Internal Conflict, Geopolitics, and State Development

Evidence from Imperial China[*]

Mark Dincecco[†] Yuhua Wang[‡]

April 28, 2020

Abstract

We analyze the conditions under which internal conflict can promote or impede state development, taking imperial China as a laboratory. We develop a novel theoretical framework in which local elites and the central state interact over the provision of internal security. Using new micro-level data across more than half a millennium, we show evidence that, traditionally, the central state improved its monopoly over violence in response to mass rebellion. The mid-nineteenth century marked the start of a new geopolitical era in China, when the central state's costs of external military defense rose greatly. In turn, we find evidence for a shift in leadership toward local elites in response to internal conflict. This change reduced the central state's monopoly over violence and promoted state failure. Our study implies that the relationship between internal conflict and state development depends on the cost conditions for a given geopolitical context.

Keywords: Mass Violence, External Threats, State Capacity, Elite Action, Kin Networks, China

---

[*]We greatly thank Carles Boix, Dean Dulay, James Fenske, Benjamin Fordham, Vicky Fouka, Jeffry Frieden, Philip Hoffman, Ruixue Jia, Timur Kuran, Alexander Lee, John Londregan, Debin Ma, Anil Menon, Jack Paine, Jared Rubin, Daniel Slater, Tuan-Hwee Sng, John Wallis, Daniel Ziblatt, and seminar participants at Duke University (PEW; State Capacity Workshop), Harvard University, the London School of Economics, McGill University, UCSD, the University of Michigan (CCS; CPRD), the University of Oxford, the University of Virginia, the University of Warwick, Yale University, the IR Conference at the University of Chicago, the PIEP Conference at Indiana University, and the State Capacity Conference at Princeton University for feedback; Shuo Chen, James Kung, and Tuan-Hwee Sng for data-sharing; and Yunpeng Bai, Yuqian Chen, Cheng Cheng, Yusi Du, Maggie Huang, Jialu Li, Shiqi Ma, Jia Sun, Patricia Sun, Yihua Xia, Siyao Zheng, and Ziwen Zu for research assistance. We gratefully acknowledge the support of the Alexander and Diviya Maguro Peer Pre-Review Program at the Institute for Quantitative Social Science at Harvard University, and thank Associate Director Chase Harrison and three anonymous pre-review referees.

[†]Department of Political Science, University of Michigan, Ann Arbor, MI 48109; dincecco@umich.edu

[‡]Department of Government, Harvard University, Cambridge, MA 02138; yuhuawang@fas.harvard.edu

1 Introduction

External warfare is a common explanation for state-making (Tilly 1975; Besley and Persson 2011; Scheve and Stasavage 2016).[1] Yet there is still relatively less consensus over the role of internal conflict. By enabling local militias to thrive, internal conflict may impede the state's monopoly over violence (Bates 2008; Acemoglu, Robinson and Santos 2013; Staniland 2015; Ch et al. 2018; Garfias and Sellars 2018). On the other hand, by inducing elite collective action, internal conflict may promote state development (Weinstein 2005; Slater 2010; Rodríguez-Franco 2016).

To make progress on this debate, we develop a novel theoretical framework that analyzes elite-state interactions in response to mass rebellion against the state. In the spirit of Weber (1946 [1918], 78) we think of state development in terms of improvements in the central state's ability to secure a monopoly over violence within its territory. Unlike Weber, however, we do not take the central state's monopoly over violence as given. Rather, we view state-strengthening as one potential outcome of a joint decision-making process over how to best provide internal security. In our framework, the central state, led by a single ruler, governs in partnership with local elites. The ruler must decide whether to strengthen the central state's leadership role as the locus of elite-state cooperative efforts against mass violence, or to "outsource" such efforts to local elites. Meanwhile, local elites must decide whether to willingly accept this shift in the locus of leadership toward the central state, or to oppose it. From the ruler's perspective, state-strengthening reduces the transaction costs of governance. This choice, however, requires the ruler to make costly new investments in state capacity to improve the center's ability to broadcast its political power. To fund such investments, the ruler must levy greater "tribute" on local elites. On the plus side, state-strengthening reduces the governance costs of local elites relative to outsourcing.

This framework implies that internal conflict will have different implications for state development depending on the cost conditions of the specific geopolitical context. First, in the benchmark case, state-strengthening is cost-effective enough to be attractive to both local elites and the ruler. Thus, we expect to observe that the central state strengthens its leadership role in elite-state cooperative efforts in response to mass violence. Second, if

[1] Recently, Queralt (2019) shows evidence that wars financed via domestic taxation have promoted long-run improvements in state capacity, but that wars financed via external loans have not.

there is a significant change in the geopolitical context, such as a severe new external threat that induces the ruler to undertake new investments in military defense, then elites may now oppose state-strengthening, for two reasons. The ruler will expect them to pay a much higher tribute, and yet they will still have to pick up a good deal of slack in local governance, since the ruler must devote a greater portion of resources to military defense, rather than help contribute to non-military public goods. Given new external threats, then, we expect to observe a shift in leadership in elite-state cooperation toward the local elites and away from the central state in response to mass violence. Depending on the severity of the central state's loss over the monopoly over violence, then state failure may eventually become more likely.

We evaluate this framework by analyzing the long-run dynamics of state development in imperial China. Historically, the majority of military conflicts fought in China were internal. Furthermore, China has very well-documented historical records. Imperial China is therefore an ideal laboratory to evaluate the debate about the relationship between internal conflict and state development. By taking a long-term view, our study contributes to this debate both theoretically and empirically, by documenting the conditions under which internal conflict may improve – or reduce – the central state's monopoly over violence.

We construct original data from a variety of archival and contemporary documents. In imperial China, the clan was an important provider of local public goods including security vis-à-vis the central state, and local elites played a major role in the organization of collective action by the clan. Our main database exploits new panel data on internal conflict and clan activity (and thus the extent of outsourcing of elite-state cooperative efforts) at the local level (i.e., 25x25 km grid cells) in imperial China that span more than half a millennium, from the start of the Ming Dynasty to the downfall of the Qing. To proxy for the local extent of internal conflict, we identify the geographical locations of all major recorded mass rebellion battles over this period. To proxy for the local extent of clan activity, we build off of the precedent in Greif and Tabellini (2017, 5-7) and focus on the compilation of genealogy books. Namely, we identify the geographical locations of all such books on record. Given historical data limitations, this approach to measure clan activity is both feasible and systematic. We complement this panel database with two cross-sectional databases for which historical data on state development outcomes are available. The first of these includes new geocoded data on both military garrisons and courier routes during the Ming Dynasty, while the second

includes land tax data during the Qing Dynasty. Finally, to proxy for state failure, we employ data on local independence declarations from the imperial Qing state.

Our empirical analysis proceeds in several parts. First, we perform a panel regression analysis of mass rebellion and clan activity in the traditional geopolitical context in imperial China (i.e., pre-1850). To help control for unobservable features that may bias our results (e.g., initial demographic and economic conditions, trends in cultural conservatism), this analysis includes grid cell and period fixed effects and county-specific time trends. Furthermore, we perform an instrumental variables analysis that exploits exogenous variation in the frequency of droughts. We find a negative, statistically significant, and robust relationship between mass rebellion and clan activity. This result is consistent with our empirical prediction that, traditionally, internal conflict weakened any leadership role that local elites may have played in elite-state cooperative efforts.

We next provide evidence that mass rebellion was traditionally associated with new state capacity investments that improved the center's ability to broadcast political power. Our cross-sectional regression analysis reveals a positive and significant relationship between internal conflict and the construction of military garrisons and courier routes, respectively, even after controlling for local fixed effects. Similarly, we find a significant positive correlation between internal conflict and total land taxation. These results are consistent with our empirical prediction that, traditionally, state development took place in response to mass rebellion.

Defeat by Britain in the First Opium War (1839-42) marked the start of a new geopolitical era in imperial China. In this new context, the central state's costs of military defense rose greatly. Thus, according to our theoretical framework, mass rebellion would no longer strengthen the central state's leadership role vis-à-vis local elites in the provision of internal security, as was traditionally the case. Rather, our theoretical framework predicts that we should now observe greater outsourcing of elite-state cooperative efforts to local elites, and thus greater clan activity, in response to mass rebellion.

The exact timing of China's first military defeat by Britain was somewhat unpredictable and subject to chance. The development of superior military technology in Western Europe was an idiosyncratic process in which several factors co-evolved over hundreds of years (Hoffman 2015, 19-66). A priori, it was not obvious at which specific point in time this

indigenous process would bear fruit and enable the West to truly threaten imperial China's sovereignty. We exploit this "plausibly exogenous" timing to study the relationship between internal conflict and security provision in the new geopolitical context.

Namely, we extend the panel regression analysis described above to the entire 1350-1900 period. Relative to the traditional era, we find that, in the new geopolitical context, the relationship between mass rebellion and clan activity turns positive in sign, and is significant. This evidence is consistent with our empirical prediction that, following the shift in leadership in elite-state cooperation toward local elites, then we should observe greater clan development in response to mass rebellion. Falsification exercises provide further evidence that the positive post-1850 relationship between mass rebellion and clan activity stands out relative to previous shocks of domestic upheaval, and is not driven by unobserved features specific to the nineteenth century.

Finally, we evaluate the extent to which greater clan activity in the new geopolitical context influenced subsequent decisions to declare independence from the imperial state. Here, we find a positive correlation between increased local clan activity and independence declarations. This evidence is consistent with our empirical prediction that the shift in leadership in elite-state cooperation toward local elites may have eventually promoted state failure.

In summary, our study shows how, depending on the cost conditions for a given geopolitical context, elite-state cooperation in response to internal conflict can improve or reduce the central state's monopoly over violence. As described at the outset, it thereby enhances our understanding of the relationship between internal conflict and state-making.

In addition, our study contributes to the literature on long-run state development. Much of this literature centers on state-making in medieval and early modern Europe (Tilly 1975; Ziblatt 2006; Stasavage 2011; Blaydes and Paik 2016; Cox 2016). We still know relatively less about this historical process across other, late-modernizing parts of Eurasia (Kuran 2018, 1353).[2] Imperial China displays several features common to late modernizers: an agricultural economy, recurrent civil conflict, and strong kin networks. Thus, the insights into the state development process that we glean from this case should generalize to a variety of other non-European contexts. We will discuss this point in further detail in the conclusion of the paper.

[2]One exception is Blaydes and Chaney (2013), who identify a divergence in pre-1500 political stability between Western Europe and the Muslim world.

Finally, our study improves our knowledge of state development within China (Shue 1988; Hui 2005; Sng 2014; Sng and Moriguchi 2014; Bai and Jia 2016; Mattingly 2017; Koyama, Moriguchi and Sng 2018; Ma and Rubin 2019). Much of this literature focuses on a particular dynasty or historical time period. Our analysis complements this literature by taking a long time span, enabling us to track the evolution of the central state's monopoly over violence in imperial China in terms of a single theoretical and empirical framework.

We proceed as follows. Section 2 develops our theoretical framework. Section 3 provides historical background. Section 4 describes the data that we will employ. Section 5 presents the empirical strategy, main results, and robustness checks for the traditional geopolitical context (i.e., pre-1850), and Section 6 for the new geopolitical context (i.e., post-1850). Section 7 provides concluding remarks.

2 Theoretical Framework

Does internal conflict improve or reduce a central state's monopoly over violence? This topic remains open to scholarly debate. Slater (2010, 5-7), for example, argues that mass threats of redistribution can induce elites to form a broad-based coalition that supports greater state strength as an institutional safeguard against popular revolt. In the same vein, Weinstein (2005, 14-26) argues that internal conflict may produce effective governance when victorious actors face high survival threats during conflict, coupled with a large revenue imperative and few external resources.

Another strand of this literature argues that internal conflict can promote state failure. According to Bates (2008, 2), for example, elites may form private local militias for security against mass threats during conflict. Afterward, central governments may have political incentives to partner with private local militias, thereby reducing the central government's capacity to establish a monopoly over violence (Acemoglu, Robinson and Santos 2013, 7; Staniland 2015, 771).

We now develop a simple theoretical framework that draws on both of the above viewpoints. Namely, we argue that whether internal conflict promotes or impedes state development is dependent on the specific cost conditions of the geopolitical context in question. Here, we follow a common approach in the literature and think of internal conflict in terms of mass rebellion against the state. To save space, we leave the technical details to Appendix A.

2.1 Setup

Our framework takes the form of a sequential-move game in which Nature moves first. There is widespread evidence for the link between climate and conflict (Burke, Hsiang and Miguel 2015). Thus, we postulate a simple relationship whereby a negative climate shock automatically increases the odds of mass violence. In line with the above premise, Nature determines whether there is a negative climate shock, and therefore whether the peasantry decides to rebel against the imperial state. We think of the imperial state in terms of an institutional structure in which the central state governs in partnership with local elites (Skocpol 1979, 48-9). Given our focus on elite-state interactions in response to mass violence, to be described ahead, we treat the central state as a unitary actor led by a single ruler such as the emperor. In response to mass violence, the ruler must decide whether to strengthen the central state's leadership role as the locus of elite-state cooperative efforts, or to "outsource" this governance task to local elites. If the ruler decides to strengthen the central state's leadership role, then local elites must decide whether to willingly accept this shift in the locus of leadership in elite-state cooperation toward the central state, or to oppose it.

2.2 Optimal Decisions

To make optimal decisions, both the ruler and local elites weigh their benefits against their costs. We assume that each actor receives a benefit from putting down mass violence, regardless of the specific locus of leadership in elite-state cooperation. If the ruler decides to strengthen the central state's leadership role as the locus of elite-state cooperation, and local elites accept this shift, then the ruler receives an additional benefit, which we may think of in terms of a reduction in transaction costs in governance. First, the central state may have a comparative advantage in putting down mass violence. Second, stronger leadership from the center may reduce the potential for opportunistic behavior by local elites. Alternatively, we may think of this benefit as a "prestige" payoff to the ruler for strengthening the central state's leadership role. On the other hand, state-strengthening requires the ruler to make new state capacity investments to improve the center's ability to broadcast its political power across its territory (and to help contribute to local non-military public goods such as grain reserves). To fund such investments, the ruler must levy greater "tribute" (e.g., donations, taxes). Given the subsistence income of the peasantry, we assume that local elites must cover much if not all of this new cost. Alternatively, we may think of this cost as a loss in

political autonomy by local elites to the central state. In addition, local elites must pay their part of the governance cost added (including help in the provision of local non-military public goods), even under state-strengthening. However, this governance cost is lower than if the locus of leadership in elite-state cooperation is outsourced to local elites. Finally, if local elites oppose the ruler's state-strengthening choice, then the ruler's ability to exact greater tribute and undertake new state capacity investments is (at least somewhat) curtailed. In this case, the ruler must pay a "rejection" cost for being rebuffed by local elites, thereby jeopardizing the stability of the elite-state governance partnership (vis-à-vis the peasantry). Similarly, local elites must pay a penalty for their opposition.

