The Great Revenue Divergence

Alexander Lee* Jack Paine†

October 7, 2021

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 having a long history of fiscal innovation. Only over the next half century did Western revenue levels permanently diverge. We argue that fiscally strong states require both a pressing demand for revenue and an existing supply of effective bureaucratic institutions. Neither factor in isolation will lead to sustained high levels of revenue extraction. We formalize this insight in a formal model in which governments can choose among low-effort, legibility-intensive, and crony-favoring strategies for raising revenues. Empirically, our theory accounts for low taxation and reliance on indirect taxes in periods of low demand (nineteenth-century West) or low bureaucratic supply (twentieth-century former colonies), and for subsequent 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 such discrepancies 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 (Levi, 1989; Besley and Persson, 2014; Dincecco, 2017).

We demonstrate that major discrepancies in state revenue collection emerged relatively recently. On the eve of World War I, South American countries and some export-oriented colonies collected per capita revenue levels similar to those in the West despite being poorer; and often exceeded Western revenue collection when accounting for income differences. Many Asian and African countries lagged Europe, but even 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. These changes created a revenue gap between Western and non-Western countries that persisted, and in fact widened, even after former European colonies gained independence. Thus, the *great revenue divergence* is a twentieth-century phenomenon.

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 (1998) with historical exchange rates, gold prices, and population. 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 century only (or mainly) for European countries (Beramendi, Dincecco and Rogers,

[1] Calculated by authors using data from the International Monetary Fund (2017).

2019; Dincecco, 2011, 2015; Karaman and Pamuk, 2013; Scheve and Stasavage, 2016), or only the late twentieth century (International Monetary Fund, 2017; Queralt, 2019). By combining depth and breadth, our dataset is uniquely suitable for analyzing comparative historical trends in revenue levels.[2] We supplement our measure of revenues per capita with data on taxes/GDP from Andersson and Brambor (2019), and demonstrate a qualitatively similar pattern of revenue divergence on a truncated sample.

Existing theories of fiscal extraction, taken in isolation, cannot explain the great revenue divergence. Some scholars propose *fiscal demand* explanations that emphasize how some states have greater revenue extraction needs than others, focusing mainly on international wars. By contrast, *fiscal supply* explanations focus on the bureaucratic institutions used to gather information about the population. By making production legible to the state, bureaucracies help states to target their taxes and extract revenues.

Yet these families of 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. European countries were clearly distinguished from others in many important ways by the turn of the twentieth century that relate to fiscal supply (Tilly, 1992). They had invented sophisticated debt instruments to finance expensive wars (Brewer, 1990; Stasavage, 2011), enacted modern fiscal devices such as income taxes (Mares and Queralt, 2015), collected voluminous information about their populations (Brambor et al., 2020), and had experienced modern industrial development (Pomeranz, 2009). By contrast, much of the rest of the world was under Western colonial rule, and England and France each collected considerably more revenue per capita than several major non-Western empires in the late eighteenth century (Karaman and Pamuk 2010, 623; Rosenthal and Wong 2011, 175; Hoffman

[2]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.

2015, 51; Dincecco 2017, 69). Why were differences in revenue collection relatively small on the eve of World War I?

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 (Herbst, 2000; Centeno, 2002; Besley and Persson, 2011). Yet many newly independent states exhibited high demand for public expenditures. Many anti-colonial activists believed that jurisdictional sovereignty would engender higher levels of public spending by aligning the government's incentives with their citizens rather than with European bondholders and civil servants (Naoroji, 1901; Furnivall, 2014), and 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, which created additional demands for public expenditures.

To unravel the puzzle of the great revenue divergence, we develop a theoretical framework that highlights the interaction of fiscal demand and supply factors. In a game-theoretic model, the government chooses if, and how, they collect taxes from citizens. The government can exert low fiscal effort (e.g., rely on customs revenue), or choose either of two high-effort strategies. On the one hand, they can offer broad-based economic rights and impose an income tax on the entire citizenry. On the other 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. Bureaucratic capacity determines which high-effort strategy is more effective. We refer to the first option as *legibility-intensive* extraction because, to be effective, it requires high knowledge about the exit option for a broad strata of citizens—that is, high societal legibility. Yet the second, *crony-favoring* strategy, can be lucrative precisely because the government can concentrate economic gains among highly legible cronies.

We demonstrate that states choose legibility-intensive extraction only if *both* fiscal demand and supply are high. Absent high demand, states are unwilling the face the social costs associated with heavy extraction, even if latent fiscal capacity is high. Yet contrary to implications from existing models such as Besley and Persson (2011), high demand does not necessarily engender "common value" states that do not predate their economy. States facing a demand shock cannot develop an effective bureaucracy overnight because levels of social legibility are highly persistent, at least in the short term (Scott, 1998; Lee and Zhang, 2017; Blaydes, 2018; Stasavage, 2020). Low legibility makes states ineffective at collecting direct sources of taxation, and thus high fiscal demand causes them to raise revenue by favoring cronies. Only states with high bureaucratic supply will respond to high demand with legibility-intensive extraction, which yields superior revenue intake than crony-favoring policies in legible societies. Over time, this mechanism can generate a divergence in revenue intake. If fiscal demand is initially low, states with competent bureaucracies will distinguish themselves in terms of revenue intake only later, when fiscal demand becomes high. Their later advantages are accentuated because we allow bureaucratic capacity to grow over time via a learning-by-doing effect for states that choose legibility-intensive extraction. By contrast, even when fiscal demand is expected to be high in the future, low initial bureaucratic supply creates path dependence and encourages crony-favoring policies. Thus, our main theoretical implication is that high levels of revenue intake require a conjunction of high demand and high supply.

This implication fits the facts well. By the nineteenth century, Western countries had amassed important advantages in latent fiscal capacity relative to other countries, and continued to build bureaucratic capacity throughout the century. But the relative lack of intra-European wars between 1815 and 1914—a period that also predated modern welfare states—limited the demand for revenues, and thus tax collection. 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. They restructured their economies to fight total war, which unleashed permanently higher demand for social spending because of ensuing franchise expansion and the creation of welfare states. Their existing stock of fiscal institutions enabled them 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 legibility. Easy-revenue strategies, such as promoting primary product exports or collecting direct taxes locally to balance the budget, may have been adequate from the perspective of colonial elites, but were wholly inadequate for meeting heightened demand after countries gained independence. Many post-colonial states in which fiscal demand was particularly high turned toward crony-favoring strategies rather than legibility-intensive collection because prior underinvestment in bureaucracies hindered their ability to collect direct taxes effectively. The main exceptions were East Asian states, in particular Japan. Intense geopolitical pressure combined with a history of bureaucratic government enabled large increases in direct taxation.

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 compare 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

To construct our main measure, central government revenue per capita in gold grams, we performed the following steps.

1. We use data on *central government revenues* from Mitchell (1998). 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 (1998) to calculate *revenue per capita*. 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 which no census occurred within two decades). For this reason, revenue per capita cannot be estimated before the date of the first census, even when revenue data is available from an earlier date.[3]

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.[4] We were unable to perform this step for country-years with

[3]One exception is that for Africa we incorporated Frankema and Jerven's (2014) data for 1850–1960, which prevents much of this data from being dropped from the sample. We also incorporated additional population data for Russia that we discuss in Appendix A.7.

[4]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. However, COW does not include data before 1870 or from colonies (although most colonies used the mother country's currency). Additionally, 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%. We further supplemented the COW data using Denzel (2010) and Officer (2016).

non-convertible currencies, and thus we drop such observations even if revenue and population data are available.[5]

4. Finally, we converted revenue per capita in British pounds into *gold grams* using gold ounces from Officer (2016). This step is unnecessary for cross-national comparisons, but reduces problems in data visualization stemming from the volatility of the pound.

Our revenue variable advances existing quantitative data on state revenue in both geographical and 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 fiscal 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 end of the Bretton Woods system. For this reason, we only analyze data through 1969. Appendix Figure A.1 plots revenues over time for each territory in the dataset.

Although we are not the first to use the Mitchell revenue data for historical analysis, our approach to weighting the data points enables us to incorporate considerably more information than used in existing studies. For example, Besley and Persson (2014) incorporate information only from eighteen rich countries and compute an unweighted average over time. Consequently, they do not calculate revenue collection for poorer countries in the nineteenth or early twentieth centuries, nor do they make time-series cross-section comparisons across a broad country and time sample. Mann (1993, 358-401) and Karaman and Pamuk (2010, 2013) use Mitchell data from the nineteenth century, but only for European countries. Cagé and Gadenne (2018) and Beramendi, Dincecco and Rogers (2019) use Mitchell and other sources to construct a sample that is expansive globally after 1945,

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

but confined to Western Europe, Japan, and the Southern Cone in the nineteenth century.

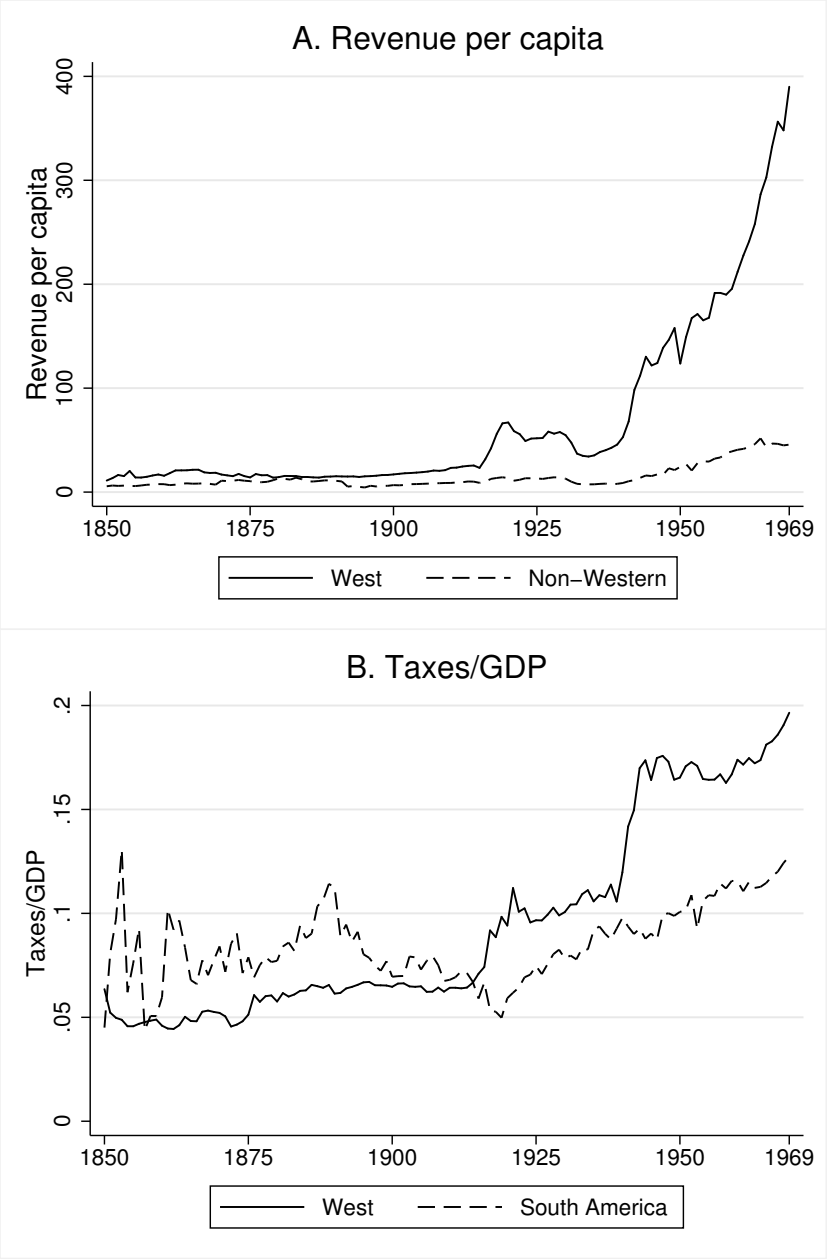
Despite clear advantages of expansive country and time coverage, our approach to measuring state revenue intake also has drawbacks. Although we follow existing work on historical revenue collection by expressing revenue in gold (Dincecco, 2011, 2017) or silver (Karaman and Pamuk, 2010, 2013), research 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 changes in societal income mostly explain differences in revenues. We address this concern in two ways. First, we analyze patterns for taxes as a fraction of GDP using data from Andersson and Brambor (2019), 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 summarizes 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 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 quite low by contemporary standards. It is also mostly driven by extremely low revenues in many African countries under colonial rule. 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.

