

The Paradox of Power: Principal-Agent Problems and Fiscal Capacity in Absolutist Regimes*

Debin Ma[†]

Jared Rubin[‡]

Abstract

Tax extraction is often low in absolutist regimes. Why are absolutists unable to convert power into revenue? Supported by evidence from Imperial China, we explain this puzzle with a principal-agent model which reveals that absolutists, unconstrained by rule of law and unable to commit to not predated on their tax-collecting agents (and the masses), may find it optimal to settle for a low wage-low tax equilibrium, while permitting agents to keep extra, unmonitored taxes. Our analysis suggests that low investment in administrative capacity is a conscious *choice* for an absolutist since it substitutes for credible commitment to refrain from confiscation.

Keywords: administrative capacity, fiscal capacity, state capacity, principal-agent problem, monitoring, credible commitment, absolutism, limited government, taxation, China, Europe, Qing Empire

JEL classifications: N45, N43, H20, P48, P51

* We are grateful for comments received from Avner Greif, Bishnu Gupta, Phil Hoffman, Noel Johnson, Jin Li, Omer Moav, Thomas Rawski, Jean-Laurent Rosenthal, Matthew H. Sommer, Yu Hao, and participants at the 2015 CEPR NYU Abu Dhabi economic history workshop, 2015 World Economic History Congress in Kyoto, 2016 Economic History Society Meetings in Cambridge, 2016 Economic History Association Meetings in Boulder, 2016 SUFE economic history workshop in Shanghai, the Harvard University economic history seminar, and the University of Warwick economics department seminar. All errors are ours.

[†] London School of Economics; d.ma1@lse.ac.uk

[‡] Chapman University; jrubin@chapman.edu

I. Introduction

Tax extraction in Qing China was low relative to the leading Western European powers of the early modern period. This is paradoxical: China was much more absolutist and had stronger rights over property and people than any European power. Why did the Qing not convert their absolute power into revenue? Did they face inherent impediments to enhancing their fiscal capacity, or was their limited fiscal capacity a conscious strategy informed by their history of governing an empire without any form of constitutional constraint?

Limited fiscal capacity is not unique to early modern China. In the contemporary world, tax collection is inversely related to economic development. Besley and Persson (2014) give a host of reasons why this is the case: the presence of a large informal economy, foreign aid, resource dependence, a non-compliance culture, and weak state capacity. Historically, many states, especially autocratic ones such as early modern Spain and the Ottoman Empire, had difficulty collecting taxes in spite of their rulers ostensibly having absolute power (Tilly 1990; Dincecco 2009, 2014; Karaman and Pamuk 2013; Irigoien and Grafe 2013).

One key reason autocratic institutions are associated with low tax collection is that autocrats face a particularly vexing problem of *administration*: what we view as a set of “paradoxes of power”. First, given that unconstrained, arbitrary power is the very foundation of absolutist rule, how can an autocrat incentivize a tax collection agent to collect taxes, when the agent might be subject to the whim of discretionary power and the fear of arbitrary confiscation of wealth?¹

We suggest that an unconstrained ruler’s promise to refrain from preying on his agents is only credible when he has little capacity to find and extract the agent’s hidden, unmonitored wealth. In such a case, there is little to gain from predation since the agent has little for the ruler to confiscate. An absolutist ruler can commit to refraining from predation—tying his own grabbing hands—when administrative capacity is weak and thus monitoring costs are high. This idea is related to Greif (2005, p. 755), who proposed that rulers can commit to protecting rights by creating an ineffective administration, which increases the cost of confiscation and thus makes property

¹ Myerson (2015) argues that the ruler-agent problem can be solved via institutions that constrain rulers from arbitrarily firing their agents, so long as contracts are structured to pay off heavily near the end of their duration (which gives agents the incentive to act honestly). Yet, such a solution requires institutionalized monitoring as well as the capacity to build institutions which enable the ruler to credibly commit to upholding his promises. Building such institutions is not always possible, especially for an autocrat who stands to lose significant privileges by constraining himself.

more secure.² Our analysis therefore suggests that low investment in administrative capacity can be a rational *decision* for an absolutist ruler.

But this solution engenders the second paradox: in the absence of a monitoring infrastructure, how could autocrats prevent their agents from extracting too much taxation from the masses, which could in turn threaten imperial rule via revolt. We argue that in the trade-off between these two paradoxes emerges a “low wage-low tax” equilibrium, where rulers pay their agents low wages in return for low levels of taxes remitted to the center, supplemented with informal taxation that is *de jure* extralegal but *de facto* permitted. Rulers cannot offer high wages since such wages are not secure from their own grabbing hands, while they cannot request high levels of taxation because they must leave surplus for the agent to collect as an “extra-legal” wage (that is hidden and thus free from confiscation).³

Although our model focuses on the ruler-agent relationship, our analysis captures the fundamental uncertainty in property rights and contract enforcement issues pervasive in parts of the world ruled by arbitrary and unconstrained rule.⁴ Our model therefore sheds light on a number of stylized facts from the pre-modern era. First, it explains why the Qing rarely invested in administrative capacity and total Chinese taxation remained stationary despite population and territorial expansion. Rather than taking Confucian ideology or the Chinese ruling logic of light taxation themselves as explaining the low-taxation outcome (Rosenthal and Wong 2011), our model suggests that low tax revenues may have been an endogenous outcome of an absolutist regime with no formal constitutional constraint, legitimated and reinforced by the Confucian ideology itself. This outcome contrasts with the constitutionally constrained regimes of Europe, which had incentive to make investments in administrative capacity and were therefore able to collect much higher taxes per capita (Dincecco 2009; Karaman and Pamuk 2013; Brandt et al.

² González de Lara, Greif, and Jha (2008) and Greif (2008) argue that constraint on executive power arises when administrators are sufficiently powerful to constrain rulers, where their power derives from the fact that rulers rely on them to implement policy.

³ Absolutist rulers may also desire a low level of taxation for a related, but different reason: they fear their administrative agents growing too strong and posing a threat to the center. See Haiwen Zhou (2012). Karaman (2009) made a similar argument for the case of the early modern Ottoman Empire, where the sultan employed the judiciary to monitor tax farmers.

⁴ Confiscations could lead to an escalation of extraction at all levels of the bureaucracy. As extractions throughout the hierarchy would in the end fall on the masses at the bottom, this in turn would heighten the risk of insurrection in the long run. Although our paper is restricted to the ruler-agent relationship, insecurity in property rights due to the absence of credible commitment was a pervasive phenomenon, especially in times of political crisis, for absolutist states like Qing China (Chen 1992, ch. 7; see North and Weingast 1989 for the case of England).

2014). Second, most absolutist regimes contain an element of informal or extra-legal taxation above the official target that is permitted de facto by the ruler.⁵ We show that the ruler's inability to commit from confiscating the *known* income of the agents entails that, in order to keep the agent in the system, it must turn a blind eye to the agent collecting informal, *unknown* (to the ruler) income—permitting de facto what is illegal de jure. In other words, extra-legal taxation—often viewed as corruption—is a necessary component for an absolutist's tax administration to work properly.

Our model is consistent with these insights while *also* explaining why absolutist regimes such as the early modern Chinese Qing Empire invested so little in administrative capacity.⁶ Olson's (1993) well-known model of roving and stationary banditry has largely taken administrative capacity as given and rulers as simple revenue maximizers. Building on these insights, Tilly (1990), Besley and Persson (2009, 2010), Acemoglu (2005) and Hoffman (2015) argue that investments in fiscal capacity—and hence administrative capacity—arise endogenously precisely because of common interests in the provision of protection against external enemies.⁷ Our model adds to this literature by suggesting that a certain type of administrative capacity arose in Europe only in the presence of *both* an external threat *and* an internal institutional arrangement which consisted of some degree of credible commitment based on representative institutions or alternative arrangements. State capacity in these cases was often marked by the development of a fiscal-financial-legal nexus, the maturing of a market for public debt (Stasavage 2011; Besley and Persson 2011), and the establishment of a transparent taxation system where governments paid civil servants good wages in return for tax revenue. But this solution only worked if governments could *credibly commit* to not confiscating the agents' wealth in times of crisis. These insights imply

⁵ Tax farming also has features similar to informal or extra-legal taxation in that the ruler has little capacity to monitor the actual amount of taxes collected by tax-farmers. See Balla and Johnson (2009) and Johnson and Koyama (2014). Besley and McLaren (1993) report numerous cases, mainly in the developing world, where tax collectors regularly take bribes above the official level. Their model suggests that when enough tax collecting agents are dishonest and monitoring is expensive, permitting such bribery (and paying the agent a low, “capitulation wage”) incentivizes agents to remit the highest amount of taxes.

⁶ For excellent overviews of recent contributions to the literature on state capacity in economics and political science, see Dincecco (2014) and Johnson and Koyama (2017).

⁷ Dincecco (2009) stresses the role representative institutions played in generating fiscal capacity through increased taxation and lower sovereign credit risk. Karaman and Pamuk (2013) argue that the connection between representative institutions, war, and fiscal capacity is dependent on the economic structure of the regime: the interests of representative assemblies align with rulers with respect to war in urban settings but not in rural settings, where local control over coercive power dominates. Gennaioli and Voth (2015) take this argument one step further, arguing that once fiscal and state capacity becomes important for war-making, a divergence arises between internally cohesive states and those without cohesion, with the latter set of states dropping out of existence.

that absolutist rulers faced a fundamental paradox where their unchecked power was ultimately constrained by a weak state and low administrative capacity. This paper therefore has implications for the long-run economic divergence between China and northwestern Europe, where the issue of fiscal and state capacity has not been given adequate attention.⁸

Finally, our paper contributes an explanation of the contrast between the structure of a singular large and dominant state in Qing China versus that of multiple smaller-state in Europe. Our model indicates that incentives behind the territorial expansion of an absolutist is more related to internalizing an external threat than mobilizing the resources necessary to dispel a sudden attack—an absolutist is unable to collect high levels of taxes from newly-conquered territory—whereas the incentives of a limited ruler for expansion are likely to be the reverse. This helps explain why early modern European states sought colonial empires while the Qing expanded on their frontier.

The question of contrasting state formation and state structures has received much attention: a large literature exists on the *causes* of the relative size of states (Friedman 1977; Alesina and Spolaore 1997, 2005; Turchin 2009; Levine and Modica 2013; Ko, Koyama, and Sng 2017) and its *consequences* for economic development. The “consequences” literature considers the relationships between fractionalization and state capacity (Tilly 1990; Hoffman 2015; Gennaioli and Voth 2015), state size and monitoring capacity (Stasavage 2011; Sng 2014; Sng and Moriguchi 2014), and state size and scientific development (Diamond 1997; Lagerlöf 2014; Mokyr 2016). The China-Europe comparison has been the focus of a few of these studies. Ko, Koyama, and Sng (2017) argues that the unidirectional threat faced by China from steppe raiders encouraged political centralization, while Europeans faced attacks from multiple directions, which discouraged political centralization. Rosenthal and Wong (2011) argue that the relative fractionalization of Europe is in part responsible for the “Great Divergence” as frequent warfare in Europe, caused by fractionalization, encouraged an urban bias in manufacturing and more generally, institutions and relative prices more conducive to long-distance commerce in Europe than in China. Sng (2014) argues that the historically low levels of Chinese tax collection were a result of its size, which

⁸ This is not to imply, of course, that the “Great Divergence” literature spurred by Pomeranz (2000) is devoid of insight into the mechanisms underlying the divergence between China and Western Europe. For instance, as noted previously, Rosenthal and Wong (2011) focus on the interaction between the quality of institutions and relative prices in China and Western Europe, emphasizing the contribution of informal arrangements to supporting long-distance commerce. Greif and Tabellini (2017) compare the kin-based Chinese culture, with its clan-based enforcement institutions, and the “generalized morality” cultures of Europe, which coevolved with corporate, group-independent institutional forms. Also see Ma (2011a, 2011b, 2014) and Brandt et al. (2014).

made tax collection difficult far from the center.⁹ As in our model, Sng (2014) suggests that the Chinese kept taxes low in order to prevent over-exploitation of the masses. But our paper goes beyond this by providing an explanation on the plausible causes of the size of state. Although our paper is not the first to address these stylized facts, it is the first to our knowledge to address them in one, unifying framework.

The paper proceeds as follows. Section II overviews the nature of absolutist power in Imperial China and lays out the main historical phenomena underlying the features of low and stationary formal governmental taxation co-existing with an entrenched element of extra-legal taxation above the official level. Section III presents a model which makes sense of these features. It suggests they are equilibrium outcomes in an absolutist regime with little monitoring power and no constitutional constraint. Section IV presents further historical narratives and empirical evidence based mainly on the Qing to illustrate the model's insights. It extends the model to shed light on the long-run evolution of the Chinese imperial political structure and the comparative fiscal and financial development of other types of regimes that prevailed in early modern Europe. Section V concludes.

II. The Chinese Model of Absolutism

The origin of Chinese imperial absolutism can be traced to China's first unification under the Qin dynasty in 221 BC. At the time of the unification, the first Qin Emperor Qin Shi Huang opted against a feudal type of political arrangement and instead implemented an empire-wide prefectural system with household registration. In this model of Chinese absolutism, the emperor held ultimate power, commanding property rights over all factors of production including land and labor. At the bottom of the socio-economic hierarchy were the masses (farmers and peasants), who were nominally the tenants and cultivators of land owned by the emperor.¹⁰ The administration of the empire—tax collection, suppression of violence, and some minimal provision of public goods—was governed by direct imperial rules executed by an impersonal bureaucracy. This led to the rise of the ideology of nomenclature in Imperial China. That is, the legitimacy of bureaucrats to rule within a jurisdiction was rooted in their appointment from within the imperial administrative

⁹ Sng and Moriguchi (2014) extend these insights in a comparison of pre-industrial China and Japan.

¹⁰ The imperial ownership of land was expressed by the traditional notion of 'Wang-tu wang-min (王土王民, king's land, king's people)', which appeared in The Book of Songs compiled during the age of Warring States (403-221 B.C.) and persisted throughout the imperial period. For more, see Kishimoto (2011).

hierarchy. This political arrangement implied the dominance of a single imperial household, along with the eventual disappearance of the hereditary aristocracy and the weakening of autonomous political units and civil society.