2.2.1 Benchmark

Our theoretical framework implies that mass violence will have different implications for state development depending on the cost conditions of the specific geopolitical context. Say first that the reduction in governance costs to local elites under state-strengthening by the ruler is relatively high (vis-à-vis outsourcing to local elites via clans as the locus of leadership). Similarly, say that the tribute cost that local elites must pay to fund new state capacity investments is relatively low (vis-à-vis the cost of opposition to the ruler's state-strengthening efforts). Finally, say that the payoff to the ruler from the reduction in the transaction costs in governance due to state-strengthening exceeds the cost of new investments in state capacity to improve the center's ability to broadcast its political power. In this context, the ruler will opt for state-strengthening, and local elites will accept this choice. We view the spirit of this outcome, by which the central state strengthens its leadership role vis-à-vis local elites as the locus of elite-state cooperation, as the "benchmark."

2.2.2 New External Threat

Now say that the ruler must undertake new investments in (external) military defense in response to a severe new external threat. To cover this new cost, and to continue to cover the cost of internal security to put down mass violence, local elites are expected to pay a much higher tribute than under the benchmark. At the same time, the ruler must devote a greater portion of total revenue to (external) military defense, thereby reducing his traditional contribution to local non-military public goods (e.g., grain reserves). The reduction in governance costs to local elites under state-strengthening is therefore relatively low, since they may still have to pick up a good deal of the slack. Thus, there is a role reversal. Local

elites will now oppose state-strengthening by the ruler.

2.2.3 State Failure

Extrapolating from our (i.e., static) theoretical framework, in the new geopolitical context in which the central state partly loses its monopoly over violence and contributes relatively less to non-military public goods, we may think that mass violence will eventually become more likely. Namely, there may be an "unraveling" over time, in which the new external threat reinvigorates local elites as the locus of leadership in elite-state cooperation, while subsequently making mass violence more likely (since local elites may lack a comparative advantage in putting it down), which then further reinvigorates the role of local elites, thereby further reducing the central state's monopoly over violence. In this manner, state failure may eventually become more likely.[3]

2.3 Predictions

Summarizing, our theoretical framework gives rise to two main predictions and two ancillary predictions, which we will use to help guide our empirical analysis.

1. *Greater Monopoly over Violence*. In the benchmark case, we expect to observe that the central state strengthens its leadership role in elite-state cooperative efforts in response to mass violence.

2. *Loss of Monopoly over Violence*. Given new external threats, we expect to observe a shift in leadership in elite-state cooperation toward local elites and away from the central state in response to mass violence.

A1. *Taxation*. Given new external threats, we expect to observe an increase in taxation of local elites (or at least an attempt at) by the central state.

A2. *State Failure*. If the central state's loss of the monopoly over violence is severe enough due to a shift in leadership in elite-state cooperation toward local elites, then state failure may eventually occur.

[3] Garfias and Sellars (2018, 5-8) describe how the delegation of authority to local elites may actually enhance the central state's monopoly over violence, since the ruler can extend his control over a larger swath of territory without having to invest in the infrastructure of direct rule. This result depends on whether local elites have enough incentive to cooperate with the central state to put down mass violence. In the context described in the main text above, local elites must pay a high tribute cost, and the ruler's contribution to traditional non-military public goods is reduced. Thus, the incentive of local elites for such cooperation may fall short, promoting state failure.

3 Historical Background

We now provide a brief historical overview of violent conflict, elite-state interactions, and state development outcomes in imperial China in terms of our theoretical framework.

3.1 Mass Violence

More than 65 percent of military conflicts fought in imperial China were internal (Dincecco and Wang 2018, 343). Exogenous climate shocks (e.g., droughts) were important catalysts for peasant rebellions (Perry 1980, 3; Kung and Ma 2014; Jia 2014*b*). Typically, local revolts began when peasant groups plundered affluent neighbors (Perry 1980, 4), prompting the local arm of the central government to put them down. Yet local revolts sometimes erupted into major internal conflicts, compelling central forces to directly intervene (Kuhn 1970, 52).[4] Major mass rebellion battles between central government forces and mass rebel groups (and not local revolts, which generally went unrecorded at the national level) will form the basis of our empirical measure of internal conflict in Section 4.

3.2 Elite-State Interactions

Given radical redistributive demands (e.g., Spence 1996, 160), mass rebellions posed serious threats to the property and lives of traditional land-owning local elites, known collectively as the gentry.[5] When nearby peasants took up arms, the gentry had two basic options. The first was to rely on cooperative efforts led by the central state. From the Song Dynasty (960-1279) onward, most central governments placed their military garrisons at or near major urban centers. In times of peasant revolt, the gentry could flock to these "walled safe havens" for temporary refuge (Rowe 2007, 29). During mass rebellion, the gentry often gave greater "tribute" (e.g., donations, taxes) to the state to help it build up its defense infrastructure (Chang 1955, 83). In this manner, internal conflict (and threats thereof) could transfer resources from local land-owning elites to the central government, strengthening its monopoly over violence.

The gentry's second option was to collaborate with kin members to protect themselves.

[4] Battlefield success by the peasantry was very unlikely. Between 1350 and 1900 only one mass rebellion actually succeeded: the Zhu Yuanzhang Rebellion of the mid-fourteenth century.

[5] Chang (1955, 3) defines the gentry as all holders of academic degrees under the imperial civil service exam system. Although such a degree made an individual qualified for office, most degree holders were not office-holders, and dwelled in their home districts.

If the central government found it too costly to provide security on its own, then the emperor might (partially) outsource leadership in elite-state cooperative efforts to local elites (e.g., Kuhn 1970, 135-52). Outsourcing leadership, however, threatened the central state's monopoly over violence. Thus, the central government was reluctant to allow elites to take on a greater leadership role. If the emperor did agree to outsource leadership, then the gentry could keep more of its resources for themselves, rather than transfer them to the central government, thereby undermining its strength.

3.3 Clan Activity

The clan was the primary means through which the gentry organized cooperation among kin members (Rowe 2007, 65-6; Greif and Tabellini 2017).[6] During mass rebellion, the gentry could conscript their clan to provide temporary shelter, build a fortress, or establish a private militia, if the central government decided not to (or could not) provide sufficient protection.

Prior to the mid-nineteenth century, mass rebellion often weakened kin networks, since the gentry typically fled to walled urban centers, and poor clan members were left to fend for themselves. Describing the Yao clan of central China during the turbulent dynastic change from the Ming to the Qing, for example, Beattie (1979, 94) writes that "with the weakening of ties respect for the lineage had become considerably eroded."

By contrast, once the gentry began to depend more heavily on the clan for security in the mid-nineteenth century, mass rebellion often strengthened kin networks. Rowe (2007) provides a detailed account of how clan militias functioned in Macheng County in central China. Noting that "lineage consciousness was probably the most fundamental element in Macheng people's personal identity," Rowe describes how clans organized collective defense (Rowe 2007, 65-6). For example, military forts in Macheng County were sometimes lineage-specific, including the Yu clan's Cloud Dragon Fort and Xia clan's Stonewall Fort (Rowe 2007, 205).

The compilation of genealogy books was one common kin-oriented response to violent conflict. Rowe (2007, 70-1) argues that genealogical compilation by the gentry during the late imperial period was a reaction to "the plebeian mass around them" and took place "in the wake of social crisis." Similarly, Szonyi (2016, 444) states of this era: "The scope of

[6] Freedman (1958, 3-4) defines the clan as a lineage organization that includes all male descendants within five generations of a common male ancestor.

genealogical compilation typically expands whenever social instability generates a strong sense of need for new and broader social networks."

Thus, building off of the precedent in Greif and Tabellini (2017, 5-7), we believe that genealogy books can serve as a meaningful proxy for the extent of local clan activity and elite strength vis-à-vis the central state, and in particular in response to violent conflict. In Section 4, we will describe in detail the genealogy data that we will exploit for this purpose.

3.4 New Western Influence

Traditionally, external threats came primarily from Steppe nomads (Bai and Kung 2011, 975; Hoffman 2015, 70-2; Ko, Koyama and Sng 2018, 290-2).[7] There were two cases in which the nomads actually overtook the imperial state. In each case, the nomads kept the imperial system intact (Wakeman 1975, 85; Skocpol 1979, 67). Following the Manchu conquest of the Ming Dynasty in 1644, for example, the new Qing government put down a mass rebellion in central China, convincing the local gentry that "they could and should work as partners with the alien dynasty" (Rowe 2007, 157). Still, in Section 5, we will show quantitative evidence that the Ming-Qing dynastic change was not a critical shock to the viability of the imperial system.

Britain's victory in the First Opium War (1839-42), along with the resulting Treaty of Nanjing, marked the start of a new geopolitical era in imperial China (Rosenthal and Wong 2011, 221; Koyama, Moriguchi and Sng 2018, 181). Treaty concessions created an "unprecedented financial crisis" (Shi and Xu 2008, 55). In accordance with the Treaty's most-favored nation clause, any concession granted by China to one foreign power was automatically extended to other treaty signatories (Wakeman 1975, 138). Thus, Britain's military victory did not turn the imperial Qing state into a European colony.

Several "downstream" outcomes followed China's military defeat in the First Opium War, including the Taiping Rebellion (1850-64). To endure the financial crisis, the Qing government enacted new taxes, the bulk of which fell onto the gentry (Chang 1962, 151; Duara 1988, 3, 219; Shi and Xu 2008, 107). The large-scale arrival of missionaries, moreover, promoted the spread of Christianity (Wakeman 1975, 143). Finally, droughts in Guangxi – the region in which the Taiping Rebellion originated – led to famine conditions (Spence 1996,

[7] Kang (2020) argues that tributary trade relations reduced the likelihood of warfare between China and other powerful East Asian states such as Japan, Korea, and Vietnam.

81). All three events created new socioeconomic discontent.

In 1853, the Taiping rebels captured the city of Nanjing, declaring it the capital of the Taiping Heavenly Kingdom. The property and lives of the gentry were severely threatened by the Taiping rebels (e.g., Platt 2012, 18). Due to the weak fiscal position of the imperial Qing state, its military forces were not typically paid on time and were in poor fighting shape (Kuhn 1970, 10; Shi and Xu 2008, 58-60; Platt 2012, 118). Reluctantly, the central state allowed the gentry to take on greater leadership in local governance matters. In particular, the gentry began to play a more prominent role in organizing local defense, both in terms of finance and leadership (Kuhn 1970, 89-92; Yang 2012, 335). With help from the gentry leader Zeng Guofan and a coalition of private militias, the imperial Qing state finally put down the Taiping Rebellion by 1864, and other mass rebellions by 1869.

This victory brought a period of stability and reform to the Qing government. Between 1849 and 1885, central government revenue grew from 43 to 77 million taels (Rosenthal and Wong 2011, 201-2). While greater revenue enabled the imperial Qing state to begin to respond to new foreign and domestic challenges, it was still not enough (Rosenthal and Wong 2011, 201; von Glahn 2016, 380-2). Furthermore, the state was forced to reduce its traditional provision of non-military public goods (Shi and Xu 2008, 50, 232; Rosenthal and Wong 2011, 200).

3.5 State Failure

By granting the gentry a prominent leadership role, the Qing's endorsement of local governance during the Taiping Rebellion may have eventually tipped the balance of power (Kuhn 1970, 211-25; von Glahn 2016, 380-1). The gentry were now formally involved in both local defense and local administration. Thus, political power shifted from the hands of the central officials into those of local elites (Kuhn 1970, 211).

The Wuchang Uprising, followed by declarations of independence by local military forces throughout China, prompted the downfall of the imperial Qing state in 1911. With respect to the deeper roots of Qing state failure, Wakeman (1975, 228) highlights the longer-term shift in the power balance toward local elites and away from the central government that had began more than a half-century before.

4 Data

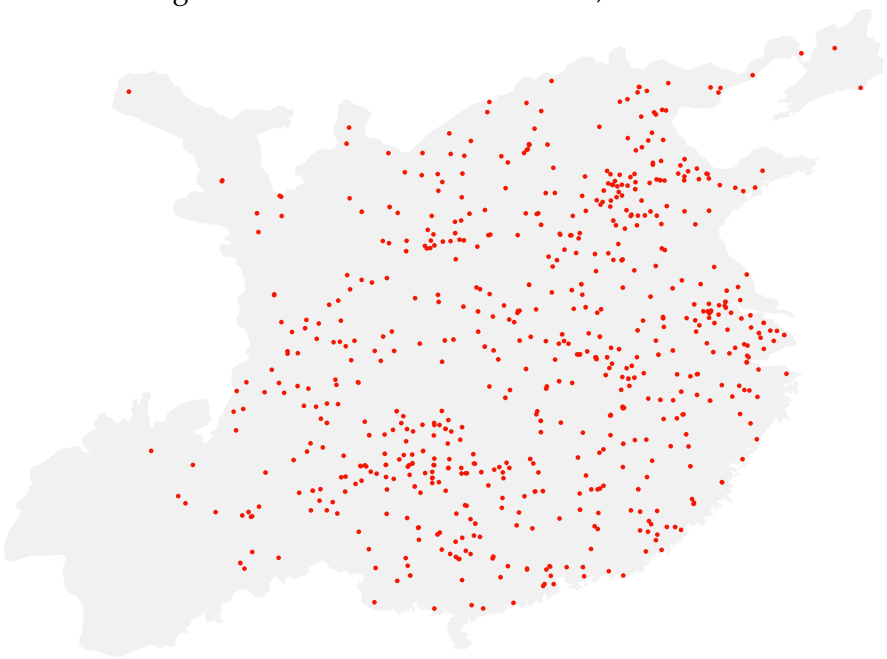
4.1 Violent Conflict

To construct the historical conflict data, we rely on the *Catalog of Historical Wars* produced by the Nanjing Military Academy (2003). This catalog contains detailed information including dates, locations of individual battles, and leaders for each major internal and external conflict that took place in China from approximately 1000 BCE to the downfall of the Qing Dynasty in 1911. The *Catalog* derives this information from China's official historical books, known as the "twenty-four histories." Traditionally, each dynasty in China compiled a standardized history of its predecessor, typically based on official court records. The official historical books produced as a result of this process are among the most important sources of systematic data on Chinese history (Wilkinson 2000, 501).

Given the historical nature of these data, there may be measurement error. In our view, selection bias is unlikely to be severe, since each official book was written by relatively contemporaneous historians whose main task was to provide the available facts and draw lessons for the incoming dynasty. For this reason, we are confident that the main historical conflicts in imperial China are well-represented in the *Catalog*. The official history books did not record casualty totals, limiting our ability to distinguish between the magnitudes of different conflicts. Presumably, however, all recorded conflicts were judged by historians to be above a certain threshold of significance, thereby justifying their inclusion in the official history books. Thus, we are confident that recorded data enable us to make "apples-to-apples" comparisons between conflicts. Nevertheless, the quality of the conflict data coverage may differ by place and time. To account for this possibility, our regression analysis ahead will employ grid cell and period fixed effects and county-specific time trends, along with a variety3 of robustness checks.