Figure 1: The Great Revenue Divergence



Notes. The lines show 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).

These patterns changed drastically after 1914. Many countries participated in the two world wars, a worldwide depression, geopolitical competition during the Cold War, and spent increased sums on redistributive policies following franchise expansion. Western countries experienced a massive

expansion in state revenue collection in the following decades, consistent with existing characterizations (Lindert, 2004). Between 1913 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 457%. Thus, gains outside the West were 70% smaller than those among Western countries. The patterns are largely similar among different subsets of non-Western countries: there was a 359% increase in South America and the Caribbean, 428% in Africa, and 564% in Asia (Middle East, South Asia, Southeast Asia, East Asia excluding Japan).[6] 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) 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.

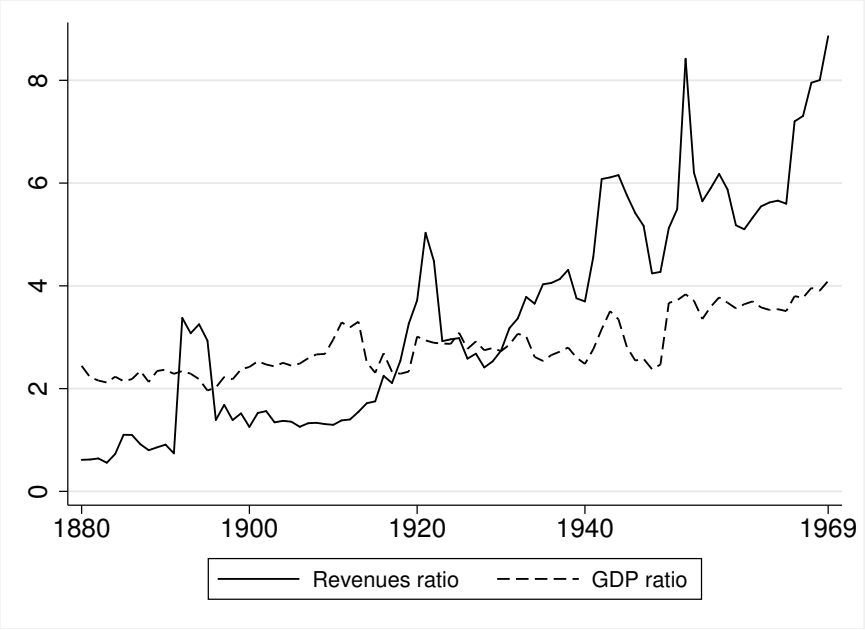
The divergence is also stark when assessing taxes as a fraction of GDP. As Panel B shows, South America outpaced the West on this measure throughout the nineteenth century. In 1913, Britain lagged Brazil, Uruguay, Chile, and Argentina. However, between 1913 and 1969, taxes rose from 6.4% to 19.7% of 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 57% smaller than those in the West. We primarily analyze the taxes/GDP data from Andersson and Brambor (2019) because, compared to alternatives, it (a) has lesser missingness relative to our core dataset (48% of total country-years) and (b) uses natural units. However, in Appendix A.1, we demonstrate qualitatively similar trends when analyzing alternative datasets that account for differences in GDP. We analyze taxes/GDP from Beramendi, Dincecco and Rogers (2019), which contains only 31% of the total country-years in our core dataset, and 62% fewer country-years for non-Western countries compared to Andersson and Brambor (2019). We also constructed a

[6]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 a constant basket of countries.

panel of normalized revenue data. Despite better data coverage (58% 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 (2018) update of Angus Maddison.

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 usually refer to the divergence in per capita economic output between Western and non-Western countries (Pomeranz, 2009). Although scholars debate the timing and causes of this divergence, they agree it occurred no later than the mid-nineteenth century, after the Industrial Revolution had spread 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.

Figure 2: Comparing Income and Revenue Divergence



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

1.3 Additional Robustness Checks

In the appendix, we analyze the robustness of our core pattern. One concern with our main measure of revenues per capita 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. Appendix A.2 explains two ways in which our main measure guards against these concerns, and presents intra-imperial comparisons (hence territories using the same currency or a highly stable peg). We also cannot directly account for differences in purchasing power, nor do we directly measure *tax* intake, although 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 revenues 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 disaggregate 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 (Tilly, 1992; Brewer, 1990). Other authors make the converse argument that less intense geopolitical competition in many ex-colonies in Sub-Saharan Africa and Latin America has undermined state-building efforts (Herbst, 2000; Centeno, 2002; Thies, 2004).

Preparation for and participation in an external war raises the state's need for revenue to pay and deploy soldiers for the conflict. Mass mobilization wars in particular can create political consensus for egalitarian taxation systems (Scheve and Stasavage, 2016), and these conditions may persist in a post-war ratchet effect (Peacock and Wiseman, 1961). States need to service debt accumulated during the conflict, and may also face pressure to sustain programs of social redistribution—which require high taxes to fund—that emerged during the war (Lindert, 2004). Such concerns are particularly pressing when participation in warfare engenders franchise expansion to groups that prefer an active state and redistribution. Besley and Persson (2011) formalize the core mechanism for the war-demand logic. The key choice in their model is whether the incumbent government invests 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. They propose external threats as the most natural interpretation of a high value of their public goods parameter (46-7, 58).[7]

[7]War is not the only possible source of high demand for revenues. Broad democratic franchises create incentives for greater taxation to finance redistribution and other welfare policies. Demand for revenue also varies within limited-participation polities because elites in different social groups diverge in their desire for public spending (Karaman and Pamuk, 2013; Mares and Queralt, 2015; Saylor and Wheeler, 2017; Beramendi, Dincecco and Rogers, 2019; Schenoni, 2021). However, even these alternative sources of high demand relate in part to conflict: the world wars contributed

Yet numerous counterexamples raise questions about the universality of the war-demand logic as well as the importance of related demand factors. 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 (Barnett, 1992) and nineteenth-century South America (Centeno, 2002) also failed to engender strong fiscal apparatuses. Earlier in European history, warfare often generated crippling debt and encouraged irresponsible actions such as debasing the currency, as opposed to promoting fiscal systems that could generate consistent tax revenues over the longer term.

2.2 Fiscal Supply

Other scholars focus on the "supply" of bureaucratic 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" (Scott, 1998; Lee and Zhang, 2017; Blaydes, 2018; Stasavage, 2020). 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 considerably more information-intensive. Although the concept of bureaucratic supply is inherently multi-faceted, recent research measures key components of states' information-collection abilities across broad comparative samples. Brambor et al. (2020) collected data on civil regis-

to episodes of mass franchise expansion in many countries, and involvement in external wars is typically a major point of contention amid intra-elite disagreements.

tration systems and state statistical offices dating back to the eighteenth century. 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 (2017) and D'Arcy and Nistotskaya (2017) compiled data on the effectiveness of censuses in the twentieth century, which correlates strongly with public goods provision.

Importantly, *fiscal capacity* differs from *revenue intake*. States can collect information about production and life events (birth, death, marriage) without using them for taxation. They can also govern a literate population capable of filling out written tax forms, but might not require them to do so. In such scenarios, fiscal capacity is *latent* and ready to be employed when the ruling group wishes.

One, but not the only, factor that influences societal legibility is participation in warfare. European history provides numerous examples of states enacting political reforms to gain an advantage at fighting. For example, Britain introduced the Bank of England in 1694 during the Nine Years' War with France, which created a major financing advantage (Brewer, 1990). Later, Britain responded to dire fiscal pressure during the Napoleonic Wars to impose the world's first modern income tax. Mann (1993, 444-78) examines five great powers (Britain, France, Prussia, Austria, United States) and argues that warfare was the main stimulus to bureaucratic reforms before the French Revolution that constituted the early stages of modern statehood. These reforms introduced at least minimal standards for hiring and promotion, and shifted toward salaried rather than office-owning state officials.

If fiscal supply was solely a function of short-term fluctuations in fiscal demand, this would be problematic for our conceptual and empirical distinction between demand and supply factors. However, we present three arguments for why supply varies independently of demand.

First, bureaucratic capacity is highly persistent. Consequently, states facing a demand shock are rarely able to rapidly and dramatically improve societal legibility. Dincecco's (2017) 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 (see also Stasavage 2020). Conversely, once created, bureaucracies tend to be self-perpetuating even when fiscal demand is temporarily low. Some legibility reforms can persist without any spending at all. For instance, Scott (1998) discusses the introduction of last names and addresses as being crucial for states to find citizens and distinguish them from each other. Once a state has forced its citizens to adopts last names and addresses, this information provides the basis for future revenue extraction, even if not used immediately.

Second, numerous factors influence legibility that are independent of war, or of state policy. In agricultural societies, the structure of agricultural production influences the feasibility of creating an effective bureaucracy. In ancient China, intensive agriculture made production legible to the central government (Stasavage, 2020, 80-83), and in modern Africa, production by co-ethnics is more legible to the government than is production by other groups (Kasara, 2007). The type of terrain also influences legibility. For example, peasants can easily escape the reach of the state when surrounded by mountains, but not when surrounded by deserts (Scott, 2009). Geographic size, high linguistic diversity, external rule, and low literacy are other characteristics that undermine legibility. Thus, countries such as Sweden with flat terrain, low linguistic diversity, a long-standing indigenous state, and high levels of literacy are inherently better situated to collect considerable information about their citizens than are countries such as the Democratic Republic of the Congo, which enjoys none of these advantages.

Third, even in Europe, wars cannot explain the key nineteenth-century bureaucratic reforms that preceded unprecedented tax collection in the twentieth century. We discuss this important point in depth in the empirical analysis.

Despite adding another important piece, analyses of fiscal supply also offer incomplete explanations of the great revenue divergence. In particular, these accounts cannot explain why a large and permanent divergence did not occur earlier. For most of the nineteenth century, Western Europe and the United States outpaced the rest of the world in terms of collecting information about their citizens and educating their population. As noted, Britain even imposed the world's first successful

income tax during the Napoleonic Wars. However, it suspended its income tax after the wars and did match its 1810 per capita revenue record until 1915. Britain's high fiscal capacity remained largely latent throughout the nineteenth century.

3 A Formal Model of Revenue-Extraction Strategies

To address the shortcomings of theories of fiscal demand and fiscal supply when considered in isolation, we develop a theoretical framework that formally models their interaction. We establish two main findings. First, we explain how demand and supply affect the optimal revenue-raising strategy in a single period of play. If demand is low, then societal opposition to high levels of taxation yields low revenue intake regardless of the state's latent fiscal capacity. The onset of war or franchise expansion creates demand for greater revenue extraction, but states can respond in different ways. Demand shocks may propel the incumbent government to pursue *legibility-intensive* extraction. This entails imposing broad-based taxes, like income taxes, that require information about citizens. However, only states with high bureaucratic capacity (i.e., high supply) can generate high yields from income taxes. The government can instead choose *crony-favoring* extraction, which promotes monopolies and other forms of easy rent extraction that require less information about citizens. This is optimal for states with low legibility, even if such strategies yield less revenue compared to legibility-intensive extraction in states with higher bureaucratic supply.