Despite challenges and reversals, both the concept and practice of centralized rule with a hierarchical bureaucracy survived as enduring characteristics of Chinese political history. This system developed to its greatest maturity under the relatively unitary imperial structures of the Ming (1368-1644) and Qing (1644-1911). Under these dynasties, a nationwide three-tier (capital-province-county) Civil Service Examination, open to nearly all males, was used to recruit official bureaucrats inculcated in the Confucian ruling ideology. Successful examination candidates were appointed to bureaucratic posts based on a system of 3-5 year empire-wide rotation and the rule of avoidance (which precluded appointees from serving their home county). By granting life-long privileges of tax-exemption and legal immunity to varying levels of civil service examination candidates, the system generated a class of career officials having no autonomous territorial or functional power base.¹¹

This model of Chinese autocracy was ruler-centered, with no formal or external institutional constraints placed on the power of the Imperial rulers and their agents, except perhaps the vaguely defined "Mandate of Heaven".¹² There was a system of checks against bureaucratic abuses of power, but only within the administrative hierarchy, with the emperor being the final arbiter. The Imperial Chinese legal system, which was largely subsumed under the administrative hierarchy, contrasted greatly with European civil and common law, which evolved to be bound by a fixed set of precedents, rules, and procedures, and were interpreted and enforced by a neutral third party.¹³ The primary constraint rulers faced from within was the threat of insurrection: if pushed below subsistence by excessive imperial or bureaucratic abuses, the masses might resort to violent rebellion to overthrow imperial power. The well-known Confucian adage that "water can float as

¹¹ Gentries with no official posts often resided in their home villages, extending the informal power of imperial rule beneath the official bureaucratic structure. See Chang Chung-li (1955) for the role of gentry.

¹² The problem of the absence of formal constraints on the emperor is succinctly summarized by Ray Huang's study of the Ming (1368-1644) imperial system, the heyday of Chinese imperial despotism: "...Final authority rested in the sovereign, bureaucratic action was limited to remonstrance, resignation, attempted impeachment of those who carried out the emperor's orders, and exaggeration of portents as heaven-sent warnings to the wayward emperor. When all these failed, there was no recourse left" (Huang 1974, p. 7).

¹³ For the nature and problem of the so-called "disciplinary mode of justice" in traditional China, see Ma (2011a) and Stephens (1992). Under the disciplinary mode of justice, where the objective of political stability often took precedence over that of justice, the severity of punishments were determined not just by the nature of offense but also by political and other non-legal considerations.

well as overturn a boat,” just like masses could do to their rulers, is a vivid characterization of this insurrection constraint. Indeed, rebellions and insurrections were an enduring feature of Chinese history, which was marked by periodic political fragmentation and dynastic strife (Ma 2012).

The Chinese form of absolutism displayed features of benevolent authoritarianism, secular rationalism, and meritocracy, which, through the writings of Jesuits from the 16th century, won the admiration from some of the greatest 18th-century European enlightenment thinkers such as Voltaire and Leibniz (de Barry 1993; Fukuyama 2011). But for some of China’s own intellectual-cum-bureaucrats who lived within the system, like Huang Zongxi (1610-1695) and Gu Yanwu (1613-1682), much of what they witnessed in their lifetime and condemned was despotism, persecution, corruption, and excessive concentration of imperial power (de Bary 1993). In particular, the term “Huang Zongxi rule,” coined by the modern historian Qin Hui (2000), describes how numerous landmark fiscal reforms from at least the Tang dynasty (618-907 AD), which aimed at reducing fiscal burdens by sticking to a fixed official tax target and abolishing or consolidating all additional (extralegal) fees, all ended up achieving exactly the opposite outcome despite some initial temporary success.¹⁴ The problem, as Huang pointed out, was that once the previously extralegal taxation had been consolidated into the formal tax target, the rulers’ promise to stay within the bounds of the stipulated target could be quickly broken by the imposition of extralegal fees beyond the target. This led to a vicious cycle of ever increasing broken promises and tax extractions over time. Huang surmised that this vicious circle deteriorated towards the 16th and 17th centuries with escalating internal and external threats, possibly leading to the downfall of the Ming in 1644. Hence, as Huang lamented, the so-called tax consolidation was in fact a tax increase in disguise (de Bary 1993, pp.134-8; Qin 2000).

While Huang’s railing against excessive extraction followed a long-standing Confucian call for benevolent rule, his attack on the intractable problem of extralegal and arbitrary taxation reflected what we would now call a credible commitment problem on the part of the ruler. Indeed, having lived through the dark repressive years of late Ming, Huang—dubbed as China’s John Locke—and Gu Yanwei advocated some form of constraint on the power of the emperor, the

¹⁴ Among the vociferous critiques of the ills of nearly every aspect of the imperial regimes under the late Ming and early Qing, Huang Zongxi’s iconoclastic attack on China’s imperial fiscal regime recently turned into a rallying cry against state fiscal extraction in contemporary China thanks to the popular writings by historian Qin Hui. The Huang Zongxi rule came into prominence when it invoked by the then Chinese premier Wen Jiabao on the intractable problems of excessive extralegal fees and dues on Chinese farmers in the 1990s. See Qin (2000).

strengthening of local autonomy—or what Gu would advocate as restoring some elements of China’s pre-Qin decentralized feudal institutions—and the relative independence of civil service and public education. Following the rise and consolidation of absolutist power under the Qing from the 17th century onward, the ideas of Huang and Gu were at best marginalized until the late 19th century when they became the new enlightenment inspirations for China’s modernizers in their endeavor to embark on modern constitutional reform (de Barry 1993).

Turning to the context of Chinese absolutist rule in the 18th century, a well-known fiscal reform under the high Qing offers insight into the dilemma raised by Huang Zongxi. To mitigate the problem of insufficient local revenue, which became a major source of abuse and arbitrary taxation at the local level, the energetic Qing emperor Yongzheng (1723-35) increased surcharges to land taxes by essentially legitimizing what was previously “illegally” levied local extractions. These surcharges, now included in the official tax revenue system, were used to support local administration and provide additional salary supplements for bureaucrats, the so-called *Yang Lianyin*, which is literally translated as “honesty nourishing silver”. The reform achieved initial success during the short Yongzheng reign by boosting local taxation and provision of public goods. But by the Qianlong reign (1735 – 1796), it was largely abandoned. There are multiple possible reasons the Qianlong abandoned this provision. First, almost in a nod to the warnings of Huang Zongxi and the Confucian ideology of benevolence, the much more conservative Qianlong was wary that levying *Yang Lianyin* would raise the official tax target. But more importantly, the *Yang Lianyin*, once designed for local administration, was by then increasingly taken out of the hands of local magistrates and re-allocated by the upper hierarchy, often in times of fiscal stress. The outcome was that local governments were forced to turn to a new round of “extralegal taxation” (Zelin 1984, ch. 7).¹⁵ Hence, a reform that had at one point provided a basis for formal local taxation and curbed corruption ended up exposing previously hidden revenue to possible extraction from the upper level officials. In the end, local officials found the extra-legal taxation—being outside the official purview—as the most secure source of local finance, and many even cynically resisted having their “salaries,” or *Yang Lianyin*, raised (Kaske 2016, p. 13; Zelin 1984, p. 301).

Ironically, some of the intended objectives of the Yongzheng fiscal reform may have come into fruition in late 19th and early 20th century China in a completely different political context. Western colonial administrators such as Robert Hart and Richard Dane reformed and reconstructed

¹⁵ See Hao and Liu (2016) for a preliminary empirical test on the impact of the Yongzheng fiscal reform.

traditional Chinese fiscal institutions, such as the Maritime Customs and Salt Revenue, to modern centralized national level bureaucracies with taxes levied according to transparent and pre-set rules and staffs paid good wages with stable tenure and relatively little corruption. The critical elements of success were its autonomy in personnel appointment and the freedom from central and other local governmental predations on both the tax revenue stream and the income of the staff. Both trade and tax revenue increased rapidly under the new administrations, to the benefit of both the government and private business. The pretext for Western colonial intrusion into these Chinese tax institutions was precisely because these revenues were often used as collateral for Chinese government debt in the late 19th and early 20th centuries. The rise of public debt and secondary markets during the early 20th century led to a remarkable financial transformation (Adshead 1970, van de Ven 2014, Ma 2016).

We illustrate the early 20th century Chinese fiscal reform not to condone Western imperialism but to reveal an unusual natural experiment regarding how rule of law or some form of what we later refer in the model as “limited rule” need not necessarily arise from some certain grand political events such as the 1688 Glorious Revolution.¹⁶ But once implemented, even in a Chinese cultural context, it could open a new path towards the establishment of a well-functioning fiscal bureaucracy paid with formal efficiency wages and proper monitoring infrastructure (Besley and McLaren 1993, Besley and Persson 2014). In the context of imperial Chinese absolutism, the inability of the rulers within the hierarchy to *commit* to their ex-ante promise to not confiscate their agents’ income once their incomes were known and transparent reflected a general and pervasive problem of insecure property rights and government opportunism. Hence, information asymmetry embedded in informal taxation—itsself a source of corruption and bureaucratic abuse—served as a form of substitute for credible commitment to tie the grabbing hands of the absolutist regime, albeit with a huge cost.

Next, we will turn to a model that captures these seemingly contradictory stylized facts regarding Qing taxation in one consistent and coherent framework. These facts include: a low

¹⁶ It is well-known that rule of law and constraints on the ruler existed well before the rise of modern parliamentary democracy. For instance, the Magna Carta, the rise of independent legal profession in medieval Europe, and the rise of common law in England all placed constraints on rulers. This allowed rulers to credibly signal their intent to upholding the rule of law, at least for elites (North, Wallis, and Weingast 2009). In our model, we are interested in equilibrium outcomes under different regimes, so we do not endogenously model credible commitment. Instead, we simply assume that credible commitment exists or it does not (perhaps due to a constitution (North and Weingast 1989; Myerson 2008) or public declarations of legitimacy (Greif and Rubin 2017)), and solve the model therein.

statutory tax rate (i.e., taxes remitted to the center); “off-book” taxes taken by bureaucratic administrators; de facto (though not de jure) acceptance by the central government of these off-book taxes; and persistently weak administrative capacity.

III. The Model

In this section, we model the relationship between a ruler and a tax-collecting agent. The ruler needs taxes to fend off exogenous threats (e.g., invasion, natural disaster), and he relies on the agent to collect the taxes. In our model, the players face three key constraints: i) the ruler sets a tax rate subject to the constraint that high taxes encourage the masses to revolt; ii) the ruler must offer a high enough reward to the agent to satisfy the agent’s participation constraint; iii) the agent must choose its actions subject to the constraint that he may be monitored and punished by the ruler.

III.1. Setup

Consider an economy in which a ruler (R) and a tax-collecting agent (A) interact for one period consisting of three stages.¹⁷ The ruler desires tax revenue to deal with an exogenous threat, $X \in \{0,1\}$, which occurs at the end of the period with probability $\theta \in [0,1]$, as well as an endogenous threat from the (unmodeled) population, which arises when they are taxed too heavily. The agent’s primary job is to collect taxes, $T \in \mathbb{R}^+$, from the exogenous, taxable surplus produced by the population.¹⁸ The agent’s goal is to maximize the expected value of his income stream minus tax collection costs. The ruler aims to maximize the probability of staying in power, which is threatened by both the exogenous and endogenous threats.

The period consists of three sequential stages. In the first stage, the ruler sets a statutory tax $t \geq 0$ for the agent to collect and to be remitted to the ruler, offers the agent a wage $w \geq 0$ for the agent to collect, while also deciding whether to monitor the agent, $M \in \{0,1\}$, at cost $m \in \mathbb{R}^+$. The wage, w , is not a transfer from the ruler to the agent, but instead is the level that the ruler deems

¹⁷ In a previous version of the paper we solved a model in which the ruler and agent act over infinite periods. The key results hold in the one period model, which has the benefit of being much simpler and more straight-forward. Hence, we present the one period model in this paper. The infinite-horizon model is available upon request.

¹⁸ We assume that the surplus is large enough that a corner solution does not exist (i.e., where the agent collects all of the surplus in taxes).

permissible for the agent to collect and keep for himself.¹⁹ The crucial feature of statutory taxes and the agent’s wage is that they are “on-the-book” and can be tracked within the official accounting system.

In the second stage, the agent collects taxes $T \geq 0$ from the population at cost $c(T)$, where c is smooth, continuous, and increasing and convex in its argument. c is interpreted as the non-monetary costs of tax collection (e.g., time cost, opportunity cost, engendering hatred from the populace). The agent may choose to collect 0 taxes if doing so does not yield his reservation utility, $\underline{u} \in \mathbb{R}^+$. If the agent chooses $T > 0$, he pays statutory taxes and his wage from T , entailing that $T \geq t + w$. However, he may also choose to take extra taxes (i.e., residuals) above the level set by the ruler, which we denote $t^O = T - t - w$. These “off-book” taxes are not visible to the ruler unless he invests in building a monitoring infrastructure and incurs a monitoring cost. The crucial tradeoff is that the ruler can always adjust t (i.e., “statutory taxes”) to attain the lowest possible probability of the exogenous threat succeeding, but when the he asks for a higher level of t , he encourages the agent to take a higher level of “off-book” taxes, which heightens the probability of the endogenous revolt succeeding.

In the third stage, the ruler decides whether to fire (and punish) the agent, $F \in \{0,1\}$, at cost $f \in \mathbb{R}^+$. Firing the agent is costly because it involves a search cost for a replacement, damage to bureaucracy morale, and so on.²⁰ A fired agent keeps none of the tax revenue he collected—on-book or off-book—reflecting the idea that rulers can punish agents harshly. A fired agent receives his reservation utility \underline{u} .

After the third stage, the exogenous threat is realized with probability θ . Figure 1 summarizes the progression of game play, and Table 1 summarizes the model’s key variables and parameters.

Figure 1: Game Play

<u>Stage 1 (Ruler)</u> R chooses statutory tax t , wage w , and monitor, M	<u>Stage 2 (Agent)</u> A collects tax T and remits t	<u>Stage 3 (Ruler)</u> R decides whether to fire agent, F
---	---	--

¹⁹ In our context, w could denote all the formal revenue the ruler permits to be retained for local use, which could cover agent’s wage income and cost of tax collection. We therefore refer to w as wage income for exposition purposes.

²⁰ If we were to imagine a repeated game setting, the cost of firing is a simplification for the cost to the ruler of gaining a (bad) reputation for firing agents undeservedly.

Table 1: Summary of Key Variables and Parameters

<u>Choice Variables</u>		<u>Parameters</u>	
t	Statutory tax set by ruler	m	Monitoring cost
w	Wage permitted by the ruler	f	Firing cost
T	Total taxes collected by agent	X	External event (0/1) w/ probability θ
M	Ruler's decision to monitor (0/1)	\underline{u}	Agent's reservation utility
F	Ruler's decision to fire (0/1)		

We interpret m to reflect the administrative capacity of the ruler. For example, $m = 0$ entails that the ruler has the administrative capacity to costlessly monitor the agent, while $m \rightarrow \infty$ indicates that it is impossible (i.e., too costly) for the ruler to monitor. When $m = 0$, the ruler can perfectly monitor T , and the so-called off-book taxes simply form part of the statutory taxes and wage income ($T = t + w$). This is close to the ideal of modern governance institutions where a professional and impartial bureaucracy collects and remits taxes.