For the purposes of this study, we focus on mass rebellion, defined as a violent conflict between a central government force and a mass rebel group (e.g., peasants, artisans). Here, we identify a rebel group as a mass organization so long as its leadership did not hold any official government positions according to the *Catalog*. This coding method is similar in spirit to Jia (2014*b*, 96), and yields broadly similar patterns of mass rebellion over time and place. Unlike Jia, however, we geocode the specific location of each conflict event (including

Figure 1: Mass Rebellion Locations, 1350-1900



Notes. This figure shows the location of each recorded mass rebellion battle in China between 1350-1900 within Ming-era borders.

not only mass rebellion, but also elite rebellion and external war, to be described ahead), and our database spans a longer time period. The Li Zicheng Rebellion in the late Ming era and the Taiping Rebellion in the late Qing era (which we will analyze in Section 6) are two examples of mass rebellion included in the *Catalog*.

Our sample data consist of 887 individual battles linked to 453 recorded mass rebellions between 1350 and 1900. This period approximately spans the start of the Ming Dynasty to the downfall of the Qing. We focus on this period for several reasons. First, the central state did not establish a political equilibrium of internal spatial integration until the late thirteenth century (Rosenthal and Wong 2011, 12-13, 23-4). Second, as we will describe ahead, historical data for local state development outcomes that we can partner with the clan activity data are not widely available prior to the Ming era. Third, we have an accurate shapefile of the external borders in place during the Ming Dynasty, but not for its predecessor (i.e., the short-lived Yuan Dynasty). Figure 1 maps the mass rebellion battle locations, which took place all throughout China.

The above definition excludes rebellions led by local governing elites (e.g., the Revolt of the Three Feudatories led by Wu Sangui in the early Qing era), since such rebellions did

not typically pose significant redistributive threats. Rather, the goal of such elite rebels was generally to gain regional independence. Nonetheless, we control for elite rebellions in our regression analysis ahead. Appendix Figure B-1 maps the elite rebellion locations, which were far less common than mass rebellions.

Following the seminal work by Chen (2007 [1940], 3), we define external warfare as a violent conflict between a China-based dynasty and a non-Han state or state-like power.[8] There were 541 individual battles linked to 334 recorded external wars during our sample period. Appendix Figure B-2 provides a map of these external battles. Consistent with the evidence cited in Section 3, the geographical pattern suggests that many external conflicts were fought against nomads from the Eurasian Steppe. In our regression analysis ahead, we will control for external warfare.

Appendix Table C-1 breaks down the distribution of conflict types in our sample.

4.2 Genealogy Books

Genealogy records are a rich, yet under-analyzed, source of demographic information in Ming-Qing China, a historical period for which alternative data sources are not widely available (Shiue 2016, 459). Our genealogy data draw on Wang (2008), who has cataloged roughly 51,200 genealogy books from the end of the first millennium to the present day in a print registry. This effort represents the most comprehensive registry of known Chinese clan genealogies to date (Greif and Tabellini 2017, 2). Wang's team spent eight years gathering genealogy records from all known sources, including local and national archives and libraries, private holdings, and overseas collections (Wang 2008, 8-9).[9]

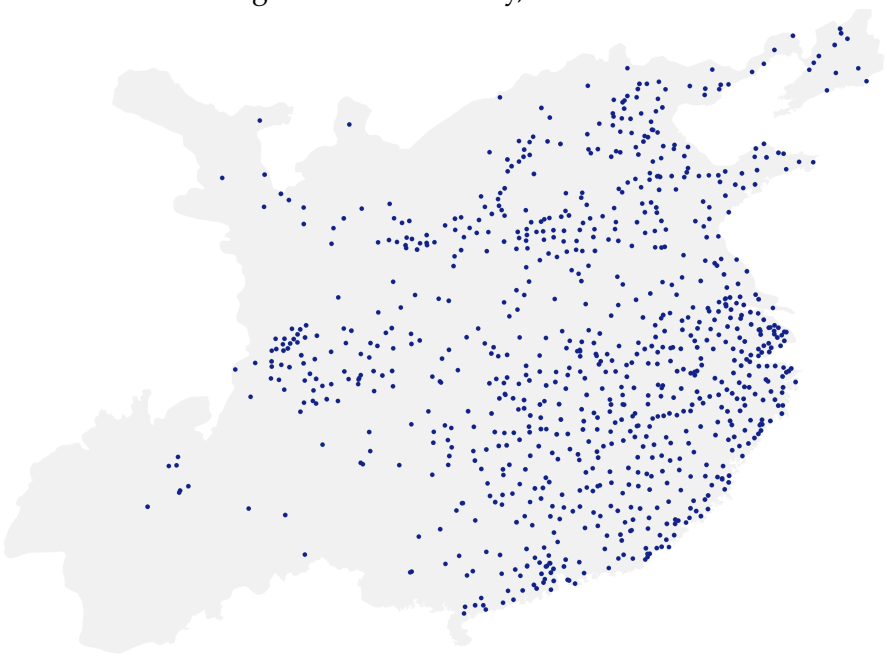
Each entry in Wang's registry reports a record of a clan's genealogy book, including the year in which it was compiled. A clan may have had multiple registry entries. For example, the Li clan based in the city of Taiyuan compiled its first genealogy book in 1701 (entry 1), which it then updated in 1754 (entry 2) and 1802 (entry 3), for a total of three genealogy books. Each entry also includes information on the clan's surname and current location.

We digitized this entire print registry, and geocoded each genealogy book based on its

[8]Thus, for example, battles between the Manchu invaders and the imperial Ming state were categorized as "external," while those between the subsequent imperial Qing (i.e., Manchu) state and mass rebel groups were categorized as "internal."

[9]Wang's registry includes all 10,000 microfilmed genealogy records archived by the Genealogical Society of Utah – the largest overseas collection of Chinese genealogy.

Figure 2: Clan Activity, 1350-1900



Notes. This figure shows the location of each recorded genealogy book written down in China between 1350-1900 within Ming-era borders.

reported location. To the best of our knowledge, this geocoding is the first such effort of its kind. We first used optical character recognition software to read the entire registry into a Microsoft Excel file. Next, with the help of research assistants, we manually checked each entry in order to ensure accuracy. Finally, we relied on the *China Historical Geographic Information System* (2018) for latitudes and longitudes for the purposes of geocoding. Figure 2 maps the locations of recorded genealogy books written down for our sample period. Consistent with previous qualitative evidence (e.g., Freedman 1958, 129), the geographical pattern suggests that historical clan activity was more prevalent in the South than in other parts of China.

We believe that the genealogy book data provide the most systematic and best available proxy – even if imperfect – for documenting long-run trends in clan activity in imperial China. In this respect, our approach builds off of the precedent in Greif and Tabellini (2017, 5-7). As described in Section 3, moreover, historians argue that genealogical compilation by the gentry was one common kin-oriented response to violent conflict. Thus, we are confident that locales that produced more genealogy books did in fact experience greater clan activity.

Still, data concerns including measurement error remain. The compilation of genealogy

books may have been sensitive to the availability of printing materials, changing macroeconomic conditions, and migration patterns. Our regression analysis ahead will account for such potential confounders through grid cell and period fixed effects and county-specific time trends.

Furthermore, elites may have found it difficult to compile genealogy books during internal conflicts, and these books may have been less likely to survive and be cataloged. According to the available narrative evidence, it appears that the gentry financed the compilation of genealogy books in proportion to the relative size of their land holdings (Beattie 1979, 94-5). Since most clan members did not hold much if any land, the per capita cost of book production should have been low. Thus, whether a clan decided to compile its genealogy records should have depended primarily on the gentry's need to formalize kin networks, rather than the cost to the average clan member. Greater investments in defense preparedness in response to internal conflicts, therefore, did not automatically require a clan to forgo investments in genealogy books.

Additionally, as our regression analysis ahead will show, we find a *positive* and significant relationship between elite rebellions – which, like mass rebellions, could be physically destructive – and the number of genealogy books. Similarly, we find a significant *increase* in the number of genealogy books in zones that experienced more mass rebellions in the post-1850 geopolitical context. If internal conflicts necessarily made the compilation and maintenance of genealogy books less likely, then we would not expect to observe either result.

4.3 State Development

There is somewhat of a lack of systematic historical data available on the central government's local strength. Given this limitation, we are still able to unearth several measures of imperial state power.

We assess the central state's infrastructural reach (Mann 1984) in two ways. Upon taking power, the imperial Ming state embarked on an ambitious garrison construction plan, in great part to help suppress mass revolt (Downing 1992, 50). To connect the capital of Beijing with the military garrisons, as well as to link the garrisons themselves, the Ming constructed a network of courier routes. Both the military garrisons and the courier routes were built primarily during the Ming's first century of rule (1368-1467). We thus geocode data on the location of each new military garrison and courier route over the 1368-1467 period accord-

ing to Yang (2006) and the *China Historical Geographic Information System* (2018). Appendix Figures B-4 and B-5 map the garrison and courier route locations, respectively.

Second, to assess the central state's fiscal reach (Besley and Persson 2011), we employ data on the extent of local land taxation in 1820 under the imperial Qing state from Sng (2014).[10] The land tax, collected by local bureaucrats and remitted to the central government, was the most important source of tax revenue in imperial China.[11] The year 1820, moreover, is the cross-section of these data that historians regard to be the most systematic and accurate (Liang 2008, 555-71). Appendix Figure B-6 displays total land taxation across Qing prefectures in this year.

Due to the dearth of historical data, the state development outcomes above are all cross-sectional in nature. Nevertheless, since the infrastructural power data are from the Ming era, while the fiscal capacity data are from the Qing era, our empirical analysis ahead can still provide novel insights into whether the relationship between internal conflict and state development held across different time periods of traditional imperial rule.

4.4 State Failure

Finally, to proxy for state failure at the start of the twentieth century, we geocode the location of each local military group that declared independence from the imperial Qing state in 1911 according to Guo (2015). Appendix Figure B-7 shows the locations of these military groups.

4.5 Database Construction

4.5.1 Main Database

We use the data gathered above to construct several novel databases. Our main database is a panel that conjoins the geocoded data on mass rebellions and genealogy books at the grid cell level between 1350 and 1900. We first divide mainland China's territory into several thousand grid cells. The virtue of this approach is that we can "exogenously" impose grid cells on imperial China's territory, unlike internal political borders (e.g., 1990 county bor-

[10] We focus on total land taxation (rather than per capita land taxation) since the Qing state did not take into account local population size when deciding the extent of local land taxation. Following the bundling of all local taxation into the land tax in the 1720s and 1730s by the Yongzheng Emperor, both the local tax quota and tax rate were wholly determined by local soil fertility and topographical conditions (Wang 1973, 32).

[11] The Kangxi Emperor froze the labor service tax quota in 1712. However, he left the land tax quota unfrozen (Spence 2002, 178). Revenue from land taxation continued to grow over the eighteenth century, particularly after the incorporation of labor service surcharges and taxes into the land tax under the Yongzheng Emperor during the 1720s and 1730s (Zelin 2002, 202).

ders), which may have been endogenously determined. We choose 25 km x 25 km grid cells as the benchmark because this size is relatively close to the average size of the township, the smallest administrative unit in China. Appendix Figure B-3 maps these grid cells. For robustness, we employ an alternative grid size of 50 km x 50 km. As a further alternative, we employ 1990 county borders. The main results remain similar across both alternative demarcations (Appendix Tables D-6 and D-7).

For the panel analysis, we restrict the grid cells to those which fall within Ming-era external borders for three reasons. First, given that the start year of this analysis (i.e., 1350) corresponds with the approximate establishment of the Ming Dynasty, these are the "initial" borders. Furthermore, an accurate shapefile of the Ming-era external borders is in fact available (*China Historical Geographic Information System* 2018). Finally, the imperial Ming state was small relative to its successor (i.e., the Qing Dynasty), and thus serves as an "Inner China" that forms a continuous spatial core across the Ming and Qing eras. Still, we show that our main results are robust to other border configurations, including the use of all grid cells within the Qing-era external borders (Appendix Table D-8).

We then divide the 1350-1900 period into 50-year periods. This interval length makes sense since genealogy books were typically revised roughly every half-century, ceteris paribus. Traditionally, there was the expectation that a clan should update its genealogy book every three generations (Feng 2006, 67). Given that males typically married and had their first child in their late teens (Chaffee 1989, 345), three generations translates into $3 \times 18 = 54$ years, or roughly one half-century. Furthermore, this interval length provides a good deal of variation in both the number of mass rebellion battles and genealogy books written down.

Overall, the unit of analysis in our main database is grid cell-period, with $N = 5,803$ (25 km x 25 km grid cells) and $T = 11$ (50-year periods).

4.5.2 Other Databases

As described above, the next two databases are cross-sectional in nature. The second database conjoins the geocoded data on mass rebellions with the infrastructural power data (i.e., military garrisons and courier routes) at the grid cell level during the first century of Ming rule. Once more, $N = 5,803$ (25 km x 25 km grid cells). The third database conjoins the geocoded conflict data with the fiscal capacity data (i.e., land taxation) in 1820 at the prefectural level under Qing rule. Here, we take Qing-era prefectures as the unit of analysis rather than grid

cells, since the land tax data are available at this level ($N = 318$).

Finally, we conjoin the geocoded data on clan activity in the two decades prior to the downfall of the imperial Qing state with the data on independence declarations from the Qing in 1911 at the grid cell level. For this part of the analysis, it makes sense to use the Qing-era external borders (which were significantly larger than during the Ming era), increasing the sample size to $N = 15,103$ (25 km x 25 km grid cells). However, we show that these results remain robust if we restrict the sample to the Ming-era borders (Appendix Table F-1).

5 Mass Rebellion and State Development Before 1850

To systematically analyze the relationship between mass rebellion and state development in imperial China, we now undertake a regression analysis. In this section, we focus on the traditional geopolitical context prior to Britain's victory over China in the First Opium War. Here, our theoretical framework predicts that we should observe that the central state strengthens its leadership role vis-à-vis local elites (via clans) in elite-state cooperative efforts in response to mass rebellion. We now provide two main types of evidence in support of this empirical prediction. First, we show evidence for a negative relationship between mass rebellion and clan activity over time. Second, we show evidence for a positive relationship between mass rebellion and state development outcomes. Taken together, we view this evidence as consistent with the argument that, traditionally, mass rebellion weakened any leadership role that local elites may have played in elite-state cooperative efforts and strengthened the leadership of the state.

5.1 Clan Activity

5.1.1 Methodology

We estimate the following benchmark OLS specification:

$$ClanActivity_{i,t} = \alpha + \beta Rebellion_{i,t} + \mu_i + \lambda_t + \epsilon_{i,t}. \quad (1)$$

The dependent variable $ClanActivity_{i,t}$ reflects clan activity in each 25 km x 25 km grid cell i over 50-year period t as proxied by the number of genealogy books written down there.[12] The variable of interest $Rebellion_{i,t}$ measures the number of mass rebellions in each grid cell per 50-year period.[13] μ_i and λ_t are grid cell and period fixed effects, respectively. $\epsilon_{i,t}$ is a random error term. All standard errors are robust, clustered at the cell level to account for any within-cell serial correlation in the error term. However, the main results continue to hold if we cluster the standard errors at higher levels of aggregation including the county or prefectural level (Appendix Tables D-9 and D-10). Appendix Table C-2 displays the summary statistics for all the regression variables used in our analysis.

