Elite Capture, Meritocracy and Social Stability: Evidence from the Chinese Imperial Examinations

Jingjing Chen
Jianlei Han
Shing-Yi Wang
Yongxiang Wang*

August 24, 2026

Abstract

This paper examines reforms in China's imperial examination system during the Northern Song dynasty that anonymized exam submissions and ended local officials' discretion to recommend candidates. Using prefecture-level data from 960–1127 and a difference-in-differences design, we compare areas with and without prominent elite families, where corruption and elite capture were more likely. The reforms increased the share of successful candidates from commoner backgrounds and narrowed their performance gaps with elites in scholarship and political achievement, indicating better selection of commoner talent. Treated prefectures also saw less social unrest. Overall, the reforms strengthened meritocracy, broadened political access, and promoted stability.

*Jingjing Chen is from the Macquarie Business School, Macquarie University, jingjing.chen5@hdr.mq.edu.au; Jianlei Han is from the Macquarie Business School, Macquarie University, jianlei.han@mq.edu.au; Shing-Yi Wang is from the Wharton School, University of Pennsylvania, was@wharton.upenn.edu; Yongxiang Wang is from the Shanghai Advanced Institute of Finance (SAIF), Shanghai Jiao Tong University, yxwang@saif.sjtu.edu.cn. We are grateful to Roshan Kannan and Rosemary Zhang for excellent research assistance. This paper has benefited from comments from Santosh Anagol and various seminar participants. All errors are our own. Yongxiang Wang would like to thank the National Natural Science Foundation of China (Grant No. 72572103) for financial support. Jianlei Han would like to thank the Australian Research Council (Grant No. DP220102985) for financial support.

1 Introduction

In most countries, modern civil service recruitment relies on formal examination systems designed to select candidates on the basis of merit and ability.[1] A central purpose of these systems is to limit the influence of nepotism, political patronage, and family background in determining access to government office. When recruitment is corrupted by favoritism, however, the consequences may extend beyond the misallocation of talent in bureaucratic positions. Unfairness in access to office can exclude qualified but unconnected candidates, preserve the influence of established elites, and generate grievances among educated aspirants who view institutional channels of advancement as closed. Reforms that make recruitment more meritocratic may therefore affect not only the composition of the bureaucracy, but also broader social stability. Scholars have emphasized the importance of opportunities for entry into the elite for social stability (Acemoglu and Robinson, 2008; North et al., 2009), yet rigorous empirical work on this topic remains limited.

This paper studies these issues in the context of one of the earliest and most influential meritocratic civil service examination systems in the world, which occurred in imperial China.[2] Our data cover the Northern Song dynasty from 960 to 1127. In particular, we study how two sets of reforms surrounding the examination process affect the types of individuals who pass the exam. In addition, we examine the broader impacts of the reforms on social unrest. To our knowledge, we provide the first causal evidence on how reforms to the internal design of an existing civil-service examination system — aimed at limiting discretion and corruption — affect who is selected and shape longer-run social unrest.

The examination system was formally meritocratic, but its initial implementation still left room for corruption and bias. The first stage of the selection process involved an examination at the prefecture level; candidates who passed that stage went on to central-level examinations. At the prefecture level, candidates submitted application information to verify eligibility, but this information also revealed the candidates' family background to the examiners. Thus, how the

[1] Over 75% of countries use examinations in the hiring process of some of their public sector workers (Teorell et al., 2012).

[2] The imperial examination system was the central institution through which the Chinese state recruited administrative talent to become part of the ruling elite.

examiners graded the submissions may have reflected either active discrimination or corruption with the knowledge of the identity of the test taker, or subconscious biases about commoners versus elites.[3] Furthermore, officials in local prefectures were given discretion in making recommendations and could recommend candidates based on subjective judgment in addition to relying on the examination score.

We examine two sets of reforms that reduced the scope for discrimination and corruption in this recruitment process. In 1033, the state introduced anonymous grading for the prefecture-level examinations, which removed candidates' names and personal information from exam submissions. In 1037, the state further anonymized submissions by requiring exam papers to be transcribed before grading, preventing examiners from identifying candidates by handwriting. Finally, in 1044–1045, the state abolished discretionary local recommendation and made examination performance the sole criterion for prefecture-level selection. Together, these reforms reduced identity-based favoritism, examiner discretion, and the ability of local elites to influence access to bureaucratic office.

Historians have long discussed the potential of the imperial examination reforms to alter talent selection and increase social stability. In 1227, Wang Yong wrote, "Therefore, our sage dynasty opened the gates of the civil service examinations widely, allowing everyone to retain the hope of advancement, so that they would not become bandits or lawbreakers" (Wang, 1981). More recently, Liu (2006) wrote, "By implementing the sealing and transcription system to ensure 'fair competition' in the civil service examinations, the state could select outstanding talents capable of governing the country and advancing social development; it could ease the anger and resentment of unsuccessful candidates, thereby maintaining social stability."

Our empirical strategy exploits the fact that the reforms were likely to matter most in prefectures where pre-reform favoritism was most severe. We compare prefectures where powerful great families resided to other prefectures before and after the reforms. Historical accounts suggest that these great families exerted influence over examination outcomes and that candidates from humble

[3]This ties into a literature that identifies discrimination in the hiring selection process (Goldin and Rouse, 2000; Behaghel et al., 2015). The other papers in this literature, like ours, find strong evidence of discrimination on the basis of observable characteristics.

backgrounds were disadvantaged as a result (Li, 2004; Hartwell, 1982).

We therefore interpret great-family prefectures as places with greater pre-reform scope for elite capture in examination selection. Using a prefecture-term panel from 960 to 1127, we estimate difference-in-differences specifications that compare changes in bureaucratic selection and social unrest across treated and control prefectures before and after the reforms. This analysis is made possible by our manual assembly of several novel historical datasets that document who entered the bureaucracy and long-run social stability.

This setting provides a useful context for studying the relationship between bias in the selection process and the identity of civil servants. The examination selection operated under quotas for each prefecture. As a result, favoritism toward connected candidates had direct allocative consequences: admitting a connected candidate reduced the opportunity for another candidate to advance. Bias in selection can therefore distort the allocation of scarce bureaucratic entry slots.

Our main measure of talent selection is the proportion of individuals who pass the exam who are from commoner backgrounds in each prefecture-term.[4] We interpret this outcome as capturing the extent to which candidates outside established elite families were selected into the prestigious civil service. This measure is central to our argument because commoner candidates were precisely the group most likely to be disadvantaged by identity-based favoritism and elite recommendation networks before the reforms. If the reforms were successful in improving meritocratic selection, we should observe an increase in the share of commoner candidates who succeeded in the examination system, especially in prefectures where great-family influence had previously been stronger.

We find that the reforms substantially increased commoner access to bureaucratic office. Prefectures with great families experienced an approximately 18-percentage-point increase in the share of successful candidates from commoner backgrounds after the reforms. The estimates show no evidence of differential pre-trends between treated and control prefectures before the reforms. After the reforms, however, prefectures with great families experienced a sharp and persistent rise in commoner representation among exam passers. The results are robust to including prefecture

[4]We define individuals as either commoners or elites based on their family background and kinship ties. Elites encompass more than great families to include any clan or family ties to those with official positions. This is defined in greater detail in Section 4.1.

characteristics interacted with post-reform indicators, using a propensity-score-matched sample, and measuring treatment intensity by distance to the nearest great-family prefecture.

We next examine whether increased selection of commoners came at the expense of candidate quality. On average, commoner exam passers were less likely than elite exam passers to produce high-quality written work and less likely to reach high official ranks, suggesting that elite family background continued to confer other career advantages through wealth, education, networks, and support. However, the gap in these outcomes between commoner and elite exam passers narrowed significantly after the reforms. This pattern suggests that the post-reform commoner entrants were positively selected relative to earlier commoner cohorts. Among those who were promoted to prefectural governors, we find no evidence that commoner officials performed worse than elite officials across measures of local governance, including school construction, famine, and social unrest. These results suggest that the reforms expanded access to office without reducing the quality of officials who ultimately governed.

Finally, we test whether increasing meritocracy in the selection of talent had broader consequences on social unrest. There are several mechanisms through which this may occur. First, the population may be happier with the government as a result of the government combating corruption and promoting fairness and less likely to agitate. Second, if favoritism in bureaucratic recruitment generated resentment among educated candidates who were excluded from office, reforms that made selection more meritocratic may have reduced their grievances and improved social stability.[5] Third, changes in the identity of who is in government may alter policies that in turn drive the propensity for unrest.

We find that treated prefectures experienced a decline in social unrest after the reforms in the examinations. The decline is concentrated in unrest led by scholars. This pattern is consistent with the mechanism that reforms to the examination system reduced grievances among educated aspirants, the group most directly affected by unfair access to bureaucratic office. By contrast, the effects on unrest led by non-scholars are weaker, suggesting that the reforms did not simply

[5]The first mechanism is about the general population's satisfaction with government policy, while the second one emphasizes the role of social mobility highlighted by historians as far back as 1227 (Wang (1981), Liu (2006)). Wang and Yang (2025) have documented a correlation between the emergence of the imperial examination system in China, social mobility and political stability.

reduce all forms of local disorder but specifically affected unrest connected to educated political discontent. Given the lack of observable changes in the policy outcomes associated with commoner prefectural governors, this provides evidence against the third mechanism that policy changes were the vehicle for less social unrest.

This paper is unique in bridging the literatures on selection into public service and social unrest in the historical context of China. First, it contributes to the literature on selection into public service. A large theoretical literature argues that corruption and rent seeking can distort the allocation of talent across occupations and activities, with potentially important consequences for productivity, state capacity, and long-run development (Murphy et al., 1991; Shleifer and Vishny, 1993; Acemoglu, 1995). A growing empirical literature documents how corruption, patronage, and social connections affect selection into valuable positions. For example, politically connected individuals are more likely to obtain public-sector jobs, often at the expense of competence (Colonnelli et al., 2020). Corrupt hiring can distort the allocation of bureaucratic positions (Weaver, 2021; Xu, 2018). Related work on public-sector personnel shows that recruitment rules and incentives can shape the honesty, ability, and motivation of individuals entering the state (Hanna and Wang, 2017; Dal Bó et al., 2013; Finan et al., 2017; Barfort et al., 2019). We extend this literature by showing that changes in the selection process can affect not only the allocation of jobs or honors, but also the composition of high-level governing elites and the stability of the political order.

Second, we contribute to the literature that identifies the causes of social unrest. The focus of this existing literature has been on economic causes of social unrest.[6] For example, papers have shown that food prices (Bellemare, 2015), labor-saving technology (Caprettini and Voth, 2020), economic downturns (Manacorda and Tesei, 2020; Campante and Chor, 2012), credit contractions (Braggion et al., 2020) and government austerity measures (Ponticelli and Voth, 2020) increase social unrest.[7] To our knowledge, we are the first to provide empirical evidence testing the causal link between meritocratic recruitment and social unrest. Conceptually, this is related to theories of grievance and political opportunity: when institutional channels of advancement are perceived as closed or unfair, excluded groups may turn to contentious forms of political expression instead

[6]There is also a related literature that links economic causes and civil conflict (Miguel et al., 2004).
[7]Cantoni et al. (2024) find that year-to-year economic fluctuations do not correlate with protests.

(Hirschman, 1972; Eisinger, 1973).

Finally, the paper contributes to the existing economics literature focusing on the historically important civil service examinations in China. Chen et al. (2020) show that the Chinese civil service examination system during the Ming-Qing period (1368–1905) has persistent effects on educational outcomes in modern times. Yuchtman (2017) explores the historical wage premiums associated with passing the Chinese civil service exams. Bai and Jia (2016) and Hao et al. (2022) link the abolition of the civil service examination system in China in 1905 to the probability of social uprising and protests. The focus of our paper is on these earlier merit-based reforms in the civil service examination system, and we are the first to document that anonymizing the grading of exams and removing discretion in recommendations changes selection and reduces the probability of social unrest.

The next section of the paper provides institutional background on the historical context we are studying. In Section 3, we present the identification strategy. Section 4 describes the data. The results are presented in Section 5. The paper then concludes.

2 Institutional Background

2.1 Imperial Examination System

The Imperial Examination System was a nationwide selection process for government officials in Imperial China that lasted nearly 1,400 years. From its inception, the system served as a prestigious pathway to government office, providing entry into the ruling elite for both elites and commoners.

During the Northern Song dynasty (960-1127), the examinations comprised three levels: an examination administered by the local prefectural government, a central-level examination, and the final *jinshi* examination, also known as the palace examination, which was held in the imperial palace of the capital under the supervision of the emperor.