Second, we analyze multiple periods to explain how states can diverge over time in revenue collection. The preceding result anticipates a straightforward mechanism. If demand is initially low, then states with greater legibility will not be distinct in revenue collection. Later, differences will emerge over time when demand becomes high.

We focus the dynamic analysis on a more strategically interesting—and substantively relevant—setting in which bureaucratic capacity can increase over time via learning by doing. Throughout most of history, the degree of societal legibility needed to effectively collect an income tax was

unattainable. This was true for most Western states in the early nineteenth century, when we begin our empirical analysis.[8] In the model, states can build bureaucratic capacity over time even when demand is low, and the initial level of legibility determines whether they do so in equilibrium. States endowed with intermediate bureaucratic capacity in period 1 will forgo their optimal short-term strategy by choosing legibility-intensive extraction, which enables them in period 2 to gain from the learning-by-doing effect for bureaucratic capacity. For this reason, over time, states that started with very low legibility (which makes it optimal to not to invest in future bureaucratic capacity) will fall behind intermediate-legibility states whose bureaucratic capacity grows over time. This implication contrasts with, for example, Besley and Persson's (2011) proposition that high fiscal demand always engenders common-interest states that invest in bureaucratic capacity.[9] For states with very low initial bureaucratic capacity, the gains from investing to boost future capacity are not worth the effort even if demand is known to be high in the future because crony-favoring strategies yield more revenue—but also perpetuate low bureaucratic capacity.

Overall, we highlight the fundamental interaction between demand and supply factors. Our mechanism explains why high levels of both are needed to generate high revenue intake. Later we apply this mechanism to help explain the twentieth-century great revenue divergence.

3.1 Setup

We model an interaction between a government and citizens, which unfolds over two periods. Time is denoted by $t \in \{1, 2\}$. The continuum of atomless citizens in society has a mass of $N \in \left(1, \overline{N}\right)$, with an upper bound $\overline{N} > 1$ defined in the analysis. Citizens are indexed by i. In both periods, each citizen produces an output worth Y_i (see below) and has an exit option that enables them to

[8] Britain was a partial exception, although as we discuss below, even here, Britain lacked the level of social consensus needed to retain the income tax following the Napoleonic Wars.

[9] The two alternative revenue-extraction strategies in our model instead more closely resemble those in Queralt (2015).

consume a fraction $e_i \in (0,1)$ of their output. 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 $[0,1]$.[10]

In each period, the government is endowed with customs revenues worth $R^{\text{cus}} > 0$, which require low effort to extract. The government can rely solely on customs revenues to fund expenditures. Alternatively, the government can exert high effort to collect additional taxes, which requires sinking a cost to boost revenues. High-effort extraction comes in two forms.

First, under a *legibility-intensive* strategy, the government grants basic economic rights (e.g., legal rights to participate in the formal economy) to all N citizens. This choice yields output of $Y_i = 1$ for each citizen, which is subject to taxation. However, due to limitations in bureaucratic capacity, the government does not know the value of the exit option for each citizen. A fraction $l_t < 1$ of citizens are *legible*, and the government perfectly knows the value of e_i for legible citizens. The remaining fraction $1 - l_t$ of citizens are illegible and the government knows only the prior distribution of possible values of e_i for such citizens. Choosing legibility-intensive extraction entails a fixed cost $F \in \left(0, \overline{F}\right)$, with an upper bound $\overline{F}$ defined in the analysis.[11]

Second, under a *crony-favoring* strategy, the government favors a subset of legible citizens, normalized to mass 1. This could involve limiting economic production to a specific set of 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 income produced by the mass $N-1$ of remaining citizens outside the reach of the state. To make the tradeoffs non-trivial, we assume that the crony-favoring strategy bolsters the income of favored citizens relative to legibility-intensive extraction $\left(Y > 1\right)$. However, crony-favoring extraction diminishes the mass of citizens subject to taxation $\left(1 < l_t \cdot N\right)$ as well as total output that is

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

[11] The government pays this cost only in periods it chooses this strategy. For example, choosing legibility-intensive extraction in both periods would entail paying F in both periods, whereas it would only pay F once if it chose legibility-intensive extraction in a single period.

potentially taxable $\big(Y < N\big)$. This revenue-extraction strategy 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 creates costs for a government to actively intervene to distort market competition. To isolate the role of legibility in distinguishing legibility-intensive from crony-favoring extraction, we assume the fixed cost F is the same (see also the stipulations in footnote 11).

The fraction of legible citizens, l_t, is a function of bureaucratic capacity. We assume that 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 greater learning-by-doing in developing fiscal capacity. By contrast, if the government chooses low effort or crony-favoring extraction in period 1, then $l_2 = l_1$.[12]

After choosing if (and how) to exert high effort, the government proposes an individual-specific tax rate $\tau_i \in [0, 1]$ to each citizen. Each citizen simultaneously responds by complying and consuming $(1 - \tau_i) \cdot Y_i$, or exiting and consuming $e_i \cdot Y_i$.

Total revenues in each period, R_t, equal the customs endowment R^{cus} plus any additional taxes collected from pursuing either high-effort extraction strategy. For the government, the marginal benefit equals 1 for any revenue intake up to an exogenously determined amount R_t^{dem}. For higher amounts, the marginal benefit of the endowed customs revenues is 0, and is negative for taxes garnered from high-effort extraction. We conceptualize societal demand for expenditures as R_t^{dem}. The government is rewarded for increasing revenue to get closer to the amount of expenditures de-

[12] Assuming a deterministic relationship between revenue extraction strategies and bureaucratic development enhances tractability, although the results would be qualitatively similar under alternative assumptions. For example, we could assume that initial bureaucratic capacity atrophies upon pursuing crony-favoring policies (e.g., Democratic Republic of the Congo). 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).

manded, and penalized for setting taxes such that revenue exceeds the desired amount. We assume that Nature draws R_t^{dem} independently each period 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. These terms satisfy $0 < R_{\text{low}}^{\text{dem}} < R_{\text{high}}^{\text{dem}}$, and below we impose additional threshold values that make the difference between the low and high draws strategically meaningful. The government additionally pays the fixed cost F if they pursue either high-effort extraction strategy. Consumption for each citizen is as described above: complying with taxation yields $(1 - \tau_i) \cdot Y_i$, and choosing the exit option yields $e_i \cdot Y_i$.[13]

3.2 Discussion of the Government's Menu of Choices

The model incorporates important elements from existing theories of fiscal demand and supply. Here we briefly elaborate upon the substantive motivation for the menu of choices available to the government. We conceptualize the revenue endowment R^{cus} as *customs* taxes, which resonates with our empirical examples. Collecting customs taxes requires relatively few agents at one or several major ports. Thus, these indirect taxes are easy to collect if the economy is already organized in a manner to facilitate international trade (Hinrichs, 1966). 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.

Crony-favoring revenue sources require significant state involvement in and restructuring of the economy. This justifies the fixed cost F to implement crony-favoring extraction. 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 often creates a symbiotic political relationship whereby the government has easier

[13] Citizens that lack legal rights in a crony economy simply consume their exit option, although these actors are not strategically relevant.

access to information about the firm's production, and the firm gains economic advantages from its political access. This justifies the assumption that the output of favored citizens increases to $Y > 1$. Collectivized agriculture in the Soviet Union provides an extreme example of this tactic. More typical cases are ones like Egypt and India in which the government actively intervenes in the economy to create a "captive tax base" (Waterbury, 1993, 134). As Chaudhry (1993, 252) 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." Related strategies, such as selectively allocating permits for international trade or requiring licenses to engage in certain economic activities also constitute the crony-favoring strategy (Haber, Maurer and Razo, 2003). Such strategies require less restructuring of the economy than the aforementioned tactics. However, they match our conceptualization of crony-favoring extraction by concentrating economic gains while hindering overall output, similar to common tactics earlier in history such as selling offices or granting monopolies to guild members and nobles.

Legibility-intensive extraction requires information about the broad population. Direct taxes such as income taxes and value-added taxes provide the primary source of revenue for most modern states. Efficient collection of such taxes requires detailed information about the identities and productivity of citizens, as well as complex bureaucracies to collect and process this information. Consequently, as we show below in the analysis, if governments with low fiscal supply need to immediately raise additional revenues, they must turn to revenue sources that do not require high levels of legibility to implement.

We also adopt the ideas from the literature that bureaucratic capacity is generally persistent, but subject to change with concerted effort. Hence, initial legibility, l_1, is a fixed parameter that may take the same value in period 2. However, we also allow bureaucratic capacity to evolve between the two periods if the government chooses legibility-intensive extraction in period 1. This is a natural assumption about learning-by-doing, and also resembles the fiscal-capacity investment decision

in Besley and Persson (2011).

3.3 Short-Term Strategies: Analysis of Period 2

We solve backwards to characterize subgame-perfect strategies. In period 2, the government cares solely about short-term revenue intake. High demand induces the government to choose whichever high-effort strategy maximizes net revenues in period 2. This is the legibility-intensive strategy if bureaucratic supply is high enough, and the crony-favoring strategy otherwise. Alternatively, if demand is low, then the state underutilizes its fiscal potential.

If fiscal demand is low, then the government can fund all desired expenditures while exerting low effort at tax collection. Specifically, if the demand for public expenditures, R_2^{dem}, is low relative to the endowed customs revenues, R^{cus}, then the government does not seek additional tax revenues and underutilizes its fiscal capacity.

If instead fiscal demand is higher, then the government exerts effort to boost revenue intake.[14] The following explains whether they choose legibility-intensive or crony-favoring extraction. For each legible citizen, the government sets the individual-specific tax rate to make the citizen indifferent between complying and exiting, $\tau_i^* = 1 - e_i$. This proposal induces every legible citizen to comply. 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 more revenue from each citizen that complies, but decreases the fraction of citizens that comply. Consequently, the optimal tax rate on each illegible citizen solves $\hat{\tau} \equiv \arg\max_{\tau \in [0,1]} \int_0^{1-\tau} \tau \cdot dH(e_i)$. In equilibrium, only illegible citizens with low-valued exit options comply, $e_i < \hat{\tau}$, yielding a fraction $1 - \hat{\tau}$ that comply.

Crony-favoring policies yield a unit mass of favored citizens that each produce $Y_i = Y$ and are

[14] We assume for now that the government seeks to maximize revenues, and later we present the conditions under which this behavior is optimal.

perfectly legible, and the remaining mass of $N-1$ citizens are beyond the reach of the government. Legibility-intensive policies enable all citizens to legally produce $Y_i = 1$, but only a fraction l_2 are legible. Thus, expected revenues under each strategy are:

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

$$R^{\text{leg}}(l_2) = N \cdot \Bigg[\underbrace{l_2 \cdot \int_0^1 (1 - e_i) \cdot dH(e_i)}_{\text{Legible citizens}} + \underbrace{(1 - l_2) \cdot \int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)}_{\text{Illegible citizens}} \Bigg]. \tag{2}$$

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:[15]

$$l_2 \geq \bar{l} \equiv \frac{\frac{Y}{N} \cdot \int_0^1 (1 - e_i) \cdot dH(e_i) - \int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)}{\int_0^1 (1 - e_i) \cdot dH(e_i) - \int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)} \in (0, 1). \tag{3}$$

Thus, maximum revenues in period 2 are:

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

[15]The following explains why $\bar{l} \in (0, 1)$ for appropriately set $\overline{N}$. The denominator is strictly positive because expected revenues from legible producers exceed those from illegible producers. To see this formally, rearrange the denominator to $\int_0^{1-\hat{\tau}} (1 - e_i - \hat{\tau}) \cdot dH(e_i) + \int_{1-\hat{\tau}}^1 (1 - e_i) \cdot dH(e_i)$. The bounds of the first integral assume $e_i < 1 - \hat{\tau}$, thus establishing the claim. Given this, the numerator is also strictly positive for low enough N, and hence we set the upper bound of N as $\overline{N} \equiv Y \cdot \frac{\int_0^1 (1-e_i) \cdot dH(e_i)}{\int_0^{1-\hat{\tau}} \hat{\tau} \cdot dH(e_i)}$. Finally, because $Y < N$, both terms in the numerator are strictly bounded between 0 and 1. Thus, because $Y > 1$ and $N > 1$, the preceding results establish that the difference between these terms is strictly less than 1.