The genealogy data increase in mean and variance across time, particularly before and after 1850. We thus take the inverse hyperbolic sine (IHS) of the number of genealogy books as the dependent variable.[14] This transformation reduces the range of the mean and variance of $ClanActivity_{i,t}$, and allows us to make use of all observations, since it is defined at zero. However, the main results remain robust if we take $\ln(1 + GenealogyBooks_{i,t})$ rather than the IHS, or keep the dependent variable in its original linear form (Appendix Table D-1). Beyond this transformation, we exclude the post-1850 period from our current analysis (we will include it in Section 6). Finally, given that the mean and variance of $Rebellion_{i,t}$ does not display any obvious increase over time, we keep this variable in its original linear form.

Unobserved local or temporal features may have affected both mass rebellions and clan activity alike. For example, rough terrain may have promoted internal conflict by enabling peasant rebels to hide from state forces, and by allowing the gentry to evade the rebels themselves (Fearon and Laitin 2003, 84). To help account for such features, our analysis always include grid cell and period fixed effects. Grid cell fixed effects help control for local initial conditions, such as the initial demographic and economic conditions of a locality, as well as local features that are time-invariant including local geography, such as soil quality, terrain ruggedness, natural resources, distance from major rivers, distance from the coast, and distance from Beijing. Similarly, period fixed effects help control for widespread shocks spe-

[12] As we will describe ahead, the combination of grid cell fixed effects and county-specific trends help account for size differences in local populations. Thus, $ClanActivity_{i,t}$ should proxy for actual clan activity *net* of any local population size differences.

[13] To account for the fact that the influence of mass rebellion on clan activity may not be contemporaneous, we repeat the main analysis for the lagged variable of interest $Rebellion_{i,t-1}$. The results remain similar to the main ones (Appendix Table D-5).

[14] Namely, $ClanActivity_{i,t} \equiv \ln[GenealogyBooks_{i,t} + (GenealogyBooks_{i,t}^2 + 1)^{1/2}]$.

cific to each 50-year interval, such as changing demographic or macro-economic conditions, trends in cultural conservatism, developments in the technology of state repression, and the possibility of dynastic cycles.

Nonetheless, unobserved time-varying features may still affect the results. To address this possibility, we add county-specific linear time trends, which help account for unobservable features specific to each county that may have changed linearly over time. Such features include local demographic trends, along with local trends in the economy, urbanization, and culture. The counties are larger than the 25 km x 25 km grid cells. On average, there are roughly 3 grid cells per county, and thus 1,547 counties overall (i.e., 1,547 county-specific trends).[15] We further test the robustness of the results in a variety of ways, which we will describe ahead.

5.1.2 Main Results

Benchmark Specification

Columns 1 to 4 of Table 1 show the main estimation results for the relationship between mass rebellion and clan activity in China between 1350 and 1850. The benchmark specification controls for time-invariant local features and widespread time-specific shocks through fixed effects (column 1). The coefficient estimate for $Rebellion_{i,t}$ is negative in sign and significant, with value -0.068.

County-Specific Trends

To help account for unobserved changes over time in local features (e.g., demographic patterns), column 2 adds county-specific trends to the benchmark specification. The coefficient estimate for $Rebellion_{i,t}$ remains significant, with value -0.046.

Elite Rebellion and External Warfare

Similarly, to account for other time-varying observable features, we control for elite rebellion and external warfare battles in a given grid cell in column 3. The coefficient estimate for mass rebellion is once more negative in sign and significant. By contrast, there is a significant *positive* relationship between elite rebellion and clan activity. This result is consistent with the logic that local elites strengthened their kinship networks in response to elite rebellion.

[15] Given data limitations, we rely on the Qing-era county shapefile from *China Historical Geographic Information System* (2018) to identify historical county borders.

Finally, the coefficient estimate for external warfare is negative and significant. This result suggests that, in line with the literature on state-making (e.g., Tilly 1975), external warfare helped weaken any leadership role by local elites (vis-à-vis the central government) in elite-state security cooperation.

Ming-Qing Dynastic Change

The main analysis accounts for widespread shocks by including period fixed effects. Still, the dynastic change from the Ming to the Qing during the mid-1600s may have been a critical shock to the viability of the imperial system. Appendix Figure B-8 depicts the relationship between mass rebellion and the Ming-Qing dynastic change. It indicates that mass rebellions were highest in the late Ming era between 1620 and 1649 (top panel). Most of these battles pitted the imperial Ming state against the Li Zicheng rebels. However, there was no enduring increase in clan activity either during or after the 1620-49 period of rebellions. As we will describe in Section 6, this result stands in contrast to the consequences of the mid-nineteenth century Taiping Rebellion, in which there was large and persistent increase in clan activity (Figure 3). To systematically analyze clan activity in response to the Ming-Qing dynastic change, we interact $Rebellion_{i,t}$ with the period fixed effect for 1600-50 in column 4 of Table 1. The interaction term is not significant. Taken together, this evidence further suggests that, in line with the historical account in Section 3, the Ming-Qing dynastic change did not fundamentally alter the viability of the imperial system.

Summary

Overall, the Table 1 results indicate that, traditionally, there was a negative and significant relationship between mass rebellion and clan activity. According to the coefficient estimate in column 2, each additional mass rebellion battle was associated with a 5 percent average *decrease* in clan activity (as proxied by the number of genealogy books written down) per 50-year period between 1350 and 1850.

Table 1: Mass Rebellion and Clan Activity

Dependent variable:	Genealogy Books (IHS)						
	1350-1850				1350-1900		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mass rebellion	-0.068***	-0.046***	-0.035***	-0.037***	-0.072***	-0.051***	-0.063***
	(0.014)	(0.011)	(0.011)	(0.012)	(0.012)	(0.012)	(0.014)
Mass rebellion × Post-1850					0.533***	0.321***	0.525***
					(0.065)	(0.068)	(0.065)
Elite rebellion			0.211***			0.228***	
			(0.035)			(0.038)	
Elite rebellion × Post-1850						0.144***	
						(0.013)	
External war			-0.034**			-0.039**	
			(0.015)			(0.015)	
External war × Post-1850						0.044	
						(0.109)	
Mass rebellion × 1600-50 (Ming-Qing)				-0.027			-0.030
				(0.026)			(0.030)
Mass rebellion × 1750-1800 (White Lotus)							0.049
							(0.050)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-specific trends	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.045	0.295	0.302	0.295	0.337	0.359	0.337
Observations	58030	58030	58030	58030	63833	63833	63833

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period in columns 1-4 is 1350-1850. Sample period in columns 5-7 is 1350-1900. Dependent variable in all columns is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest in all columns is number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

5.1.3 Robustness Checks

Instrumental Variables

To instrument for mass rebellion, we exploit exogenous variation in the frequency of droughts. This approach is similar in spirit to Kung and Ma (2014) and Jia (2014*b*), both of whom document a significant correlation between negative climate shocks and peasant rebellions in Chinese history.

Ideally, we would like drought data that span all of Ming-Qing China over the entire sample period at the grid-cell level. The best available data, however, are from the State Meteorological Society (1981), which provides annual information on climate conditions for 88 prefectures in Eastern and Southern China from the late fifteenth century to the 1970s. This source gathered data from local gazetteers, official documents, and weather stations. Our main instrument computes the number of years in each 50-year period between 1500 and 1850 in which a prefecture experienced extreme drought. Alternatively, we re-compute the instrument for mild *or* extreme drought. We next match the drought data at the prefectural level to each 25 km x 25 km grid cell according to the *China Historical Geographic Information System* (2018). On average, grid cells in this sub-sample experienced nearly 3 years of extreme drought and roughly 9 years of mild or extreme drought each half-century (Appendix Table C-2).

Appendix Table D-2 reports the results of the IV analysis. There is a positive and significant relationship between drought and mass rebellion at the first stage. The Kleibergen-Paap F-statistics exceed 10, indicating instrument strength. The second-stage coefficient estimates for $Rebellion_{i,t}$ are positive and significant. Here, the IV estimates are larger than the OLS estimates. One possible explanation for this difference is that sub-sample with historical drought data are concentrated in the parts of China where genealogical books were more prevalent, strengthening the magnitude of the estimated relationship between mass rebellion and clan activity.

Lagged Dependent Variable

The main analysis controls for unobserved regional patterns over time by including county-specific trends. Nonetheless, clan activity in period t may still influence the chance of mass rebellion in period $t+1$. To explicitly account for the role of previous clan activity, Appendix Table D-3 includes the lagged dependent variable $ClanActivity_{i,t-1}$ as an indepen-

dent regressor.[16] The coefficient estimates for $ClanActivity_{i,t-1}$ are positive and significant, indicating that clan strength in period t was partly a function of previous clan activity. While the coefficient estimates for $Rebellion_{i,t}$ fall somewhat in magnitude relative to the main results, they remain negative in sign and significant across both specifications.

Exclude Provinces and Periods

To further account for unobserved features across place and time, we exclude each province and each period, respectively, one by one (Appendix Figures D-1 and D-2). No single province or period drives our results.

Spatial Spillovers

To help account for potential spatial spillovers, we control for mass rebellions in neighboring grid cells in Appendix Table D-4. The main results remain robust. The coefficient estimates for $RebellionNeighbor_{i,t}$ are also negative and significant. Their magnitudes are very small, however, suggesting that mass rebellion within a given grid cell was of far greater importance to local clan activity than that in neighboring cells.

Other Checks

Finally, as described previously, we also test: (1) two alternative specifications of the dependent variable (Appendix Table D-1); (2) the use of the lagged variable of interest $Rebellion_{i,t-1}$ (Appendix Table D-5); (3) the alternative grid cell size of 50 km x 50 km (Appendix Table D-6); as well as (4) 1990 county borders (Appendix Table D-7); (5) the inclusion of all 25 km x 25 km grid cells within Qing-era external borders (Appendix Table D-8); and (6) the alternative clustering of standard errors (Appendix Tables D-9 and D-10).[17] The results are robust across all the above checks.

5.2 State Development

The results above support the empirical prediction that, traditionally, mass rebellion weakened any leadership role that local elites may have played in elite-state cooperative efforts. To complement this evidence, we now analyze the relationship between mass rebellion and

[16] Including the lagged dependent variable induces asymptotic bias of order $1/T$. Given that $T = 10$ in our pre-1850 panel, however, Nickell bias should be relatively small.

[17] For the tests for the 50 km x 50 km grid cells and 1990 county borders, we replace the county-specific trends with province-specific trends (i.e., a higher-level administrative unit), given the larger sizes of these zones relative to the benchmark 25 km x 25 km grid cells.

state development in the traditional (i.e., pre-1850) context.

5.2.1 Military Garrisons and Courier Routes

We first evaluate the relationship between mass rebellion and the central state's infrastructural power, as proxied by the construction of military garrisons and courier routes during the first century of Ming rule. We use OLS to estimate:

$$Infrastructure_{i,1368-1467} = \alpha + \beta Rebellion_{i,1368-1467} + \gamma_j + \epsilon_j. \quad (2)$$

The dependent variable $Infrastructure_i$ measures one of two outcomes. For reasons similar to those described above, the first takes the inverse hyperbolic sine of the number of garrisons built in each 25 km x 25 km grid cell i between 1368-1467.[18] The second is a binary indicator that equals 1 for the presence of a courier route in a grid cell over the same time period. Our variable of interest $Rebellion_{i,1368-1467}$ is the number of mass rebellions in grid cell i over the 1368-1467 period.

Given that Equation 2 uses cross-sectional data, we address the possibility of omitted variable bias by including county fixed effects γ_j, which help control for institutional features common to each administrative unit in which sub-groups of grid cells fall into. To address the potential interdependence of grid cells within the same county, we cluster standard errors at the county level in the main analysis, and at the prefectural level for robustness (Appendix Table E-2).

Columns 1 and 2 of Table 2 display the results of this analysis when the infrastructural power outcome is military garrisons. Column 1 shows the bivariate correlation between mass rebellion and the establishment of Ming-era garrisons, while column 2 adds county fixed effects. The coefficient estimate for $Rebellion_{i,1368-1467}$ is positive in sign and significant across both specifications, with values between 0.161 and 0.152. The magnitude of the column 2 estimate suggests that each additional mass rebellion battle during the first century of Ming rule was associated with an increase in the likelihood of having a military garrison of 16-17 percent.

Similarly, columns 3 and 4 show the results for Ming-era courier routes. There is a positive and significant relationship between mass rebellion and the presence of Ming-era

[18] The results remain robust, however, if we take $\ln(1 + Garrisons_{i,1368-1467})$, or keep this dependent variable in its original linear form (Appendix Table E-1).

Table 2: Mass Rebellion and Ming State Development

Dependent variable:	Garrisons (IHS)		Courier Route	
	(1)	(2)	(3)	(4)
Mass rebellion (1368-1467)	0.161***	0.152***	0.148***	0.110**
	(0.039)	(0.044)	(0.035)	(0.044)
County FE	No	Yes	No	Yes
R^2	0.024	0.422	0.006	0.539
Observations	5147	5147	5147	5147

Notes. Estimation method is OLS. Unit of analysis is 25 km x 25 km grid cell. Sample period is first century of Ming Dynasty (1368-1467). Dependent variable in columns 1-2 is early Ming military garrisons as proxied by the inverse hyperbolic sine (IHS) of the number of garrisons. Dependent variable in columns 3-4 is binary indicator of Ming-era courier route. Variable of interest in all columns is number of mass rebellions over 1368-1467. Robust standard errors clustered at county level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

courier routes across both specifications. The coefficient value for $Rebellion_{i,1368-1467}$ in column 2 suggests that each additional mass rebellion battle during the Ming era was associated with an increase in the likelihood of having a courier route of 11 percent.

Overall, the results for the Ming-era measures of infrastructural power are consistent with our empirical prediction that, traditionally, there was a positive and significant relationship between mass rebellion and state development.

5.2.2 Land Taxation

We next evaluate the relationship between mass rebellion and the central state's fiscal capacity, as proxied by Qing-era land taxation. We again use OLS to estimate:

$$Taxation_{i,1820} = \alpha + \beta Rebellion_{i,1644-1819} + \omega PopDensity_i + \zeta_z + \epsilon_z. \quad (3)$$

The dependent variable $Taxation_i$ takes the inverse hyperbolic sine of the total land tax (in silver kg) in prefecture i in 1820, enabling us to incorporate a few null observations.[19] Our variable of interest $Rebellion_{i,1644-1819}$ is the number of mass rebellions in prefecture i over the 1644-1819 period.