The prefecture-level examination was a written exam testing poetry, composition, argumentative essays, policy essays, classics, and Confucian texts. The examination was designed to assess

classical education, literary sophistication, and political judgment. The prefecture-level examination operated under a quota system, with quotas assigned by the central Department of Rites to each province, and the provincial governor further assigning those quotas to each prefecture. The pass rate was quite low: for example, in one local gazetteer in 1090 in Fuzhou prefecture, about 1.33% passed (He, 1629).[8] Those who passed became "recommended scholars" (*juren*), and were eligible for the central-level exams. The *juren* degree was not just an academic achievement but also a symbol of social status.[9]

The content and style of the central-level examination were similar to the prefecture-level examination. The acceptance rate of candidates advancing to the palace examination was fairly stable around 10.5% during the Northern Song dynasty (Chaffee, 1985). The final palace examination was held in the imperial court and was a written exam. Its contents differed from the prior examinations: while the prior examinations were more comprehensive, the palace examination was usually more limited, requiring candidates to complete poetry and an essay in a single day. The acceptance rate for the palace examination varied across different periods, ranging from at least 60% in the period from 998 to 1021 and peaking at 100% in the period from 1086 to 1127 (Chaffee, 1985). Those who passed the palace exam were conferred with the highly prestigious *jinshi* ("advanced scholar") degree. During the Song dynasty, individuals with a *jinshi* title constituted approximately 45% of the total bureaucracy (Chaffee, 1985).

The prefectural examination is central to our study for two reasons. First, it was the first formal gatekeeping stage in the entire examination hierarchy. Second, the prefectural examination operated under local quotas. Because each prefecture had a limited number of *juren* slots, favoritism toward connected candidates had direct allocative consequences: admitting a connected candidate reduced the probability that an unconnected candidate could advance. Bias at this stage therefore affected not only the fairness of grading, but also the allocation of scarce opportunities for entry into the governing elite.

[8] Specifically, around 3,000 candidates participated in the prefecture-level exam and 40 passed.

[9] *Juren* scholars were also exempted from state service obligations and had the right to convert certain punishments into fines (Chaffee, 1985).

2.2 Great Families and Their Influence on Imperial Examinations

Great families were politically influential lineages with deep roots in particular prefectures. Their influence was not limited to social prestige; many of these families specialized in government service and maintained ties to high-ranking officials. Because the prefectural examination was administered locally and involved both grading and recommendation, these families had opportunities to influence which candidates advanced. Their presence therefore provides a useful proxy for prefectures where pre-reform elite capture in examination selection was likely to be more severe.

One clear indication of the power of the great families was their presence among the top ruling elite of the Song dynasty (Hartwell, 1982). For example, the Wang family in Taiyuan Prefecture, a prominent great family from the Tang dynasty, remained in the top tier of the ruling class during the Northern Song dynasty. Among the members of this family, Wang Dan (957-1017) served as Chancellor under Emperor Zhenzong for more than a dozen years in the early 11th century until his resignation due to illness in 1017 (Tuotuo, 1977). Wang Dan's grandson-in-law, Han Jiang (1005-1088), also served as a Chancellor during the mid-Northern Song period under Emperor Shenzong (r. 1068-1085) (Hartwell, 1982).

The duration of influence of great families is subject to debate among historians. However, even David Johnson, the most prominent historian who believed that the influence of medieval great families decreased during the late Tang and Five Dynasties period (the 9th-10th centuries), acknowledged that when it came to attaining official positions in the Song dynasty, coming from one of these great families remained a significant advantage. As he put it: "In practice, it was a formidable advantage to come from a powerful family" (Johnson, 1977).

Given their enduring social and political influence, it is not surprising that these great families leveraged their status to exert undue influence over the examination process. Members of these families often received favoritism in grading from the examiners at the prefecture and central levels, and in the extreme cases, they might initially fail the exam but still be selected as *juren* scholars. Historical records from the Song period document accounts of such manipulations and interventions by great families in imperial examinations (Chaffee, 1985).

One famous Song scholar named Liu Kai (948-1001) openly questioned the fairness of the

imperial examination system in the mid-to-late 10th century. He wrote, "If someone is born into a great family, with a powerful father and distinguished brothers, their honor or disgrace is enough to draw immense attention or cause a stir in society, and their extensive network of relationships could compel everything to fall in their favor. When such a person comes to claim a *jinshi* title, would the officials truly dare to resist his overwhelming power?" (Liu, 2015). His critique highlights the unfairness of the system but also implicitly calls for reforms that would level the playing field for all candidates.

This influence of the great families on imperial examinations was even a concern of the ruling house. In one of the Northern Song dynasty imperial edicts, Emperor Taizu, the founder of the Northern Song dynasty, explicitly stated that "*jinshi* degrees were predominantly awarded to powerful great families, which obstructed the path for the humble and impoverished" (Li, 2004).

2.3 Two Types of Reforms Targeting Bias in Imperial Examinations

The government attempted to crack down on the bias and corruption in the imperial examination system through two types of reforms. The first type of reform focused on *identity* concealment. The second type of reform focused on *discretion* and removing personal recommendations from the process.

In 1033, the state introduced blind marking (*huming*) in prefectural examinations. Candidate names and personal information were removed from exam submissions before grading. In 1037, the state added copied transcription (*tenglu*) under which exam papers were recopied before grading so that examiners could not identify candidates by handwriting.[10] These reforms targeted favoritism in grading by making it harder for examiners to observe a candidate's identity, family background, or social connections.

Despite these two reforms, local officials could still bypass exam scores and directly recommend candidates, reinstating those who had failed the examination onto the *juren* list to proceed for the central exam (Zhu and Fu, 2012). Indeed, officials in local prefectures often prioritized family

[10]The new 1033 and 1037 policies were met with resistance. The resistance to these reforms peaked in 1043, when the new examination policies were reversed as part of the Qingli Reform (Lee, 1979). The two reforms aimed at concealing the identity of test takers were repealed in favor of reinstating the original local recommendation system in 1043-1044 (Li, 2004). However, in 1045, both reforms concealing identity of examinees were fully reinstated.

reputation over exam performance in these recommendations. In an official statement issued in 1044, the Department of Rites explicitly criticized the recommendation system's lack of objectivity (Li, 2004).[11]

The second type of reform involved the removal of discretion. In 1044–1045, the state abolished discretionary local recommendation and limited prefectural officials to verifying candidates' documents and conduct. Grading authority was transferred to professional examiners appointed by higher-level administrative offices. After this reform, examination performance became the sole criterion for prefecture-level selection. Together, the reforms reduced both identity-based favoritism in grading and discretionary favoritism in recommendation.

3 Identification Strategy

To evaluate the effects of these reforms, we employ a difference-in-differences framework. In addition to looking in the years before and after the implementation of the reforms, we use the residency prefectures of the great families as the treatment group. The prefectures without the presence of the great families serve as the control group.

This strategy assumes that prefectures where great families are located had more influence over the prefecture-level imperial exams prior to the reforms than in prefectures in which they were not directly residing and working. We are not arguing that the control prefectures are free of corruption and bias in imperial examinations. There were also politically connected candidates in these prefectures. However, our identification strategy relies on the key assumption that there was *more* bias in the prefectures in which great families were located. Indeed, the presence of corruption and bias in our control prefectures would lead to an underestimation of the true effects of the reforms. The key identifying assumption is that, absent the reforms, treated and control prefectures would have followed parallel trends in the outcomes of interest.

[11] A debate then took place among the central-government officials and the emperor on whether prefecture-level selection should be based purely on examination scores or a mixture of exam scores and recommendations by local officials from 1043 to 1045 (Xu, 2014).

Our baseline specification takes the form:

$$Y_{pt} = \beta(Treat_p \times Post_t) + \delta_p + \mu_t + \epsilon_{pt} \tag{1}$$

where Y_{pt} denotes an outcome in prefecture p and term t, $Treat_p$ is an indicator for whether prefecture p was a great-family residency prefecture, and $Post_t$ indicates the post-reform period. Prefecture fixed effects δ_p absorb time-invariant differences across prefectures, including geography, historical educational traditions, and persistent elite presence. Term fixed effects μ_t absorb shocks common to all prefectures in a given period. Standard errors are clustered at the prefecture level.

The coefficient β provides the key estimate of interest, capturing the different effects of the reforms on treated prefectures compared with control prefectures. Because the reforms applied nationally, our estimates of β do not capture the aggregate effect of the reforms relative to no reforms at all. Instead, they capture whether the reforms had larger effects in prefectures where pre-reform elite capture was likely more severe.

We estimate several versions of this specification. First, we use an indicator for the period after 1033, when anonymous grading was introduced. Second, we use an indicator for the period after 1045, when discretionary local recommendation was abolished and the anonymity reforms were reinstated. Third, we allow the effects to differ between the intermediate period after the identity-concealment reforms and the period after the discretion reform. This allows us to examine whether the main effects occurred after identity concealment or after the removal of local recommendation discretion.

To address the possibility that treated and control prefectures differed in ways that generated different post-reform trends, we also have a specification that includes interactions between post-reform indicators and predetermined prefecture characteristics, including population, land size, distance to major rivers, distance to the border, historical Confucian educational resources, distance to publishers, and climate conditions. In addition, we implement a propensity-score-matched difference-in-differences design that compares treated prefectures to observationally similar control prefectures. Finally, we use the distance from each prefecture to the nearest great-family prefecture as an alternative measure of exposure to pre-reform elite influence. Together, these exercises

test whether the results are robust to alternative ways of accounting for differential exposure to great-family influence and observable prefecture characteristics.

4 Data

This section describes the four historical datasets collected for the analysis. First, we construct a prefecture panel of successful examination candidates and classify each *jinshi* by family background to measure commoner access to elite office. Second, we identify the residency prefectures of great families, which define exposure to pre-reform exam influence. Third, we construct a prefecture panel of social unrest events. Fourth, we collect individual-level career outcomes for *jinshi* and prefecture governors.

4.1 Data for Advanced Scholars (*Jinshi*)

To understand the types of individuals who pass the examinations, we gather data on *jinshi* during the period 960-1127, sourced from the China Biographical Database (CBDB).[12] The database includes information on their name, the year their *jinshi* title was granted, their hometown, and their kinship ties.

While the majority of *jinshi* obtained their title through imperial examination, a small minority received the title by imperial grace. Since our study focuses on reforms to the examination system at the prefecture-level, we exclude individuals whose titles were granted by the discretion of the emperor. Excluding 310 titles granted through imperial grace, our final sample consists of 9,454 *jinshi*.

We then manually check each *jinshi*'s family background and kinship ties to determine whether he is a commoner or from an elite family. Note that our use of the term "elite family" is not limited to the 778 great families; rather, it includes any clans or families with official backgrounds.[13] To do this, we search in several sources, including the CBDB and the China Essential Ancient

[12]The database is jointly developed by Harvard University, Academia Sinica, and Peking University. It can be accessed through https://projects.iq.harvard.edu/cbdb.

[13]Great family status defines prefecture-level treatment. Commoner versus elite family background is an individual-level classification for our outcome of interest.

Books Database by Airusheng.[14] The China Essential Ancient Books Database, developed by the Beijing Airusheng Digital Technology Research Center, collected and digitalized approximately 10,000 Chinese texts dating from the pre-Qin period (11th century BC).[15] It is among the most authoritative and comprehensive full-text databases of premodern Chinese literature.

We classify a *jinshi* as commoner if he did not come from an identified elite family and if none of his patrilineal elders held official positions. This definition is intentionally conservative: individuals are coded as elite when we find evidence of official family background or elite lineage ties. Following this process, we identify 69.3% of the *jinshi* in our dataset as commoners (Table 1).

Next, we construct a dataset of all prefectures in the Northern Song dynasty. The complete set of prefectures is compiled from "The General History of Chinese Administrative Divisions" (Zhou, 2007). Since some prefectures experienced name changes during the sample period, we take the early Northern Song dynasty (960-1000) as the reference point and assign a unique identifier to each prefecture over time. We then cross-reference each *jinshi*'s hometown information to identify its associated prefecture. Using this information, we calculate the proportion of commoner *jinshi* (denoted as $CommonerRatio$) at each prefecture and term level. Given that the examination cycle during the Northern Song dynasty varied between one and five years, we define a term as five years to ensure that at least one examination cycle is included per term, with the first reform year, 1033, marking the end of one term. In the analysis looking at the composition of those passing the exam, we have 1966 prefecture-term-level observations, covering 156 prefectures and 32 terms.[16] The left panel of Figure 1 shows the geographic distribution of the proportion of commoner *jinshi*, aggregated across the entire sample period.