We solve the model by considering two different types of rulers: a *limited* ruler and an *absolutist* ruler. The distinction between the two is that the limited ruler can *credibly commit* to not confiscating the agent's on-book wage (w) ex post, whereas the absolutist ruler cannot make such a commitment. Therefore, the distinguishing feature between a limited and absolutist ruler is that an absolutist can act like an arbitrary despot in its relations with the agent when it wishes. Either type of ruler can confiscate off-book taxes (t^0), should they be visible (i.e., when the ruler monitors the agent). Since off-book taxes are de jure illegal, confiscation does not violate our definition of credible commitment.

When the external threat is realized, the ruler can use resources to combat the threat (i.e., provide famine relief, defense against invasion). The probability the ruler successfully combats the external threat is increasing in his net revenue, π^R . The probability that the ruler successfully combats the endogenous threat is also increasing in π^R , although resources are much more urgently needed when the exogenous threat is also realized. Should the agent choose to remain in the system (i.e., $T > 0$), π^R is defined by the following equation:

$$(1) \quad \pi^R = t + F(\mathbb{I}_A w + M[T - t - w] - f) - Mm,$$

where \mathbb{I}_A is an indicator equaling one if the ruler is an absolutist (limited rulers cannot confiscate on-book wages by definition).

The ruler's survival probability, q , is a function of: i) π^R , which can be used to combat both the exogenous and endogenous threat; ii) T , since higher levels of T entail a more serious endogenous threat, and iii) X , where $X = 1$ entails that the exogenous threat is realized. We write the ex ante expected utility of the ruler as:

$$(2) \quad E[U^R] = E[q(\pi^R, T, X)],$$

where E is the expected value operator, q is smooth, continuous, and $q_1 > 0$, $q_2 < 0$, $q_{11} < 0$, $q_{12} > 0$, $q(\cdot, \cdot, 1) < q(\cdot, \cdot, 0)$, $q_1(\cdot, \cdot, 1) > q_1(\cdot, \cdot, 0)$, and $q_{12}(\cdot, \cdot, 1) < q_{12}(\cdot, \cdot, 0)$. We assume that revenue is particularly crucial when $X = 1$ (i.e., the exogenous threat is realized), meaning that $q_1(\cdot, \cdot, 1) > \bar{q}$ for some large value of \bar{q} . This highlights Olson's (1993) famous "roving bandit" idea that rapaciousness is optimal when the threat is imminent. The ruler chooses t , w , and M in the first stage and F in the third stage to maximize his expected utility subject to the agent's actions and participation constraint, or:

$$(3) \quad \max_{t, w, M} E[U^R] \quad \text{s.t.} \quad T > 0, \quad E[U^A] \geq \underline{u}.$$

Meanwhile, should the agent choose to stay in the system, his utility is his ex post income minus tax collection cost, which we write:

$$(4) \quad E[U^A] = E[F]\underline{u} + (1 - E[F])u(T - t) - c(T),$$

where $T - t$ is the agent's revenue (including his wage, w) should he keep his job, and u is smooth, continuous, and increasing and concave in its argument. The agent chooses T in the second stage to maximize $E[U^A]$, subject to the ruler's actions and his participation constraint $E[U^A] \geq \underline{u}$.

III.2. *Equilibrium and Comparative Statics*

We solve for the subgame perfect Nash equilibrium of the game described above using backwards induction. There are two equilibria of interest: one in which the ruler is absolutist and one in which he is limited. In this section we provide the intuition for the primary results. Formal proofs are available in the Appendix.

First, consider the ruler's third stage decision of whether or not to fire the agent, given t , w , M , and T . Note that the ruler's decision affects π^R but not T or X . Hence, regardless of whether the ruler is absolutist or limited, his optimal decision is the one that maximizes π^R . From (1), it is straight-forward to see that the ruler will choose $F = 1$ iff:

$$(5) \quad \mathbb{I}_A w + M(T - t - w) > f.$$

Next, consider the agent's second stage decision to collect taxes, $T \geq t + w$, given t , w , and M . The agent can also decide to leave the system—not remitting any taxes—and receive his reservation utility. The agent's choice depends on two factors: the ruler's first stage decision to monitor and whether the ruler is absolutist or limited. These two factors determine whether the ruler will fire the agent, conditional on the agent's choice of T . To solve for the equilibrium, we therefore consider the following four cases: i) $M = 0$, ruler is absolutist; ii) $M = 1$, ruler is absolutist; iii) $M = 0$, ruler is limited; iv) $M = 1$, ruler is limited.

Rather than lay out the formal equations, which are available in the proof of Proposition 1 in the Appendix, we provide the intuition behind the agent's decision. First, consider the case in which the ruler is absolutist. If the ruler chose not to monitor ($M = 0$) in the first stage, the agent knows he will be fired if $w > f$, and will therefore opt out and choose $T = 0$. However, if $w \leq f$, the ruler will not fire the agent regardless of how much tax the agent extracts. The agent therefore taxes up to the point that his marginal benefit from additional taxes equals the marginal cost of tax collection (i.e., $u'(T - t) = c'(T)$), so long as this satisfies his participation constraint. On the other hand, if the absolutist chose to monitor ($M = 1$), the agent knows that he will be fired if he collects taxes such that $T > t + f$. The absolutist ruler cannot commit to refrain from confiscating the agent's income when it is above f , and the agent's wage is $T - t$ when he collects $T \leq t + f$. Note that this decision is independent of the “on-book” wage, w , that the ruler chose in the first stage. As we will show below, this is an important point of distinction between absolutist and limited rulers, since the latter can commit to refrain from confiscating on-book wages (w) by definition. The agent in the absolutist regime therefore chooses to opt out if choosing $T = t + f$ does not yield his reservation utility, while he chooses $T = t + f$ otherwise.²¹

If the ruler is limited, the agent knows he will never be fired if the ruler chose not to monitor. Hence, the agent taxes up to the point that his marginal benefit from additional taxes equals the marginal cost of tax collection (i.e., $u'(T - t) = c'(T)$), so long as this satisfies his participation constraint. On the other hand, if the limited ruler chose to monitor, the agent knows he will be fired if he chooses to collect taxes $T > t + w + f$. In this case, the agent's on-book income, w , is free

²¹ Formally, there exists the possibility that the agent chooses an interior solution such that $T < t + f$. However, it is straight-forward to show that the ruler will never choose a value of t in which the interior solution is chosen in equilibrium. To keep the analysis straight-forward, we focus on equilibrium path actions in this section.

from confiscation by definition, and hence the ruler only fires the agent if the agent takes more in off-book taxes ($T - t - w$) than it costs to fire him. The agent in the limited regime therefore chooses to opt out if choosing $T = t + w + f$ does not yield his reservation utility, while he chooses $T = t + w + f$ otherwise.

Moving to the ruler's choice of t , w , and M in the first stage, first consider the choice made by an absolutist. If he chooses to monitor, he cannot commit to firing the agent unless the agent's total income is less than the cost of firing, f . If this income level is less than the agent's reservation utility, the agent will not collect taxes. Even if a tax level t does exist in which the agent collects taxes in equilibrium, the ruler can only ask for t up to the point where the agent does not prefer his reservation utility. In other words, when the absolutist monitors, t will be small, if it exists at all. On the other hand, if the ruler does not monitor, he can ask for higher t , which the agent will be willing to collect because he can collect off-book taxes that are free from confiscation. This comes with the tradeoff of higher T , and thus a higher probability the endogenous threat will succeed.

Whether the absolutist chooses to monitor or not is therefore a function of the cost of monitoring (m) and the probability the exogenous threat will be realized (θ). In this section, we focus on how monitoring costs affect equilibrium actions, since this parameter is central to our historical narrative. Our first, and most straight-forward, result is that the absolutist will monitor if and only if $m < m^A$, for some sufficiently high level of monitoring cost $m^A \geq 0$.²² In other words, the size of the parameter space over which the ruler monitors is decreasing in the cost of monitoring. For a large part of the parameter space, $m^A = 0$; i.e., the absolutist does not monitor even when monitoring is *free*. This is because when the absolutist monitors, it can only commit to not confiscating from the agent if the agent's total income is less than the cost of confiscation, f . If this income yields the agent less utility than his reservation wage, the agent will leave the system and not collect any statutory taxes. In this case, the absolutist is better off not monitoring and permitting the agent to take his income off-book, while remitting some positive level of statutory taxes.

Next, consider the limited ruler's first stage choices. Unlike the absolutist, the limited ruler can increase the agent's income by offering him on-book wage w , which is secure from confiscation.

²² We formally define m^A and m^L in the Appendix.

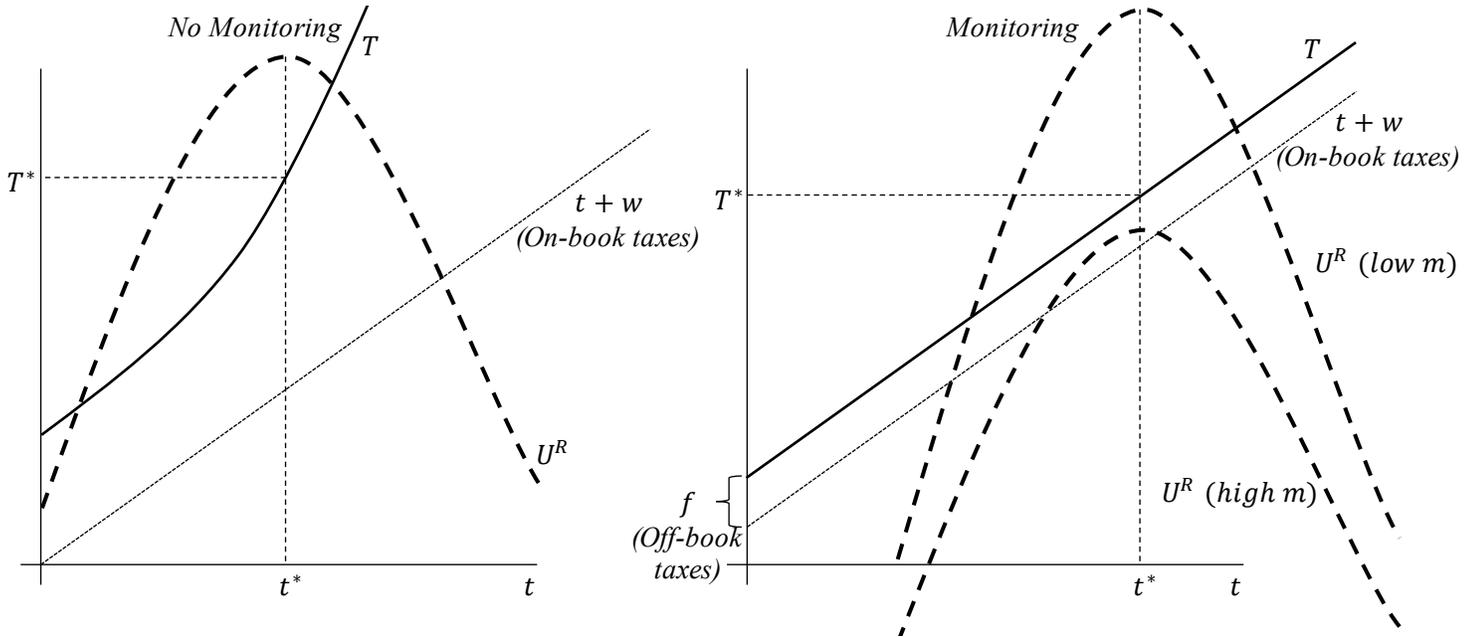
The limited ruler therefore chooses a combination of t and w that gives the agent at least his reservation utility while also optimizing the tradeoff between higher t and higher T . On the other hand, if the limited ruler does not monitor, he cannot commit to ever firing the agent, and the agent therefore collects taxes up to the point where his marginal benefit from revenue equals his marginal cost of tax collection. This may allow the ruler to collect a higher level of t relative to when he monitors, but it comes with the tradeoff of a higher level of T . Like the absolutist, whether the limited ruler chooses to monitor or not is therefore a function of m and θ . Since the intuition is the same, we simply note that the limited ruler will monitor if and only if $m < m^L$, for some $m^L \geq 0$.

We attempt to clarify this logic in Figure 2, which shows (qualitatively) the tradeoff the ruler faces between statutory taxes remitted to the center (t) and total taxation collected by the agent (T), and how this tradeoff affects the choice made by the ruler. When the (absolutist or limited) ruler does not monitor (left panel) a higher level of t comes with the tradeoff of a higher level of T , since the agent takes taxes up to the point that $u'(T - t) = c'(T)$. A higher level of t benefits the ruler since it increases the probability an exogenous threat will be dealt with successfully, while a higher level of T harms the ruler, since it makes the endogenous revolt more likely to succeed. At low levels of t the former effect dominates while at higher levels of t the latter effect dominates. Hence, there is some internal optimal combination $\{t^*, T^*\}$ that maximizes the ruler's utility.²³ Meanwhile, a similar tradeoff between t and T faces the ruler when he monitors (recall that $T^* = t^* + f$ in an absolutist regime and $T^* = t^* + w^* + f$ in a limited regime, and hence T^* is linear in t). Note that the agent takes off-book revenue equal to f in equilibrium; since the ruler monitors, he will fire (and confiscate income) of an agent that collects more than it costs to fire. We superimpose two utility functions onto the figure (right panel), one with a high monitoring cost (m) and one with a low monitoring cost. The high m curve is merely the low m curve shifted down by m . Comparing the two figures, it is straight-forward to see that the ruler prefers to monitor when the monitoring cost is low but prefers not to monitor when the monitoring cost is high (i.e., his utility when monitoring is higher in the former case but not in the latter case). We formalize this intuition in Proposition 1.

²³ Since $w = 0$ under the absolutist and is folded into the agent's overall income under a limited ruler, we simplify this figure by showing the case where $w = 0$ regardless of whether the ruler is absolutist or limited.

Proposition 1. In equilibrium, the absolutist monitors if and only if $m < m^A$, while the limited ruler monitors if and only if $m < m^L$.

Figure 2: t and T tradeoff and Ruler's utility, with and without monitoring



Our primary interest is the *comparison* between absolutist and limited regimes. Specifically, our motivating questions include: Which regime collects more on-book taxes (including both statutory taxes, t , and on-book wage, w)? Which one is more likely to monitor? Which one permits more off-book taxes? The remainder of this section is dedicated to addressing these questions.