These results establish the interactive effect of demand and supply factors on the government's optimal revenue-raising strategy. If demand $\left(R_2^{\text{dem}}\right)$ is low, then the government forgoes additional revenues. If demand is high, only if legibility $\left(l_2\right)$ is also high does the government choose legibility-intensive extraction. Otherwise, crony-favoring extraction yields more revenue. Thus, existing arguments are correct that high demand stimulates governments to collect more revenues, but do not simultaneously incorporate the supply side of revenues—which may imply that distorting the economy yields greater revenue intake. However, given the reasonable premise that crony-favoring extraction limits the total *potential* taxable output $\left(Y < N\right)$, in equilibrium, states that collect income taxes will bring in more revenues than states than crony-favoring extraction.

Proposition 1 presents a subgame perfect Nash equilibrium strategy profile. The only considerations not covered in the preceding discussion are (a) whether the government maximizes revenue extraction and (b) the precise thresholds at which the government chooses each strategy. The government may want more revenues than R^{cus}, but not all the way up to $R^{\text{cus}} + R_2^{\text{max}}$. 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 $R^{\text{cus}} + R_2^{\text{max}}$, then the government pays the fixed cost for either legibility-intensive or crony-favoring extraction, but intentionally sets taxes to collect less than the maximum possible revenues.[16] To eliminate strategically uninteresting cases, we restrict the upper bound on the fixed costs such that $\overline{F} < R^{\text{crony}}$.

[16] All equilibria are payoff equivalent. There are a continuum of equilibria because the government is indifferent between legibility-intensive and crony-favoring extraction when either is sufficient to fund the less-than-maximal extraction amount.

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

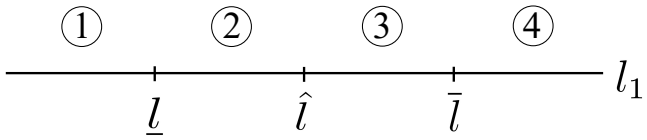
- ***Low demand.***
  - *If $R_2^{dem} \leq R^{cus} + F$, then the government does not pursue additional revenues: $R_2 = R^{cus}$.*
  - *If $R^{cus} + F < R_2^{dem} < R^{cus} + R^{crony}$ and $l_2 < \bar{l}$, then the government chooses crony-favoring extraction but underutilizes its fiscal potential: $R_2 = R_2^{dem}$.*
  - *If $R^{cus} + F < R_2^{dem} < R^{cus} + R^{leg}(l_2)$ and $l_2 \geq \bar{l}$, then the government chooses legibility-intensive extraction but underutilizes its fiscal potential: $R_2 = R_2^{dem}$.*
- ***High demand and low supply.*** *If $R_2^{dem} \geq R^{cus} + R^{crony}$ and $l_2 < \bar{l}$, then the government chooses crony-favoring extraction and uses its full fiscal potential: $R_2 = R^{cus} + R^{crony}$.*
- ***High demand and high supply.*** *If $R_2^{dem} \geq R^{cus} + R^{leg}(l_2)$ and $l_2 \geq \bar{l}$, then the government chooses legibility-intensive extraction and uses its full fiscal potential: $R_2 = R^{cus} + R^{leg}(l_2)$.*

3.4 Long-Term Strategies: Analysis of Period 1

In period 1, the government cares not only about how its fiscal strategy affects contemporaneous revenues, but also intake in period 2. If initial bureaucratic supply is intermediate, then the government "invests" in future fiscal capacity, that is, incurs the fixed cost of legibility-intensive extraction in period 1 even if that is not the optimal revenue-raising strategy in the short term. Boosting fiscal capacity for period 2 enables them to collect higher tax revenues in response to a demand shock. Counterfactually, had they not invested in fiscal capacity in period 2, they would instead choose crony-favoring extraction in response to a demand shock in period 2. Overall, the core implication is unaltered: only governments with high fiscal demand and supply choose legibility-intensive extraction. However, the threshold for "high" supply is lowered because the shadow of the future heightens incentives to pursue legibility-intensive extraction. We also derive a dynamic implication: initial differences in bureaucratic capacity that do not manifest as higher revenues in the short term can engender long-term divergence in revenue intake.

Formally, we demonstrate that there are four ranges of initial legibility, l_1, that determine revenue-extraction strategies across the two periods. To focus the analysis on strategically interesting considerations, we ask: If fiscal demand is high in period 2, does the government choose legibility-intensive extraction in equilibrium?[17] Figure 3 summarizes the different parameter ranges. If legibility is either particularly high or low, then the analysis is identical to above. In region 4, initial legibility is very high, $l_1 > \bar{l}$ (see Equation 3 for this threshold). Regardless of the government's actions in period 1, legibility is high enough in period 2 that legibility-intensive extraction yields more revenues than crony-favoring extraction. Conversely, in region 1, initial legibility is very low, $l_1 < \underline{l} \equiv \frac{\bar{l}}{\Delta}$. Even if the government gets the learning-by-doing boost to bureaucratic capacity in period 2, crony-favoring extraction would yield more revenues than legibility-intensive extraction.

Figure 3: Initial Legibility and Revenue Extraction Over Time



The intermediate regions demonstrate how the shadow of the future changes the government's calculus. In both cases, legibility is below the threshold of $\bar{l}$ at which legibility-intensive extraction is optimal if there is no shadow of the future. However, because were are examining the period 1 calculus, the government also takes into account how investments in fiscal capacity in period affect the outcome in period 2. Suppose that demand is low in period 1. If the government chooses legibility-intensive extraction in period 1, then its total expected consumption across the two peri-

[17] To simplify the number of cases to examine without qualitatively altering the insights, we place bounds on the low and high draws for fiscal demand. For a low draw of fiscal demand, the government does not seek additional revenues in period 2 beyond the endowment. For a high draw of fiscal demand, the government seeks maximum extraction. These are formalized as $R^{\text{dem}}_{\text{low}} < R^{\text{exo}} + F$ and $R^{\text{dem}}_{\text{high}} > R^{\text{exo}} + R^{\text{leg}}(1)$, respectively. Proposition 1 provides the rationale for these thresholds. For the upper bound, note that $\max\{R^{\text{crony}}, R^{\text{leg}}(l_t)\}\big|_{l_t \in [0,1]} = R^{\text{leg}}(1)$.

ods is:

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

If instead the state invests in bureaucracy in period 1, then its expected utility is:

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

Comparing these two thresholds shows that the government invests in bureaucratic capacity in period 1 if and only if, in period 2, legibility-intensive extraction (after getting the learning-by-doing boost) yields sufficiently more revenues than crony-favoring extraction. This threshold takes into account the fixed costs paid in period 1 and the probability that demand is indeed high in period 2. Equating the two preceding expressions enables us to define an implicit threshold $\hat{l} > \underline{l}$ that determines whether the government makes this investment: $R^{\text{leg}}\big(\Delta \cdot \hat{l}\big) = R^{\text{crony}} + \frac{F}{p_{\text{high}}}$. Thus, the degree of legibility still determines the government's optimal revenue-raising strategy. The difference is that the government chooses legibility-intensive extraction for a greater range of parameter values, any $l_1 \geq \hat{l}$, because of the long-run gains from bureaucratic investments. It is also straightforward to see from this equation that higher p_{high} increases the state's willingness to invest in bureaucratic capacity in period 1. Proposition 2 formalizes this intuition.[18]

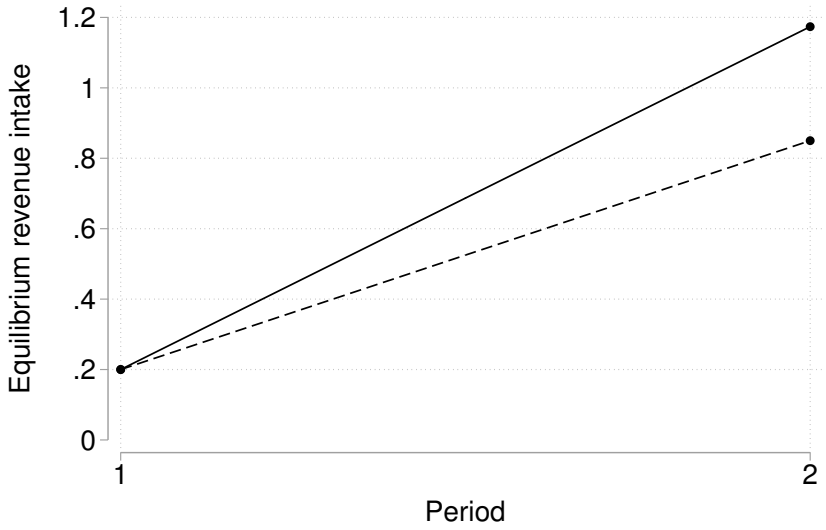
Proposition 2 (Equilibrium legibility in period 2)**.**

- *If $l_1 < \underline{l}$, then $l_2 < \hat{l}$ regardless of the government's action in period 1.*
- *If $l_1 \in \big(\underline{l}, \hat{l}\big)$, then the government does not invest in fiscal capacity in period 1, and $l_2 < \hat{l}$. Counterfactually, had the government chosen legibility-intensive extraction in period 1, then $l_2 > \hat{l}$.*
- *If $l_1 \in \big(\hat{l}, \bar{l}\big)$, then the government invests in fiscal capacity in period 1, and $l_2 > \hat{l}$. Counterfactually, had the government not chosen legibility-intensive extraction in period 1, then $l_2 < \hat{l}$.*
- *If $l_1 > \bar{l}$, then $l_2 > \hat{l}$ regardless of the government's action in period 1.*

[18] We omit a formal proposition for the full strategy profile in Period 1, in which the number of cases to check obscures the core theoretical insight that we highlight here.

This intuition explains revenue divergence over time. 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 $l_1 \in \left(\hat{l}, \bar{l}\right)$. We assume that Nature draws low fiscal demand for both in period 1. Consequently, neither government seeks to raise revenue beyond their endowed customs taxes. However, the state with intermediate initial bureaucratic capacity (solid line) nonetheless sinks the costs associated with legibility-intensive extraction. This choice boosts its bureaucratic capacity in period 2, in the anticipation that future demand will be high (for example, when Britain reformed its tax system, in particular the income tax, during the nineteenth century despite keeping taxes low overall amid a period of low warfare). The low-legibility state (dashed line) does not invest in future fiscal capacity despite facing the same probability of high demand in period 2. Because its initial legibility is so low, the learning-by-doing effects from investing in bureaucracy are not sufficient to turn the state away from crony-favoring extraction.

Figure 4: 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, we have $\underline{l} = 0.122$, $\hat{l} = 0.167$, 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.

To yield divergence, we also assume that Nature draw high fiscal demand for both in period 2. The state with higher bureaucratic capacity chooses legibility-intensive extraction whereas the low legibility state chooses the crony-favoring strategy. Counterfactually, the state choosing legibility-

intensive extraction would not have done so had it not invested in bureaucratic capacity in period 1. This contrasts with the main implication from Besley and Persson (2011) that states anticipating high demand in the future necessarily invest in their bureaucracy to capitalize on demand shocks. In that alternative model, there would be no revenue divergence in period 2 because the states have identical expectation for high fiscal demand in period 2.