4.2 Prefecture Residency of the Great Families

The primary sources of information about the great families are genealogical works, which document their lineage and social standing.[17] In this study, we use the latest compilations, *Shizu Lun*

[14] See Figure A.1 as an example.
[15] This database can be accessed at `http://er07.com`.
[16] Our analysis omits prefectures with no one passing the exam over time (i.e. $CommonerRatio$ not defined).
[17] The wealth of genealogical studies reflected the efforts of great families themselves in consolidating their identity, maintaining social status, and legitimizing their political power.

(Treatise on Great Families) and *Xinji tianxia xingwang shizu pu* (Newly Compiled Genealogical Register of Great Families and Prestigious Lineages of the Realm), of the great families prior to the Song dynasty.[18] These are recognized as the most reliable sources in this regard by historians (Johnson, 1977). We identify 778 great families located in 78 prefectures. A total of 41.1% of the observations are from prefectures in which great families reside (treatment group). The right panel of Figure 1 shows the treated prefectures, corresponding to the residency prefectures of the great families. The complete list of the great families and their residency prefectures is presented in Table A.1.

For a robustness check, we also use the geographical distance between each prefecture and the nearest great family's location (denoted as *Distance*) as a proxy for the degree of pre-reform corruption and bias, given that great families may influence their neighboring prefectures. To calculate geographical distance, we use the coordinates provided by the China Historical Geographical Information System (CHGIS), developed by the Center for Geographic Analysis at Harvard University.

4.3 Performance Outcomes of Exam Passers

Academic performance. We measure a *jinshi*'s academic performance using the number of Song prose works recorded in "Quan Song Wen" (Zeng and Liu, 1988), a comprehensive collection of widely circulated prose from the Song dynasty. This collection contains more than 170,000 works authored by more than 10,000 individuals.[19] We search for the prose works attributed to each *jinshi* and construct a dummy variable equal to one if a *jinshi* scholar has at least one prose work included in "Quan Song Wen" ($HasArticle$). Within our sample, we identified 16.5% of the *jinshi* who had authored at least one piece of prose published within "Quan Song Wen" (Table 1).[20]

Political achievement. To measure the political career achievement of a *jinshi*, we use the rank of the highest official title they obtained. We extract official position histories for each *jinshi* from the CBDB. We then manually identify the rank of each official title by referencing Hucker (1985),

[18] See Figure A.2 for an example of how we coded this data from these historical records.
[19] See Figure A.4 as an example.
[20] While almost all *jinshi* would have written essays, official documents and policy discussions, the works in the "Quan Song Wen" capture a smaller subset that were considered higher quality or having greater political influence.

which offers the most comprehensive guide of official titles in the traditional Chinese government. Official titles in traditional China are classified into nine grades, where grade 1 is the highest and grade 9 is the lowest. Each grade is further divided into two classes, "a" and "b", with "a" being superior.[21] The governor of a prefecture (*zhizhou*) is ranked as 5b.[22] We then construct two measures to proxy for political career achievement. One is *AnyOfficial*, which is a dummy variable that equals one if the *jinshi* obtained an official title within the government. The other variable is *HigherOfficial*, which is a dummy variable used to indicate whether a *jinshi* achieved an official title equal to prefecture governor or higher. As shown in Table 1, 33.4% of *jinshi* obtained an official title, and 21.8% of *jinshi* achieved an official title ranked at or above prefecture governor.

We measure the governance performance of those who attained the official position of prefecture governor or higher. By examining the official career histories of all *jinshi*, we identify 856 prefectural governors, along with their governing prefectures and years. Of these 856, 217 were commoners. We construct a prefecture-year panel based on their governing tenures and define a dummy variable, $CommonerHigherOfficial$, which equals one if the higher official is a commoner and zero if elite. This occurs in 22.9% of the prefecture-terms (Table 1). The resulting dataset consists of 3,974 observations.

Governance performance. We measure prefecture governance performance of prefectural governors using three indicators. First, we construct a dummy variable, $NewSchool$, indicating whether a new county school was built or an existing one rebuilt in a prefecture-year during the prefecture governor's tenure. These schools are free of charge and target commoners. We compile historical records on county schools, including their locations and years of construction during this period, from Zhou (1996) and match them to prefecture-year observations.[23] New schools occur in 3.2% of prefecture-years. Second, we construct a dummy variable, $Famine$, indicating whether a famine occurred in a given prefecture-year. We manually collect the historical records of famine from "History of the Song" (Tuotuo, 1977), "Song Huiyao Jigao" (Xu, 2014), and "A Chronology of Natural and Human Disasters in Chinese History" (Chen, 1986).[24] Famines occur in 1.9% of

[21] In our coding, a 5a would correspond with 5.0 and 5b with 5.5.
[22] See Figure A.5 for examples of the mapping.
[23] See Figure A.6 as an example.
[24] See Figure A.7 as an example.

prefecture-years. Third, we measure the incidence of a social unrest event in a prefecture-year as $HasUnrest$, which is discussed in detail in the next subsection.[25]

4.4 Social Unrest

We construct an original dataset documenting social unrest events that occurred during the Northern Song dynasty. The events are manually collected from "A Compilation of Historical Records on Peasant Rebellions during the Song" (He, 1976), "Chronological Table of the Wars in China" (Fu et al., 2003), and "History of the Song" (Tuotuo, 1977).[26] Conflicts and wars involving foreign entities are excluded from our analysis. The final dataset comprises 404 social unrest events, each recorded with its starting year and associated prefectural location(s). We then construct a panel dataset at the prefecture and five-year term levels, where the dummy variable, $HasUnrest$, indicates whether any social unrest started in the prefecture-term. As shown in Table 1, social unrest occurs in 4% of prefecture-term observation.

To classify each social unrest according to whether it was led by educated individuals, we manually examined historical texts. Specifically, an event is coded as educated-led if the historical record indicates that its leaders or core organizers were Confucian scholars (*juren*), private secretaries (*muliao*), or other literati. We also categorized the social unrest as educated-led if the uprising exhibited literati-style political practices, such as the issuance of manifestos (*xiwen*), or political essays (*celun*). Using these criteria, we identify 81 social unrest events as being led by educated individuals. We then construct a dummy variable, $HasUnrestEducated$. Less than 1% of prefecture-terms have an episode of social unrest led by educated individuals.

4.5 Control Variables

We use prefecture characteristics to address the concern that treated and control prefectures may have differed in ways that generated different post-reform trends. These controls include climate shocks, population, land size, distance to major rivers, distance to the border, historical

[25] Note this is the same measure discussed in the next subsection but looking at the prefecture governor's location rather than the location of origin of *jinshi*.

[26] See Figure A.3 as an example.

Confucian educational resources, and distance to publishing houses. In the main specifications, these characteristics are interacted with post-reform indicators rather than included directly, since prefecture fixed effects absorb time-invariant levels.

Climate: We obtain natural climate hazard records from the "Atlas of Regional Extreme Droughts and Floods in China Over the Millennium" (Yang, 2024), which documents the locations and years of droughts and floods in China from 1000 to 2020. The atlas assigns a severity rating to each natural hazard, ranging from level 1 to level 5. We use these severity levels (*Climate*) ranging from 0 to 5 where 0 corresponds to no severe hazard to control for the impact of natural hazards within a prefecture. We control for extreme climate because it can significantly affect agricultural productivity and economic stability, which in turn affect educational opportunities and resources for exam preparation. The average value of this variable is 0.89, reflecting the preponderance of zeros in this variable (Table 1).

Demography: We measure population size (*Population*) using the estimated number of households for each prefecture based on Ge (2000).[27] We take the estimate in year 976 for our analysis. The average number of households per prefecture is 21,443 (Table 1).

Geography: We control for three geographical characteristics. First, we measure the land area of a prefecture in square kilometers (*Landsize*). It is obtained from prefecture-level shapefiles provided by CHGIS (2016) and supplemented with numbers from the "Universal Geography of the Taiping Era 976-983". The average prefecture in our sample is 13,406 square kilometers (Table 1).

Second, we calculate the distance to a major river as *DistanceRiver*.[28] Using shapefiles of all major rivers in China from "China Rivers, 1820", provided by CHGIS (2016), we compute the distance from the central coordinates of each prefecture to the nearest major river. The distance is measured in kilometers. On average, as shown in Table 1, the prefectures are 65.8 kilometers from a major river. We control for distance to major rivers because regions closer to rivers typically exhibit better economic conditions. Additionally, proximity to waterways facilitates

[27] For each prefecture, Ge (2000) provides three estimates in years 976, 1078, and 1102.

[28] Major rivers are defined by their discussion in historical texts as watercourses (*chuan*), waterways (*shui*), large rivers (*jiang*), or rivers (*he*), rather than brook (*xi*), ravine streams (*jian*), or gullies (*gou*). During the period of our analysis, there are 525 total major rivers.

easier access to information and transportation, which may positively influence performance in imperial examinations.

Third, we measure the distance of each prefecture to the national border as *DistanceBorder*. It is calculated using prefecture-level shapefiles provided by CHGIS (2016). We control for distance to the border to account for exposure to foreign conflicts.[29] The prefectures are on average 93 kilometers from the border (Table 1).

Education: To account for the possibility that initial levels of education had differential impacts over time, we measure the prevalence of Confucian scholars in each prefecture. Using records from the "Old Tang Book" (Liu, 1975) and "New Tang Book" (Ouyang, 1976), we compile a list of all Confucian philosophers during the Tang dynasty and count the number of philosophers (denoted as *PhilosopherNum*) associated with each prefecture. In our sample, we identify a total of 59 Confucian philosophers. Regions with more Confucian philosophers may have higher education levels for exams. The average prefecture has 0.19 Confucian scholars (Table 1).

Cost of Education: Movable type was invented around 1040 and increased the spread of printed materials, giving commoners better access to texts and study materials. Thus, we control for the proximity of each prefecture to a publishing house. Using data from Zhang (2006), which documents the history of printing in China, we collect the locations of official publishing houses established during the Five Dynasties and Northern Song periods (907-1127). We then calculate the geographical distance from the central coordinate of each prefecture to the nearest publishing house as *DistancePublisher*. The average distance is 397 kilometers (Table 1).

5 Results

5.1 Descriptive Evidence

Before proceeding to formal analysis, we present suggestive evidence by plotting the national proportion of commoner *jinshi* over the Northern Song period, which is calculated for each five-year term for treated and control prefectures in Figure 2. The red vertical dashed line indicates

[29] There were 111 foreign conflicts during this period.

the start of the blind marking reforms in 1033, while the blue vertical line indicates the second reform focused on discretion in 1045. The proportions are calculated separately for the two groups in each five-year term, with each group demeaned at its average level prior to the year 1000 to facilitate comparison. The figure indicates that both the treated and control groups exhibit a similar, relatively flat trend prior to the first reform year in 1033.[30] This provides support for parallel pre-reform trends. After 1033, the treated group experiences an upward trend in the share of commoner *jinshi*, diverging significantly from the control group.

5.2 Exam Reforms and Commoners Passing

We begin with the outcome of the proportion of commoners among *jinshi* recipients. The estimates are shown in Table 2 where the odd columns show the parsimonious effects while the even columns show the estimates with the time-varying impact of prefecture-level controls. In columns (1) and (2) of Table 2, we focus on the period after the blind marking reform began (1033). The estimates indicate a 10.4 to 19.3 percentage point increase in the proportion of commoner *jinshi*, which is over a substantial 30% increase with respect to the sample mean shown in Table 1 (0.641). The estimates are significant at the 5 and 1% levels, respectively.

In columns (3) and (4), we examine the period after 1045 when the recommendation discretion reform began. The estimates, at 8.8 to 16.6 percentage points, are slightly lower than the average in the period after 1033. These estimates are significant at the 10% and 1% levels, respectively.

Finally, in the last two columns of Table 2, we allow for separate effects of the two types of reforms. In column (5), we see that the blind marking reforms were effective at increasing commoner representation among *jinshi* by 9.1 percentage points. There is a slightly larger increase after 1045 (10.5 percentage points). This suggests that the largest effects came from the stage of anonymizing the exams, and the discretion reform increased the proportion of commoners as *jinshi* by an additional 1.4 percentage points. However, we note that these are not statistically different from each other. The estimates with the additional controls are larger in magnitude but also show similar results that suggest the largest effects are coming from the blind marking reforms with a

[30] Note that while the first reform was announced in 1033, given the timing of the exams, the first class for which the implementation of the reform is relevant would be one selected in 1034.

small additional change potentially associated with the recommendation discretion reform.