We first address the question of statutory tax collection. The logic laid out above indicates that in the part of the parameter space where both types of rulers choose not to monitor, i.e., $m \geq \max\{m^A, m^L\}$, both types collect the same level of statutory taxes. In this case, their problems are identical: the agent will collect taxes up to the point where, conditional on t , their marginal benefit from additional revenue equals their marginal cost of collection. This is represented by the $\{t^*, T^*\}$ combination in the left panel of Figure 2. On the other hand, in the part of the parameter space where both types of rulers choose to monitor, i.e., $m < \min\{m^A, m^L\}$, the limited ruler collects more statutory taxes. In this case, the absolutist cannot offer the carrot of a higher wage to incentivize tax collection, since that wage is not secure from confiscation. The limited ruler can offer such a wage, however, which incentivizes the agent to collect more taxes (up to the point he

reaches his reservation utility). Therefore, when both types of rulers choose to monitor, the limited ruler collects more statutory taxes.

Since the limited ruler can collect more statutory taxes when he monitors, monitoring is more attractive to the limited ruler. Hence, $m^L \geq m^A$, ceteris paribus. Moreover, in the part of the parameter space in which the limited ruler monitors and the absolutist does not monitor, the limited ruler collects more statutory taxes than the absolutist when the exogenous threat is highly probable, i.e., $\theta \geq \bar{\theta}$, for some $\bar{\theta} \geq 0$. To see this, note that a limited ruler *could* act like an absolutist and not monitor; it would derive the same level of both statutory taxes and total taxation. The fact that he chooses to monitor indicates that he is better off from doing so. When θ is sufficiently large, this entails choosing a higher tax level to fend off the exogenous threat (remember, we assume that $q_1(\cdot, \cdot, 1) > \bar{q}$ for some large value of \bar{q}). We summarize these insights in Propositions 2 and 3.

Proposition 2. If $\theta \geq \bar{\theta}$, a limited ruler collects (weakly) greater statutory tax revenue than the absolutist, ceteris paribus.

Proposition 3. The size of the parameter space over which the ruler chooses to monitor is (weakly) larger for the limited ruler than the absolutist.

III. 3. Endogenous administrative capacity

The analysis in the previous section indicated that absolutist rulers suffer less than limited rulers when administrative capacity is weak (i.e., m is large), since failing to monitor their agents gives absolutists a mechanism through which they can commit to not confiscating their agents' income. Indeed, we noted in the previous section that the absolutist may choose to not invest in monitoring even when it is *free* (i.e., $m = 0$). This indicates that an absolutist ruler may decline to invest in improved administrative capacity even if doing so is inexpensive. In this section, we briefly extend the model to capture this insight.

Suppose the ruler plays the same game with the agent as described above, except that monitoring cost m is a choice variable. Specifically, assume that prior to the first stage the ruler can choose an investment in administrative capacity which allows him to monitor at cost $m \geq 0$. Investment in administrative capacity costs the ruler $d(m)$, where $\lim_{x \rightarrow \infty} d(x) = 0$ and $d' < 0$. The

intuition espoused above and formalized in Proposition 3 suggests that limited rulers are more likely to invest in monitoring capacity than absolutists. By investing in administrative capacity, limited rulers gain the ability to reduce taxes collected by the agent (T) and thus reduce the probability of an endogenous threat succeeding. The limited ruler therefore weighs the benefit of a lower probability of the endogenous threat succeeding with the cost of investing in monitoring capacity to choose an optimal level of investment.

The absolutist faces a different set of incentives from the limited ruler. If the absolutist monitors, he cannot commit to refraining from confiscating the agent's income above the cost of firing, f . Hence, the benefit of monitoring is lower for the absolutist than the limited ruler, since the level of statutory taxes the absolutist collects is lower. Therefore, the marginal benefits of investing in lower m are lower for the absolutist, while the marginal costs are the same. Proposition 4 summarizes this intuition.

Proposition 4: An absolutist ruler invests in (weakly) less administrative capacity than a limited ruler, *ceteris paribus*.

Combined, Propositions 2, 3, and 4 indicate that low levels of statutory taxation, monitoring, and investment in administrative capacity—the “Qing equilibrium” described in Section 2—may be a rational outcome for a ruler who cannot credibly commit to constraining himself in time of need.

III. 4. Endogenous size of empire

In this section, we extend the model to account for the size of empire. The model is conducive to such an analysis because one of the key difficulties of extending empire is administrating the newly-conquered territory. The further the new territory is from the ruler's capital, the more costly it is to monitor. There are numerous reasons why a ruler may want to conquer a new territory: personal glory, co-opting a potential rival, or increased tax revenue or resources (Alesina and Spolaore 2005; Ko, Koyama, and Sng 2017; Zhou 2012). Indeed, the model suggests the possibility that limited and absolutist rulers have different motivations for conquering new territory. A limited ruler can build a monitoring infrastructure which permits him to extract resources from the new territory. Meanwhile, although an absolutist will have a difficult time extracting revenue, precisely for this reason he has incentive to co-opt a potential rival, who poses

a costly threat should it attack. Which one of these factors dominates the other depends on the monitoring costs each ruler faces, as we spell out below.

Suppose the ruler plays the same game with an agent in the territory he currently rules as laid out in Section III.1. In addition, suppose the ruler is considering conquering a new territory. For simplicity, assume that the ruler will conquer the territory with certainty if he spends resources r (and he will not conquer it otherwise), and the cost of monitoring an agent in the new territory is \hat{m} , which is greater than m because the new territory is further from the capital and thus more costly to monitor. Assume the new territory has some surplus which the ruler can extract via a new agent, subject to the same constraints in the previous game. Finally, the ruler faces the exogenous threat with probability $\hat{\theta}$ if he conquers the territory. We assume $\hat{\theta} < \theta$, since the ruler is co-opting a potential outside threat. Therefore, the benefits of conquering new territory are greater access to revenue as well as a lower probability of facing an external threat. The primary cost to the ruler is the resources expended to conquer, r .

A limited ruler will be able to extract more from the new territory than the absolutist ruler for reasons laid out in Proposition 2. However, it is more attractive on the margin for an absolutist to lower the probability of an exogenous threat succeeding by co-opting part of the threat, since the probability the absolutist will succumb to the threat is greater.

This intuition, summarized in Proposition 5, indicates that when \hat{m} is small, say $\hat{m} \leq m^*$, a limited ruler has more incentive to conquer the new territory than an absolutist. At small \hat{m} , the limited ruler extracts much more revenue from the new territory than the absolutist. However, when monitoring costs in the new territory are sufficiently large, $\hat{m} > m^*$, an absolutist ruler has more incentive to conquer the new territory than a limited ruler. In this case, there is a minimal difference in revenue extraction from the new territory between the absolutist and the limited ruler, but the marginal benefit of reducing the probability of an exogenous threat is larger for the absolutist, entailing that the benefits of conquering are also greater. In other words, the driving force behind an absolutist ruler conquering new territory is the co-option of potential threats.

Proposition 5: $\exists m^*$ such that an absolutist ruler is (weakly) more likely than a limited ruler to conquer new territory if and only if $\hat{m} > m^*$, ceteris paribus.

One implication of this logic, which we will return to in Section IV, is that, in a time period of high monitoring costs, an absolutist is more likely than a limited ruler to attempt to conquer neighboring states—which are the most likely to pose a threat—whereas a limited ruler has more incentive to conquer resource-rich, but not necessarily neighboring, rivals. In the case of England or the Dutch Republic, conquering new territories was largely driven by the motivation for revenue or resource extraction for the sake of commercial interests such as the East India Company. For Imperial China, conquering new territories was more likely driven by reducing or eliminating external threats by absorbing external territories. This explains why Chinese imperial expansion often involved contiguous units which posed a potential external threat.

IV. History: Qing China and Beyond

In this section, we turn to historical narrative and empirical evidence to support our model. Our evidence centers on the case of Qing China, which displays the combined features of low and fixed official revenue (t in the model), high levels of off-book revenue taken by tax administrators ($t^o = T - t - w$ in the model), and encroachment on property rights by the central government in times of unpredictable crisis. We then extend our modelling insights to show how the Qing low-taxation equilibrium gained resilience through territorial expansion and how the initial advantages of large absolutist states turned into disadvantages compared with small European states with limited rule.

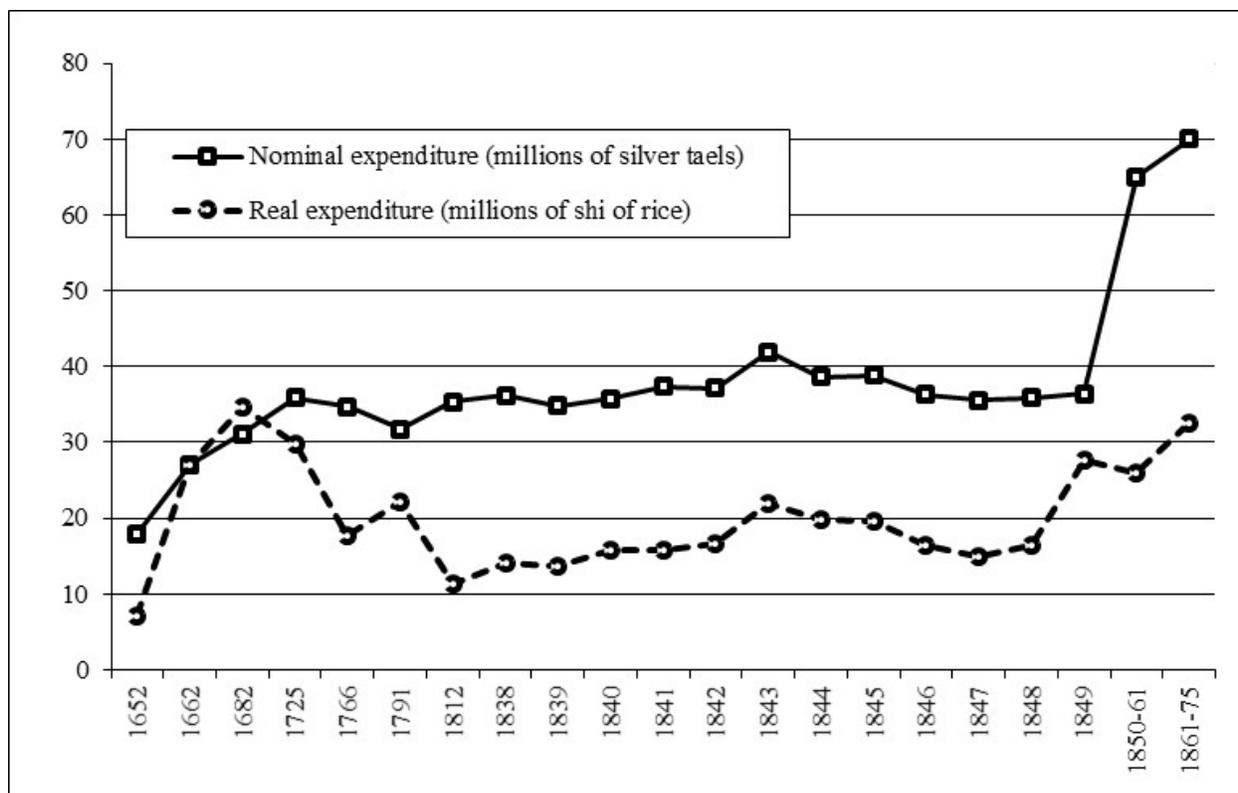
IV.1. Taxation in Qing China

The Qing fiscal regime reflected the nature of a highly centralized and hierarchical political system. Underpinning the regime was an elaborate accounting and reporting system cross-cutting the three layers of government administration at the central, provincial, and county level. In principle, the use and allocation of almost every budgetary item had to be reported and matched with detailed imperial rules and regulations. Although taxes were collected at the county level from the highly decentralized producing units across a giant empire, almost all revenues were in principle under the purview of the Board of Finance.²⁴ There was no officially recognized local-level finance, although the Qing did distinguish between remitted taxes (equivalent to t in the

²⁴ See Iwai (2004) and Shi and Xu (2008) for a description of Qing fiscal institutions. The Qing Imperial court had its own source of revenue and expenditure under the office of the so-called Nei-wu-fu (内务府). Overall, the share of Nei-wu-fu budget was small relative to that of the Board of Finance. For a recent in-depth study on Nei-wu-fu economic activities, see Lai (2014).

model) and retained ones (equivalent to the broader definition of w in the model) with the latter often recognized as the local cost of tax collection, which formed part of the de facto local administrative budget. Remitted taxes were either directly transferred to Beijing or other regions in China facing revenue deficits.²⁵

Figure 3: Government Expenditure in Qing China



Source: Ma (2014).

Figure 3 reconstructs scattered series of official expenditure under the direct purview of the Qing Imperial Board of Finance, which is roughly the revenue equivalent of $t + w$ in the model (i.e., statutory taxes plus all local retained revenue, including agents' wages). It shows a largely trendless expenditure series with an average of about 36 million silver taels but a standard

²⁵ By the 18th century, the land tax accounted for 70% of total Qing taxation with the remainder coming from commercial taxes. The revenue system was largely monetized with silver supplemented by an in-kind tax in the form of grain tribute amounting to a little more than a 20% share of the total fiscal revenue (Wang 1973, p. 80). On the expenditure side, the Qing spent about 50% of revenue on direct payments to soldiers and another 17% on the salaries of officials and bureaucrats. Expenditure on public goods such as maintenance of river transport or famine relief was only slightly above 10% (Shi Zhihong 2008, p. 68; Iwai 2004, p. 32).

deviation of only 3.2 million for the period between 1662 and 1849. The series began to rise from the mid-19th century when China was forced to open by Western Imperialism, but even then it remained almost stationary in real terms. Yet, these trendless series belie the fact that *per capita* revenues declined dramatically under the Qing.²⁶ During their two and one-half century reign, the population nearly tripled and the territory under their control nearly doubled.²⁷

Panel C of Table 2 shows that Qing official per capita fiscal revenue (roughly equal to t in the model) amounted to just over two days' earnings of an urban unskilled worker in the early-18th century, and dropped further by the late-18th century, reflecting the combined effect of a fixed revenue target accompanied by explosive population expansion. Table 2 suggests that absolutist regimes in general collected less per capita tax revenue than constrained regimes, although China's level of tax collection was especially low. The Qing collected anywhere from 1/3 to 2/3 of the per capita taxes of the Ottoman Empire from 1650-1800, while Ottoman tax collection efforts paled in comparison to their Western European rivals. Meanwhile, the constrained regimes in England and the Dutch Republic collected many times the per capita taxes of the Chinese. The disparity in fiscal capacity was so great that the total taxes collected by the Qing in the latter half of the 18th century were only about 3.5 times of those collected by the Dutch despite the Chinese population being about 143 times greater.²⁸

²⁶ The 18th century Qing emperors seemed to take pride in their relative frugality. K'ang-xi himself testified that "in our Dynasty, the total sum of military and civil expenses is about the same as that of the Ming period. But speaking of the Court expenses, the aggregate amount spent by the Court is even less than that for one palace of the Imperial Concubines. The accumulated sum of the past 36 years is less than that spent in one year's time during the Ming." (quoted in Ma 2011b).