To help control for common institutional features among sub-groups of prefectures, we

[19]Following Sng (2014, 115), we exclude seven prefectures near Beijing (i.e., Baoding, Chengde, Shuntian, Tianjin, Yongping, Xuanhua, and Zunhua), since many farmers there directly paid rents to the central state rather than taxes. However, the results are robust if we include them.

Table 3: Mass Rebellion and Qing State Development

Dependent variable:	Total Land Tax (IHS)		
	(1)	(2)	(3)
Mass rebellion (1644-1819)	0.248*	0.234**	0.187**
	(0.122)	(0.086)	(0.070)
Provincial FE	No	Yes	Yes
Population density	No	No	Yes
R^2	0.017	0.487	0.560
Observations	258	258	258

Notes. Estimation method is OLS. Unit of analysis is Qing prefecture. Sample period is start of Qing Dynasty to 1820. Dependent variable is the inverse hyperbolic sine (IHS) of total land tax in silver kg in 1820. Variable of interest is number of mass rebellions over 1644-1819. Robust standard errors clustered at provincial level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

include provincial fixed effects ζ_z.[20] In particular, provincial fixed effects help control for local initial economic conditions, including initial fiscal capacity. Similarly, to help account for local demographic features, we control for the population density of each prefecture according to Sng (2014). We cluster standard errors at the provincial level to address the potential interdependence of prefectures within the same province. Here, we exclude all provinces located in peripheral zones of Qing China for which land tax data were not available (Appendix Figure B-6).

Table 3 displays the results. Column 1 shows the bivariate correlation, column 2 adds provincial fixed effects, and column 3 adds population density. Across all three specifications the coefficient estimate for $Rebellion_{i,1644-1819}$ remains positive in sign and significant, with values between 0.187 and 0.248. The magnitude of the estimate in column 3 suggests that each additional mass rebellion battle during early nineteenth-century Qing rule was associated with 0.21 silver kg increase (≈$124) in total land taxation.

Overall, these results suggest that there was an increase in the imperial Qing state's fiscal capacity in response to mass rebellion.

6 Mass Rebellion and State Development After 1850

The results in the previous section support our first empirical prediction that, traditionally, mass rebellion was associated with less clan activity and greater state development in im-

[20]To identify Qing-era administrative units, we use the shapefile from *China Historical Geographic Information System* (2018).

perial China. We now turn to our second empirical prediction of our theoretical framework, namely that in the new geopolitical context (i.e., in the aftermath of China's defeat by Britain in the First Opium War and the resulting Treaty of Nanjing), we should observe a shift in leadership in elite-state cooperation away from the central state and toward local elites via clans in response to mass rebellion.

There are several phenomena specific to the nineteenth century that we must account for. Traditionally, woodblock printing was standard, with printing centers located near raw materials such as bamboo and pine (Brokaw and Chow 2005, 10, 79). The introduction of modern printing technology in the 1870s, however, led to the establishment of new printing centers (e.g., Shanghai, Tianjin), potentially reducing the costs of publishing genealogy books (Brokaw and Reed 2010, 6-7). Similarly, the nineteenth century saw the arrival of Protestant missionaries, who helped spread mass printing, newspapers, and mass education, each of which could affect mass rebellion and clan activity alike (Woodberry 2012). Finally, Jia (2014*a*) shows that treaty ports conceded by the Qing to the West over the nineteenth century experienced faster subsequent growth. In the analysis ahead, we rely on period fixed effects and county-specific linear time trends to help control for nineteenth-century phenomena including modern printing, missionary activity, and treaty ports. We also perform a falsification exercise (to be described ahead).

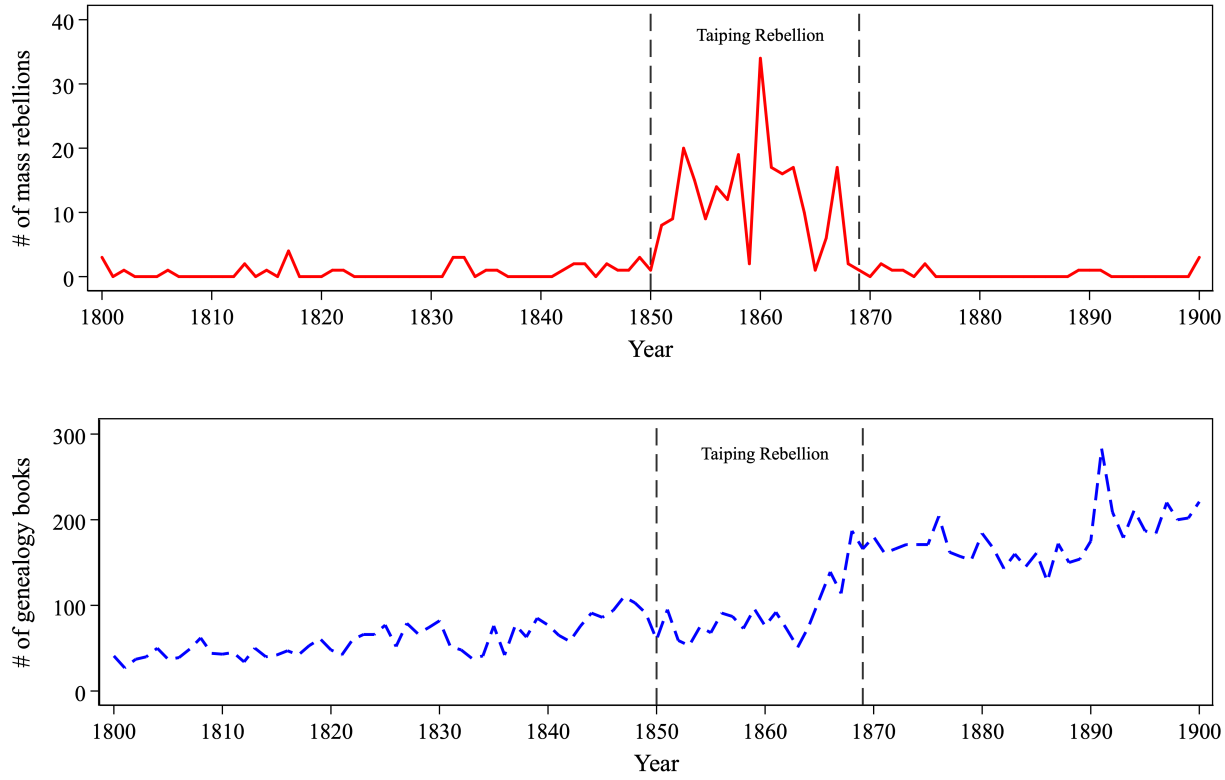
6.1 Main Results

The last three columns of Table 1 repeat the regression analysis based on Equation 1 for the entire 1350-1900 period. In column 5, we interact $Rebellion_{i,t}$ with the period fixed effect for 1850-1900. Relative to the constitutive term for (i.e., pre-1850) mass rebellion, the interaction term $Rebellion \times Post$-1850 switches signs: in the new geopolitical context, the relationship between mass rebellion and clan activity is *positive* and significant. The magnitude of the coefficient estimate for this interaction term suggests that each additional post-1850 mass rebellion battle was associated with a 59 percent *increase* in clan activity.

Figure 3 helps illustrate the above result. It indicates that mass rebellions peaked between 1850 and 1869 (top panel). There were 230 mass rebellion battles over this period, of which nearly 60 percent involved the Taiping.[21] Similarly, there was a level-step increase in clan activity in the aftermath (bottom panel). The number of genealogy books rose from

[21]The remainder involved other peasant rebel groups including the Nian.

Figure 3: Mass Rebellion and Clan Activity, 1800-1900



Notes. This figure shows the annual number of mass rebellion battles (top panel) and genealogy books written down (bottom panel) in China between 1800 and 1900. The (red) solid line indicates the number of mass rebellions per year, and the (blue) dashed line indicates the number of genealogy books.

less than 100 per year before 1850 to nearly 200 by 1870. This descriptive evidence provides further support for our empirical prediction that renewed clan development took place in response to mass rebellion in the new geopolitical context.

6.2 Falsification Exercises

To test the robustness of the column 5 result, we perform two sorts of falsification exercises. The first exercise (column 6 of Table 1) interacts the control variables for elite rebellion and external warfare battles, respectively, with period fixed effects for 1850-1900. Both the constitutive term for elite rebellion and the interaction term $EliteRebellion \times Post$-1850 are positive and significant. That is, unlike the interaction term for post-1850 mass rebellion, there is no sign switching. The interaction term $ExternalWar \times Post$-1850 does in fact switch signs relative to the constitutive term for external warfare. Unlike the interaction term for

post-1850 mass rebellion, however, this interaction term is small in magnitude and not significant. This exercise offers additional evidence that the positive and significant interaction term for post-1850 mass rebellion is not simply due to unobserved features specific to the nineteenth century (e.g., the introduction of modern printing).

The second exercise (column 7 of Table 1) includes interaction terms for two pre-1850 shocks of domestic upheaval: (1) the Ming-Qing dynastic change as described in Section 5.1; and (2) the White Lotus Rebellion (1796-1804), the suppression of which significantly reduced the central government's revenue reserves (Ma and Rubin 2019, 289).[22] Neither interaction term is significant, and the main result for $Rebellion \times Post$-1850 remains unchanged. This exercise provides further evidence that the positive relationship between mass rebellion and clan activity after 1850 stands out relative to previous shocks of domestic upheaval.

Overall, the results in this section are consistent with our empirical prediction that, following the new geopolitical context, mass rebellion was associated with a shift in leadership in elite-state cooperation away from the central state and toward local elites via clans.

6.3 Independence Declarations

A final (i.e., ancillary) empirical prediction of our theoretical framework is that greater clan development in the new geopolitical context could eventually prompt state failure. As described in Section 3, greater reliance on local elites as the locus of elite-state cooperation, coupled with renewed clan activity, altered the traditional security balance that favored a prominent leadership role by the central state, since now the gentry began to mobilize greater resources and increase their political autonomy.

Recall from Section 4 that we proxy for state failure at the start of the twentieth century by geocoding the location of each local military group that declared independence from the imperial Qing state in 1911. We use this information to create a binary indicator variable $Independence_{i,1911}$ that equals one if there was at least one independence declaration by a local elite group within the borders of (25x25) grid cell i.

To evaluate the extent to which local independence declarations were a function of re-

[22] Specifically, to account for the White Lotus Rebellion, we interact $Rebellion_{i,t}$ with the period fixed effect for 1750-1800.

Table 4: Clan Activity and Qing State Failure

Dependent variable:	Declaration of Independence in 1911	
	(1)	(2)
Genealogy Books (IHS)	0.045***	0.044***
	(0.007)	(0.009)
County FE	No	Yes
R^2	0.054	0.382
Observations	15103	15103

Notes. Estimation method is OLS. Unit of analysis is 25 km x 25 km grid cell. Dependent variable is binary indicator of formal declaration of independence from imperial Qing state in 1911. Variable of interest is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books, between 1890-1909. Robust standard errors clustered at county level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

newed post-Taiping clan activity, we use OLS to estimate:

$$Independence_{i,1911} = \alpha + \beta ClanActivity_{i,1890-1909} + \gamma_j + \epsilon_j. \quad (4)$$

Here, our variable of interest is post-Taiping clan activity $ClanActivity_{i,1890-1909}$ between 1890-1909. Given that this test uses cross-sectional data, we address the possibility of omitted variable bias by including county fixed effects γ_j.

Table 4 displays the results of this analysis. Column 1 shows the raw bivariate correlation, and column 2 adds county fixed effects. Bai and Jia (2016) argue that counties that had higher quotas for the imperial civil service exam were more likely to experience revolutionary uprisings once this system was abolished in 1905. Here, the county fixed effects control for civil service exam quotas. The coefficient estimate for $ClanActivity_{i,1890-1909}$ is positive in sign and significant across both specifications, with values from 0.044 to 0.045. This result is similar if we restrict the sample to Ming-era external borders (Appendix Table F-1), or cluster standard errors at the prefectural rather than the county level (Appendix Table F-2).

Overall, this analysis provides additional support for our framework's final predication, that greater clan development in the new geopolitical context could eventually prompt state failure.

7 Conclusion

In this study, we have analyzed the long-run dynamics of internal conflict, elite-state cooperation, and state development outcomes in China. We have argued that whether internal conflict improves or reduces the central state's monopoly over violence depends on the geopolitical context. Traditionally, the public provision of security was cost-effective enough to be attractive to the gentry, and the cost of new public investments in defense was low enough for the central state to make them. Thus, state development was likely to occur in response to mass rebellion. In the new geopolitical context (i.e., in the aftermath of China's defeat by Britain in the First Opium War and the resulting Treaty of Nanjing), however, the cost of new defense investments greatly increased, making greater reliance on local elites as the locus of elite-state cooperation more attractive to the central state. Clan development therefore became more likely in response to mass rebellion. The central state's loss of the monopoly over violence, however, could eventually promote state failure.

To evaluate the predictions of our argument, we have exploited a variety of novel databases for imperial China at the micro level. We have shown evidence that, traditionally, mass rebellion was associated with a greater leadership role by the central state in elite-state cooperative efforts in response to mass rebellion. This dynamic changed, however, in the new geopolitical context. We have found that, in the new geopolitical context, mass rebellion was associated with movement toward local elites via clans. Finally, we have shown evidence for a positive relationship between renewed clan activity in the post-Taiping era and eventual state failure.

To the best of our knowledge, this study is among the first to provide both a theoretical logic and systematic evidence that, depending on the specific cost conditions of elite-state cooperation, internal conflict can improve or reduce the central state's monopoly over violence. Our study suggests that the trajectory of state development in imperial China proceeded in the same direction as that of Western Europe – with central governments in both parts of Eurasia gaining a greater monopoly over violence – into the nineteenth century. Only over the second half of the nineteenth century did the central state in China began to lose control over its monopoly over violence, while nation-states in Western Europe steamed ahead, exacerbating East-West political divergence.

One may ask why major external threats in European history such as the risk of invasion

by Napoleon at the start of the nineteenth century did not engender a similar shift in the long-run power balance within states toward local elites. We believe that a key difference was the unique family structure in medieval and early modern Europe. There, church actions undercut traditional kinship groups, and the nuclear family became dominant (Greif 2006, 308-9). Thus, in response to violent conflict, elites in early modern Europe did not have the same option to collaborate with kin members to protect themselves as in imperial China (or, at least not to the same extent). Rather, the main option in the European context was to rely on cooperative efforts led by the central state. We therefore believe that our framework can account for this case. Finally, as described above, the European nuclear family structure was exceptional. Strong kin networks were common across many late modernizers. Thus, in our view, the insights that we glean from the case of imperial China should generalize to a variety of other non-European contexts.

References

Acemoglu, Daron, James Robinson and Rafael Santos. 2013. "The Monopoly of Violence: Evidence from Colombia." *Journal of the European Economic Association* 11(S1):5–44.

Bai, Ying and James Kai-sing Kung. 2011. "Climate Shocks and Sino-Nomadic Conflict." *Review of Economics and Statistics* 93(3):970–981.