5.2.1 Parallel Trends and Dynamic Effects

The key assumption of our identification strategy is the parallel trends assumption. We have already shown descriptive evidence to support this in Figure 2. We also extend our baseline model equation 1 by replacing the post dummy with a set of new dummy variables that indicate five-year intervals before and after 1033 –$Pre_{(1019,1023)}$, $Pre_{(1024,1028)}$, $Pre_{(1029,1033)}$, $Post_{(1034,1043)}$, $Post_{(1044,1048)}$, $Post_{(1049,1053)}$, $Post_{(1054,1058)}$, and $Post_{1059}$. These variables indicate each of the terms related to the reform events. We can also examine whether the results persist over time or if the elite find other ways to keep commoners from succeeding. In Column (1) of Table A.2, we can see that the coefficients of the pre-period dummies that interact with Treat are nonsignificant and fluctuate noisily around zero. The results confirm that the treated and control groups exhibit parallel trends in the pre-reform periods, validating the parallel trends assumption. Column (2) explores the dynamic effects of these reforms. Overall, we see similar, substantial and almost immediate positive impacts of these two reforms during different post-reform terms, comparable to what is observed in the baseline result in magnitude, suggesting the long-lasting effect of these two reforms. In Column (3), we divide both the pre- and post-reform periods into different terms and illustrate the full dynamics of the effects of these two reforms, and similar results are observed.

5.2.2 Propensity Score Matching

Although our difference-in-differences design is supported by the parallel trends evidence and includes controls for time-varying observable differences, treated and control prefectures may still differ in their characteristics. As an additional robustness check, we implement propensity score matching (PSM) to construct a comparison group of control prefectures that is more similar to the treated prefectures. We then re-estimate our DiD specification using the matched sample to assess whether the estimated treatment effects are robust to improved covariate balance between treated and control prefectures.

Using all the prefecture-level control variables, we match each treated prefecture to a control

prefecture through nearest neighbor one-to-one matching without replacement. After the matching process, we obtain a sample that includes a total of 112 prefectures with 1,572 observations. Panel A of Table A.3 presents a comparison of the characteristics between the treated prefectures and their matched counterparts. For most of the six prefecture-level variables, the differences in means between the two groups are statistically insignificant. The exception is the number of Confucian philosophers, which is relatively concentrated in a small number of prefectures. In Panel A of Table 3, we present the estimates using the matched sample. In all columns, the coefficients are positive and similar in magnitude to our baseline estimates. The estimates are significant at the 5% level or higher. These results suggest that our results are not driven by differences between the treatment and control groups.

5.2.3 Alternative Measure of Treatment

In addition to creating the treatment group versus the control group on the basis of the presence of great families, we use the geographical distance between each prefecture and the nearest great family's location (denoted as *Distance*) as a proxy for the degree of pre-reform corruption and bias. In Panel B of Table 3, the coefficients on $Distance \times Post_{1033}$ and $Distance \times Post_{1045}$ are both negative and significant. This indicates that prefectures closer to great families experience a greater increase in the share of commoners passing the imperial exams after the reforms.

5.2.4 Potential Confounding Events

We consider one potential confounding shock, the invention of ceramic movable-type printing around 1040 (Needham, 1954).[31] This technological breakthrough introduced movable components to reproduce textual elements efficiently. The new technology increased the dissemination of printed materials, including educational texts and examination preparation resources, and reduced prices of printed materials. Thus, it had the potential to increase the chance of success for commoners who might not have had the financial resources to buy these materials before such an invention.

[31] We discuss other events that occur around the same time as the reforms of our focus in Appendix A.

In our main analyses above, when we control for the interaction between a prefecture's distance to the nearest publishing house ($DistancePublisher$) and our reform time dummies, the coefficients are larger. Furthermore, if the new printing technology did increase the likelihood of commoner candidates, those who were closer to the publishing house would benefit more. However, in Table A.4, we present the full regression output. We do not find a significant positive effect of $DistancePublisher$.

5.3 Career Achievements of Commoner and Elite *Jinshi* Scholars

Following these reforms, commoners were more likely to be admitted to "the preeminent status hierarchy in Chinese society, which usually entailed power and wealth" (Chaffee, 1985). In this section, we consider whether commoner *jinshi* scholars achieve better or worse academic and political performance than elite *jinshi* scholars. This can help shed light on whether the examination reforms affected talent misallocation. For example, it could be that giving discretion for promotion rather than relying solely on exam scores improved selection of productive individuals.

The regression specification changes to reflect that we are now looking at the long-run career outcomes of those who succeeded in becoming a *jinshi*. We estimate the following equation:

$$CareerY_i = \beta_1 CommonerJinshi_i + \beta_2 CommonerJinshi_i \times Select(1033, 1045)_{it} + \beta_3 CommonerJinshi_i \times SelectPost1045_{it} + \mu_t + \epsilon_{it} \quad (2)$$

where $CareerY_i$ is a long-run career outcome of *jinshi* i. $CommonerJinshi_i$ is an indicator for whether the *jinshi* was from a commoner background versus elite. The subscript t here denotes the term that the *jinshi* passed the exam. The variables $Select(1033, 1045)_{it}$ and $SelectPost1045_{it}$ indicate whether the term that the *jinshi* passed the exam was between the two reform years or after the second reform. The coefficient β_1 tells us whether commoners have different career outcomes relative to elites. The estimates on β_2 and β_3 indicate whether the correlations between career outcomes and commoner status changes for those selected in the period after the first reform (but before the second reform) and after the second reform, respectively.

We first look at the correlation in our sample of commoner versus elite status on having a prestigious academic article ($HasArticle$). In Table 4, column (1) shows that on average, *jinshi* from non-elite backgrounds have a 29.1 percentage point lower probability of producing high-quality prose. This estimate is significant at the 1% level. In general, the writings of commoner *jinshi* are less likely to be considered important enough to be included in "Quan Song Wen." This reflects lower quality in commoner *jinshi's* academic work relative to elite *jinshi*. Column (1) also shows a positive coefficient of 0.070 on the interaction term $CommonerJinshi \times SelectPost1045$ and a positive coefficient of 0.079 on $CommonerJinshi \times Select(1033, 1045)$. Both estimates are significant at the 5% level or higher. This suggests that the gap in academic performance between commoner *jinshi* and elite *jinshi* narrowed by more than one-quarter following the reforms.

We also investigate the political achievements of elite and non-elite *jinshi* scholars. We first examine whether they obtained a government official title ($AnyOfficial$) in column (2), then whether they became higher-level government officials ($HigherOfficial$) in column (3). In column (2) of Table 4, the patterns for political achievements are very similar to academic achievements. Overall, commoners are much less likely to become officials or higher officials. The coefficients are large in magnitude, suggesting commoners are about 50 percentage points less likely to become either an official or a higher-level official relative to elites. Both estimates are significant at the 1% level. However, as shown by the positive and significant coefficients on the interaction terms $CommonerJinshi \times SelectPost1045$, the gap significantly narrows for both political outcomes after the reforms. Overall, the reforms corresponded with a slight improvement in both the academic and political performance of commoner *jinshi* relative to elites but did not close the gap entirely.

In terms of general negative performance of commoners relative to elites as *jinshi*, we conjecture that their limited access to economic resources and political connections might have hindered their advancement in the long term. Another possibility is that while commoner *jinshi* scholars might have been skilled in taking imperial examinations, they might not have the ability to write high-quality prose or to excel in the civil service.

This provides suggestive evidence that the selection of commoners changed such that those

who achieved *jinshi* titles after the reforms (and who would not have passed prior to the reforms) were higher quality in both academics and politics than the average commoners who attained *jinshi* before the reform.

5.4 Governance Performance of Commoner and Elite Officials

Many civil service exams include both ability-focused examinations and interviews that assess other characteristics, such as honesty and integrity. Indeed, Fan Zhongyan, the Vice Chancellor at the time, opposed the *huming* reform in 1043, arguing that *jinshi* selection should rest on moral virtue rather than anonymized examination performance (Fan, 2007). Such discretion may incorporate traits that we think are important for civil servants beyond exam performance (Liu and Zhang, 2025). However, it simultaneously creates opportunities for favoritism and corruption. Whether examination reforms improve the efficiency of talent selection therefore depends on which of these forces dominates. If the discretion effect dominates the bias and corruption effect, we should observe that prefectural governors (*zhizhou*) drawn from commoner *jinshi* perform worse in governing a prefecture than those drawn from elite *jinshi*.

We focus on a specific group of *jinshi* - the subset appointed as the prefectural governors. This allows us to examine the three governance performance metrics that we discussed in subsection 4.3: $NewSchool$, $HasUnrest$, and $Famine$. $NewSchool$ captures whether the official was successful in getting new schools built or renovated, and is correlated with the level of education in the prefecture over time. $Famine$ is correlated both with the level of economic development of the prefecture and the prefectural governor's ability to respond to disaster and procure relief grain. $HasUnrest$ reflects the prefectural governor's capacity to maintain social stability and respond to crises.

We estimate a regression model as follows:

$$\begin{aligned} GovernanceY_{pt} = \beta_1 CommonGov_{pt} + \beta_2 CommonGov_{pt} \times Select(1033, 1045)_{pt} + \\ \beta_3 CommonGov_{pt} \times SelectPost1045_{pt} + \beta_4 Select(1033, 1045)_{pt} + \beta_5 SelectPost1045_{pt} \\ + \alpha_1 X_{pt} \times Select(1033, 1045)_{pt} + \alpha_2 X_{pt} \times SelectPost1045_{pt} + \delta_p + \mu_t + \varepsilon_{pt} \end{aligned} \tag{3}$$

where $GovernanceY_{pt}$ is one of the three governance performance outcomes in a prefecture p in year t: $NewSchool$, $Famine$, and $HasUnrest$. $CommonGov_{pt}$ is a dummy variable which equals one if the governor is a commoner *jinshi*, and zero if the governor is an elite *jinshi*. $Select(1033, 1045)_{pt}$ indicates whether the governor received the *jinshi* degree between 1033 and 1045. $SelectPost1045_{pt}$ is a dummy variable which equals one if the governor obtained his *jinshi* degree after 1045, and zero otherwise. X_{pt} are the control variables. Prefecture fixed effects and year fixed effects are also included.

The estimate of β_1 reflects what happens to the governance performance outcomes when the prefecture's governor changes from elite to commoner, or vice versa. The coefficients β_4 and β_5 reflect what happens to governance performance when the prefectural governor changes over time regarding whether the prefectural governor got his degrees between 1033 and 1045 and after 1045, respectively. The coefficient β_2 gives the impact of a change in prefectural governor from elite to commoner (or commoner to elite) between the two reforms relative to before them. The coefficient β_3 gives the impact of a change in prefectural governor from elite to commoner (or commoner to elite) after the 1045 reform relative to before both reforms.

Table 5 presents the regression results. First, in looking at the coefficients on $CommonGov$, none of the three outcome measures indicates that commoner governors show significantly poorer governance performance. Only one of the six interaction terms is significant. In Column (1), where the dependent variable is $NewSchool$, the coefficient on $CommonGov \times Select(1033, 1045)$ is negative, and significant at the 5% level. This finding suggests that governors who were commoner *jinshi* and selected between 1033 and 1045 achieved lower performance in educational development. Overall, the results suggest that governors who were commoners are not weaker in governance than those who were elites. Therefore, the costs of removing discretion in talent selection does not outweigh the benefits of the reforms.

5.5 Examination Reforms and Social Unrest

We consider whether the success of the reforms in reducing corruption in the imperial examination process translated into reduced social unrest. Table 6 presents the results corresponding

to equation 1 where the dependent variable, $HasUnrest_{pt}$, is an indicator variable representing whether any social unrest occurred in the prefecture-term. In columns (1) and (2), the coefficients on $Treat \times Post_{1033}$ and $Treat \times Post_{1045}$ both show the unrest declined by 2.8 percentage points, which represents 55% of the sample mean before the reform. Both estimates are significant at the 5% level. When we include both the interactions with the post dummies $Between_{(1033,1045)}$ and $Post_{1045}$ in the same regression in column (3), we see that the coefficient on the interaction term with $Post_{1045}$ shows a 3 percentage point decline in social unrest and is significant at the 5% level. The coefficient for the interaction term with $Between_{(1033,1045)}$ is smaller in magnitude and not statistically significant. However, in column (3), the coefficient on the interaction with $Between_{(1033,1045)}$ is not statistically different from the interaction with $Post_{1045}$.