²⁷ Despite this growth, Qing administrative units hardly expanded: there were only 1,360 counties under the Qing compared to 1,180 under the Han and 1,230 under the Song (Skinner 1977, p. 19). Similarly, the size of the 18th-century Qing standing army of about 800,000 was lower in absolute number than during the Ming and Song (Iwai 2004, p. 33).

²⁸ While reliable GDP estimates for China in the 18 and 19th centuries are unavailable, guestimates place the official fiscal revenue of 36 million taels in the range of a mere 1-2% of GNP even in the 1910s (Wang 1973, p. 133). Wang's result also seems broadly consistent with the daily wage conversion in Table 2. This again contrasts with Britain, whose total tax revenue rose 17-fold from 1665 to 1815, with its share in national income surging from 3% to 18% between 1688 and 1810 (O'Brien 1988, p. 3). Unlike Qing taxation system, the surge in British tax receipts came disproportionately from indirect taxes such as customs and particularly on excise duties, which together accounted for nearly 70% of total revenue towards the end of the 18th century (O'Brien 1988, p. 9-10; Dauntton 2012, p. 119).

Table 2: Qing Central Government Annual Revenue in International Comparison

Panel A. Aggregate Revenue (tons of silver)

	<u>Absolutist Regimes</u>					<u>Constrained Regimes</u>	
	China	Ottoman	Russia	France	Spain	England	Dutch Rep.
1650-99	940	248		851	243	239	
1700-49	1,304	294	155	932	312	632	310
1750-99	1,229	263	492	1,612	618	1,370	350
1800-49	1,367					6,156	
1850-99	2,651					10,941	

Panel B. International comparison of per capita tax revenue (grams of silver)

	<u>Absolutist Regimes</u>					<u>Constrained Regimes</u>	
	China	Ottoman	Russia	France	Spain	England	Dutch Rep.
1650-99	7.0	11.8		46.0	35.8	45.1	
1700-49	7.2	15.5	6.4	46.6	41.6	93.5	161.1
1750-99	4.2	12.9	21.0	66.4	63.1	158.4	170.7
1800-49	3.4					303.8	
1850-99	7.0					344.1	

Panel C. Per capita revenue expressed in days' wages for unskilled workers

	<u>Absolutist Regimes</u>					<u>Constrained Regimes</u>	
	China	Ottoman	Russia	France	Spain	England	Dutch Rep.
1650-99		1.7		8.0	7.7	4.2	13.6
1700-49	2.3	2.6	6.4	6.7	4.6	8.9	24.1
1750-99	1.3	2.0	8.3	11.4	10.0	12.6	22.8
1800-49	1.2					17.2	
1850-99	2.0					19.4	

Sources: Brandt et al. (2014); Dincecco (2009) for absolutist/constrained distinction

With a largely fixed level of annual revenue at a time of rapid population and territorial expansion and a near absence of any officially designated local finance, the Qing regime, particularly at the local level, could not survive without the so-called informal, unofficial revenue incurred beyond the official accounting system, equivalent to $t^O = T - t - w$ in the model. It is well-known the local officials were often punished and made personally liable for any revenue shortfall that may have occurred under their watch. As the official tax revenue allocated to local administration fell far short of the requirements of normal administration—often insufficient to cover the salaries of official bureaucrats let alone their expenses and support staffs—various levels

of bureaucrats relied on the infamous extralegal surcharges beyond the official level. The sources of these revenues ranged from the levying of surcharges, manipulation of weights, measures, and currency conversion in tax collection, falsifying reports, shifting funds across fiscal seasons, retaining commercial tax revenue, hoarding tax revenue from newly claimed land, and exacting contributions and donations from local farmers or merchants. Provincial level officials and their “unofficial” staffs relied on the extraction of gifts and contributions from the lower level officials and engaged in practices such as skimming funds in purchases and allocations (i.e., buying at a low price but reporting a high price). Reliance on informal local taxation and the employment of unofficial staffs for public administration often led to the privatization of public services.²⁹

A seminal study by Chang Chung-li on Chinese gentry income put non-official income (t^0) extracted above the reported level (i.e., excluding income earned through business or other activities) at 19 times of official income. The total unofficial income for officials above the provincial level were, according to Chang, 63 million taels—which was 81% of the total official tax quota around 1884 (Chang 1962, ch. 1). Hence, it was the informal, unregulated, and arbitrary nature of these extractions that help explain the apparent contradiction of a low official tax rate and the rapacious image of the Qing. The logic presented in the model helps resolve this contradiction. Because the center could not commit to refrain from confiscating the *known* wealth of its administrative agents (w), it could not pay them a high formal wage. It could also not ask its agents to remit a high level of taxes (t), because in order to incentivize the agents to remit any taxes whatsoever, it had to leave enough for the agents to collect “off book” (t^0). In the end, the masses faced tax levels that not only exceeded the stated official target but were also arbitrary in nature, even though the center only saw a fraction of those taxes.

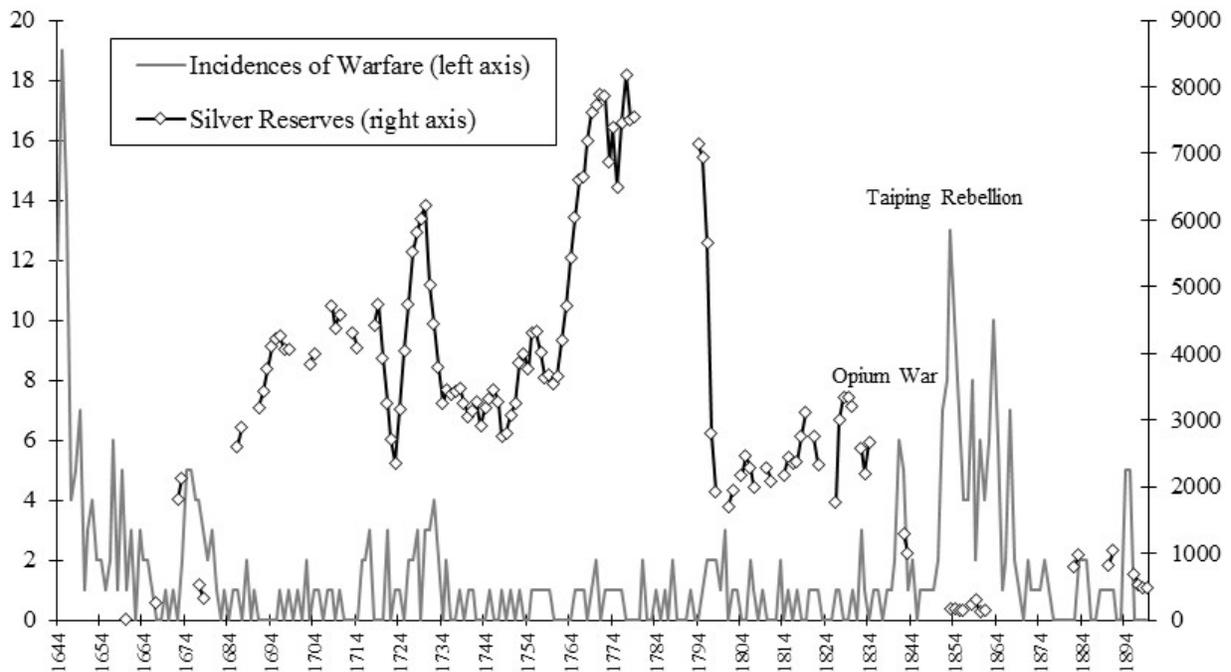
IV.2. Shocks, Conflicts, and Government Expenditure under the Qing

To examine how the combination of a largely fixed level of official revenue (t) alongside extralegal taxation above the official level (t^0) allowed the Qing to cope with recurrent episodes of political instability and exogenous shocks, we would need a relatively continuous data series that reveals data on both ordinary and extraordinary expenditure in times of relative peace and

²⁹ See Zelin (1984, p. 46-71). Sometimes staffs kept duplicate set of account books, with the set for local use marked by secret codes impenetrable from the official examination. These special types of account books circulated informally within a wide area (Zelin 1984, p. 240). Official collusion could backfire in unexpected ways. Often, the extralegal nature of these surcharges forced the parties involved to pay blackmail (Iwai 2004, p. 3-4).

instability. In the absence of such data, we can make use of an alternative data series constructed by Shi Zhihong (2008) on the annual warehouse (银库) receipts from the Board of Finance, which recorded the actual inflows (usually consisting of tax remittances from the provinces) and outflows (governmental payments for various expenditures from the warehouse). In the absence of organized public debt, the cumulative stock of silver reserves at the Board of Finance warehouse was equivalent to the cumulative stock of government savings. Figure 4 plots the annual stock of silver reserves against episodes of warfare. This conveys a fuller and more telling portrayal of the relationship between Qing fiscal policy and political stability.

Figure 4: Annual Recorded Incidences of Warfare (left axis) and Silver Reserves (in ten thousand taels; right axis) in Qing China (1644-1911)



Source: Ma (2014).

In its early years of military conquest in the 1660s, Qing silver reserves were minimal. As a non-Han minority ruler of China, the early Qing relied on Chinese generals and military force to suppress former Ming loyalists, which led to the build-up of relatively autonomous power bases in Southern China. This ended in 1683 when Emperor Kangxi (1661-1722) quashed the rebellion of the so-called “three feudatories” and annexed their territories into the Qing centralized

administration. Two years later, Kangxi broke the resistance of the rebellious naval kingdom of Zheng Chenggong and officially integrated the island of Taiwan into the Qing administration. In the final decades of the 17th century, the Qing contained the threat from an expansionary Russia by signing the Treaty of Nerchinsk in 1689 and conquered China's North-western territory in 1696. In 1720, the Qing attained control of Tibet with the installation of a new Dalai Lama. By the early 18th century, the Qing successfully consolidated power and established monopoly rule over China's largest ever territory.³⁰ It was in this context that we understand the Kangxi's emperor's open declaration in 1712 that there would not be any new additional taxes on newly added population and land. This commitment was supported by the deliberate abandonment of any new land or population surveys, and a conscious non-investment in monitoring infrastructure (Wang 1973).

As seen in Figure 4, the Qing gradually built up their reserves and entered into a prolonged phase of silver reserve accumulation during the 18th century, which was a period of political stability and infrequent warfare. Reserves peaked at over 70 million by the 1790s. Yet, despite a century of relative peace and little expenditures, reserves peaked at only (roughly) two years of total tax revenue, reflecting the limited fiscal capacity of the Qing regime. Indeed, the suppression of the White Lotus rebellion around the turn of the eighteenth century, towards the end of the Qianlong rule, led to a sharp drop in silver reserves. This heralded a turning point where the value of silver persistently increased as Chinese silver outflows regularly outpaced inflows. The Qing enacted desperate measures to replenish their dwindling silver stocks: the sale of government offices and titles increased sharply in 1804, 1827 and 1834, reaching over 10 million taels—one-third to one-half of the annual central government revenue (Ma 2014).

The silver reserve data reported in Figure 4 are partially corroborated by Chen Feng's (1992, p. 275) meticulous yet incomplete calculation of the extraordinary expenses the Qing incurred in the face of various shocks. It shows a highly uneven inter-temporal pattern, ranging from several tens of thousands of taels in the mid-18th century to a peak of nearly 150 million taels for suppressing the White Lotus Rebellion of 1796-1804. Given that the Qing Board of Finance, even at its peak, had 70 million taels in its coffers—equivalent to no more than 3 or 5% of GDP³¹—the Qing had little room to maneuver within the normal fiscal framework to weather these shocks. As

³⁰ See Jonathan Spence (1990) for the standard narrative.

³¹ These figures are from Wang Yeh-chien's (1973) calculation.

a result, the Qing resorted to numerous sources of extraordinary revenue raised to cover military expenditure, including on-site confiscation and predation, advanced collection of land taxes, temporary but arbitrary surcharges on existing categories of taxation, an increased share of remitted revenue at the expense of retained revenue for the local government, deductions of formal salaries, forced contributions from wealth holders, and the sale of government offices and titles (Chen 1992, ch. 7). In the devastating mid-19th century Taiping Rebellion, the desperate Qing eventually succumbed to monetary debasement (Chen 2008, ch. 11).

IV.3. Long Live the Emperor

Despite challenges and periodic deteriorations of public finance and private property rights, Chinese imperial regimes survived, and often even thrived, until the 20th century. We argue that Qing absolutism combined with a weak monitoring infrastructure (high m in our model) represented an institutional balance that may have helped sustain dynastic longevity through a combination of incentives and discipline. Indeed, by leaving their agents residual taxation (i.e., t^o , or off-book income), the Qing managed to better align the short-term incentives of their agents with their long-term objective of dynastic tenure and helped rein in the worst form of decentralized corruption often witnessed in more fragmented authoritarian regimes (Bardhan 1997). By allowing information asymmetry between agents and tax-payers—such as Kangxi’s openly declared abandonment of regular population and land surveys—the Qing may have provided some form of de facto property rights that allowed economic growth and population expansion under the most stringent form of absolutism. Chinese imperial rulers’ eventual acquiescence and accommodation of local corruption and extra-legal taxation was in fact a rational compromise, as suggested by the model, to reconcile the inherent contradiction between the discretionary power of the state and the ideological commitment to a fixed revenue target (in terms of t , or on-the-book, statutory tax) as the foundation of Qing legitimacy. In this regard, the 18th century European enlightenment thinkers were not completely delusional.

For the system to work, the discretionary power of the absolutist may also be critical. The Qing emperors’ undisputed power to reward and punish agents was indispensable to the survival of the regime, especially in times of crisis. A legal system controlled within the administrative hierarchy may not deliver the best justice, but it may be far more effective in disciplining agents and pacifying the populace by targeting, often selectively, what were seen as excessive bureaucratic

abuses that could imperil imperial stability and dynastic survival.³² Moreover, the power to confiscate gave the Qing the capacity to rapidly mobilize resources and combat short-run shocks. As shown in our model, even though limited rulers can generate more *on-book* revenue (t) than absolutists during periods of exogenous threat, the fact that an absolutist can resort to discretionary confiscations during periods of particularly acute threats means that the *total* amount of revenue available to the ruler—the on-book taxes (t) plus the one time confiscation ($w + Mt^0$)—may be significantly greater.³³ Hence, in the absence of an alternative mechanism such as public debt, absolutists may have a better expenditure smoothing mechanism than limited rulers, at least in the short run.