4 Empirical Evidence for Theoretical Implications

Our main theoretical implication is that high revenue intake requires the conjunction of high fiscal supply and high fiscal demand. This fits the evidence well. Although Western countries enjoyed an advantage in bureaucratic supply across the entire period, demand was low until World War I. This enabled non-Western primary product exporters and some agrarian empires to keep pace or catch up with the West. Later, the conjunction of high fiscal demand and high fiscal supply explains the revenue spike in the West, and hence the emergence of huge fiscal discrepancies relative to non-Western countries. After independence, their colonies 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 supply), hence exhibiting a positive interaction effect.

4.1 Low Fiscal Demand in the West Before World War I

Between the conclusion of the Napoleonic Wars and the start of World War I, revenue intake was low in Western countries. Figure 5 shows a zoomed-in version of Figure 1. Britain imposed the

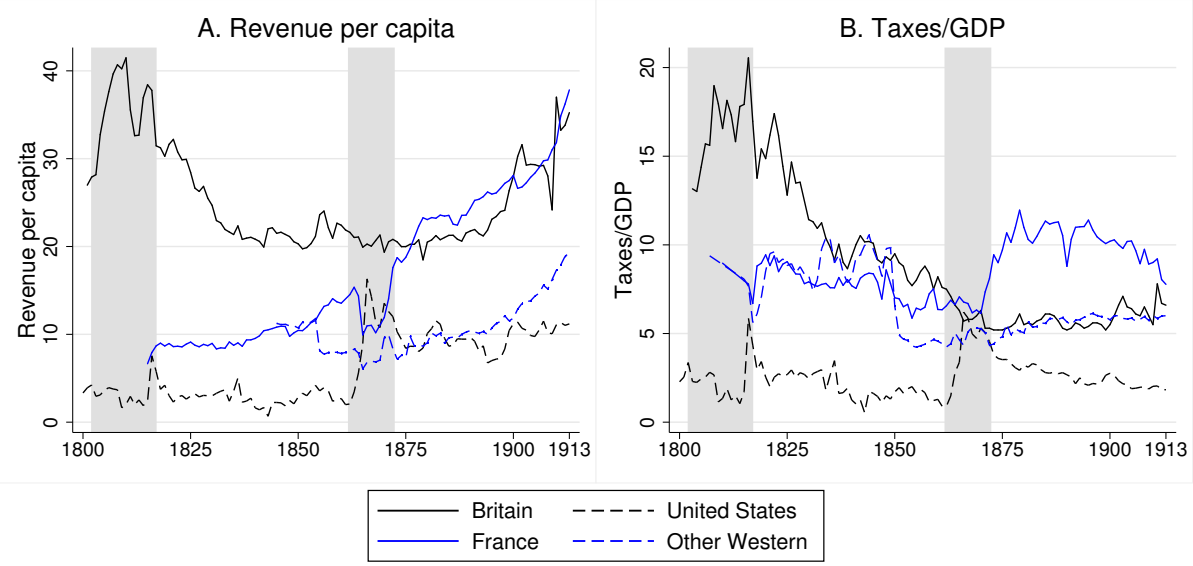
world's first modern income tax during the Napoleonic Wars, a period we highlight in gray.[19] However, per capita revenue intake declined afterwards, and this decline is even more pronounced when accounting for Britain's strong economic growth. Even when Britain reimposed an income tax in 1842, the 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 government was a "state of courts and parties," and as of the 1870s, customs revenues from the Port of New York accounted for more than half of all federal revenues (Skowronek, 1982, 24, 61). 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 figure to not obscure the main pattern), which throughout the century 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 considerably 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 years between 1815 to 1914. 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. Both Britain's upper class political leadership saw little personal advantage to high levels of taxation, and faced no demands from their middle-class electorate to boost expenditures. Across the region, demand for welfare provisions was low throughout the nineteenth century (Lindert, 2004).

By contrast, legibility was high and growing throughout this period. This observation is consistent

[19] See Aidt and Jensen (2009, 172); Dincecco (2017).

Figure 5: Western Revenue Intake Before World War I



Notes. See note for Figure 1.

with the idea in the model that states may make concerted effort to improve their bureaucratic capacity even when fiscal demand is temporarily low. The two main theoretical scope conditions for such behavior fits European states well. First, initial bureaucratic supply was already relatively high. Second, they anticipated that demand would be high in the future, given the extensive history of warfare among European great powers.

Two main factors made Western societies more legible in the twentieth century than they had been before the French Revolution. First, every Western country experienced some industrialization prior to World War I. The state was actively involved in projects such as building railroads and expanding public education. Moreover, ruling elites made a conscious decision to permit new economic activities initiated by private actors, despite the generic fear of weakening their grip on power (Acemoglu and Robinson, 2012). Second, European states engaged in extensive bureaucratic reforms, a process that Mann (1993) describes for five major powers. Indicative of these reforms, all ten countries that introduced registration systems for births and deaths before 1850 are in Western Europe or its offshoots (Brambor et al., 2020). Daunton (2001) describes the re-

lated process by which Britain reformed its tax system in the nineteenth century. Public frustration with elevated taxation necessitation by high spending through the Napoleonic wars led to major retrenchment afterwards. Although *overall* revenue intake dropped throughout the nineteenth century, the reintroduction of the income tax in 1842 eventually facilitated greater public trust by making the tax system more equitable, which also bolstered quasi-voluntary compliance (Levi, 1989).

These increases in fiscal capacity also demonstrate that fiscal supply and fiscal demand varied independently. European countries boosted their bureaucratic capacity during an extended period of low warfare among major powers. Unsurprisingly, given generally low fiscal demand, Brambor et al. (2020, 202) find that participation in warfare is uncorrelated with information-capacity levels; in fact, their coefficient estimate is negative. Qualitatively, Mann (1993) stresses the importance of warfare for bureaucratic reforms during the eighteenth century. However, between 1780–1850, he highlights industrialization and pressure from outside the ruling class as equally important factors, and the most important factors after 1850 (and attributes no role for warfare in this latter period). In concurrent work, Goenaga, Sabate and Teorell (2021) show that warfare was not a major driver of fiscal expansion in European countries during the long nineteenth century.

4.2 Customs Revenues in Primary Product Exporters

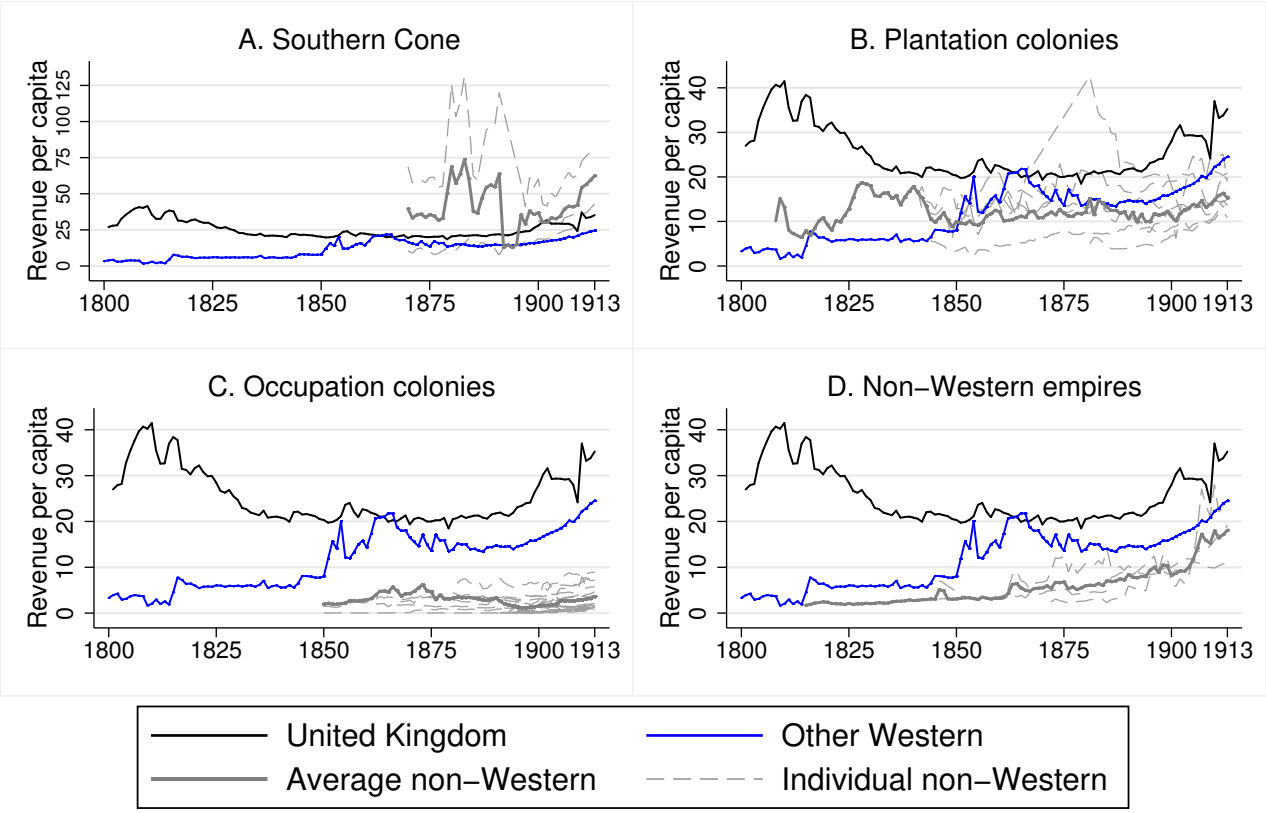
Even with low demand for public expenditures in the West, we might still expect these states to collect more revenue than states currently or previously under Western colonial rule. European colonial rule was typically based on predatory extraction, which perpetuated low fiscal supply. Colonizers shaped the fiscal systems of dependencies to reflect the needs of the metropole. Most colonies made extensive use of either coercive labor institutions or local intermediaries, both of which tended to reduce the central government's level of 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 (Van Waijenburg, 2018), and considerable financial decentralization. Britain collected head, hut, and other direct taxes in Africa through Native Authorities acting on the state's behalf (Gardner, 2012). Throughout Africa, the general principle was to impose "hegemony on a shoestring" by co-opting local indigenous institutions and collecting enough taxes to balance the budget (Berry, 1992). 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 faced great difficulty in raising the rates that they assessed (Lee, 2019).

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 levels of fiscal effort. Although Europe's military and economic dominance in this period was unchallenged, it was based on superior military technology (Hoffman, 2015), scientific innovations and economic development (Pomeranz, 2009), and epidemiological advantages (Diamond, 1999), as opposed to high levels of taxation. In Figure 6, 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 line.

Countries in the Southern Cone of South America gained independence in the first half of the nineteenth century. On average, their revenue intake was quite 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 taxes in Britain and among the average of other Western countries were 6.4% of GDP, the corresponding figure for Southern Cone countries is 9.2%. Revenues were particularly high in Chile, which reflected a boom in nitrate mining. Between 1900–13, customs taxes constituted, on average, 71.3% of Chile's total revenues. In addition to the

Figure 6: Comparative Revenue Intake Before World War I



Notes. The lines show estimated central government revenue per capita in gold grams, converted at nominal exchange rates (see Figure 1). 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.

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 (Schenoni, 2021, 418-19).

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.[20] 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 69% more than Britain, and 2.2 times more than other Western European colonizers.[21]

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 the discrepancy was eight-fold for other Western countries. Yet once again, these magnitudes were small by modern standards, and differences in GDP account for most of the discrepancy. When normalizing by revenue intake, Britain collected only two times more in revenue than occupation colonies, and the rest of the West collected only 65% more than occupation colonies.[22]

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

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

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

[22]The sample of non-Western countries for which we have GDP data at this time is much smaller than those for which we also have revenue data (only India, Indonesia, and Sri Lanka). However, from 1900–13, these three colonies collected, on average, 31% more in per capita revenues than the entire group of occupation colonies. This suggests that differences in GDP explain most of the gap in per capita revenue intake.