5.5.1 Mechanisms for Changes in Social Unrest

There are three mechanisms through which the examination reforms might have reduced social unrest. First, perceptions that government institutions are improving may reduce popular protest. Second, people may perceive greater opportunities for social mobility through the imperial exam system, and this reduction in systematic inequality may reduce social unrest. Third, the mechanism may work through changes in policy.

The first potential mechanism, the possibility that public satisfaction increases when the government cracks down on corruption and elite capture, is explored in Appendix B, where we examine an anti-corruption reform in 1009 that is unrelated to the civil service examination system. There are no significant changes associated with commoners passing the exam or social unrest, suggesting that the main mechanism of the reforms that we focus on in this paper is through the selection process changing the likelihood that commoners are selected and those people being less likely to lead social unrest.

Next, we test whether the mechanism of the decline in protests is because educated commoners were less likely to lead protests against the government if they felt that they had viable pathways for mobility and changing policy by entering the civil service rather than through social unrest. In writing published in 1227, Wang Yong emphasized this idea in describing one of main rebellion

leaders at the time, Wang Xianzhi, as a person who failed the *jinshi* examination and his followers "were likewise men whose ambitions had been frustrated in the *jinshi* examination" (Wang, 1981). We estimate equation 1 using $HasUnrestEducated$ and $HasUnrestUneducated$ as dependent variables. $HasUnrestEducated$ is an indicator for social unrest led by those with education while $HasUnrestUneducated$ equals one if there was unrest in the prefecture led by those without education.

Panel A of Table 7 reports the results for the outcome, $HasUnrestEducated$, while Panel B reports the results for $HasUnrestUneducated$. In Panel A, when the dependent variable is $HasUnrestEducated$, the interactions are negative and significant at the 10% level or higher in all four of the interactions with Post1045, suggesting that the unrest declined in the long run rather than immediately after 1033. In Panel B, when the dependent variable is $HasUnrestUneducated$, the coefficients on the interaction terms are also negative in most cases but statistically insignificant at the standard levels. The magnitude of the coefficients relative to the mean of the dependent variable is also much larger for unrest led by the educated than for unrest led by the uneducated. Taken together, these results indicate that the reform reduced the probability that educated commoners led protests against the government. This provides suggestive evidence that the impacts on social unrest are driven by educated commoners having alternative means of trying to institute policy changes besides leading social unrest. The lack of results on unrest led by non-scholars also provides evidence against the first mechanism; if the reforms had increased satisfaction among the general population, we would also expect to see a decline in the unrest led by the uneducated.

On the last mechanism, where declines in social unrest are driven by policy changes, the results in Table 5 suggest that policies of commoner governors were not significantly different from the policies of elite governors. Furthermore, the timing of the changes shown in Figure 3 indicates that the declines in social unrest began within five years of the reform. Given that it takes time for newly appointed *jinshi* to rise in the ranks enough to change policy, we would expect longer lags if policy changes were the primary mechanism. Overall, the evidence indicates that any changes in social protests are unlikely to be primarily driven by policy changes.

5.5.2 Parallel Trends and Dynamic Effects

We further test for the validity of the parallel trends assumption and dynamic effects by replacing the post dummies with a set of dummy variables as follows: $Pre_{(1019,1023)}$, $Pre_{(1024,1028)}$, $Pre_{(1029,1033)}$, $Post_{(1034,1043)}$, $Post_{(1044,1048)}$, $Post_{(1049,1053)}$, $Post_{(1054,1058)}$, and $Post_{1059}$ in equation 1, and using $HasUnrest$, $HasUnrestEducated$, and $HasUnrestUneducated$ as dependent variables. Figure 3 (and Table A.5) presents the results. For all three dependent variables, the coefficients on the interaction terms between the treatment indicator and pre-period dummies are insignificant, fluctuating around zero. These results support the validity of the parallel trends assumption. Take Column (2) of Table A.5 as an example, when the dependent variable is $HasUnrestEducated$, the coefficients of the post-reform dummy interactions begin to be significantly negative with $Post_{(1049,1053)}$, and remain stably negative thereafter, including $Post_{(1054,1058)}$ and $Post_{1059}$. A similar pattern emerges for the other two dependent variables, but the effects are less statistically significant for $HasUnrest$ and insignificant for $HasUnrestUneducated$. These results provide further evidence that the reforms have a persistent negative effect on the frequency of social unrest in the treated prefectures, while the effect is stronger for those led by educated individuals. This effect appears to be a long-term consequence, emerging approximately five years after the second reform.

6 Conclusion

The selection process in public-sector recruitment can have important consequences for the identity of officials who enter into government service. When access to office is shaped by favoritism and family background, such corruption and lack of meritocracy can generate political grievances among those excluded from advancement as well as those who view the system as unfair. This paper studies the link between elite capture in the recruitment process, bureaucratic selection, and social stability in the context of reforms to the Northern Song imperial examination system.

Using newly constructed historical datasets and a difference-in-differences design, we show that these reforms substantially altered the composition of successful candidates. Prefectures

more exposed to great-family influence experienced a large increase in the share of successful candidates from commoner backgrounds after the reforms. This finding suggests that procedural safeguards in selection can matter even when formal examinations already exist.

We also examine whether expanded access for commoner candidates came at the expense of bureaucratic quality. Commoner *jinshi* were, on average, less likely than elite *jinshi* to achieve high levels of scholarly and political success, suggesting that elite family background continued to confer advantages even after entry into the examination system. However, these gaps narrowed after the reforms, indicating that the reforms improved the selection of commoner talent. Among those who later served as prefectural governors, we find no evidence that commoner officials performed worse than elite officials across three measures of local governance.

Finally, we find that the reforms had broader social consequences. Treated prefectures experienced a decline in social unrest after the reforms, especially unrest led by educated individuals. This pattern is consistent with the idea that elite capture in access to public office can generate grievances among educated aspirants whose paths to advancement are blocked by favoritism. By expanding credible access to bureaucratic office, anti-corruption reforms may have helped incorporate talented outsiders into the state and reduce incentives for contentious political action.

Given that we have no evidence that government policies at the level of prefecture governors changed in favor of commoners after the reform, the mechanism through which social unrest changed is unlikely to have been changes in policy. Rather, our evidence supports the idea that educated commoners were less likely to lead social protests when they had the opportunity to enter into government service. Finally, our evidence suggests that the main mechanism is unlikely to be that the general population was happier with the government in response to the anti-corruption nature of the reforms.

Taken together, reducing discretion and concealing identity in a high-stakes selection process broadened access to elite office, improved the selection of non-elite talent, and reduced social unrest. The findings highlight a broader lesson: the design of recruitment institutions can shape both the composition of the state and the stability of the society it governs.

References

Acemoglu, Daron, "Reward Structures and the Allocation of Talent," *European Economic Review*, 1995, *39* (1), 17–33.

_ and James A Robinson, "Persistence of power, elites, and institutions," *American Economic Review*, 2008, *98* (1), 267–293.

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

Barfort, Sebastian, Nikolaj A Harmon, Frederik Hjorth, and Asmus Leth Olsen, "Sustaining honesty in public service: The role of selection," *American Economic Journal: Economic Policy*, 2019, *11* (4), 96–123.

Behaghel, Luc, Bruno Crépon, and Thomas Le Barbanchon, "Unintended effects of anonymous resumes," *American Economic Journal: Applied Economics*, 2015, *7* (3), 1–27.

Bellemare, Marc F, "Rising food prices, food price volatility, and social unrest," *American Journal of agricultural economics*, 2015, *97* (1), 1–21.

Bó, Ernesto Dal, Frederico Finan, and Martín A Rossi, "Strengthening state capabilities: The role of financial incentives in the call to public service," *The Quarterly Journal of Economics*, 2013, *128* (3), 1169–1218.

Braggion, Fabio, Alberto Manconi, and Haikun Zhu, "Credit and social unrest: Evidence from 1930s China," *Journal of Financial Economics*, 2020, *138* (2), 295–315.

Campante, Filipe R and Davin Chor, "Why was the Arab world poised for revolution? Schooling, economic opportunities, and the Arab Spring," *Journal of Economic Perspectives*, 2012, *26* (2), 167–188.

Cantoni, Davide, Andrew Kao, David Y Yang, and Noam Yuchtman, "Protests," *Annual Review of Economics*, 2024, *16* (1), 519–543.

Caprettini, Bruno and Hans-Joachim Voth, "Rage against the machines: Labor-saving technology and unrest in industrializing England," *American Economic Review: Insights*, 2020, *2* (3), 305–320.

Chaffee, John W., *The Thorny Gates of Learning in Sung China: A Social History of Examinations*, Cambridge: Cambridge University Press, 1985.

Chen, G., *A Chronology of Natural and Human Disasters in Chinese History*, Beijing: Beijing Library Press, 1986.

Chen, Ting, James Kai sing Kung, and Chicheng Ma, "Long Live Keju! The Persistent Effects of China's Civil Examination System," *The Economic Journal*, 2020, *130* (631), 2030–2064.

CHGIS, "China Historical GIS (CHGIS)," Dataset 2016. Fairbank Center for Chinese Studies (Harvard) and Center for Historical Geographical Studies (Fudan University).

Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso, "Patronage and selection in public sector organizations," *American Economic Review*, 2020, *110* (10), 3071–3099.

Eisinger, Peter K, "The conditions of protest behavior in American cities," *American political science review*, 1973, *67* (1), 11–28.

Fan, Zhongyan, *The Complete Collection of Fan Zhongyan*, Sichuan University Press, 2007. Originally written in 1043.

Finan, Frederico, Benjamin A Olken, and Rohini Pande, "The personnel economics of the developing state," *Handbook of economic field experiments*, 2017, *2*, 467–514.

Fu, Z. et al., *Chronological Table of the Wars in China*, Chinese People's Liberation Army Press, 2003.

Ge, Jianxiong, *The History of Population in China*, Fudan University Press, 2000.

Goldin, Claudia and Cecilia Rouse, "Orchestrating impartiality: The impact of "blind" auditions on female musicians," *American economic review*, 2000, *90* (4), 715–741.

Hanna, Rema and Shing-Yi Wang, "Dishonesty and selection into public service: Evidence from India," *American Economic Journal: Economic Policy*, 2017, *9* (3), 262–290.

Hao, Yu, Zhengcheng Liu, Xi Weng, and Li-An Zhou, "The making of bad gentry: The abolition of Keju, local governance, and anti-elite protests, 1902–1911," *The Journal of Economic History*, 2022, *82* (3), 625–661.

Hartwell, Robert M, "Demographic, Political, and Social Transformations of China, 750–1550," *Harvard Journal of Asiatic Studies*, 1982, *42* (2), 365–442.

He, Qiaoyuan, *Min Shu [Book of Fujian]* 1629. 154 juan, 1629 edition.

He, Z., *A Compilation of Historical Records on Peasant Rebellions during the Song*, Zhonghua Book Company, 1976.

Hirschman, Albert O, *Exit, voice, and loyalty: Responses to decline in firms, organizations, and states*, Harvard university press, 1972.

Hucker, Charles O, "A Dictionary of Official Titles in Imperial China," 1985.

Johnson, David, *The Medieval Chinese Oligarchy*, Westview Press, 1977.

Lee, T. H. C., "Several Issues on Education and the Civil Service Examination System in the Song Dynasty," *Journal of Chinese Studies*, 1979, *10*, 323–348.

Li, X., *Xu Zizhitongjianchangbian*, Zhonghua Book Company, 2004. Originally written in 1183.

Liu, H., *The End of the Imperial Examination System and the Rise of Imperial Examination Studies*, Wuhan: Central China Normal University Press, 2006.

Liu, K., *The Collected Works of Hedong*, Zhonghua Book Company, 2015. Originally written in 1001.

Liu, Kevin Zhengcheng and Xiaoming Zhang, "Discretion, talent allocation, and governance performance: Evidence from China's imperial bureaucracy," *Journal of Development Economics*, 2025, *172*, 103391.

Liu, X., *Old Tang Book*, Zhonghua Book Company, 1975. Originally written in 945.

Manacorda, Marco and Andrea Tesei, "Liberation technology: Mobile phones and political mobilization in Africa," *Econometrica*, 2020, *88* (2), 533–567.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti, "Economic shocks and civil conflict: An instrumental variables approach," *Journal of political Economy*, 2004, *112* (4), 725–753.

Murphy, Kevin M, Andrei Shleifer, and Robert W Vishny, "The Allocation of Talent: Implications for Growth," *The Quarterly Journal of Economics*, 1991, *106* (2), 503–530.

Needham, Joseph, *The Shorter Science and Civilization in China*, Cambridge University Press, 1954.

North, Douglass Cecil, John Joseph Wallis, and Barry R Weingast, *Violence and social orders: A conceptual framework for interpreting recorded human history*, Cambridge University Press, 2009.