Unconstrained coercive power also allowed the absolutist Qing to exert political control, resulting in limits on migration and the use of ideological propaganda.³⁴ These tools eventually yielded the Qing a comparative advantage in suppressing internal rebellion within its own jurisdiction relative to limited rulers. In this regard, the Qing's successful territorial expansion, particularly during the 18th century, aimed much more at the reduction of external threats and interstate competition rather than resource or revenue extraction, as our model predicts (Proposition 5).

The case of Taiwan further exemplifies this insight. As an important strategic trading ground that lured Portuguese, Dutch, and Japanese merchants throughout the early modern era, it attracted little political attention until it became a power base for anti-Qing Ming royalists led by Zheng Chenggong in the 17th century. To defeat Zheng, Kangxi instigated a trade blockage by forcibly removing residents from China's southeast coast, leading to a major economic depression on China's southeast coast. After the surrender of Zheng, the Qing incorporated Taiwan into China's administrative system with direct rule and closed it off for trade. The logic is clear: political stability trumped economy or trade.

³² See He Ping (1998, p. 293-5) for the periodic and selective capital punishment on the so-called “economic crime meted out to high level government officials.” Huang (1974, p. 13-14) counted in detail the sorry fate of all the 89 ministers of Revenue under the Ming from 1380.

³³ In the context of the model, if the exogenous threat occurs with varying severity, i.e., requiring a varying amount of resources to combat, then the absolutist ruler may not be able to commit to refrain from confiscations when the threat is sufficiently stark. While our model does not address this question directly, it is easy to see how the addition of minimal complexity provides insight. By “minimal complexity” we mean complexity to the model's setup, not its solution. What we are proposing here adds significant complexity to the model's solution without altering its basic insights.

³⁴ For a case of an absolutist state with massive monitoring capacity and coercive power in the early Ming dynasty during the 14th century, see Guanglin Liu (2005). Also see Xue and Koyama (2017) on intellectual persecution, and Lai (2014) on the Qing's use of Tibetan Buddhism to control Central Asia.

These insights help explain the apparent contradiction of the Qing's vast territorial expansion and increasingly inward looking mindset, while at the same time reinforcing the "Confucian equilibrium" of low tax extraction and weak administrative capacity. At a time when the Qing vastly expanded its territories into China's northwestern frontiers, which had historically been the source of military conflicts, it restricted its overseas trade to the port of Canton under a monopoly trading cartel and banned Chinese migration to Southeast Asia, which traditionally posed no military threat (Brandt, Ma, and Rawski 2014). When visited in 1792 by the British trade mission led by Lord George Macartney, who aimed to show off the best of Western trade and technology, the Qianlong emperor famously replied that "Our Celestial Empire possesses all things in prolific abundance and lacks no product within its borders. There was therefore no need to import the manufactures of outside barbarians in exchange for our own produce."³⁵

Hence, absolutism carried significant costs: rents and privileges needed to keep the agents in the system, corruption, market distortions, short-term extractions, periodic trampling of property rights, and possibly the underdevelopment of formal contract and information-intensive institutions and markets. But an absolutist regime endowed with improved monitoring infrastructure (low m in our model) may result in even *worse* outcomes, particularly for the masses, agents, and even the ruler. Given the highly varied personalities of the rulers throughout Chinese history, the combination of massive monitoring infrastructure with the unconstrained coercive power of the state could turn this infrastructure into a most lethal weapon of political repression and fiscal extraction, even in the absence of exogenous threats.³⁶ Hence, weak monitoring infrastructure was an effective commitment device against the emergence of capacious rulers, thereby ensuring a softer version of absolutism far more "benevolent" and hence resilient than the more ominous totalitarian version.³⁷

³⁵ Quote from <http://www.history.ucsb.edu/faculty/marcuse/classes/2c/texts/1792QianlongLetterGeorgeIII.htm>, accessed August 30, 2017.

³⁶ For another case of a command economy built under Mao Zedong in the 1950s, which partly accounted for the 1959-61 Great Leap Famine, see Chen and Kung (2011).

³⁷ If an absolutist state strengthens its monitoring infrastructure, it creates perverse incentives for dynamism and innovation. For a model of how slavery was usually most effective in an economy with simple and easily measurable tasks, see Dari-Mattiacci (2013). Indeed, Chinese reform in the post-1978 era was successful in part because the government retreated from a strict monitoring of many sectors of the economy (Bai et al. 1999).

IV.4. Comparing Chinese Absolutism to European Limited Governance

We now apply our modelling insights to the historical prototypes of fiscal regimes in Western Europe. With credible commitment, a limited ruler commits to not arbitrarily raise the statutory tax quota without prior consent from the agent. Even with high monitoring costs, the agent may not need to hide the amount of taxes he collects, as he is secure from extraction from the ruler.

We view the deal between medieval/early modern European rulers and the autonomous, free cities as a case of limited ruler with weak monitoring capacity (high m) at the center but with credible commitment. In exchange for a fixed sum of tax revenue (t), European rulers granted cities a secure charter which guaranteed their rights. Since autonomous cities could collect taxes (formal and informal) within their jurisdiction without fear of extraction from the ruler, their accounting and revenue systems became increasingly open. In other words, it may well be possible that autonomous cities could develop an efficient taxation system within the urban administration even though there was poor monitoring from the Imperial center. The fiscal capacity thus developed helped trigger a financial transformation of public debt, particularly on the secondary market, which became more viable in Western Europe because it could be securitized with a relatively stable and transparent revenue stream. Thus it comes as no surprise that free cities were often the vanguard of Western financial innovation and market.³⁸

In a similar vein, post-Glorious Revolution England was a prototypical case of a limited ruler endowed with strong fiscal monitoring infrastructure (low m). Even though the Crown knew both statutory taxes (t) as well as off-book taxes (t^o) once it built its strong monitoring capacity, the Crown's credible commitment not to raise arbitrary taxes or confiscate ex post provided its agents a secure and large present value of future income, all of which were visible. Hence, the value to the agent of staying within the bureaucracy and being honest remained high. Indeed, a modern civil bureaucracy in the form of an excise bureau, where bureaucrats are paid high, transparent wages in order to deter short-term dishonest behavior, came into shape in post-Glorious Revolution Britain. Here, modern bureaucrats valued long-term tenure over entrepreneurship (Daunton 2012; North and Weingast 1989). Similarly, the Crown met one-time revenue demands in times of crisis through public debt—a sustainable public debt required credible commitment—or through parliamentary consent (Johnson and Koyama 2014). In this regard, the credible commitment

³⁸ For much more on the joint growth of sovereign debt and limited rule in medieval and early modern Europe, see Stasavage (2011, 2016).

mechanism reinforced through the rule of law offered the political regime a way out of the cycles of confiscations and extended the time horizon of political rule beyond the tenure of individual rulers or a coalition of rulers. Of course, this only worked where rulers were credibly constrained. The Habsburg Spanish Empire provides a useful counterexample. King Philip II (r. 1556-98) was notorious for defaulting on his debt; it was only because his Genoese lenders could coordinate to cut off loans to him that he was able to access loans in the first place (Drelichman and Voth 2011). Under such absolutist conditions, it is not surprising that the Spanish found it difficult to create a sustainable public debt.

Turning back to China, many of the advantages of absolutist rule turned out to be impediments in the long run. The dominance of a single and centralized absolutist state ended up as a serious obstacle to the growth of a viable market for public debt. While absolutist regimes in Europe such as the Habsburgs faced difficulties in the development of a domestic market for public debt given the absence of credible commitment, they could still float public debt abroad or in international markets within a multi-state framework. Such an option was not viable for the Qing, given that the entire territory was under the reign of a single centralized state. Off-book income, which made up a majority of the wage of Qing tax collecting agents, could not be used as collateral for government bonds. As Max Weber hinted at one point, the Chinese may not have achieved the same type of fiscal and financial capacity as their European counterparts partly due to the combined absence of internal constitutional constraint and external political fragmentation (Weber 1951, p. 103-4). As long as the state could resort to coercive means of extracting resources without facing external competition, it was less inclined to develop state capacity based on market forces such as long-term debt.³⁹ Ironically, by granting certain grants to agents and the masses, the limited ruler ended up developing in the long run a far more powerful expenditure smoothing mechanism based on public debt. When the Qianlong emperor gloated of the Qing savings of 70 million silver taels—a few percentage of its total GDP—the English cumulative debt had reach nearly two times of its GDP, a truly great divergence in fiscal and financial capacity between the two types of states.

³⁹ See Ma (2016) on the rise of public debt in early 20th century China under a different institutional set-up, which offers some insights into why public debt and the related secondary market was hard to come by in early modern China.

V. Concluding Thoughts

Motivated by the case of the absolutist Qing Empire, our paper seeks to explain why fiscal capacity (the ability to collect taxes from the masses) and administrative capacity (the ability to monitor and punish administrators tasked with carrying out the ruler's policies) are frequently weak in absolutist regimes, despite the ruler's absolute power over people and property. We suggest that the absolutist's unconstrained power and inability to refrain from confiscation could turn out to be a weakness in the long run compared with constitutionally constrained regimes. The lack of credible commitment means that absolutists are only able to encourage their administrators to collect and remit taxes when the ruler lacks the ability to inexpensively monitor and punish his agents. Under such circumstances, agents value "staying in the system" because they have a steady stream of income via the collection of extra-legal taxation, which is secure from the ruler because the ruler does not know it exists. This logic entails that an absolutist's failing to build up administrative infrastructure is an equilibrium *outcome* in the situation where the ruler faces few constitutional constraints. Moreover, it entails that extra-legal taxation—so common in the Qing Empire and other absolutist states—is a *necessary* component of an equilibrium where an agent collects and remits taxes to the center.

This paper also brings to the fore the issue of fiscal, financial, and legal capacities in the long-run economic divergence between China and Northwestern Europe. In particular, it suggests that once the rulers of England and the Dutch Republic were subject to sufficient constitutional constraint—following the Dutch Revolt in the Dutch Republic and the Glorious Revolution in England—they were incentivized to build administrative infrastructure, develop public debt markets, and increase the level of taxation (see Table 2). On the other hand, absolutist regimes—both Qing China as well as the absolutist regimes of Europe including France, Spain, and the Ottoman Empire—often sought other means to consolidate their power rather than build such infrastructure and markets based on some form of credible commitment mechanism.⁴⁰ It is no surprise, therefore, that the Dutch Republic and England—both of whose territories and population size were minuscule compared with the Chinese, Russian, or Ottoman Empires—became the

⁴⁰ For the case of Spain, see Irigoien and Grafe (2013). Like the Qing, the early modern Spanish Empire had little administrative capacity. The Spanish Crown was nonetheless able to extract significant revenue, mainly from the New World, by outsourcing fiscal functions to private individuals. Irigoien and Grafe argue that Spanish coercive power helps explain this observation. Their model suggests that fiscal capacity is a function of coercive power and it follows an inverted-U shape; initial investments in coercion pay off well, but eventually diminishing returns kick in.

world's two leading economies of the early modern period and for better or worse, constructed the world's most dynamic modern global empires.⁴¹

In this regard, our paper sheds new light on the Great Divergence debate. A critical point put forward by Pomeranz is the importance of New World resources in the rise of the West versus China (Pomeranz 2000, ch. 6). We argue that China's lack of overseas explorations is best understood by a lack of motivation rather than lack of access: the main motivation of Qing territorial expansion was the reduction of external threats rather than the control or extraction of economic resources. It was not just simply constitutional constraint or the government's credible commitment to repay public debt that mattered. Instead, limited governance directly contributed to the build-up of fiscal, administrative, and financial institutions, all of which are hallmarks of the modern economy.

References

- Acemoglu, Daron. 2005. "Politics and Economics in Weak and Strong States." *Journal of Monetary Economics* 52(7): 1199-1226.
- Adshead, S.A.M. 1970. *The Modernization of the Chinese Salt Administration, 1900-1920*. Cambridge: Harvard University Press.
- Alesina, Alberto and Enrico Spolaore. 1997. "On the Number and Size of Nations." *Quarterly Journal of Economics* 112(4): 1027-56.
- Alesina, Alberto and Enrico Spolaore. 2005. "War, Peace, and the Size of Countries." *Journal of Public Economics*, 89(7): 1333-54.
- Bai, Chong-En, David D. Li, Yingyi Qian, and Yijiang Wang. 1999. "Anonymous Banking and Financial Repression: How Does China's Reform Limit Government Predation without Reducing Its Revenue?" Working paper.
- Balla, Eliana, and Noel D. Johnson. 2009. "Fiscal Crisis and Institutional Change in the Ottoman Empire and France." *Journal of Economic History* 69(3): 809-45.
- Bardhan, Pranab. 1997. "Corruption and Development: a Review of Issues." *Journal of Economic Literature* 35(3): 1320-46.
- Besley, Timothy and John McLaren. 1993. "Taxes and Bribery: The Role of Wage Incentives." *Economic Journal* 103(416): 119-41.

⁴¹ For more on the rise of the Dutch in the early modern period, see de Vries and van der Woude (1997). van Zanden, Buringh, and Bosker (2012) recognize the Dutch and English "little divergence" as associated with the divergence in parliamentary development between northwestern Europe and southern Europe between 1500 and 1800. Rubin (2017) adds to this insight, arguing that the little divergence in parliamentary development can be explained by the Dutch and English adopting the Reformation, which necessitated that rulers legitimate their rule and capacity to collect taxes via parliaments.