Southern Cone extracted large amounts of revenue, nor why occupation colonies in Africa and Asia extracted small amounts of 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 (1982, 119) argue that the empire is best characterized as "a redistribution of income within the United Kingdom than as a transfer from the empire to the mother country." 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 (Frankema and Booth, 2019, 6-8). 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" (5; emphasis in original).

4.3 Reforms in Non-Western Empires

In the final panel in Figure 6, we compare the West to major non-Western empires. Many scholars highlight a large gap in revenue intake between the West (in particular Britain) and major non-Western empires by the end of the eighteenth century (Karaman and Pamuk 2010, 623; Rosenthal and Wong 2011, 175; Hoffman 2015, 51; Dincecco 2017, 69). Despite this early mini-divergence, we show that this gap was relatively small at the beginning at the twentieth century for some of these empires, which we attribute to high fiscal demand.

We have data for three major non-Western states before World War I: Egypt, Japan, and Russia.[23] Like several other empires (China, Ethiopia, Ottoman, Siam), these states engaged in defensive

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

modernization programs to resist Western encroachment. Their ruling elites perceived high demand for centralized revenues, even in years that these states were not actively participating in war. Reforms in Japan followed two centuries of isolation and decentralized rule under the Tokugawa Shogunate, in which demand for public expenditures was low. Japan enjoyed a long history of domain-level taxation and a professional state service (Sng and Moriguchi, 2014), 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.[24]

Russia and Egypt highlight how crony-raising 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 gap between the West and Russia at the conclusion of the Napoleonic wars is even larger than the discrepancies listed by Dincecco (2017, 69) in the 1780s, which were 6.6 and 3 for Britain and France, respectively. However, the gap closed 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.[25] Between 1900–13, Russia collected 55% of the per capita revenue of Britain, and 78% of other Western countries. Although we lack GDP data during this period, it is

[24]There are discrepancies among our datasets on the latter point. Although Japan raised less on Andersson and Brambor's (2019) taxes/GDP measure, it raised more on our alternative measures: Beramendi, Dincecco and Rogers's (2019) taxes/GDP measure and our normalized revenue measure.

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

likely that this relatively small gap is entirely explained by income differences.

Finally, 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 (Ralston, 1990, 84, 91). In the 1870s, Egypt's per capita revenue intake averaged 62% of that in the West, and slightly exceeded the average Western country in normalized revenue.

4.4 Rising Demand and Permanent Revenue Divergence

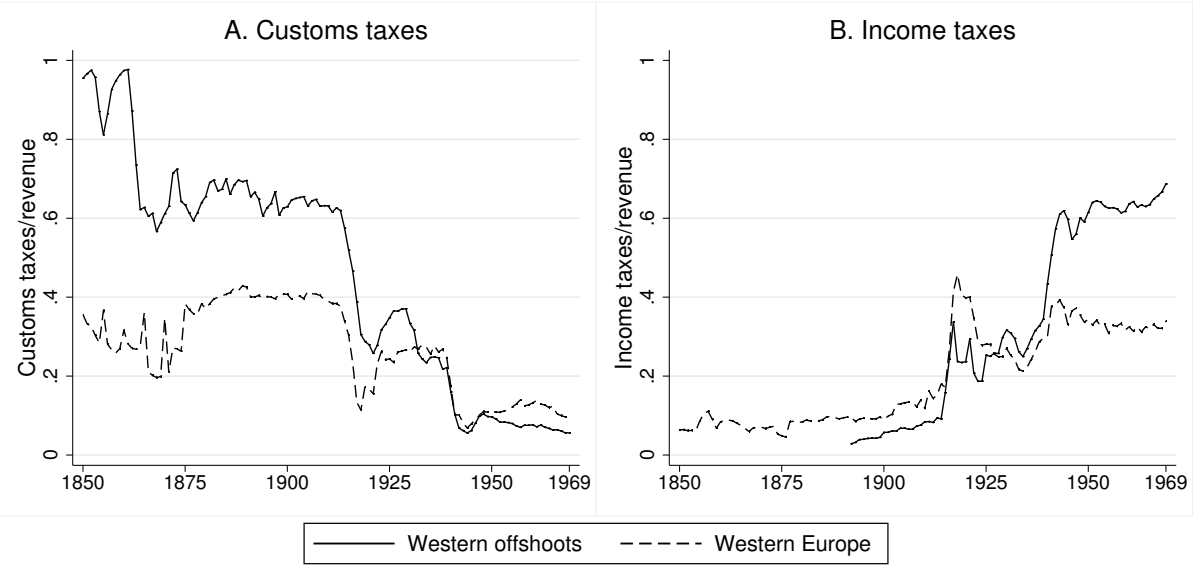
Starting with World War I, Western governments experienced permanently high demand for revenues. The two world wars required unprecedented mobilization of troops, reorganization and management of the economy to supply the war effort, and financing needs. European states not only overhauled their tax systems (Scheve and Stasavage, 2016), but also experienced pressure to expand the franchise and provide citizens with a broad array of social welfare benefits to reward their sacrifices (Lindert, 2004), even in countries that did not directly participate in the wars.[26] Although demand was the main factor that changed relative to the pre-WWI period, their prior legacy of high bureaucratic supply was crucial. "Night watchman" states in the West in the nineteenth century developed the latent capacity to raise impressive amounts of money when pressed (Briggs, 1961). Consequently, as we demonstrated in Figure 1, revenue intake in the West was considerably larger than before whether measured in per capita intake or taxes as a fraction of GDP.

Innovations in tax technology, underpinned by improvements in bureaucratic capacity, facilitated unprecedented increases in taxation. Income taxes became the predominant source of revenues in Western countries. Heavily reliant on bureaucratic competency and societal legibility, income taxes represented a major technological breakthrough in taxation capacity, with Mares and Queralt

[26] See Appendix Figure B.2 for evidence on the non-belligerents in WWI.

(2015, 1975) praising the "unprecedented revenue generating capacity" of "the most advanced fiscal instrument to date." This is also true of advanced consumption taxes, such as the value-added tax that became common in Western Europe (Steinmo, 1996). In Figure 7, we present the fraction of revenue deriving from either customs or income taxes for Western states. Whereas customs taxes once constituted the main source of revenues in Western offshoots and were also sizable in Western Europe, they had become largely unimportant by the second half of the twentieth century (in 1969, 6% of revenues in Western offshoots and 10% in Western Europe). By contrast, income taxes became the main source of revenues for Western offshoots (69%). Income taxes were less important in Western Europe (34%), but this was in large part because these countries relied more heavily on consumption taxes. In 1969, all direct taxes constituted, on average, 54% of revenues in Western European countries.[27]

Figure 7: Source of Western Revenues: Customs and Income Taxes



Notes. Data from Andersson and Brambor (2019).

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. Lead-

[27]Computed by authors by summing all categories of direct taxes from the Mitchell source data.

ing existing explanations focus on how non-European countries during this period either fought too few wars, or the "wrong" kinds of wars: limited international wars funded by debt, and civil wars (Herbst, 2000; Centeno, 2002; Besley and Persson, 2011).[28] Yet many newly independent states exhibited high demand for public expenditures. Many anti-colonial activists believed that jurisdictional sovereignty would engender higher levels of public spending by aligning the government's incentives with their citizens rather than with European bondholders and civil servants (Naoroji, 1901; Furnivall, 2014), and 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, which should create additional demand for public expenditures.

Existing explanations focused on the demand side overlook the crucial gap between the West and other states in bureaucratic institutions. Insufficient bureaucratic reforms during the colonial period offers a more compelling explanation for the general inability of non-Western states after gaining independence to converge toward Western revenue intake. The predominant strategies of taxing cash crop exports and relying on local intermediaries did not require advanced bureaucracies. For example, for direct taxation, British administrators in Africa largely relied on flat 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" (Gardner, 2012, 116). Easy-revenue sources were often sufficient to meet the limited needs of colonial states before World War II (at least relative to the costs of constructing more intensive systems), and did not greatly distinguish the colonial world because Western states exhibited such low fiscal demand.[29]

[28] Queralt (2019) analyzes the legacies of fighting the wrong kinds of wars in the nineteenth century.

[29] Western colonial rule was not the only factor that mitigated against earlier bureaucratic in-

When fiscal demand rose after independence, 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 developed a civil registration system in the late nineteenth century. At independence, India had 46 times as much census-age misreporting as the United States (Lee and Zhang, 2017). Given low supply, we anticipate that heightened fiscal demand after gaining independence did not discernibly boost revenue collection, as Lee and Paine (2019) demonstrate by estimating null differences in countries' revenue intake before and after independence.

In many post-colonial countries, low legibility has persisted long after independence. Many lack extensive written or electronic records to monitor activity, or banking intermediaries that reduce the need for government agents to meet in person to collect taxes (Moore, 2008, 40-41). In some African and Asian countries, customs revenues became *more* important in the mid-twentieth century, as 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 reform. Bates (1981) explains how many African rulers after independence undertook a classic crony-favoring strategy: using 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. Even when non-Western states have tried to impose modern direct 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.[30] One exception was South Africa

vestments. Another factor for independent Latin American countries in the nineteenth century was access to international credit at low interest rates, which enabled them to pay for wars with debt (which was often renegotiated after the wars) rather than by developing more intensive forms of domestic tax collection (Centeno, 2002; Queralt, 2019).

[30]Data for eight South American countries (and Mexico) from Andersson and Brambor (2019). Other countries are authors' calculations from the Mitchell source data: Egypt, India, Indonesia,

(51% of revenues from income taxes), which was highly effective at raising taxes within the white community (Lieberman, 2003). More recent data demonstrates that world regions exhibit drastic differences in income tax avoidance, with higher rates in Africa, South Asia, Latin America, and the Middle East than elsewhere. Losses vary from 8% of GDP in Chad to 0.16% in Finland (Cobham and Janskỳ, 2018).

Given existing demand theories, Egypt and India are "most likely" cases for strong revenue extraction. However, they instead highlight the need for legible societies, and provide examples of rulers turning to crony-favoring extraction when demand is high but supply is low. Although 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, both developed large and inefficient public sectors rather than cultivated more sustainable sources of revenues. As Waterbury (1993, 134) describes for these countries: "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." Egypt's attempt to implement a broad land reform in the 1950s and 1960s, which would have cut out large landowners as intermediaries in the tax-collection process, failed due to basic problems of bureaucratic information about land titles and related issues (Migdal, 1988). In India, the proportion of revenue collected through direct taxes fell during the twentieth century (from 28% in 1900 to 15% in 2000, with a low of 6.5% in 1987),[31] as tax-avoidance rates remained high and the government choose to raise import duties and nationalize large sectors of the economy. As of 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, 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

Iran, Pakistan, Philippines, South Africa, South Korea, Turkey.

[31] 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.

supply and demand. Japan, Taiwan, and South Korea all had long traditions of professionalized bureaucracies, in fact, longer than those in the West. Despite the brutalities of Japanese colonial rule in South Korea and Taiwan, scholars argue that Japanese institutions spurred rapid development after World War II (Kohli, 2004). These countries also 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 collected 2.3 times as much. In 1969, South Korea it collected 33% of its revenues from income taxes, which was in line with the Western European average.[32] Taiwan and South Korea converged even more closely to Western patterns in subsequent decades.

5 Conclusion

During the twentieth century, a historically unprecedented divergence in fiscal intake occurred between Western countries and the rest of the world. This divergence that occurred much later than existing theories would expect. We explain both the cross-sectional and longitudinal trends by distinguishing existing explanations in terms of "demand" and "supply" hypotheses. Whereas existing research tends to examine these in isolation, we provide a theory of how demand shocks can cause governments to engage in either legibility-intensive or crony-favoring extraction, and we show that the optimal choice depends on extant bureaucratic supply as well as the ability to boost societal legibility in the future. We then provide evidence to establish that the conjunction of high demand and high supply produces sustainably high revenue increases.