Ouyang, X., *New Tang Book*, Zhonghua Book Company, 1976. Originally written in 1060.

Ponticelli, Jacopo and Hans-Joachim Voth, "Austerity and anarchy: Budget cuts and social unrest in Europe, 1919–2008," *Journal of Comparative Economics*, 2020, *48* (1), 1–19.

Shleifer, Andrei and Robert W Vishny, "Corruption," *The Quarterly Journal of Economics*, 1993, *108* (3), 599–617.

Teorell, Jan, Carl Dahlström, and Stefan Dahlberg, "The QoG expert survey dataset," *University of Gothenburg: The Quality of Government Institute*, 2012.

Tuotuo, *History of Song*, Zhonghua Book Company, 1977. Originally written in 1345.

Wang, Erik H and Clair Z Yang, *The Political Economy of China's Imperial Examination System*, Cambridge University Press, 2025.

Wang, Yong, "Yanyi yimou lu [Records of Plans Bequeathed to Posterity]," in "Moji; Yanyi yimou lu," Beijing: Zhonghua Shuju, 1981.

Weaver, Jeffrey, "Jobs for sale: Corruption and misallocation in hiring," *American Economic Review*, 2021, *111* (10), 3093–3122.

Xu, Guo, "The costs of patronage: Evidence from the british empire," *American Economic Review*, 2018, *108* (11), 3170–3198.

Xu, S., *Song Huiyao Jigao*, Shanghai Ancient Book Press, 2014. Originally written in 1809.

Yang, Y., *Atlas of Regional Extreme Droughts and Floods in China Over the Millennium*, SinoMaps Press, 2024.

Yuchtman, Noam, "Teaching to the tests: An economic analysis of traditional and modern education in late imperial and republican China," *Explorations in Economic History*, 2017, *63*, 70–90.

Zeng, Z. and L. Liu, *Quan Song Wen*, Shanghai Lexicographical Publishing House, 1988.

Zhang, X., *The History of Chinese Printing*, Zhejiang Ancient Books Publishing House, 2006.

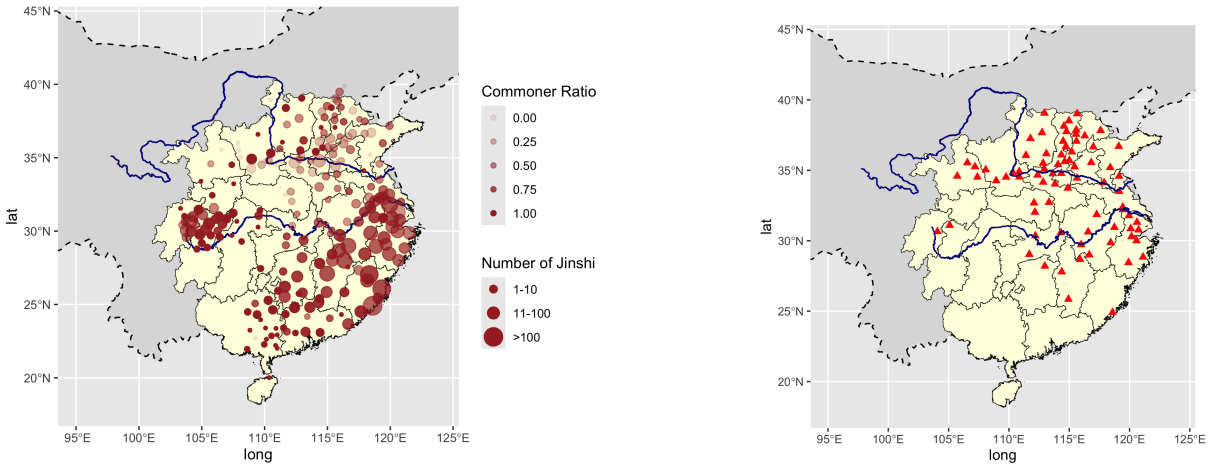
Zhou, Y., *Prefectural and County Schools in the Song Dynasty*, Taipei: National Institute for Compilation and Translation, 1996.

Zhou, Z., *The General History of Chinese Administrative Divisions*, Fudan University Press, 2007.

Zhu, Y. and X. Fu, *Quansong Biji*, Daxiang Publishing House, 2012.

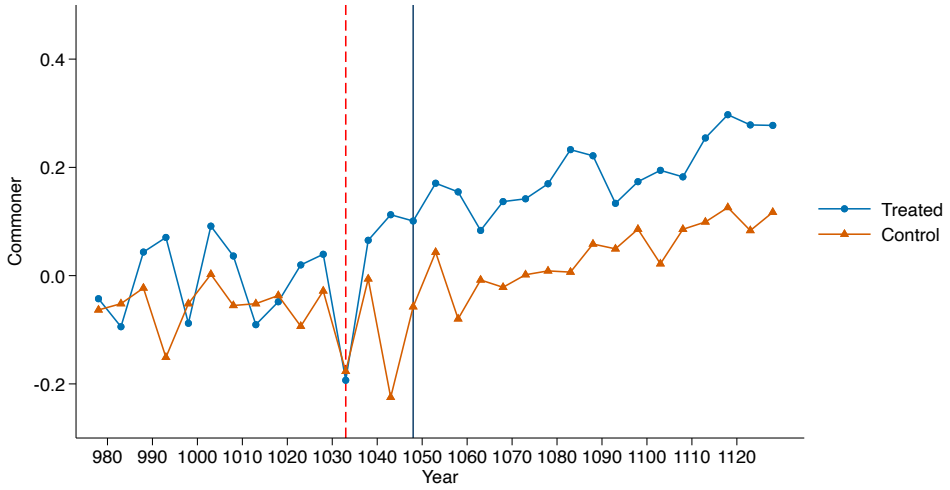
Figures and Tables

Figure 1: Geographic Distribution of Great Families and *jinshi*.



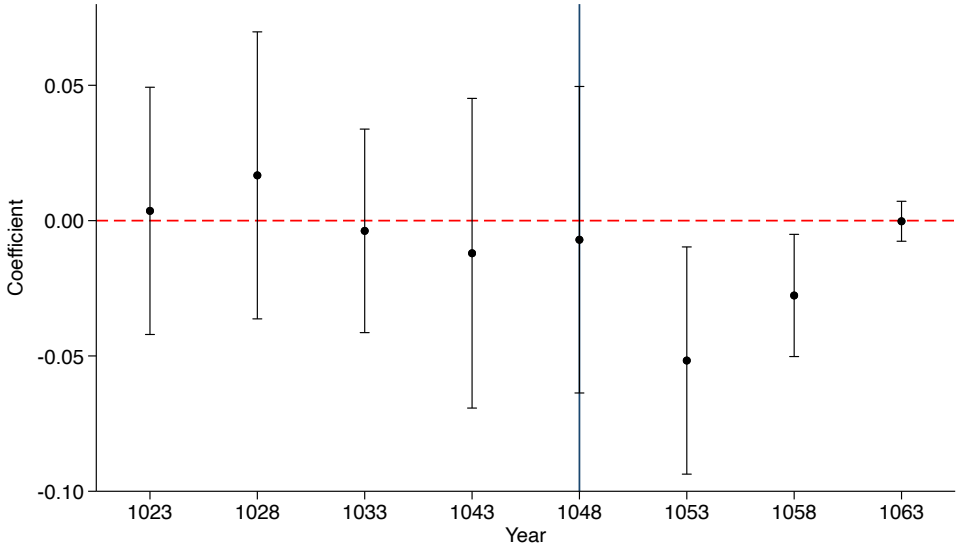
Note: The left panel shows the geographic distribution of *jinshi* during the Northern Song period (960-1127). The size of the points indicates the number of *jinshi*, whereas the color depth represents the proportion of commoner *jinshi*. The right panel shows the locations of the residency prefectures of the great families with red triangles. The light-yellow area reflects the territory of the Northern Song dynasty in 1080, which is based on the shapefile from CHGIS (2016).

Figure 2: Proportion of Commoner *jinshi* in Treated and Control Prefectures.

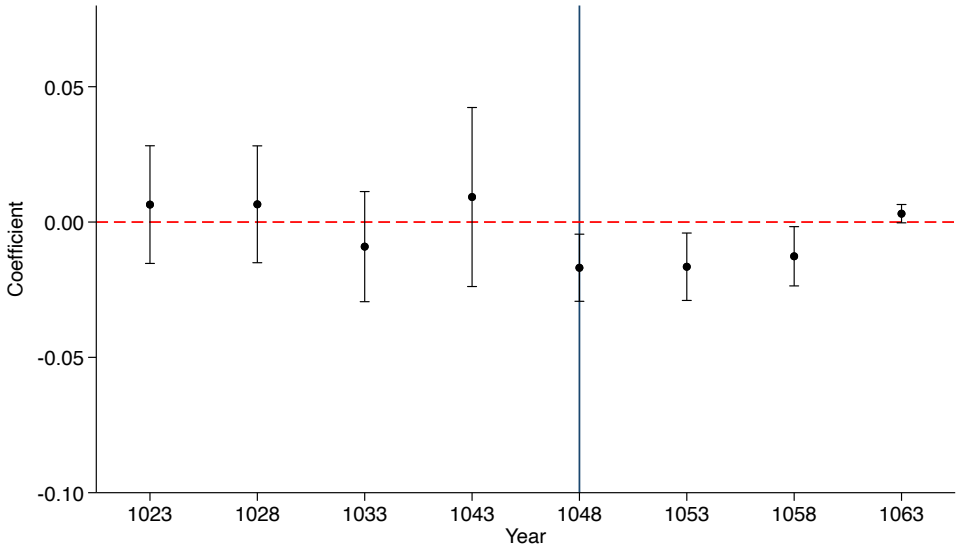


Note: This figure presents the time series trend of the proportion of commoner *jinshi* in the treated and control prefectures. The proportion is calculated for each five-year term, with each group demeaned at its average level before the year 1000 for comparison. The red vertical dashed line indicates the year of the first identity reform in 1033, whereas the blue vertical line indicates the year of the discretion reform in 1045.

Figure 3: Examination Reforms and Social Unrest.



(a) Likelihood of Social Unrest.



(b) Likelihood of Social Unrest Led by Educated

Note: This figure presents the coefficients of the dynamic analysis of the difference in the likelihood of social unrest between treated and control prefectures. In Panel (a), the dependent variable is $HasUnrest$, while in Panel (b), the dependent variable is $HasUnrestEducated$. The bars denote the 90% confidence intervals of the coefficient. The blue vertical line indicates the second reform in 1045.

Table 1: Summary Statistics

	Mean	Std. Dev.	N
Panel A: Prefecture-Term Commoner Ratio Data			
Commoner Ratio	0.641	0.394	1966
Treat	0.411	0.492	1966
Panel B: Prefecture Level Controls			
Population (000s)	21.443	24.343	306
Landsize (000s)	13.406	14.137	306
Philosopher Num	0.190	0.787	306
Distance River	65.801	75.723	306
Distance Publisher	396.832	287.675	306
Distance Border	93.327	64.496	306
Panel C: Jinshi Individual Level Outcomes			
Has Article	0.165	0.371	9454
Any Official	0.334	0.472	9454
Higher Official	0.218	0.413	9454
Panel D: Prefecture Governor-Year Level Outcomes			
Commoner Higher Official	0.229	0.420	3974
New School	0.032	0.176	3974
Famine	0.019	0.138	3974
Panel E: Prefecture-Term Outcomes and Controls			
Has Unrest	0.040	0.196	10098
Has Unrest Educated	0.008	0.089	10098
Has Unrest Uneducated	0.032	0.176	10098
Climate	0.889	2.223	10098

Note: This table summarizes each variable's observations, mean, and standard deviation. Distance variables are measured in kilometers, and the land size variable is measured in square kilometers. The climate variable is measured on a severity scale from 0 to 5, where 0 indicates no severe hazard.