- Besley, Timothy and Persson, Torsten. 2009. "The Origins of State Capacity: Property Rights, Taxation and Politics." *American Economic Review* 99(4): 1218-44.
- Besley, Timothy and Persson, Torsten. 2010. "State Capacity, Conflict and Development." *Econometrica* 78(1): 1-34.
- Besley, Timothy and Persson, Torsten. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton: Princeton University Press.
- Besley, Timothy and Persson, Torsten. 2014. "Why Do Developing Countries Tax So Little?" *Journal of Economic Perspectives* 28(4): 99-120.
- Brandt, Loren, Debin Ma, and Thomas G. Rawski. 2014. "From Divergence to Convergence: Reevaluating the History behind China's Economic Boom." *Journal of Economic Literature* 52(1): 45-123.
- Chang, Chung-li [Zhang Zhongli]. 1955. *The Chinese Gentry Studies on Their Role in Nineteenth-Century Chinese Society*. Seattle: University of Washington Press.
- Chang, Chung-li [Zhang Zhongli]. 1962. *The Income of the Chinese Gentry*. Seattle: University of Washington Press.
- Chen, Feng. 1992. 清代军费研究 [A Study on Qing Military Expenditure]. Wuhan: Wuhan University Press.
- Chen, Feng. 2008. 清代财政政策与货币政策研究 [A Study on Qing Fiscal and Monetary Policies]. Wuhan: Wuhan University Press.
- Chen, Suo and James Kung. 2011. "The Tragedy of the Nomenclature: Career Incentives and Political Radicalism during China's Great Leap Famine." *American Political Science Review* 105(1): 27-45.
- Dari-Mattiacci, Giuseppe. 2013. "Slavery and Information." *Journal of Economic History* 73(1): 79-116.
- Daunton, Martin. 2012. "The Politics of British Taxation, from the Glorious Revolution to the Great War." In *The Rise of Fiscal States*, Bartolomé Yun-Casalilla, Patrick K. O'Brien, and Francisco Comin Comin (Eds.): 111-42.
- De Bary, W.T., 1993, *Waiting for the Dawn: A Plan for the Prince, Huang Tsung-his's Ming-i-tai-fang lu*. New York: Columbia University Press.
- de Vries, Jan and Ad van der Woude. 1997. *The First Modern Economy: Success, Failure, and Perseverance of the Dutch Economy, 1500-1815*. Cambridge, Cambridge University Press.
- Diamond, Jared. 1997. *Guns, Germs, and Steel*. New York: W.W. Norton & Company.
- Dincecco, Mark. 2009. "Fiscal Centralization, Limited Government, and Public Revenues in Europe, 1650-1913." *Journal of Economic History* 69(1): 48-103.
- Dincecco, Mark. 2014. "The Rise of Effective States in Europe." *Journal of Economic History* 75(3): 901-918.
- Drelichman, Mauricio, and Hans-Joachim Voth. 2011. "Lending to the Borrower from Hell: Debt and Default in the Age of Philip II." *Economic Journal* 121(557): 1205-27.
- Friedman, David. 1977. "A Theory of the Size and Shape of Nations." *Journal of Political Economy* 85(1): 59-77.
- Fukuyama, Francis, 2011 *The Origins of Political Order*. New York: Farrar, Straus and Giroux.
- Gennaioli, Nicola and Hans-Joachim Voth. 2015. "State Capacity and Military Conflict." *Review of Economic Studies* 82(4): 1409-48.

- González de Lara, Yadira, Avner Greif, and Saumitra Jha. 2008. "The Administrative Foundations of Self-Enforcing Constitutions." *American Economic Review* 98(2): 105-09.
- Greif, Avner. 2005. "Commitment, Coercion, and Markets: The Nature and Dynamics of Institutions Supporting Exchange." in *The Handbook of New Institutional Economics*, C. Ménard and M.M. Shirley (eds.). Norwell, MA: Kluwer Academic Publishers: 727-86.
- Greif, Avner. 2008. "The Impact of Administrative Power on Political and Economic Developments: Toward a Political Economy of Implementation." in *Institutions and Economic Performance*, E. Helpman (Ed.). Cambridge: Harvard University Press, 17-63.
- Greif, Avner and Jared Rubin. 2017. "Endogenous Political Legitimacy: The English Reformation and the Institutional Foundations of Limited Government." Stanford University Working Paper.
- Greif, Avner and Guido Tabellini. 2017. "The Clan and the Corporation: Sustaining Cooperation in China and Europe." *Journal of Comparative Economics* 45(1): 1-35.
- Hamashita, Takeshi. 2006. *Studies in China's Modern Economic History: Maritime Customs Finance and Treaty-Port Market Networks during Late Qing* (In Chinese). Nanjing: Jiangsu renmin chubanshe.
- Hao, Yu and Zhengcheng Liu. 2016. "Rationalizing Fiscal Reform, Grabbing Hand and State Capacity in Pre-Industrial China: Evidence from Fiscal Reform in Eighteenth-Century China." Unpublished paper.
- He, Ping. 1998. *清代赋税政策研究 (Research on Qing Taxation Policy)*. Beijing: Chinese Academy of Social Science Press.
- Hoffman, Philip T. 2015. *Why Did Europe Conquer the World?* Princeton: Princeton University Press.
- Huang, Ray. 1974. *Taxation and Government Finance in Sixteenth-Century Ming China*. Cambridge: Cambridge University Press.
- Irigoin, Alejandra, and Regina Grafe. 2013. "Bounded Leviathan: Fiscal Constraints and Financial Development in the Early Modern Hispanic World." In D'Maris Coffman, Adrian Leonard, and Larry Neal (eds.), *Questioning Credible Commitment: Perspectives on the Rise of Financial Capitalism*. Cambridge: Cambridge University Press.
- Iwai, Shigeki. 2004. *A Study of China's Early Modern Fiscal History* (In Japanese). Kyoto: Kyoto University Press.
- Johnson, Noel and Mark Koyama. 2014. "Tax Farming and the Origins of State Capacity in England and France." *Explorations in Economic History* 51: 1-20
- Johnson, Noel D. and Mark Koyama. 2017. "States and Economic Growth: Capacity and Constraints." *Explorations in Economic History* 64: 1-20.
- Karaman, Kamil Kivanç. 2009. "Decentralized Coercion and Self-Restraint in Provincial Taxation: The Ottoman Empire, 15th-16th Centuries." *Journal of Economic Behavior and Organization* 71(3): 690-703.
- Karaman, Kivanç and Şevket Pamuk. 2013. "Different Paths to the Modern State in Europe: The Interaction between Warfare, Economic Structure, and Political Regime." *American Political Science Review* 107(3): 603-26.
- Kaske, Elisabeth. 2016. "Austerity in Times of War: the End of Eighteenth-Century War Finance in Qing China." Working paper.

- Kishimoto, Mio. 2011. "Property Rights, Land, and Law in Imperial China." *Law and Long-Term Economic Change: A Eurasian Perspective*. Stanford: Stanford University Press.
- Ko, Chiu Yu; Mark Koyama and Tuan-Hwee Sng. 2017. "Unified China and Divided Europe." *International Economic Review*: Forthcoming.
- Lagerlöf, Nils-Petter. 2014. "Population, Technology and Fragmentation: The European Miracle Revisited." *Journal of Development Economics* 108: 87-105.
- Lai, Hui-min. 2014 乾隆皇帝的荷包 (Emperor Qianlong's Purse: Imperial Household Finance in Eighteenth Century China) Taipei: Institute of Modern History, Academia Sinica.
- Levine, David K., and Salvatore Modica. 2013. "Conflict, Evolution, Hegemony, and the Power of the State." NBER Working Paper 19221.
- Liu, Guanglin. 2005. "Wrestling for Power: The State and the Economy in Later Imperial China, 1000—1770." Ph.D. Thesis, Harvard University.
- Ma, Debin. 2011a. "Law and Economy in Traditional China: A 'Legal Origin' Perspective on the Great Divergence," in *Law and Long-Term Economic Change: A Eurasian Perspective*. Stanford: Stanford University Press.
- Ma, Debin. 2011b. "Rock, Scissors, Paper: The Problem of Incentives and Information in Traditional Chinese State and the Origin of Great Divergence." London School of Economics Economic History Working Paper 152.
- Ma, Debin. 2012. "Political Institution and Long-run Economic Trajectory: Some Lessons from Two Millennia of Chinese Civilization." chapter 4 in M. Aoki, T. Kuran and G. Roland (eds.), *Institutions and Comparative Economic Development*, Basingstoke: Palgrave Macmillan: 78-98.
- Ma, Debin. 2014. "State capacity and great divergence, the case of Qing China (1644–1911)." *Eurasian Geography and Economics* 54(5-6): 484-99.
- Ma, Debin 2016 "The Rise of a Financial Revolution in Republican China in 1900-1937: an Institutional Narrative" LSE Economic History Working Paper series No. 235: <http://www.lse.ac.uk/economicHistory/workingPapers/2016/WP235.pdf>
- Mayshar, Joram, Omer Moav, and Zvika Neeman. 2013. "Geography, Transparency and Institutions." CEPR Discussion Paper 9625.
- Mokyr, Joel. 2016. *A Culture of Growth: The Origins of the Modern Economy*. Princeton: Princeton University Press.
- Myerson, Roger B. 2008. "The Autocrat's Credibility Problem and Foundations of the Constitutional State." *American Political Science Review* 102(1): 125-39.
- Myerson, Roger B. 2015. "Moral Hazard in High Office and the Dynamics of Aristocracy." *Econometrica* 83(6): 2083-2126.
- North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge: Cambridge University Press.
- North, Douglass C. and Barry R. Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49(4): 803-32.
- O'Brien, Patrick K. 1988. "The Political Economy of British Taxation, 1660-1815." *Economic History Review* 41(1): 1-32.

- Olson, Mancur. 1993. "Dictatorship, Democracy, and Development." *American Political Science Review* 87(3): 567-76.
- Pomeranz, Kenneth L. 2000. *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton: Princeton University Press.
- Qin, Hui. 2000. "Tax Consolidation Reform and Huang Zongxi Rule" Beijing: China Economic Daily 《中国经济时报》 2000, Nov. 3rd.
- Rosenthal, Jean-Laurent and R. Bin Wong. 2011. *Before and Beyond Divergence: The Politics of Economic Change in China and Europe*. Cambridge: Harvard University Press.
- Rubin, Jared. 2017. *Rulers, Religion, and Riches: Why the West got Rich and the Middle East Did Not*. Cambridge: Cambridge University Press.
- Shi, Zhihong. 2008. (史志宏) 清代戶部銀庫收支和庫存統計 (*Statistics on the Stocks, Receipts and Expenditures of the Vaults of the Qing Board of Revenue*) 福建人民出版社.
- Shi, Zhihong and Xu, Yi. 2008. *晚清財政: 1851-1894 (Public Finance in Late Qing: 1851-1894)*. Shanghai: Shanghai University of Finance and Economics.
- Skinner, G. William. 1977. "Introduction: Urban Development in Imperial China." In *The City in Late Imperial China*, edited by G. William Skinner, 3–32. Stanford: Stanford University Press.
- Sng, Tuan-Hwee. 2014. "Size and Dynastic Decline: The Principal-Agent Problem in Late Imperial China, 1700-1850." *Explorations in Economic History* 54: 107-27.
- Sng, Tuan-Hwee, and Chiaki Moriguchi. 2014. "Asia's Little Divergence: State Capacity in China and Japan before 1850." *Journal of Economic Growth* 19(4): 439-70.
- Spence, Jonathan D. 1990. *The Search for Modern China*. New York: Norton.
- Stasavage, David. 2011. *States of Credit: Size, Power, and the Development of European Polities*. Princeton: Princeton University Press.
- Stasavage, David. 2016. "What We Can Learn from the Early History of Sovereign Debt?" *Explorations in Economic History* 59: 1-16.
- Stephens, Thomas B. 1992. *Order and Discipline in China: The Shanghai Mixed Court, 1911–27*. Seattle: University of Washington Press.
- Te-Ch'ang, Chang. 1972. "The Economic Role of the Imperial Household in the Ch'ing Dynasty." *Journal of Asian Studies* 31(2): 243-73.
- Tilly, Charles. 1990. *Coercion, Capital and European States, A.D. 990-1990*. Cambridge, MA: Blackwell.
- Turchin, Peter. 2009. "A Theory for Formation of Large Empires." *Journal of Global History* 4(2): 191-217.
- van Zanden, Jan Luiten, Eltjo Buringh, and Maarten Bosker. 2012. "The Rise and Decline of European Parliaments, 1188-1789." *Economic History Review* 65(3): 835-861.
- van de Ven, Hans 2014. *Breaking with the Past: The Maritime Customs Service and the Global Origins of Modernity in China*. New York: Columbia University Press.
- Xue, Melanie Meng, and Mark Koyama. 2017. "Autocratic Rule and Social Capital: Evidence from Imperial China." Working Paper.
- Wang, Yeh-chien. 1973. *Land Taxation in Imperial China, 1750–1911*. Cambridge and London: Harvard University Press.
- Weber, Max. 1951. *The Religion of China: Confucianism and Taoism*. New York: Free Press.

- Zelin, Madeleine. 1984. *The Magistrate's Tael: Rationalizing Fiscal Reforms in Eighteenth-Century Ch'ing China*. Berkeley: University of California Press.
- Zhou, Haiwen. 2012. "Internal Rebellions and External Threats: A Model of Government Organizational Forms in Ancient China." *Southern Economic Journal* 78(4): 1120-1141.

Appendix A: Proofs of Propositions

A.1. Proof of Proposition 1

We first solve for the equilibrium. The third stage action of the ruler is given by inequality (4); i.e., the ruler chooses $F = 1$ iff (4) is satisfied. The second stage action of the agent is determined by the type of ruler (absolutist or limited) and the ruler's choices in the first stage. We note in the body of the paper that there are four cases.

In case i ($M = 0$, ruler is absolutist), the agent will be fired if $w > f$, and he therefore opts out and chooses $T = 0$. However, if $w \leq f$, the agent will not be fired and therefore collects taxes up to the point that his marginal benefit from doing so equals the marginal cost (i.e., $u'(T - t) = c'(T)$), conditional on $u(T - t) - c(T) \geq \underline{u}$. In case ii ($M = 1$, ruler is absolutist), the agent will be fired if $T > t + f$. The agent therefore opts out and chooses $T = 0$ if $u(f) - c(t + f) < \underline{u}$. Otherwise, the agent either chooses $T = t + f$ or T such that $u'(T^* - t) = c'(T^*)$, where $T^* < t + f$.

In case iii ($M = 0$, ruler is limited), the agent will never be fired. He therefore collects taxes up to the point that his marginal benefit from doing so equals the marginal cost (i.e., $u'(T - t) = c'(T)$), conditional on $u(T - t) - c(T) \geq \underline{u}$. In case iv ($M = 1$, ruler is limited), the agent will be fired if $T > t + w + f$. The agent therefore opts out and chooses $T = 0$ if $u(w + f) - c(t + w + f) < \underline{u}$. Otherwise, the agent either chooses $T = t + w + f$ or T such that $u'(T^* - t) = c'(T^*)$, where $T^* < t + w + f$.

Next, consider the ruler's first stage decision of t , w , and M . Start with the absolutist's decision. If the absolutist does not monitor ($M = 0$), he will be indifferent between choosing any $w \in [0, f]$. The statutory tax, t , he chooses must be low enough that there exists some T for which $u(T - t) - c(T) \geq \underline{u}$. In equilibrium, if the parameters are such that the agent's participation constraint holds, the absolutist's revenue is $\pi^R = t$. The absolutist therefore chooses t^A to maximize:

$$(A.1) \quad \theta q(t^A, T(t^A), 1) + (1 - \theta)q(t^A, T(t^A), 0),$$

subject to $u(T(t^A) - t^A) - c(T(t^A)) \geq \underline{u}$, where $T(t^A)$ is defined by $u'(T - t^A) = c'(T)$.

