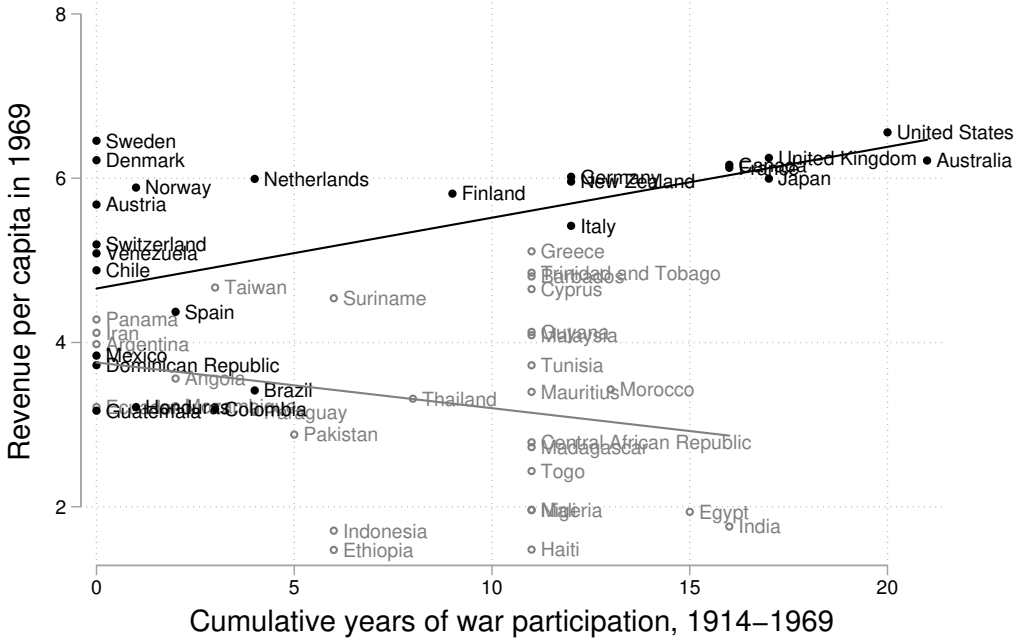
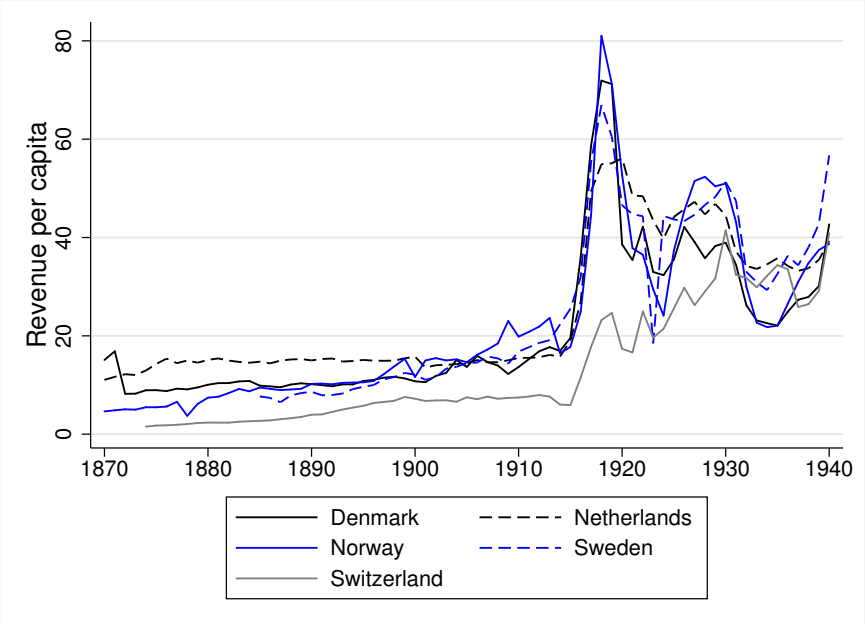