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

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 state centralization and improved fiscal capacity in some European cases in both the early modern period and the twentieth century. However, this observation does not support a simple bellicist theory in which the effect of warfare has been largely constant. The twentieth-century world wars constituted a critical juncture because the scale of European states' involvement was unprecedented (Scheve and Stasavage, 2016), which in turn facilitated unprecedented levels of revenue intake. A necessary condition for this transformation was prior bureaucratic development. Despite continual development throughout the nineteenth century, fiscal capacity remained largely latent until the world wars.

Our emphasis on the interaction of warfare with supply factors also has some resonance with earlier periods in European history despite uniformly low bureaucratic capacity. Then, differences in the incidence of parliaments is the conditioning supply variable. Although any increases in revenues during earlier periods were small by twentieth-century standards, states with limited government typically gained a greater war boost because they could borrow at lower rates and had higher levels of tax compliance, which helped to substitute for bureaucratic weakness (Dincecco, 2011; Stasavage, 2011; Karaman and Pamuk, 2013; Cox, 2016).

Our perspective also highlights the lack of bureaucratic development as central to understanding low taxation in the non-European world, shifting the focus away from the amount or type of warfare. 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 with the West. However, once demand picked up across the globe, these states were heavily restricted in their ability to raise modern sources of revenues such as income taxes. Even cases that did experience considerable international competition could not keep pace with the West. The

main exceptions were East Asian countries, such as Japan, that experienced not only high demand, but also a prior history of bureaucratic development. Outside of these exceptional cases, the institutional conditions did not exist for conflict to facilitate state building.

References

Acemoglu, Daron and James Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. Crown Business.

Aidt, Toke S and Peter S Jensen. 2009. "The Taxman Tools Up." *Journal of Public Economics* 93(1-2):160–175.

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.

Berry, Sara. 1992. "Hegemony on a Shoestring." *Africa* 62(3):327–355.

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.

Blaydes, Lisa. 2018. *State of Repression: Iraq under Saddam Hussein*. Princeton University Press.

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.

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

Briggs, Asa. 1961. "The Welfare State in Historical Perspective." *European Journal of Sociology* 2(02):221–258.

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.

Cobham, Alex and Petr Janskỳ. 2018. "Global Distribution of Revenue Loss from Corporate Tax Avoidance." *Journal of International Development* 30(2):206–232.

Cox, Gary W. 2016. *Marketing Sovereign Promises: Monopoly Brokerage and the Growth of the English state*. Cambridge University Press.

D'Arcy, Michelle and Marina Nistotskaya. 2017. "State First, Then Democracy: Using Cadastral Records to Explain Governmental Performance in Public Goods Provision." *Governance* 30(2):193–209.

Daunton, Martin. 2001. *Trusting Leviathan: The Politics of Taxation in Britain, 1799–1914*. Cambridge University Press.

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.

Diamond, Jared M. 1999. *Guns, Germs, and Steel*. WW Norton & Co.

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

Dincecco, Mark. 2015. "The Rise of Effective States in Europe." *Journal of Economic History* 75(3):901–918.

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

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* .

Haber, Stephen, Noel Maurer and Armando Razo. 2003. *The Politics of Property Rights: Political Instability, Credible Commitments, and Economic Growth in Mexico, 1876–1929*. Cambridge University Press.

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

Hinrichs, Harley H. 1966. *A General Theory of Tax Structure Change During Economic Development*. Law School of Harvard University.

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

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

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.

Kasara, Kimuli. 2007. "Tax Me If You Can." *American Political Science Review* 101(1):159–172.

Kohli, Atul. 2004. *State-directed Development*. Cambridge University Press.

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: Social Spending and Economic Growth Since the Eighteenth Century*. Vol. 1 Cambridge University Press.

Mann, Michael. 1993. *The Sources of Social Power: Volume 2, The Rise of Classes and Nation-states, 1760-1914*. 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.

Moore, Mick. 2008. Between Coercion and Contract: Competing Narratives on Taxation and Governance. In *Taxation and State-Building in Developing Countries*, ed. Odd-Helge Fjeldstad and Mick Moore. Cambridge University Press pp. 34–63.

Naoroji, Dadabhai. 1901. *Poverty and un-British rule in India*. S. Sonnenschein.

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

Peacock, Alan T. and Jack Wiseman. 1961. *The Growth of Public Expenditure in the United Kingdom*. Gregg Revivals.

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

Queralt, Didac. 2015. "From Mercantilism to Free Trade." *Quarterly Journal of Political Science* 10(2):221–273.

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: The Introduction of European Military Techniques and Institutions in the Extra-European World, 1600–1914*. 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.

Scott, James C. 2009. *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*. 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. 2011. *States of Credit: Size, Power, and the Development of European Polities*. Princeton University Press.

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.

Thies, Cameron G. 2004. "State Building, Interstate and Intrastate Rivalry." *International Studies Quarterly* 48(1):53–72.

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.

Online Appendix

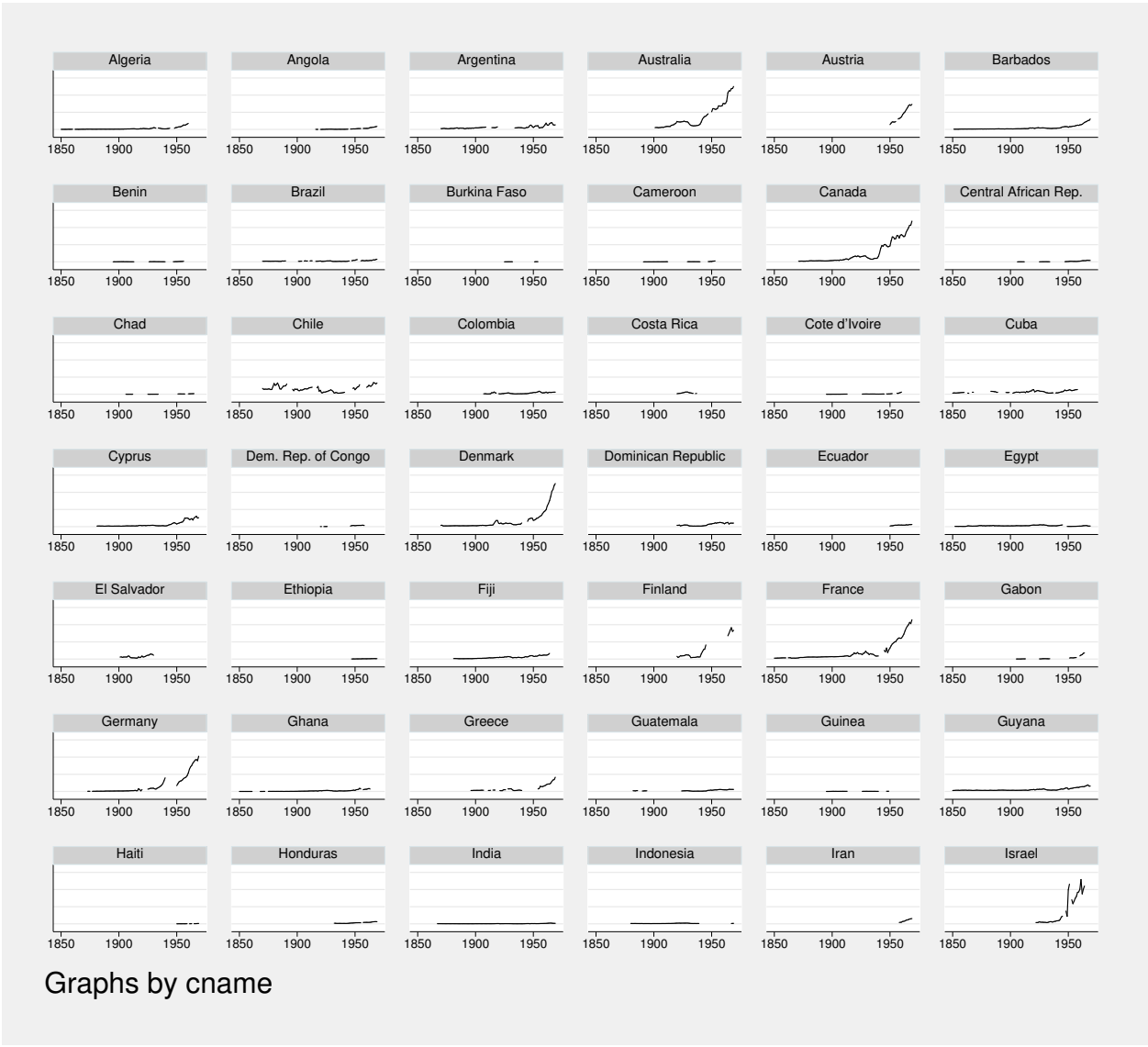
Contents

List of Tables

List of Figures

A Supporting Information on Data and Patterns

Figure A.1: Revenues Per Capita by Country, 1850–1970



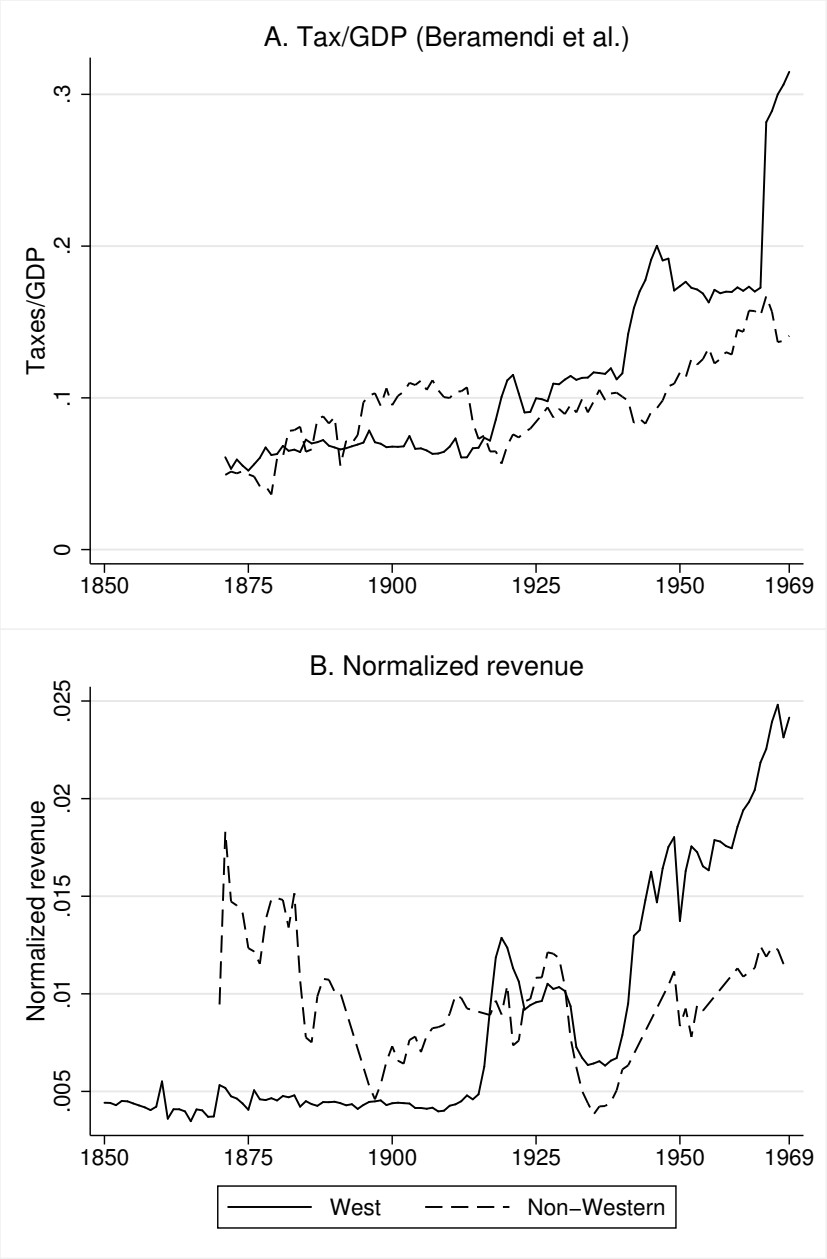
Graphs by cname

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). Among other cross-sectional historical measures of revenue that account for income levels, in the article, we discuss our preference for Andersson and Brambor (2019) because it has less missing data (relative to our core sample) than the alternatives. Here we show a similar pattern under 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, notably fewer country-years overall. For Panel B, we constructed a "normalized" revenue variable in which 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. The Mitchell source does provide some data points for nominal GDP in the local currency, although this data is unavailable for almost every non-Western country before 1950. This 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 43% 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