Table 2: Examination Reforms and Commoner *Jinshi*

	$CommonerRatio$					
	(1)	(2)	(3)	(4)	(5)	(6)
$Treat \times Post1033$	0.104**	0.193***				
	(0.048)	(0.056)				
$Treat \times Between(1033, 1045)$					0.091	0.185**
					(0.063)	(0.079)
$Treat \times Post1045$			0.088*	0.166***	0.105**	0.194***
			(0.045)	(0.055)	(0.050)	(0.058)
p-value of difference					[0.802]	[0.914]
Observations	1966	1966	1966	1966	1966	1966
Controls × Post1033	No	Yes	No	No	No	No
Controls × Between(1033,1045)	No	No	No	No	No	Yes
Controls × Post1045	No	No	No	Yes	No	Yes
Prefecture Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms on the proportion of commoner *jinshi*. The dependent variable $CommonerRatio$ is the proportion of commoner *jinshi* in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Post1033$ is an indicator that equals one if the term is after 1033. $Between(1033, 1045)$ is an indicator that equals one if the term is after 1033 and before 1045. $Post1045$ is an indicator that equals one if the term is after 1045. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The sample consists of 1,966 observations from 960 to 1127. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 3: Examination Reforms and Commoner *Jinshi*: PSM-DiD and Distance to Great Families

	CommonerRatio		
	(1)	(2)	(3)
Panel A: PSM-DiD Results			
$Treat \times Post1033$	0.178***		
	(0.058)		
$Treat \times Between(1033, 1045)$			0.188**
			(0.086)
$Treat \times Post1045$		0.145**	0.176***
		(0.057)	(0.061)
p-value of difference			[0.885]
Observations	1572	1572	1572
Panel B: Distance to Great Families			
$Distance \times Post1033$	-0.178***		
	(0.058)		
$Distance \times Between(1033, 1045)$			-0.186**
			(0.086)
$Distance \times Post1045$		-0.150**	-0.179***
		(0.060)	(0.061)
p-value of difference			[0.938]
Observations	1966	1966	1966
Controls × Post1033	Yes	No	No
Controls × Between(1033,1045)	No	No	Yes
Controls × Post1045	No	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms on the proportion of commoner *jinshi*. Panel A reports PSM-DiD results using a matched sample. Using all prefecture-level control variables, each treated prefecture is matched to a control prefecture via nearest-neighbor one-to-one matching without replacement; the matched sample consists of 1,572 observations from 960 to 1127. Panel B reports results using the full sample of 1,966 observations from 960 to 1127, where Distance is the logarithm of the geographical distance between the prefecture and its nearest great family. In both panels, the dependent variable *CommonerRatio* is the proportion of commoner *jinshi* in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Post1033$ is an indicator that equals one if the term is after 1033. $Between(1033, 1045)$ is an indicator that equals one if the term is after 1033 and before 1045. $Post1045$ is an indicator that equals one if the term is after 1045. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. Prefecture and term fixed effects are included in all specifications. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 4: Career Performance of Commoner and Elite *Jinshi*

	HasArticle	*AnyOfficial*	*HigherOfficial*
	(1)	(2)	(3)
$CommonerJinshi$	-0.291***	-0.515***	-0.496***
	(0.019)	(0.022)	(0.020)
$CommonerJinshi \times Select(1033, 1045)$	0.079**	-0.041	0.056
	(0.036)	(0.042)	(0.038)
$CommonerJinshi \times SelectPost1045$	0.070***	0.052**	0.159***
	(0.021)	(0.025)	(0.022)
p-value of difference	[0.778]	[0.013]	[0.002]
Observations	9454	9454	9454
Term Fixed Effect	Yes	Yes	Yes

Note: This table presents the analyses of the career performance of commoner and elite *jinshi*. *HasArticle* is a dummy variable that equals one if the *jinshi* has at least one prose work included in *Quan Song Wen*. *AnyOfficial* is a dummy variable that equals one if the *jinshi* had obtained an official position. *HigherOfficial* is a dummy variable that equals one if the *jinshi* ever held an official title equal or greater than zhizhou. $CommonerJinshi$ is a dummy variable that equals one if the *jinshi* is not from any great family and none of his or her patrilineal elders held official positions. $Select(1033, 1045)$ is an indicator that equals one if the *jinshi* obtained his degree after 1033 and before 1045, and $SelectPost1045$ is an indicator that equals one if the *jinshi* obtained his degree after 1045. The sample consists of 9,454 observations of *jinshi* from 960 to 1127. Standard errors are reported in parentheses. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 5: Governance Performance of Commoner and Elite Governors

	$NewSchool$	$Famine$	$HasUnrest$
	(1)	(2)	(3)
$CommonGov$	-0.002	0.001	0.010
	(0.010)	(0.008)	(0.008)
$CommonGov \times Select(1033, 1045)$	-0.053**	-0.002	-0.015
	(0.027)	(0.015)	(0.011)
$CommonGov \times SelectPost1045$	0.011	0.008	-0.012
	(0.013)	(0.009)	(0.010)
p-value of difference	[0.034]	[0.440]	[0.747]
Observations	3974	3974	3974
Controls × Select(1033,1045)	Yes	Yes	Yes
Controls × SelectPost1045	Yes	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes

Note: This table presents the analyses of the governance performance of commoner and elite *zhizhou* with *jinshi* degree. The dependent variable $NewSchool$ is a dummy variable that equals one if there is at least one new county school built or rebuilt in the prefecture-year. $Famine$ is a dummy variable that equals one if there is a famine event in the prefecture-year. $HasUnrest$ is a dummy variable that equals one if there is at least one instance of social unrest in the prefecture-year. $CommonGov$ is a dummy variable that equals one if the *zhizhou* with *jinshi* degree is not from any great family and none of his or her patrilineal elders held official positions. $Select(1033, 1045)$ is an indicator that equals one if the *zhizhou* obtained the *jinshi* degree after 1033 and before 1045, and $SelectPost1045$ is an indicator that equals one if the *zhizhou* obtained the *jinshi* degree after 1045. The main effects of $Select(1033, 1045)$ and $SelectPost1045$ are included in the regressions but not reported. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The sample consists of 3,974 observations of prefecture-year governed by *zhizhou* with *jinshi* degree from 960 to 1127. Standard errors are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 6: Examination Reforms and Social Unrest

	$HasUnrest$					
	Full Sample			PSM Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
$Treat \times Post1033$	-0.028**			-0.026*		
	(0.013)			(0.015)		
$Treat \times Between(1033, 1045)$			-0.012			0.009
			(0.024)			(0.025)
$Treat \times Post1045$		-0.028**	-0.030**		-0.032**	-0.031**
		(0.013)	(0.013)		(0.014)	(0.015)
p-value of difference			[0.433]			[0.076]
Observations	10098	10098	10098	5082	5082	5082
Controls × Post1033	Yes	No	No	Yes	No	No
Controls × Between(1033,1045)	No	No	Yes	No	No	Yes
Controls × Post1045	No	Yes	Yes	No	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms on social unrests. The dependent variable $HasUnrest$ is a dummy variable that equals one if there is at least one social unrest in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Post1033$ is an indicator that equals one if the term is after 1033. $Between(1033, 1045)$ is an indicator that equals one if the term is after 1033 and before 1045. $Post1045$ is an indicator that equals one if the term is after 1045. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The full sample consists of 10,098 observations from 960 to 1127. The PSM sample consists of 5,082 observations. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table 7: Examination Reforms and Social Unrest: Led by Educated vs Uneducated

	Full Sample			PSM Sample		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: *HasUnrestEducated*						
$Treat \times Post1033$	-0.009			-0.011*		
	(0.006)			(0.006)		
$Treat \times Between(1033, 1045)$			0.002			-0.002
			(0.015)			(0.014)
$Treat \times Post1045$		-0.011**	-0.011*		-0.012**	-0.012**
		(0.006)	(0.006)		(0.005)	(0.006)
p-value of difference			[0.362]			[0.410]
Mean dep var	0.008	0.008	0.008	0.008	0.008	0.008
Panel B: *HasUnrestUneducated*						
$Treat \times Post1033$	-0.018			-0.015		
	(0.011)			(0.013)		
$Treat \times Between(1033, 1045)$			-0.013			0.011
			(0.021)			(0.022)
$Treat \times Post1045$		-0.017	-0.019		-0.020	-0.018
		(0.011)	(0.011)		(0.012)	(0.013)
p-value of difference			[0.783]			[0.127]
Mean dep var	0.032	0.032	0.032	0.046	0.046	0.046
Observations	10098	10098	10098	5082	5082	5082
Controls × Post1033	Yes	No	No	Yes	No	No
Controls × Between(1033,1045)	No	No	Yes	No	No	Yes
Controls × Post1045	No	Yes	Yes	No	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms on social unrests. In Panel A, the dependent variable *HasUnrestEducated* is a dummy variable that equals one if there is at least one social unrest led by educated people in a prefecture-term. In Panel B, the dependent variable *HasUnrestUneducated* is a dummy variable that equals one if there is at least one social unrest led by uneducated people in a prefecture-term. *Treat* is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. *Post*1033 is an indicator that equals one if the term is after 1033. *Between*$(1033, 1045)$ is an indicator that equals one if the term is after 1033 and before 1045. *Post*1045 is an indicator that equals one if the term is after 1045. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The full sample consists of 10,098 observations from 960 to 1127. The PSM sample consists of 5,082 observations. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Appendix

A Other Potentially Confounding Events

As emperor turnovers may coincide with policy change, we also conduct a placebo test by examining the impact of emperor turnover during the Northern Song dynasty on the proportion of commoner *jinshi*. Specifically, we identify six emperor turnovers that do not overlap with the two imperial examination reforms. We then restrict our sample to include only the terms immediately before and after each emperor turnover year and define $Post_{\text{Emperor}}$ as a dummy variable that equals one if the term is after an emperor turnover and zero otherwise. We replace the post dummy in equation 1 with $Post_{\text{Emperor}}$ and re-estimate the model. Table A.6 reports the results. In Column (1), we do not include the controls, whereas in Column (2), we include them. In both specifications, the coefficients on $Treat \times Post_{\text{Emperor}}$ are statistically insignificant and small in magnitude, suggesting that emperor turnover has no effect on the proportion of commoner *jinshi*.

A confounding institutional shock occurred in 1094 (Chaffee, 1985). A new political faction took control and the court reformed the content of the civil service examinations. The reform abolished the poetry and rhapsody (*shifu*) component of the exam and shifted the focus toward Confucian Classics and policy-oriented essays. Candidates were required to specialize in selected canonical texts and to respond to questions on contemporary statecraft.

While the reform marked a departure from literary ability toward doctrinal knowledge and practical statecraft, it is unlikely to have disadvantaged aristocratic or politically connected families. We test this formally by adding a new post dummy, $Post_{1094}$, to our baseline model and report the results in Table A.7. While the coefficients on our original reform interactions remain positively significant, the coefficient on the interaction term $Treat \times Post_{1094}$ is negative but insignificant. The result is consistent with the view that the examination content reform in 1094 did not disadvantage politically connected candidates.

B An Anti-corruption Campaign in 1009

A broader anti-corruption campaign, unrelated to the examination system, was initiated in 1009. In that year, Emperor Zhenzong promulgated the *Wenwu Qitiao* (Seven Regulations for Civil and Military Officials), outlining seven principles of governance that primarily focused on countering corruption in government. During the first decade of Emperor Zhenzong's reign following the start of the anti-corruption campaign, twelve major corruption cases were investigated, involving twenty-two high-ranking officials, five of whom were sentenced to death (Li, 2004; Tuotuo, 1977).

One mechanism that we discussed that links the examination reforms to the lower rates of social unrest is that the general population is happy when the government is seen as cracking down on corruption and inequity. To test this channel, we estimate the impacts of the anti-corruption campaign of 1009. If a general response to anti-corruption efforts were driving our results, treated prefectures should exhibit comparable changes following this campaign. Restricting the sample to terms before 1033, we estimate the DiD specifications with $Post1009$ for both outcomes, $CommonerRatio$ and $HasUnrest$. Table A.8 reports the results, and we find no significant effect for either outcome. This supports our interpretation that the effects operate through commoners' greater likelihood of passing the examination rather than through the response of the general population to government crackdowns against corruption and inequity.