Figure B.2: Revenue Trends in WWI Neutrals



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 $\hat{\tau} \equiv \arg\max_{\tau \in [0,1]} \int_0^{1-\tau} \tau \cdot dH(e_i)$. Taking the first-order condition yields the implicit characterization $-\hat{\tau} \cdot h(1-\hat{\tau}) + \int_0^{1-\hat{\tau}} dH(e_i) = 0$. 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:

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

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 $\bar{l} \in (0,1)$ exists satisfying $R^{\text{leg}}(\bar{l}) = R^{\text{crony}}$:

- For the lower bound, we need $R^{\text{leg}}(0) < R^{\text{crony}}$. This holds for $N < \overline{N} \equiv Y \cdot \frac{\int_0^1 (1-e_i) \cdot dH(e_i)}{\int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)}$. 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.

$$\frac{d}{dl_2}\left[R^{\text{leg}}(l_2) - R^{\text{crony}}\right] = N \cdot \left[\int_0^1 (1-e_i) \cdot dH(e_i) - \int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)\right]$$

$$= N \cdot \left[\int_0^{1-\hat{\tau}} (1-\hat{\tau}-e_i) \cdot dH(e_i) + \int_{1-\hat{\tau}}^1 (1-e_i) \cdot dH(e_i)\right].$$

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 $N \in \left(Y,\ Y \cdot \frac{\int_0^1 (1-e_i)\cdot dH(e_i)}{\int_0^{1-\hat{\tau}} \hat{\tau}\cdot dH(e_i)}\right)$. To prove this set is non-empty for all parameter values, it suffices to demonstrate $\int_0^{1-\hat{\tau}} \hat{\tau}\cdot dH(e_i) < \int_0^1 (1-e_i)\cdot dH(e_i)$. This easily rearranges to a true statement: $\int_0^{1-\hat{\tau}} (1-\hat{\tau}-e_i)\cdot dH(e_i) + \int_{1-\hat{\tau}}^1 (1-e_i)\cdot dH(e_i) > 0$. 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 $\min\{\Delta \cdot \hat{l}, 1\}$ 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 $F < p_{\text{high}} \cdot (N - Y) \cdot \int_0^1 (1 - e_i) \cdot dH(e_i)$. Thus, combining this with the upper bound for F stated in footnote 70, we require:

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

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:

$$R^{\text{leg}}\big(\Delta \cdot \underline{l}\big) = R^{\text{leg}}\big(\bar{l}\big) = R^{\text{crony}} < R^{\text{crony}} + \frac{F}{p_{\text{high}}} = R^{\text{leg}}\big(\Delta \cdot \hat{l}\big).$$

  The first equality follows from $\underline{l} \equiv \frac{\bar{l}}{\Delta}$. 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:

$$R^{\text{leg}}\big(\Delta \cdot \underline{l}\big) < R^{\text{leg}}\big(\Delta \cdot \hat{l}\big).$$

  Because R^{leg} is a strictly increasing function, we have $\Delta \cdot \underline{l} < \Delta \cdot \hat{l}$, 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): $\hat{l} < \bar{l}$ for low-enough F.

- For uniqueness, see the first part of the proof for Proposition 3. ■

Proof of Proposition 3. Define $G(l_1) \equiv R^{\text{leg}}\big(\Delta \cdot l_1\big) - \left(R^{\text{crony}} + \frac{F}{p_{\text{high}}}\right)$. We first establish

$$\frac{\partial G(l_1)}{\partial l_1} = N \cdot \Delta \cdot \left[\int_0^{1-\hat{\tau}} (1 - \hat{\tau} - e_i) \cdot dH(e_i) + \int_{1-\hat{\tau}}^1 (1 - e_i) \cdot dH(e_i) \right] > 0.$$

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.

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

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

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

■

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*. Penn State Press.

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

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*. 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*. 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.