Bai, Ying and Ruixue Jia. 2016. "Elite Recruitment and Political Stability: The Impact of the Abolition of China's Civil Service Exam." *Econometrica* 84(2):677–733.

Bates, Robert. 2008. "State Failure." *Annual Review of Political Science* 11:1–12.

Beattie, Hilary. 1979. *Land and Lineage in China*. Cambridge.

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

Blaydes, Lisa and Christopher Paik. 2016. "The Impact of Holy Land Crusades on State Formation: War Mobilization, Trade Integration, and Political Development in Medieval Europe." *International Organization* 70(3):551–586.

Blaydes, Lisa and Eric Chaney. 2013. "The Feudal Revolution and Europe's Rise: Political Divergence of the Christian West and the Muslim World Before 1500 CE." *American Political Science Review* 107(1):16–34.

Brokaw, Cynthia and Christopher Reed. 2010. *From Woodblocks to the Internet*. Brill.

Brokaw, Cynthia and Kai-Wing Chow. 2005. *Printing and Book Culture in Late Imperial China*. University of California.

Burke, Marshall, Solomon Hsiang and Edward Miguel. 2015. "Climate and Conflict." *Annual Review of Economics* 7(1):577–617.

Ch, Rafael, Jacob Shapiro, Abbey Steele and Juan Vargas. 2018. "Endogenous Taxation in Ongoing Internal Conflict: The Case of Colombia." *American Political Science Review* 112(4):996–1015.

Chaffee, John. 1989. Status, Family and Locale: An Analysis of Examination Lists from Sung China. In *A Collection of Essays on Song History Studies (Liu Zijian boshi jinian songshou songshi yanjiu lunji)*, ed. Kinugawa Tsuyoshi. Dohosha chapter E1, pp. 341–356.

Chang, Chung-li. 1955. *The Chinese Gentry*. University of Washington.

Chang, Chung-li. 1962. *The Income of the Chinese Gentry*. University of Washington.

Chen, Gaoyong. 2007 [1940]. *Categorization and Statistics of Natural Disasters and Conflicts in Historical China (*Zhongguo lidai tianzai renhuo biao*)*. Beijing Library Press.

China Historical Geographic Information System. 2018.
URL: *http://sites.fas.harvard.edu/%7Echgis/*

Cox, Gary. 2016. *Marketing Sovereign Promises*. Cambridge.

Dincecco, Mark and Yuhua Wang. 2018. "Violent Conflict and Political Development over the Long Run: China versus Europe." *Annual Review of Political Science* 21(1):341–58.

Downing, Brian. 1992. *The Military Revolution and Political Change*. Princeton.

Duara, Prasenjit. 1988. *Culture, Power, and the State*. Stanford.

Fearon, James D and David D Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97(1):75–90.

Feng, Erkang. 2006. "On the "Self-Governance" Nature of Clans in the Qing Dynasty (Jianlun qingdai zongzu de zizhi xing)." *Journal of Huazhong Normal University* 45(1):65–70.

Freedman, Maurice. 1958. *Lineage Organization in Southeastern China*. University Athlone.

Garfias, Francisco and Emily Sellars. 2018. "From Conquest to Centralization: Domestic Conflict and the Transition to Direct Rule." *Working Paper, UCSD* .

Greif, Avner. 2006. "Family Structure, Institutions, and Growth: The Origins and Implications of Western Corporations." *American Economic Review* 96(2):308–12.

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.

Guo, Limin. 2015. *Atlas of Chinese Modern History (Zhongguo jindaishi ditu ji)*. StarMap Press.

Hoffman, Philip. 2015. *Why Did Europe Conquer the World?* Princeton.

Hui, Victoria Tin-bor. 2005. *War and State Formation in Ancient China and Early Modern Europe*. Cambridge.

Jia, Ruixue. 2014*a*. "The Legacies of Forced Freedom: China's Treaty Ports." *Review of Economics and Statistics* 96(4):596–608.

Jia, Ruixue. 2014*b*. "Weather Shocks, Sweet Potatoes, and Peasant Revolts in Historical China." *Economic Journal* 124(575):92–118.

Kang, David. 2020. "International Order in Historical East Asia: Tribute and Hierarchy Beyond Sinocentrism and Eurocentrism." *International Organization* 74(1):65–93.

Ko, Chiu Yu, Mark Koyama and Tuan-Hwee Sng. 2018. "Unified China and Divided Europe." *International Economic Review* 59(1):285–327.

Koyama, Mark, Chiaki Moriguchi and Tuan-Hwee Sng. 2018. "Geopolitics and Asia's Little Divergence: State Building in China and Japan after 1850." *Journal of Economic Behavior & Organization* 155:178–204.

Kuhn, Philip. 1970. *Rebellion and Its Enemies in Late Imperial China*. Harvard.

Kung, James Kai-sing and Chicheng Ma. 2014. "Can Cultural Norms Reduce Conflicts? Confucianism and Peasant Rebellions in Qing China." *Journal of Development Economics* 111:132–149.

Kuran, Timur. 2018. "Islam and Economic Performance: Historical and Contemporary Links." *Journal of Economic Literature* 56(4):1292–359.

Liang, Fangzhong. 2008. *Statistics on Households, Land, and Taxation in Chinese Dynasties (Zhongguo lidai hukou tiandi tianfu tongji)*. Zhonghua Book Company (Zhonghua shuju).

Ma, Debin and Jared Rubin. 2019. "The Paradox of Power: Principal-agent problems and administrative capacity in Imperial China (and other absolutist regimes)." *Journal of Comparative Economics* 47(2):277–94.

Mann, Michael. 1984. "The Autonomous Power of the State: Its Origins, Mechanisms and Results." *European Journal of Sociology* 25(2):185–213.

Mattingly, Daniel. 2017. "Colonial Legacies and State Institutions in China: Evidence from a Natural Experiment." *Comparative Political Studies* 50(4):434–463.

Nanjing Military Academy. 2003. *Catalog of Historical Wars (Lidai zhanzheng nianbiao)*. People's Liberation Army Press.

Perry, Elizabeth. 1980. *Rebels and Revolutionaries in North China, 1845-1945*. Stanford.

Platt, Stephen. 2012. *Autumn in the Heavenly Kingdom*. Knopf.

Queralt, Didac. 2019. "War, International Finance, and Fiscal Capacity in the Long Run." *International Organization* 73(4):713–53.

Rodríguez-Franco, Diana. 2016. "Internal Wars, Taxation, and State Building." *American Sociological Review* 81(1):190–213.

Rosenthal, Jean-Laurent and R. Bin Wong. 2011. *Before and Beyond Divergence*. Harvard.

Rowe, William. 2007. *Crimson Rain*. Stanford.

Scheve, Kenneth and David Stasavage. 2016. *Taxing the Rich*. Princeton.

Shi, Zhihong and Yi Xu. 2008. *Finance in the Late Qing, 1851-1894 (Wanqing caizheng)*. Shanghai University of Finance and Economics Publishing House.

Shiue, Carol. 2016. "A Culture of Kinship: Chinese Genealogies as a Source for Research in Demographic Economics." *Journal of Demographic Economics* 82(4):459–482.

Shue, Vivienne. 1988. *The Reach of the State*. Stanford.

Skocpol, Theda. 1979. *States and Social Revolutions*. Cambridge.

Slater, Daniel. 2010. *Ordering Power*. Cambridge.

Sng, Tuan-Hwee. 2014. "Size and Dynastic Decline: The Principal-Agent Problem in Late Imperial China, 1700-1850." *Explorations in Economic History* 54(October):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):439–470.

Spence, Jonathan. 1996. *God's Chinese Son*. Norton.

Spence, Jonathan. 2002. The K'ang-Hsi Reign. In *The Cambridge History of China. Volume 9, The Ch'ing Empire to 1800, Part One*, ed. Denis Twitchett and John Fairbank. Cambridge chapter 3, pp. 120–182.

Staniland, Paul. 2015. "Militias, Ideology, and the State." *Journal of Conflict Resolution* 59(5):770–793.

Stasavage, David. 2011. *States of Credit: Size, Power, and the Development of European Polities*. Princeton.

State Meteorological Society. 1981. *Yearly Charts of Dryness/Wetness in China for the Last 500-Year Period (*Zhongguo jinwubainian hanlao fenbu tuji*)*. Sino Maps Press.

Szonyi, Michael. 2016. Lineages and the Making of Contemporary China. In *Modern Chinese Religion II: 1850-2015*, ed. Vincent Goossaert, Jan Kiely and John Lagerwey. Brill chapter 10, pp. 433–89.

Tilly, Charles. 1975. Reflections on the History of European State-Making. In *The Formation of States in Western Europe*, ed. Charles Tilly. Princeton chapter 1, pp. 3–83.

von Glahn, Richard. 2016. *An Economic History of China*. Cambridge.

Wakeman, Frederic. 1975. *The Fall of Imperial China*. Free Press.

Wang, Heming. 2008. *Comprehensive Catalogue of Chinese Genealogies (Zhongguo jiapu zongmu)*. Shanghai Ancient Works Publishing House.

Wang, Yeh-chien. 1973. *Land Taxation in Imperial China, 1750-1911*. Harvard University Press.

Weber, Max. 1946 [1918]. *Essays in Sociology*. Oxford.

Weinstein, Jeremy. 2005. "Autonomous Recovery and International Intervention in Comparative Perspective." *Center for Global Development Working Paper 57* .

Wilkinson, Endymion. 2000. *Chinese History*. Harvard University Asia Center.

Woodberry, Robert. 2012. "The Missionary Roots of Liberal Democracy." *American Political Science Review* 106(2):244–74.

Yang, Guoan. 2012. *State Power and Social Order (Guojia quanli yu minjian zhixu)*. Wuhan University Press.

Yang, Zhengtai. 2006. *A Study of Courier Stations in the Ming Dynasty (Mingdai yizhan kao)*. Shanghai Ancient Works Publishing House.

Zelin, Madeleine. 2002. The Yung-Cheng Reign. In *The Cambridge History of China. Volume 9, The Ch'ing Empire to 1800, Part One*, ed. Twitchett, Denis and Fairbank, John. Cambridge chapter 4, pp. 183–229.

Ziblatt, Daniel. 2006. *Structuring the State*. Princeton.

Online Appendix for
Internal Conflict, Geopolitics, and State Development

A Model

Setup

For narrative details and assumption justifications, see Section 2.

First, Nature determines whether there is a negative climate shock, and therefore whether the peasantry rebels against the imperial state. In response to mass violence, the ruler must decide whether to strengthen the central state's leadership role as the locus of elite-state cooperative efforts, or to "outsource" this governance task to local elites. If the ruler decides to strengthen the central state's leadership role, then local elites must decide whether to willingly accept this shift in the locus of leadership in elite-state cooperation toward the central state, or to oppose it.

Each actor receives benefit $b > 0$ from putting down mass violence, regardless of the specific locus of leadership in elite-state cooperation. If the ruler decides to strengthen the central state's leadership role as the locus of elite-state cooperation, and local elites accept this shift, then the ruler receives additional benefit $c > 0$. On the other hand, state-strengthening requires the ruler to make new state capacity investments $i > 0$. To fund such investments, the ruler must levy greater "tribute" $t > 0$. In addition, local elites must pay their part of the governance cost $x > 0$, even under state-strengthening. Still, this governance cost is lower than if the locus of leadership in elite-state cooperation is outsourced to local elites. Here, this cost is $y > 0$, where $y > x$. Finally, if local elites oppose the ruler's state-strengthening choice, then the ruler's ability to exact greater tribute and undertake new state capacity investments is (at least somewhat) curtailed. In this case, the ruler must pay "rejection" cost $r > 0$ for being rebuffed by local elites. Similarly, local elites must pay penalty p for their opposition.

Optimal Decisions

We now analyze the optimal decisions by the ruler and local elites under two different contexts.

Context 1: Benchmark

Two main assumptions characterize the first context. First, let the reduction in governance costs to local elites under state-strengthening by the ruler be relatively high. Second, let the tribute cost that local elites must pay to fund new state capacity investments be relatively

low. Thus, $b - x - t > b - y - p$ for local elites. Here, we assume further that $b - x - t > 0$ and $b - y - p > 0$, since internal security provision allows local elites to maintain their privileged status in society. Finally, let the payoff to the ruler from the reduction in transaction costs in governance due to state-strengthening exceed the cost of new investments in state capacity to improve the center's ability to broadcast political power. Thus, $b + c - i > b$ for the ruler.

We solve for the subgame perfect equilibrium by backwards induction. Given the above assumptions, local elites accept state-strengthening by the ruler at the second stage, and the ruler opts for state-strengthening at the first stage. The subgame perfect equilibrium is thus (Strengthen, Accept).

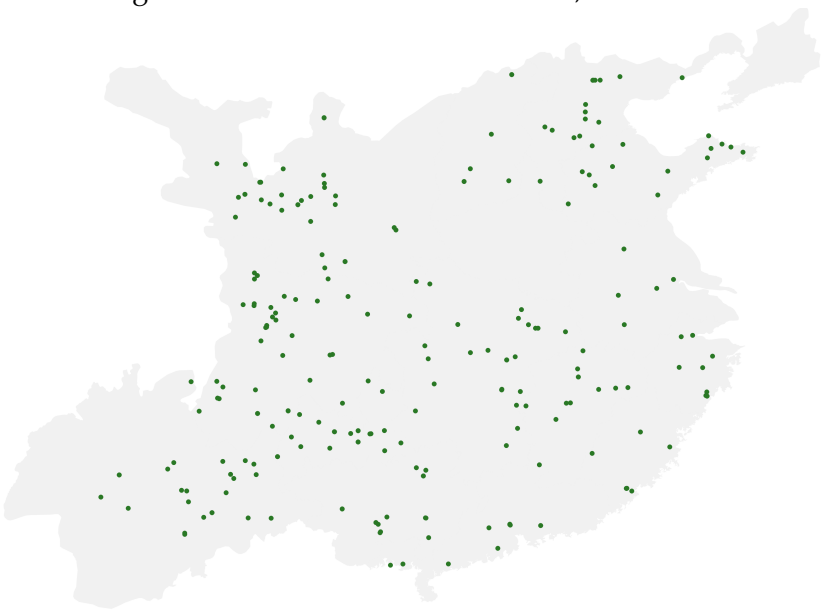
Context 2: New External Threat

Now assume that the ruler must undertake major new investments in (external) military defense in response to a severe new external threat. In turn, local elites are expected to pay a much higher tribute. Furthermore, let the reduction in governance costs to local elites under state-strengthening be relatively low. Thus, in this new context, $b - y - p > b - x - t$ for local elites.

Given the new assumptions above, local elites oppose state-strengthening at the second stage, even if the ruler still prefers to opt for it at the first stage. The subgame perfect equilibrium is thus (Strengthen, Decline).