Meanwhile, if the absolutist monitors ($M = 1$), the agent either chooses $T = t + f$ or T such that $u'(T^* - t) - c'(T^*)$, where $T^* < t + f$. Note that the ruler will never choose t that is large enough for the latter case to arise, since he could achieve this without monitoring (and paying monitoring cost m). Hence, in equilibrium, the agent chooses $T = t + f$, so long as $u(f) - c(t + f) \geq \underline{u}$. In equilibrium, if the parameters are such that the agent's participation constraint holds, the absolutist's revenue is $\pi^R = t - m$. The absolutist therefore chooses t^A to maximize:

$$(A.2) \quad \theta q(t^A - m, t^A + f, 1) + (1 - \theta)q(t^A - m, t^A + f, 0),$$

subject to $u(f) - c(t^A + f) \geq \underline{u}$.

As $m \rightarrow \infty$, the absolutist clearly does not monitor. Moreover, as m decreases, monitoring becomes more attractive (i.e., the value in (A.2) becomes larger for any t^A). In order to show that m^A exists, we must therefore show that the ruler either prefers monitoring at $m = 0$, in which case $m^A > 0$, or the ruler either does not prefer monitoring at $m = 0$, in which case $m^A = 0$. First, it is clear that $m^A = 0$ if f is small enough that $u(f) - c(f) < \underline{u}$. In this case there is no tax and wage combination the ruler can offer while monitoring—and credibly commit to not confiscating—that makes the agent better off than his reservation. If $u(f) - c(f) \geq \underline{u}$, the ruler can choose a positive tax level that the agent collects when the ruler monitors. So long as f is not too large, as $m \rightarrow 0$ the ruler will be able to collect more statutory taxes while the overall level of taxes (T) is lower when he monitors. Thus, m^A is positive as long as f is sufficiently large but not too large. If f is so large that the total level of taxes collected under monitoring is large (because the ruler cannot commit to firing in the third stage), the ruler will not monitor and $m^A = 0$.

Next, consider the limited ruler's stage 1 decisions. If the limited ruler does not monitor ($M = 0$), he will not be able to commit to firing the agent in the third stage. The statutory tax, t , he chooses must be low enough that there exists some T for which $u(T - t) - c(T) \geq \underline{u}$. In equilibrium, if the parameters are such that the agent's participation constraint holds, the limited ruler's revenue is $\pi^R = t$. The limited ruler therefore chooses t^L to maximize:

$$(A.3) \quad \theta q(t^L, T(t^L), 1) + (1 - \theta)q(t^L, T(t^L), 0),$$

subject to $u(T(t^L) - t^L) - c(T(t^L)) \geq \underline{u}$, where $T(t^L)$ is defined by $u'(T - t^L) = c'(T)$.

Meanwhile, if the limited ruler monitors ($M = 1$), the agent either chooses $T = t + w + f$ or T such that $u'(T^* - t) = c'(T^*)$, where $T^* < t + w + f$. Note that the ruler will never choose t and w such that the latter case arises, since he could achieve this without monitoring (and paying monitoring cost m). Hence, in equilibrium, the agent chooses $T = t + w + f$, so long as $u(w + f) - c(t + w + f) \geq \underline{u}$. In equilibrium, if the parameters are such that the agent's participation constraint holds, the limited ruler's revenue is $\pi^R = t - m$. The limited ruler therefore chooses t^L to maximize:

$$(A.4) \quad \theta q(t^L - m, t^L + w(t^L) + f, 1) + (1 - \theta)q(t^L - m, t^L + w(t^L) + f, 0),$$

subject to $u(w(t^L) + f) - c(t^L + w(t^L) + f) \geq \underline{u}$, where $w(t^L)$ is defined by $u(w + f) - c(t^L + w + f) = \underline{u}$.

As $m \rightarrow \infty$, the limited ruler clearly does not monitor. Moreover, as m decreases, monitoring becomes more attractive (i.e., the value in (A.4) becomes larger for any t^L). In order to show that m^L exists, we must therefore show that the ruler either prefers monitoring at $m = 0$, in which case $m^L > 0$, or the ruler either does not prefer monitoring at $m = 0$, in which case $m^L = 0$. First, note that the limited ruler can relax the agent's participation constraint by offering wage w , unlike the absolutist (who cannot commit to refrain from confiscating the wage). Hence, unlike in the case of the absolutist, the limited ruler prefers to monitor as $m \rightarrow 0$ for small values of f , since he can supplement the agent's income with w . Thus, as long as f is not sufficiently large, $m^L > 0$. It is only when f is sufficiently large enough that the equilibrium T will be large under monitoring that the ruler prefers to not monitor. In this case, $m^L = 0$.

A.2. Proof of Proposition 2

From the proof of Proposition 1, it is clear that absolutist and limited rulers have the same statutory tax revenue when $m \geq \max\{m^A, m^L\}$. In this case, neither monitors and the agent collects taxes up to the point that its marginal benefit from doing so equals its marginal cost. Moreover, it is clear that the limited ruler will collect (weakly) more tax revenue when $m < \min\{m^A, m^L\}$, since the

absolutist ruler can only collect taxes up to the point where $T = t + f$ yields the agent's reservation utility. The limited ruler can collect more taxes, if he desires, by giving the agent a higher wage, and therefore relaxing the agent's participation constraint.

It only remains to show that the limited agent collects (weakly) more taxes when $m \in (\min\{m^A, m^L\}, \max\{m^A, m^L\})$. First, we show that $m^A \leq m^L$, ceteris paribus. To see this, consider the following. If the absolutist monitors, he can only collect taxes up to the point where $T = t + f$ yields the agent's reservation utility. The limited ruler can collect taxes beyond that, *if he desires*, by giving the agent a wage w , which is free from confiscation. Hence, revealed preference indicates that there may be parts of the parameter space where the marginal benefit of monitoring for the limited ruler exceeds the marginal benefit of monitoring for the absolutist ruler. However, there is *no* part of the parameter space in which the marginal benefit of monitoring for the absolutist ruler exceeds the marginal benefit of monitoring for the limited ruler. Hence, the size of the parameter space in which the limited ruler monitors is weakly larger than that in which the absolutist monitors, entailing that $m^A \leq m^L$.

Hence, it remains to show that the limited agent collects (weakly) more taxes when $m \in (m^A, m^L)$. In this range, the limited ruler monitors but the absolutist does not. Hence, in the absolutist regime the agent collects taxes up to the point that its marginal benefit from doing so equals its marginal cost (i.e., $u'(T - t^A) = c'(T)$), conditional on t^A , $w < f$, and $u(T - t^A) - c(T) \geq \underline{u}$. In equilibrium, the absolutist chooses t^A to maximize (A.1). Meanwhile, the limited ruler monitors and therefore, from the proof of Proposition 1, the agent chooses $T = t^L + w + f$, so long as $u(w + f) - c(t^L + w + f) \geq \underline{u}$, while the limited ruler chooses t^L to maximize (A.4).

It remains to show that the limited ruler collects more statutory tax revenue when $\theta \geq \bar{\theta}$, for some $\bar{\theta}$. This is tantamount to proving the existence of $\bar{\theta}$. First, note that when $\theta \rightarrow 0$, it is possible (though not necessary) that the limited ruler chooses to collect less taxes. In this case, $E[U^R] \rightarrow q(\pi^R, T, 0)$, and thus the limited ruler may choose to collect less taxes which, while resulting in lower π^R , also results in lower T than the absolutist (remember, the limited ruler can always choose to act like the absolutist by not monitoring; the fact that the limited ruler monitors in this range suggests that he is better off from doing so). In such a case, $\bar{\theta} > 0$. To show existence it is sufficient to show that the limited ruler collects more statutory tax as $\theta \rightarrow 1$. In this case, $E[U^R] \rightarrow q(\pi^R, T, 1)$. Since we assume that $q_1(\cdot, \cdot, 1) > \bar{q}$ for some large value of \bar{q} , this indicates that the

ruler will want to collect as much statutory tax as possible when $\theta \rightarrow 1$. The absolutist can collect only up to the point where $u(T(t^A) - t^A) - c(T(t^A)) = \underline{u}$, for $T(t^A)$ defined by $u'(T - t^A) = c'(T)$. Meanwhile, the limited ruler can only collect up to the point where $u(w + f) - c(t^L + w + f) = \underline{u}$. Note that the limited ruler can collect the same amount as the absolutist by setting $w = T(t^A) - t^A - f$, which would allow him to maximize statutory tax collected at $t^L = t^A$. However, the limited ruler could simply scale back w a bit, which would ease the agent's participation constraint (since $u'' < 0$ and $c'' > 0$), allowing the ruler to collect more taxes (t^L) without breaking the agent's participation constraint. Since the ruler desires maximum tax collection as $\theta \rightarrow 1$, it therefore must be true that $t^L > t^A$. This continues to be true as θ decreases, until θ is low enough (i.e., $\theta = \bar{\theta}$) that the limited ruler prefers lower levels of T to higher levels of t^L .

A.3. Proof of Proposition 3

In the proof of Proposition 2, we showed that $m^A \leq m^L$.

A.4. Proof of Proposition 4

Consider the marginal net benefits (i.e., $MNB = MB - MC$) that limited and absolutist rulers derive from a decrease in m , absent the cost $d(\cdot)$. First, note that they set $M = 0$ at $m \rightarrow \infty$, meaning that they never monitor and thus agents extract up to the point where their MB from additional taxes equals the MC of tax collection. As m decreases, the limited ruler chooses to monitor over a (weakly) larger part of the parameter space, ceteris paribus (from Propositions 2 and 3). Hence, as m decreases, it follows that the MB of monitoring is greater for the limited ruler, and thus the MNB of a decrease in m is greater for limited rulers than absolutist rulers (since the MC is the same for both). Since the marginal cost, d' , of decreasing m is the same for both types of rulers at all values of m , the limited ruler chooses to decrease m by a greater amount than the absolutist.

A.5. Proof of Proposition 5

A ruler will conquer the new territory only if the marginal benefits from doing so are sufficiently large. We denote the expected net marginal benefits of conquering the new territory for ruler $i \in \{A, L\}$, where A is absolutist and L is limited, as:

$$(A.5) \quad [\hat{\theta}q(\hat{\pi}^{R,i}, \hat{T}^i, 1) + (1 - \hat{\theta})q(\hat{\pi}^{R,i}, \hat{T}^i, 0)] - [\theta q(\pi^{R,i}, T^i, 1) + (1 - \theta)q(\pi^{R,i}, T^i, 0)] - r,$$

where $\hat{\pi}^{R,i} (> \pi^{R,i})$ is the revenue the ruler of type i is able to collect if he conquers the territory. If the value in (A.5) exceeds 0, the ruler will choose to conquer. Hence, there is a larger set of values of r for which the absolutist chooses to conquer if and only if:

$$(A.6) \quad [\hat{\theta}q(\hat{\pi}^{R,A}, \hat{T}^A, 1) + (1 - \hat{\theta})q(\hat{\pi}^{R,A}, \hat{T}^A, 0)] - [\theta q(\pi^{R,A}, T^A, 1) + (1 - \theta)q(\pi^{R,A}, T^A, 0)] > [\hat{\theta}q(\hat{\pi}^{R,L}, \hat{T}^L, 1) + (1 - \hat{\theta})q(\hat{\pi}^{R,L}, \hat{T}^L, 0)] - [\theta q(\pi^{R,L}, T^L, 1) + (1 - \theta)q(\pi^{R,L}, T^L, 0)].$$

This can be rewritten:

$$(A.7) \quad \theta[q(\pi^{R,L}, T^L, 1) - q(\pi^{R,A}, T^A, 1)] + (1 - \theta)[q(\pi^{R,L}, T^L, 0) - (1 - \theta)q(\pi^{R,A}, T^A, 0)] > \hat{\theta}[q(\hat{\pi}^{R,L}, \hat{T}^L, 1) - q(\hat{\pi}^{R,A}, \hat{T}^A, 1)] + (1 - \hat{\theta})[q(\hat{\pi}^{R,L}, \hat{T}^L, 0) - q(\hat{\pi}^{R,A}, \hat{T}^A, 0)].$$

Note that while $\hat{\pi}^{R,L} - \hat{\pi}^{R,A} \geq \pi^{R,L} - \pi^{R,A}$, since the limited ruler can take more on the margin from the newly conquered territory, it is not necessarily true that $q(\hat{\pi}^{R,L}, \hat{T}^L, X) - q(\hat{\pi}^{R,A}, \hat{T}^A, X) > q(\pi^{R,L}, T^L, X) - q(\pi^{R,A}, T^A, X)$, since q is concave.

From Proposition 2, we know that $\hat{\pi}^{R,L} \geq \hat{\pi}^{R,A}$ and $\pi^{R,L} \geq \pi^{R,A}$ for any value of m (assuming $\theta \geq \bar{\theta}$). Moreover, as m increases, it must be true that $\hat{\pi}^{R,L} - \hat{\pi}^{R,A}$ decreases (weakly) faster than $\pi^{R,L} - \pi^{R,A}$. To see this note that $\hat{\pi}^{R,L} - \hat{\pi}^{R,A} \geq \pi^{R,L} - \pi^{R,A}$ at $m = 0$ but $\hat{\pi}^{R,L} - \hat{\pi}^{R,A} = \pi^{R,L} - \pi^{R,A} = 0$ at $m \rightarrow \infty$. Likewise, as m increases, it must be true that $\hat{T}^A - \hat{T}^L$ decreases (weakly) faster than $T^A - T^L$. To see this note that $\hat{T}^A - \hat{T}^L \geq T^A - T^L$ at $m = 0$ but $\hat{T}^A - \hat{T}^L = T^A - T^L = 0$ at $m \rightarrow \infty$.

Combined, these insights entail that $[q(\hat{\pi}^{R,L}, \hat{T}^L, X) - q(\hat{\pi}^{R,A}, \hat{T}^A, X)] - [(q(\pi^{R,L}, T^L, X) - q(\pi^{R,A}, T^A, X))]$ is weakly decreasing in m . Since $\hat{\theta} < \theta$ and $[q(\hat{\pi}^{R,L}, \hat{T}^L, X) - q(\hat{\pi}^{R,A}, \hat{T}^A, X)] = [(q(\pi^{R,L}, T^L, X) - q(\pi^{R,A}, T^A, X))]$ at $m \rightarrow \infty$, there must therefore exist some sufficiently large value of m , which we denote m^* , such that whenever $m \geq m^*$, $[q(\hat{\pi}^{R,L}, \hat{T}^L, X) -$

$q(\hat{\pi}^{R,A}, \hat{T}^A, X) - [(\pi^{R,L}, T^L, X) - q(\pi^{R,A}, T^A, X)]$ is sufficiently small (for $X \in \{0,1\}$) that inequality (A.7) holds.