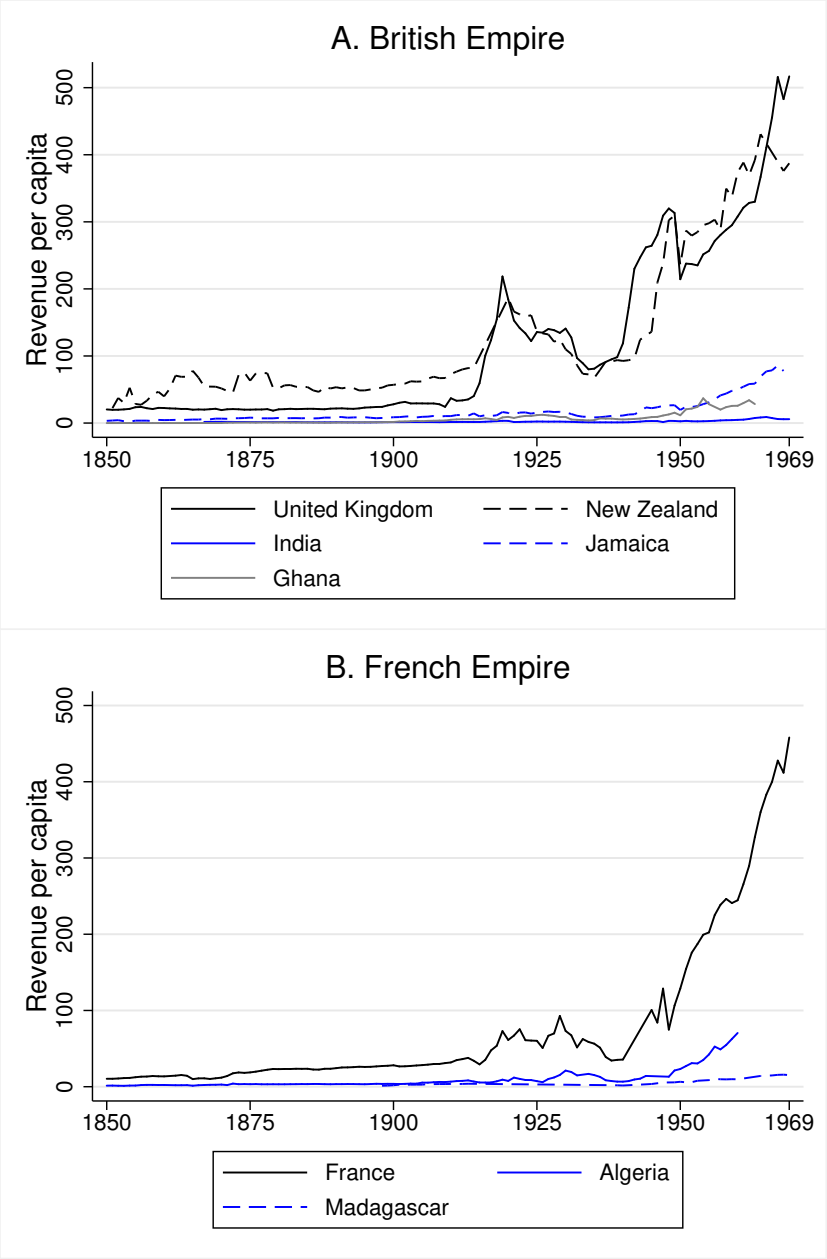
Notes. Panel A: Beramendi, Dincecco and Rogers (2019) measure of taxes/GDP. Panel B: normalized revenue. See the above text for more details on each variable.

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, but by 1968 it 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 static 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). Limited available data (i.e., only countries with PPP data in 1950 in the Penn World Tables dataset) late in our time frame show that 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 particularly 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

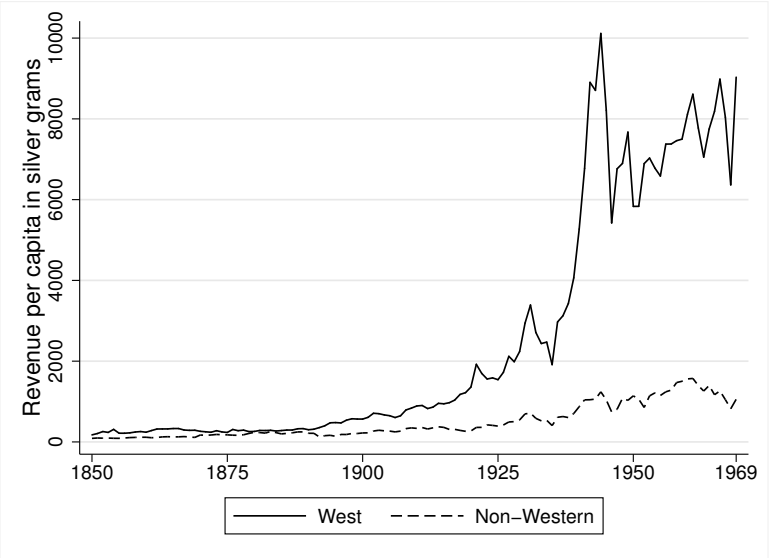
Before 1970, global currencies were typically fixed in relation to gold or silver (or both). Which precious metal is a more appropriate yardstick for value across time? This consideration is important because gold and silver prices do not move in unison. Silver prices fell by about 50% relative to gold 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 measures 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 U.S. and UK 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-1914, 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 the figures. 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 1914, 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-1914	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-1914	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-1914	0.00285		-0.0354	
	(0.00674)		(0.0553)	
West | Post-1914	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 6, 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 considerable 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.

B Regressions: Interacting Fiscal Demand and Supply

We estimate a series of regression models to assess whether the insights from the more qualitative discussion in the text generalize to a broader sample. Our main theoretical implication about combining supply and demand yields a natural statistical test, for which we provide evidence: the interaction of these variables should positively associate with revenue intake.

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.

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

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

We estimate models with two-way fixed effects to eliminate sources of heterogeneity that are constant across countries or time. Of course, decisions to participate in war and to develop a civil registration system are likely driven to some extent by country-specific factors that vary over time. However, even if so, it is not clear that this source of confounding would bias the interaction term in a positive direction. In the text we cited evidence from Brambor et al. (2020, 202) that the development of civil registration systems was not, in general, driven by participation in war. It also seems unlikely that states can usually anticipate their war needs accurately and preventively ramp up fiscal capacity. For example, although World War I eventually yielded unprecedented revenue intake in Western countries, every participant was shocked by the scale of the war effort. These

states waited several years after 1914 to impose high statutory rates on income taxes or introduce universal conscription (Scheve and Stasavage, 2016). However, given these unavoidable caveats about causal inference and data limitations, we regard these statistical associations as a plausibility probe for our theory rather than as conclusive evidence for a causal effect.

The statistical model is:

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

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

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. For the latter, it is appropriate to compare colonies with independent countries because the ability to raise revenues matters, not where the revenues are spent, although it is useful to show that such comparisons do not drive the results. Additionally, although colonized territories tended to not have civil registration systems, 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.

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 during World War I for the neutral states in Western Europe.

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

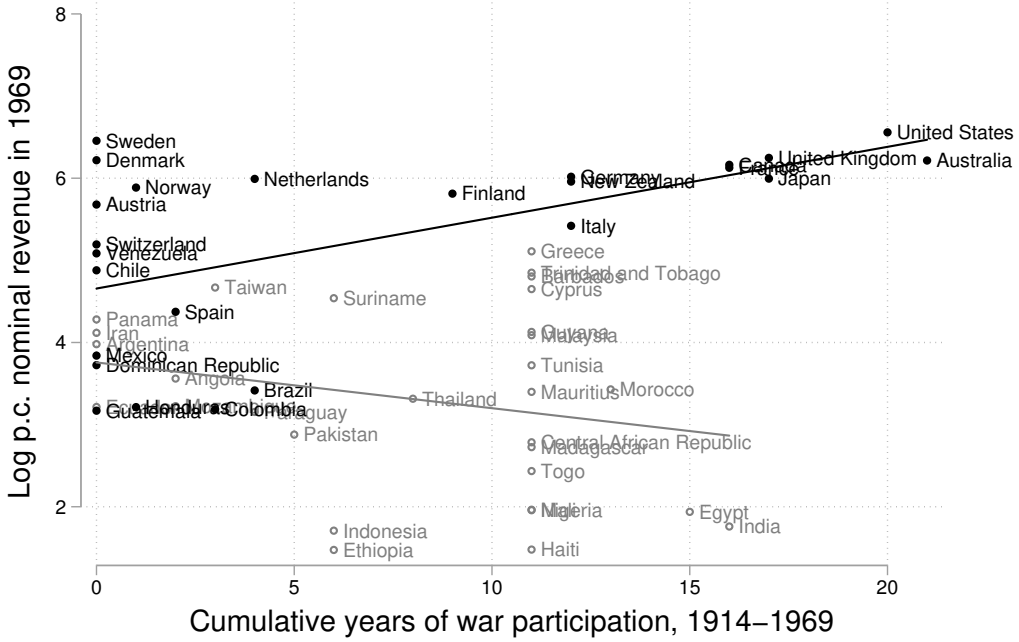
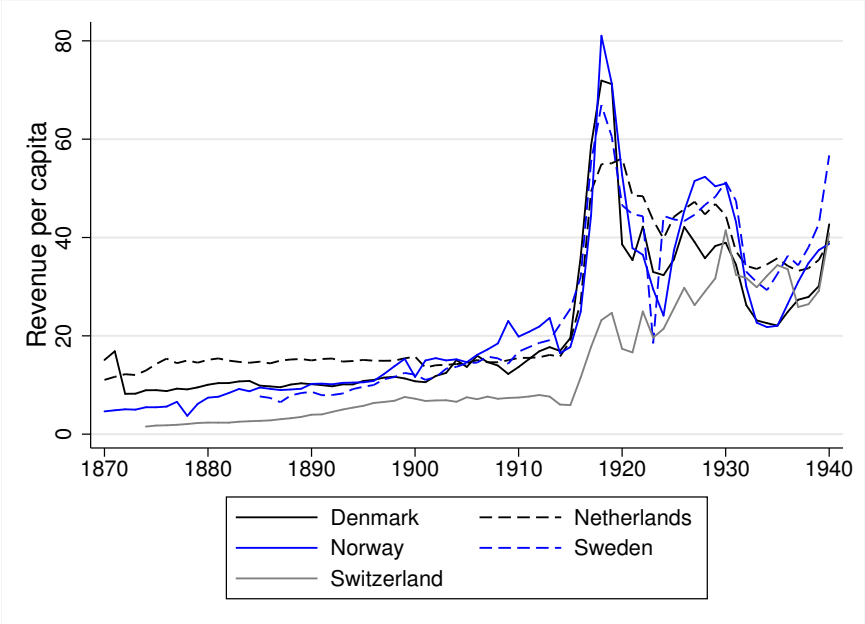


Figure B.2: Revenue Trends in WWI Neutrals



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.

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.

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.