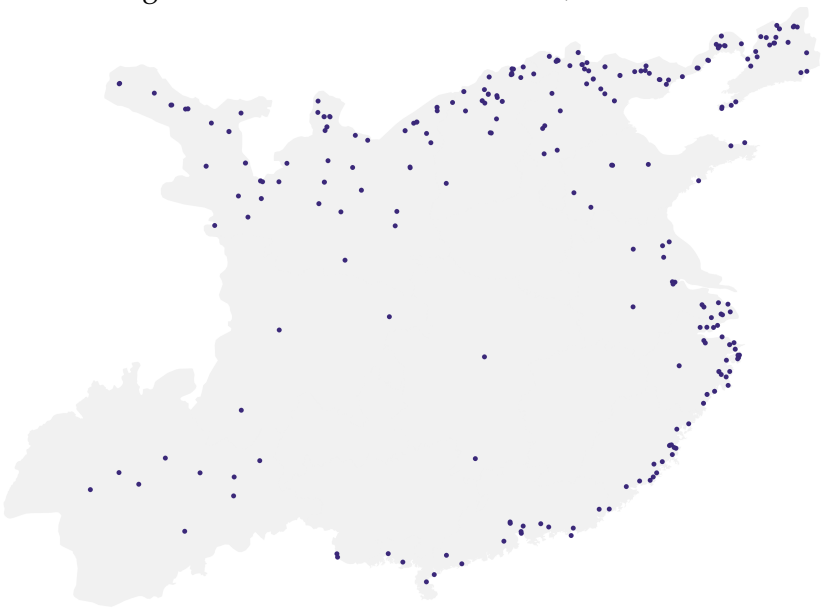
B Extra Figures

Figure B-1: Elite Rebellion Locations, 1350-1900



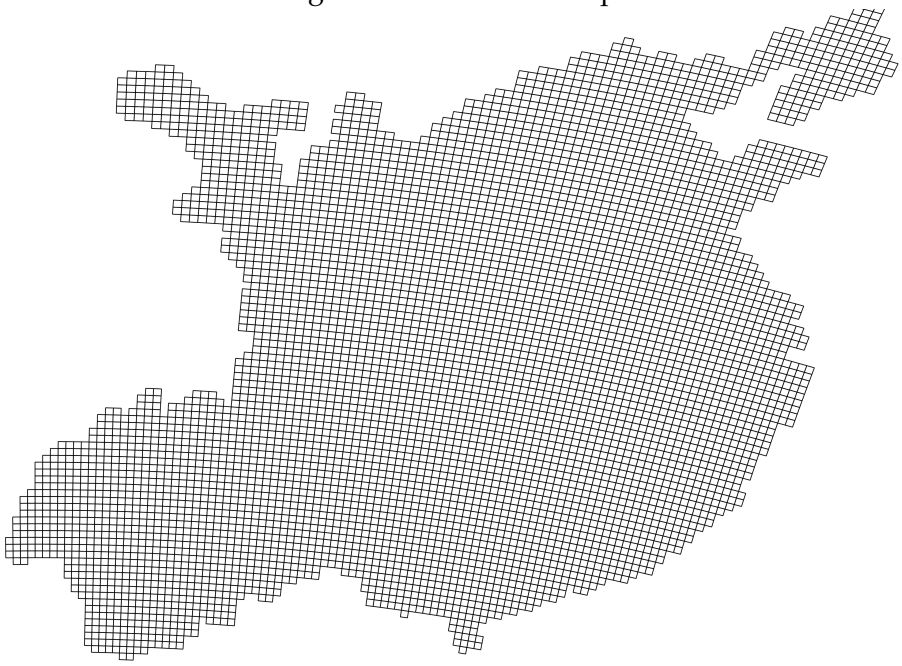
Notes. This figure shows the location of each recorded elite rebellion battle in China between 1350-1900 within Ming-era borders.

Figure B-2: External War Locations, 1350-1900



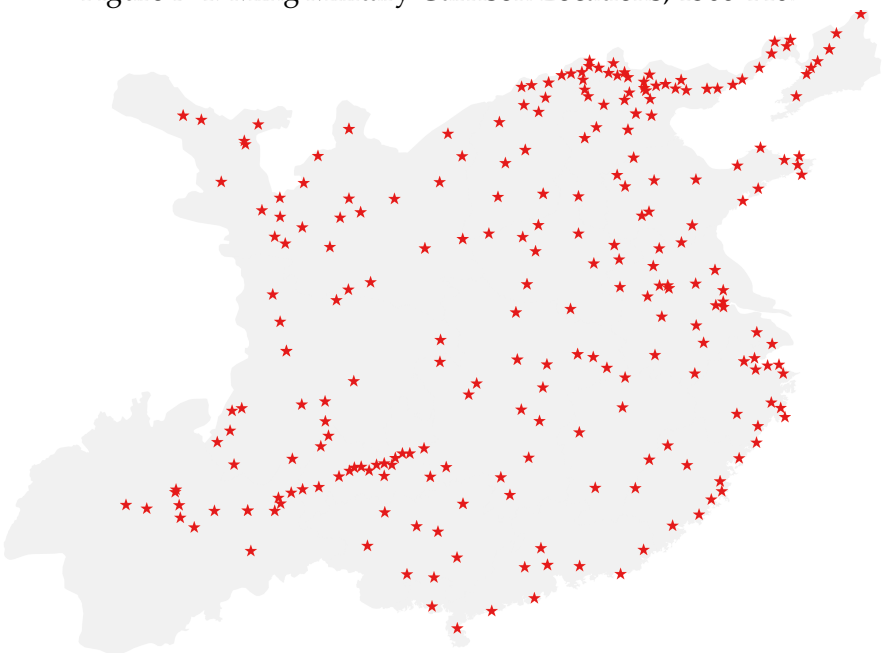
Notes. This figure shows the location of each recorded external war battle in China between 1350-1900 within Ming-era borders.

Figure B-3: Grid Cell Map



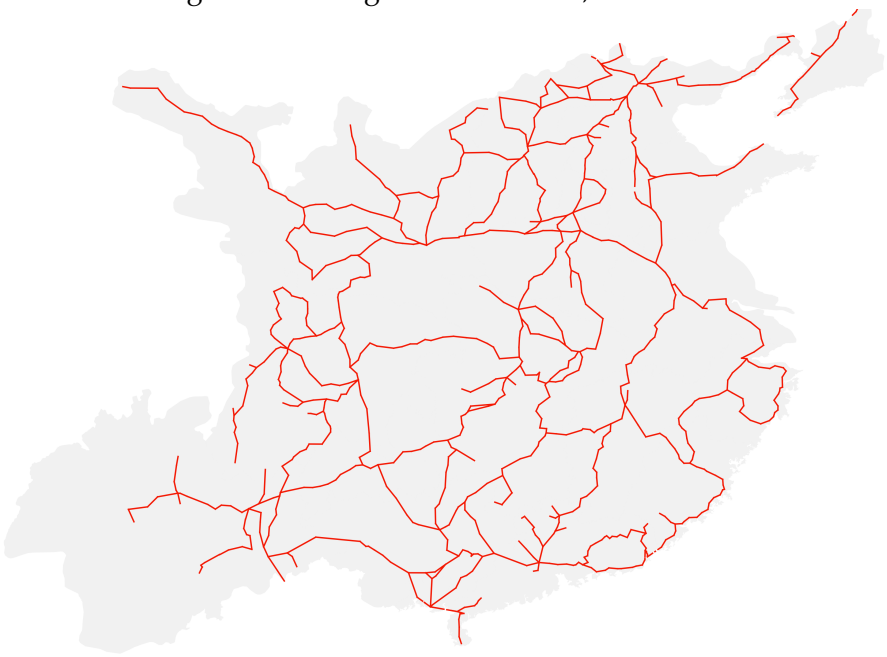
Notes. This figure shows 25 km x 25 km grid cells in China within Ming-era borders.

Figure B-4: Ming Military Garrison Locations, 1368-1467



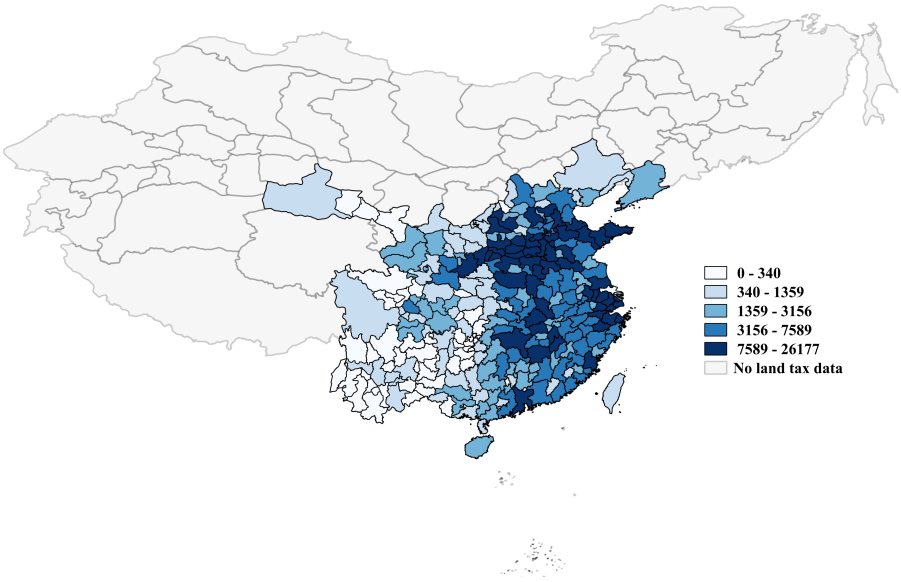
Notes. This figure shows the location of each state military garrison in China established under the first century of Ming rule (1368-1467) within Ming-era borders.

Figure B-5: Ming Courier Routes, 1368-1644



Notes. This figure shows the location of each courier route in China established under Ming rule (1368-1644) within Ming-era borders.

Figure B-6: Qing Land Taxation, 1820

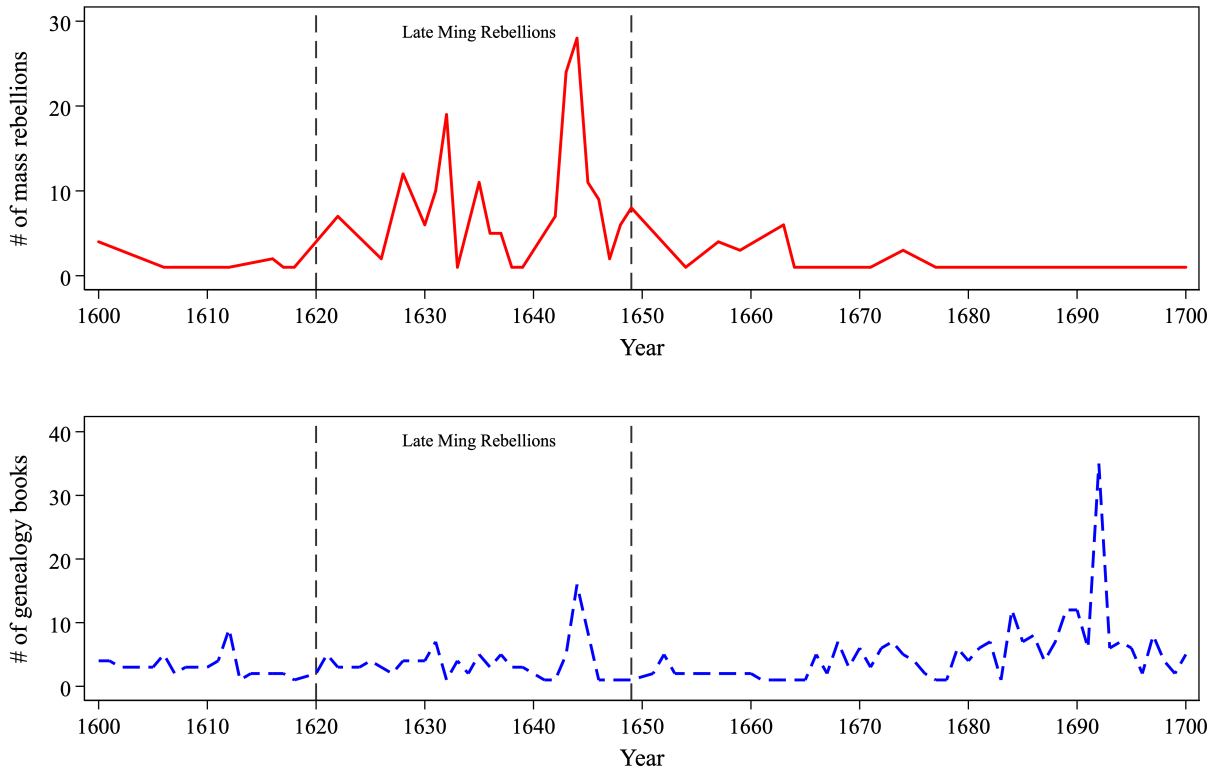


Notes. This figure shows total land taxation (in silver kg) by prefecture in China under Qing rule (1820). Prefectures are shaded by quintile, whereby those in the top quintile receive the darkest shade. Prefectural lines are for Qing-era borders.

Figure B-7: Declarations of Independence, 1911

Notes. This figure shows the locations of military groups that made a formal declaration of independence from the imperial Qing state in China in 1911. Prefectural lines are for Qing-era borders.

Figure B-8: Mass Rebellion and Clan Activity, 1600-1700



Notes. This figure shows the annual number of mass rebellion battles (top panel) and genealogy books written down (bottom panel) in China between 1600 and 1700. The (red) solid line indicates the number of mass rebellions per year, and the (blue) dashed line indicates the number of genealogy books.

C Extra Tables

Table C-1: Conflict Types, 1350-1900

	N	%
Elite rebellion	316	18.016
External war	541	30.844
Mass rebellion	897	51.140
Total	1754	100.000

Notes. See text for variable descriptions and data sources.

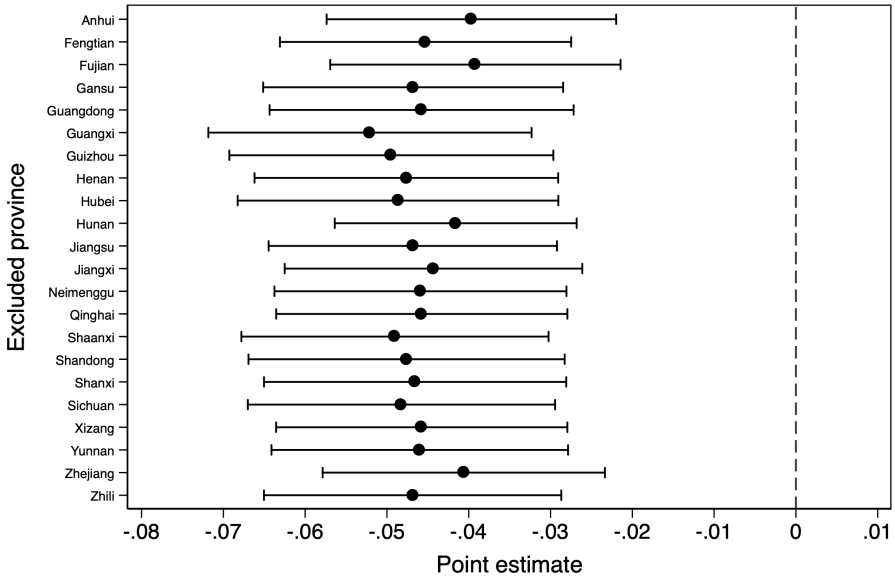
Table C-2: Summary Statistics

	N	Mean	Std Dev	Min	Max
A: Panel Analysis, 1350-1850					
Genealogy books	58030	0.094	1.660	0.000	190.000
Genealogy books (IHS)	58030	0.032	0.259	0.000	5.940
Mass rebellion	58030	0.011	0.128	0.000	7.000
B: Instrumental Variables Analysis, 1500-1850					
Genealogy books	14553	0.198	2.603	0.000	190.000
Genealogy books (IHS)	14553	0.059	0.366	0.000	5.940
Mass rebellion	14553	0.011	0.120	0.000	4.000
Extreme drought	14553	2.742	2.558	0.000	13.000
Mild or extreme drought	14553	9.376	6.506	0.000	30.000
C: Panel Analysis, 1350-1900					
Genealogy books	63833	0.199	3.621	0.000	371.000
Genealogy books (IHS)	63833	0.049	0.338	0.000	6.609
Mass rebellion	63833	0.014	0.147	0.000	9.000
D: Ming Cross-Sectional Analysis, 1368-1467					
Garrisons	5803	0.064	0.370	0.000	7.000
Garrisons (IHS)	5803	0.047	0.236	0.000	2.644
Presence of Ming courier routes	5803	0.230	0.421	0.000	1.000
Mass rebellion (1368-1467)	5803	0.021	0.223	0.000	10.000
E: Qing Cross-Sectional Analysis, 1820					
Land tax (1820) (silver kg)	260	4202.550	4996.500	0.000	26176.699
Land tax (1820) (IHS)	260	7.911	2.086	0.000	10.866
Mass rebellion (1644-1819)	311	0.444	1.011	0.000	7.000
Population density (person/km2)	260	122.995	126.853	0.435	840.042
F: Qing Cross-Sectional Analysis, 1890-1911					
Declaration of independence (1911)	15738	0.006	0.078	0.000	1.000
Genealogy books	15738	0.283	3.734	0.000	165.000
Genealogy books (IHS)	15738	0.065	0.401	0.000	5.799

Notes. See text for variable descriptions and data sources.

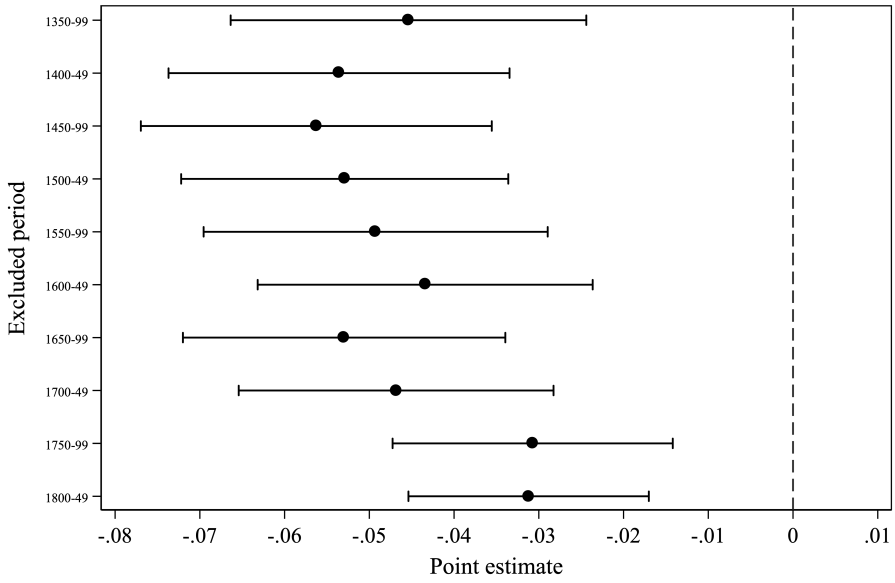
D Robustness Checks for Clan Activity Analysis, 1350-1850

Figure D-1: Mass Rebellion and Clan Activity: Exclude Provinces One by One



Notes. Sample period is 1350-1850. Each black dot represents point estimate for regression model in column 2 of Table 1 when we exclude each province one by one. Horizontal bars indicate 90 percent confidence intervals.

Figure D-2: Mass Rebellion and Clan Activity: Exclude 50-Year Periods One by One



Notes. Sample period is 1350-1850. Each black dot represents point estimate for regression model in column 2 of Table 1 when we exclude each 50-year period one by one. Horizontal bars indicate 90 percent confidence intervals.

Table D-1: Mass Rebellion and Clan Activity: Alternative Specifications of Dependent Variable

Dependent variable:	ln(1+Genealogy Books)		Genealogy Books	
	(1)	(2)	(3)	(4)
Mass rebellion	-0.055***	-0.037***	-0.345***	-0.237**
	(0.011)	(0.009)	(0.112)	(0.097)
Grid cell FE	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes
County-specific trends	No	Yes	No	Yes
R^2	0.044	0.297	0.013	0.243
Observations	58030	58030	58030	58030

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period is 1350-1850. Dependent variable in columns 1-2 is clan activity as proxied by $\ln(1 + GenealogyBooks)$. Dependent variable in columns 3-4 is clan activity as proxied by *GenealogyBooks*. Variable of interest is number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-2: Mass Rebellion and Clan Activity: IV Analysis

Panel A: First Stage

Dependent variable:	Mass Rebellion	
	(1)	(2)
Extreme drought	0.002***	
	(0.001)	
Mild or extreme drought		0.001***
		(0.000)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
R^2	0.015	0.014
Observations	14553	14553

Panel B: Second Stage

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
Mass rebellion	-2.802***	-2.610***
	(1.033)	(0.851)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
Anderson-Rubin p-value	0.000	0.000
Kleibergen-Paap Wald *rk F*-statistic	11.867	17.379
Observations	14553	14553

Notes. Estimation method is 2SLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period is 1500-1850. In Panel A (first stage), dependent variable is number of mass rebellions, while variable of interest is number of years of extreme drought in column 1 and number of years of mild or extreme drought in column 2. In Panel B (second stage), dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books, while variable of interest is number of mass rebellions, as instrumented by frequency of drought. All regressions include grid cell and period fixed effects. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-3: Mass Rebellion and Clan Activity: Lagged Dependent Variable

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
Mass rebellion	-0.028***	-0.023***
	(0.008)	(0.008)
L.Genealogy books (IHS)	0.833***	0.674***
	(0.030)	(0.033)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
County-specific trends	No	Yes
R^2	0.415	0.480
Observations	58030	58030

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period is 1350-1850. Dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest is number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-4: Mass Rebellion and Clan Activity: Spatial Spillovers

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
Mass rebellion	-0.066***	-0.047***
	(0.014)	(0.011)
Mass rebellion (neighbor)	-0.008***	0.004*
	(0.002)	(0.002)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
County-specific trends	No	Yes
R^2	0.045	0.295
Observations	58030	58030

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period is 1350-1850. Dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest is number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-5: Mass Rebellion and Clan Activity: Lagged Variable of Interest

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
L.Mass rebellion	-0.069***	-0.056***
	(0.012)	(0.010)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
County-specific trends	No	Yes
R^2	0.045	0.295
Observations	58030	58030

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period is 1350-1850. Dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest is lagged number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-6: Mass Rebellion and Clan Activity: 50 km x 50 km Grid Cells

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
Mass rebellion	-0.067***	-0.038***
	(0.017)	(0.012)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
Province-specific trends	No	Yes
R^2	0.128	0.322
Observations	15310	15310

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 50 km x 50 km. Sample period is 1350-1850. Dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest is number of mass rebellions. All regressions include grid cell and period fixed effects and province-specific time trends. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-7: Mass Rebellion and Clan Activity: 1990 County Borders

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
Mass rebellion	-0.051***	-0.047***
	(0.013)	(0.013)
County FE	Yes	Yes
Period FE	Yes	Yes
Province-specific trends	No	Yes
R^2	0.113	0.126
Observations	23720	23720

Notes. Estimation method is OLS. Unit of analysis is county-period. County is for 1990 borders. Sample period is 1350-1850. Dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest is number of mass rebellions. Regression include county and period fixed effects and province-specific time trends. Robust standard errors clustered at county level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-8: Mass Rebellion and Clan Activity: Include Grid Cells Within Qing Borders

Dependent variable:	Genealogy Books (IHS)	
	(1)	(2)
Mass rebellion	-0.067***	-0.036**
	(0.017)	(0.015)
Grid cell FE	Yes	Yes
Period FE	Yes	Yes
County-specific trends	No	Yes
R^2	0.009	0.274
Observations	359040	359040

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period is 1350-1850. Dependent variable is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest is number of mass rebellions. Regression include cell and period fixed effects and county-specific time trends. Robust standard errors clustered at grid cell level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-9: Mass Rebellion and Clan Activity: Cluster Standard Errors at County Level

Dependent variable:	Genealogy Books (IHS)						
	1350-1850				1350-1900		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mass rebellion	-0.066***	-0.044***	-0.034***	-0.035***	-0.070***	-0.051***	-0.061***
	(0.014)	(0.011)	(0.011)	(0.012)	(0.013)	(0.012)	(0.015)
Mass rebellion × Post-1850					0.527***	0.322***	0.518***
					(0.068)	(0.072)	(0.069)
Elite rebellion			0.190***			0.205***	
			(0.043)			(0.049)	
Elite rebellion × Post-1850						0.142***	
						(0.014)	
External war			-0.044***			-0.044***	
			(0.015)			(0.016)	
External war × Post-1850						0.056	
						(0.135)	
Mass rebellion × 1600-50 (Ming-Qing)				-0.027			-0.030
				(0.026)			(0.030)
Mass rebellion × 1750-1800 (White Lotus)							0.047
							(0.053)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-specific trends	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.048	0.301	0.308	0.301	0.345	0.365	0.345
Observations	51470	51470	51470	51470	56617	56617	56617

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period in columns 1-4 is 1350-1850. Sample period in columns 5-7 is 1350-1900. Dependent variable in all columns is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest in all columns is number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at county level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table D-10: Mass Rebellion and Clan Activity: Cluster Standard Errors at Prefectural Level

Dependent variable:	Genealogy Books (IHS)						
	1350-1850				1350-1900		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Mass rebellion	-0.066***	-0.044***	-0.034***	-0.035**	-0.070***	-0.051***	-0.061***
	(0.015)	(0.012)	(0.011)	(0.014)	(0.015)	(0.013)	(0.017)
Mass rebellion × Post-1850					0.527***	0.322***	0.518***
					(0.075)	(0.072)	(0.075)
Elite rebellion			0.190***			0.205***	
			(0.056)			(0.064)	
Elite rebellion × Post-1850						0.142***	
						(0.013)	
External war			-0.044***			-0.044**	
			(0.016)			(0.017)	
External war × Post-1850						0.056	
						(0.137)	
Mass rebellion × 1600-50 (Ming-Qing)				-0.027			-0.030
				(0.030)			(0.034)
Mass rebellion × 1750-1800 (White Lotus)							0.047
							(0.040)
Grid cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Period FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-specific trends	No	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.048	0.301	0.308	0.301	0.345	0.365	0.345
Observations	51470	51470	51470	51470	56617	56617	56617

Notes. Estimation method is OLS. Unit of analysis is grid cell-period. Grid cell is 25 km x 25 km. Sample period in columns 1-4 is 1350-1850. Sample period in columns 5-7 is 1350-1900. Dependent variable in all columns is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books. Variable of interest in all columns is number of mass rebellions. All regressions include grid cell and period fixed effects. Robust standard errors clustered at prefectural level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

E Robustness Checks for Ming State Development Analysis

Table E-1: Mass Rebellion and Ming Garrisons: Alternative Specifications of Dependent Variable

Dependent variable:	ln(1+Garrisons)		Garrisons	
	(1)	(2)	(3)	(4)
Mass rebellion (1368-1467)	0.125***	0.118***	0.217***	0.199***
	(0.030)	(0.034)	(0.062)	(0.066)
County FE	No	Yes	No	Yes
R^2	0.024	0.419	0.017	0.424
Observations	5147	5147	5147	5147

Notes. Estimation method is OLS. Unit of analysis is 25 km x 25 km grid cell. Sample period is first century of Ming Dynasty (1368-1467). Dependent variable in columns 1-2 is early Ming military garrisons as proxied by $\ln(1 + Garrisons)$. Dependent variable in columns 3-4 is the number of garrisons. Variable of interest is number of mass rebellions over this sample period. Robust standard errors clustered at county level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table E-2: Mass Rebellion and Ming Garrisons: Cluster Standard Errors at Prefectural Level

Dependent variable:	ln(1+Garrisons)		Garrisons	
	(1)	(2)	(3)	(4)
Mass rebellion (1368-1467)	0.161***	0.152***	0.148***	0.110**
	(0.039)	(0.045)	(0.034)	(0.043)
County FE	No	Yes	No	Yes
R^2	0.024	0.422	0.006	0.539
Observations	5147	5147	5147	5147

Notes. Estimation method is OLS. Unit of analysis is 25 km x 25 km grid cell. Sample period is first century of Ming Dynasty (1368-1467). Dependent variable in columns 1-2 is early Ming military garrisons as proxied by $\ln(1 + Garrisons)$. Dependent variable in columns 3-4 is the number of garrisons. Variable of interest is number of mass rebellions over this sample period. Robust standard errors clustered at prefectural level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

F Regression Analysis for Qing State Failure

Table F-1: Clan Activity and Qing State Failure: Ming-Era External Borders

Dependent variable:	Declaration of Independence in 1911	
	(1)	(2)
Genealogy Books (IHS)	0.044***	0.045***
	(0.007)	(0.010)
County FE	No	Yes
R^2	0.053	0.398
Observations	5887	5887

Notes. Estimation method is OLS. Unit of analysis is 25 km x 25 km grid cell. Sample is restricted to Ming-era external borders. Dependent variable is binary indicator of formal declaration of independence from imperial Qing state in 1911. Variable of interest is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books, between 1890-1909. Robust standard errors clustered at county level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.

Table F-2: Clan Activity and Qing State Failure: Cluster Standard Errors at Prefectural Level

Dependent variable:	Declaration of Independence in 1911	
	(1)	(2)
Genealogy Books (IHS)	0.045***	0.044***
	(0.006)	(0.007)
County FE	No	Yes
R^2	0.054	0.382
Observations	15103	15103

Notes. Estimation method is OLS. Unit of analysis is 25 km x 25 km grid cell. Dependent variable is binary indicator of formal declaration of independence from imperial Qing state in 1911. Variable of interest is clan activity as proxied by the inverse hyperbolic sine (IHS) of the number of genealogy books, between 1890-1909. Robust standard errors clustered at prefectural level in parentheses. ***, **, and * indicate statistical significance at 1%, 5%, and 10% level, respectively.