Table A.1: Great Families and Residency Prefectures

Prefectures	Families Surname
Kaifengfu	Ruan, He, Xie, Wei, Yin, Na, Cai, Dian, Yu, Bian, Shentu, Yi, Zhi, Qu, Quan, Cao, Ding, Lou, Dai, Xiahou, Ji, Xi, Huan, Bo, Ru
Qizhou	Bian, Shan, Dongmen, Xindu
Yizhou	Wang, Yan, Zhuge, Miao, Xu, Ge, Ai, Qian, Hui, Chang, Fu
Qingzhou	Shi, Cheng, Sheng, Ni, Gai, Tan, You, Yan, Zha, Mo, Ke, Fan, Jin, Hua, Zuo, Ning, Zhong, Yu, Ran, Fan, E, Kai, Ying, Peng, Ju, Sun, Ren, Tao, Guo, Zhangsun, Xue, Jiang, Chong, Gongsun, Gong, Yan, Fang, He, Cao
Sizhou	Guan, Yu, Shen, Qi, Gu, Guo, Pi, Hua
Songzhou	Shang, Song, Ge, Bin
Yanzhou	Tang, Li, Kong, Qi, Yu, Qu, Zai, Wan, Zai, Zeng, Zou, Xia, Che, Yan, Li, Xian, Pu, Shao, Chao, Bao, Yang, Hu, Husi, Guan, Gai, Mou, Meng, Wei, Lü, Wan, Ping, Ji
Xuzhou	Liu, Zhu, Dao, Xu, Zhuang, Wan, Zhi, Song, Zheng, Gong, Chao, Xing, , Xiao, Miao, Moqi, Duanmu
Caozhou	Ding, Bian, Jiang, Zuo, Cai, Shan, Cao, Yu
Xiangzhou	Lifei, Kuai, Fu, Qian, Jian
Dengzhou	Bai, Han, Sheng, Yue, Deng, Zong, Ye, Rang, Cen, Di, Kuang, Jing, Zhao, Ji, Qiu, Lu
Tangzhou	Yu
Henanfu	Chu, Mu, Dugu, Qiu, Zhu, Yuan, Wen, Ren, Helan, Murong, Shang, Nangong, Gu, Shan, Fang, Lin, Qing, Lüqiu, Li, Rui, Hou, Mo, Chen, Fang, Fu, Yuwen
Xuzhou	Chen, Xun, Yu, Ku, Zhong, Bai, Xu, Han, Doulu, Xianyu, Yan
Zhengzhou	Zheng, Pan, Mao, Yang, Mou, Jia
Huazhou	Chenggong, Fei, Shangguan
Chenzhou	Yuan
Yuzhou	Zhou, Yin, Jing, Xiang, Sheng, He, Xuan, Nan, Cai, Mei, Yuan, Cuan, Bei, Ying, , Ru, Yi, Yan, Chang, Lan, Du, Sha, Man, , Yu, Zhong
Damingfu	Shen, Bao, Bai, Chang, Dun, Chang
Chanzhou	Wu, Wen, Fu, Huang, Qing, Pu
Jizhou	Gao, Wu, Ouyang, Helian, Zhan, Yu, Li, Shi, Ou, Jin, Qing, Gan, Zi, Ling, Qin, Feng, Diao, Hegan, Tong, , Ji, Si, Heng, Ju, Cang, Guan, Feng, Tan
Dezhou	Shen, Jing, Mengchang, Dongfang, Shi, Nei, Yi
Beizhou	Zhang, Fang, Cui, Dai, Jin, Nie, Meng, Fu, Gai, Zhuo, Sui, Shang, Ji, Tan, Ju, Gui, Ge, Shu, Lu

Continued on next page

Table A.1 Great Families and Residency Prefectures (Continued)

Prefectures	Families Surname
Qianningjun	Lu
Yizhou	Hou, Rong, Ma, Yan, Kou, Gu
Zhendingfu	Zhen, Jiao, Lin, Zhong, Lang, Guan, Xu, Geng, Ji, Gongsun, 劉
Xiangzhou	Lu, Luo, Hu, Ku
Dingzhou	Cui, 邱, Shou, Xing, Puyang
Xingzhou	Wei, Geng, Te, Mo, Shi, Shu
Huaizhou	Sima, Shang, Jin, Xiang, He, Wang, Che, She, Song, Wen, Chunyu, Huai, Ru, 姬, Zhi, Qu, Rong
Weizhou	Qu, Sang, Wei, 韩
Mingzhou	You, Cheng, Song, Tan, Jie, 唝, 遲, Jiao
Shenzhou	Xing, Yu, Jia, Xuan, Yao, Deng, Zhan, Zhu
Zhaozhou	Li, Situ, Mu, Lang, Mie, Wen, Min
Taiyuanfu	Hong, Wang, Guo, Hao, Wen, Yuchi, Qi, Linghu, Wu, Yan, Gong, Wu, Sun, Fu, Zan, Huo, Wen, Gong, Shi, Yi, Shao, You, Liao, Yi, Long, Shao, Guang
Luzhou	Bao, Bao, Chen, Fan, Bao, Shang
Jinzhou	Wang, Juan, Yu, Cheng, Ping, Chai, Jing, Jing, Gou, Jia, Jin, Feng
Zezhou	Fan, Ba, Di, Guo, Dugu, Tang, Du, Mie, Cuan, Jiong
Daizhou	Xu, Xie, Tian, Wen, Di
Fenzhou	Jin, Bu, Song, Lin, Zhi, Xiangli, Ren, Lin, Luan, Tong
Jingzhaofu	Che, Duan, Yan, Li, Song, Qin, Zhong, Yong, Wei, Tian, Su, Yu, Mi, Leng, Zhi, Yuan, Shu, Hu, Chao, Pi, Kun, Shen, Tu, Liao, Bie, Fu, Meng, Gao, Feng, Zhu, Shi, Lun, Xing, Jin, Cheng, Gong, Di, Wu, Song, Yi, Di, Su, Ji, Su, Han, Shi, Shu, Yu, Ji, Dang, Lei, Yin, He, Li, Kou, Zhi, Feng, Gongwan, Gu
Puzhou	Pei, Liu, Xue, Chu, Pu, Wei, Nie, Ying, Lian, Mai, Hu, Hun, Man, Lang, Jia
Xiezhou	Pei
Guozhou	Yang, Tan, Qiang, Jin, Guo, Qiu
Huazhou	Yang
Binzhou	Gu, Yi, Fu, Guo
Qinzhou	Zhao, Jiang, Yin, Bie, Yan, Long, Quan, Qin, Shangguan, Li, Gui, Zhuang, Na, Pi, Shuang, Zhi, Bing, Qin, Liao, Jie
Fengxiangfu	Feng, Pang, Xuan, Yi, Dou, Ma, Zeng, Lu, Wan, Kou, Jing, Su, Hui, Ban, Fu
Weizhou	Li, Niu, Shi, Xin, Dong, Ai, Peng, Guan, Qian, Min, Wan, Fan, Bian
Jingzhou	Liang, Huangpu, Xi, Wu, Hu, An, Meng, Cheng

Continued on next page

Table A.1 Great Families and Residency Prefectures (Continued)

Prefectures	Families Surname
Chengdufu	Qie, Wen, Fei, Ren, Xi
Zizhou	Mian, Jing, Wen, Mi
Hangzhou	Fan, Cen, Chu, Sheng, Yang, Weng, Qi, Shu, Kan, Huyan, Ji, Kui, Ji, Jian
Yuezhou	Xiahou, He, Kang, Kong, Yu, Sheng, Zi, Zhong, Li, Luo, Zi, Yu, Rong, Fan, 劄, Xi, Miao, Yao, Quan, Nan
Suzhou	Zhu, Zhang, Gu, Lu, Ji
Runzhou	Gan, Ji, Na, Hong, Zuo, Xian, Yan, Guang
Huzhou	Shen, Qian, Yao, Wu, Qing, Qiu, Shi, Xuan, Meng, Jin, Yin, Yin, Xian, Niu, Mu, Qiuming
Changzhou	Jiang, Fu, Mo, Zhou
Taizhou	Qu, Leng, Jing, Tan, Yi, Ye
Chuzhou	Lao, Lai, Ye, Qu, Tan
Xiuzhou	Bhu
Yangzhou	Gao, Zhi, Qian, Sheng, Qing, Yu, Li, Dai, You, Gong, Li
Chuzhou	Qu, Chu, Gong, Nian, Hao, Jian
Haizhou	Xu, Kuang, Qi, Zhu, Yu, Guan, Jiwu, Jun, Chu, Mao
Luzhou	He, Kuang, Men, Yu
Shuzhou	Shu, Pugu
Xuanzhou	Kuang, Ben, 簡, Ju
Huizhou	Shu, Sun, Fang, Jian, Shou, Wang
Jiangzhou	Tao, Zhai, Chun, Qu, Qian, Bu
Raozhou	Rao, Rui, Duo, 婁
Hongzhou	Luo, Lei, Xiong, Chu, Qu, Chen, Hong
Nankangjun	Lai, Ye, Yin, Xun
Yuanzhou	Yuan, Peng, Yi, Shao
Quanzhou	Lin, Qiu, Hong, Shan
Jianglingfu	Xiong, Xuan, Wu, Rong, Jiu
Ezhou	Li, Huang, Cheng, Fei, Ren, Wei, Yu
Langzhou	Wu, Gong, Bu, Ran, Hua
Tanzhou	Zeng, Wu, Luo, Peng, Ru, Qin

Note: This table presents the treated prefectures with great families and the surnames of those families. For surname characters with unknown or uncertain pronunciations, we retain the original characters.

Table A.2: Parallel Trend and Dynamic Effects

	$CommonerRatio$		
	(1)	(2)	(3)
$Treat \times Pre(1019, 1023)$	0.117		0.117
	(0.140)		(0.140)
$Treat \times Pre(1024, 1028)$	-0.063		-0.063
	(0.098)		(0.098)
$Treat \times Pre(1029, 1033)$	-0.108		-0.108
	(0.123)		(0.123)
$Treat \times Post(1034, 1043)$	0.178**	0.186**	0.179**
	(0.081)	(0.079)	(0.081)
$Treat \times Post1044+$	0.187***		
	(0.063)		
$Treat \times Post(1044, 1048)$		0.218**	0.211**
		(0.103)	(0.105)
$Treat \times Post(1049, 1053)$		0.239***	0.233***
		(0.077)	(0.080)
$Treat \times Post(1054, 1058)$		0.190*	0.185*
		(0.103)	(0.105)
$Treat \times Post1059+$		0.189***	0.183***
		(0.059)	(0.064)
Observations	1966	1966	1966
Controls × Post	Yes	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms on the proportion of commoner *jinshi*. The dependent variable $CommonerRatio$ is the proportion of commoner *jinshi* in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Pre(1019, 1023)$, $Pre(1024, 1028)$, $Pre(1029, 1033)$, $Post(1034, 1043)$, $Post(1044, 1048)$, $Post(1049, 1053)$, $Post(1054, 1058)$, $Post1059+$ are dummy variables indicating the terms before and after the reform year 1033. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The sample consists of 1,966 observations from 960 to 1127. Standard errors in parentheses are clustered at the prefecture level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table A.3: PSM: Comparison between Treated Prefectures and Matched Prefectures

	Treated Mean	Matched Mean	Difference	p-value
Population (000s)	36.215	33.904	2.310	0.600
Landsize (000s)	15.881	16.691	-0.810	0.757
Philosopher Num	0.636	0.117	0.519	0.003
Distance River	46.648	55.167	-8.519	0.343
Distance Publisher	262.769	303.501	-40.733	0.203
Distance Border	92.568	88.354	4.214	0.672

Note: This table presents the impact of imperial examination reforms on the proportion of commoner *jinshi*. Using all prefecture-level control variables, each treated prefecture is matched to a control prefecture via nearest-neighbor one-to-one matching without replacement; the matched sample consists of 1,572 observations.

Table A.4: Examination Reforms and Commoner *Jinshi*: Full Model

	CommonerRatio	
	(1)	(2)
$Treat \times Post1033$	0.193***	
	(0.056)	
$Treat \times Post1045$		0.166***
		(0.055)
$DistancePublisher \times Post1033$	-0.014	
	(0.018)	
$DistancePublisher \times Post1045$		-0.015
		(0.016)
$Population \times Post1033$	-0.055**	
	(0.026)	
$Landsize \times Post1033$	-0.044	
	(0.030)	
$PhilosopherNum \times Post1033$	-0.088	
	(0.059)	
$Climate \times Post1033$	0.063	
	(0.072)	
$DistanceRiver \times Post1033$	0.008	
	(0.014)	
$DistanceBorder \times Post1033$	0.064***	
	(0.022)	
$Population \times Post1045$		-0.068***
		(0.024)
$Landsize \times Post1045$		-0.020
		(0.028)
$PhilosopherNum \times Post1045$		-0.074
		(0.057)
$Climate \times Post1045$		0.069
		(0.071)
$DistanceRiver \times Post1045$		0.002
		(0.013)
$DistanceBorder \times Post1045$		0.042*
		(0.022)
Observations	1966	1966
Prefecture Fixed Effect	Yes	Yes
Term Fixed Effect	Yes	Yes

Note: This table presents the full regression output of Columns (2) and (4) in Table 2. The dependent variable *CommonerRatio* is the proportion of commoner *jinshi* in a prefecture-term. *Treat* is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. *Post*1033 is an indicator that equals one if the term is after 1033. *Post*1045 is an indicator that equals one if the term is after 1045. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The sample consists of 1,966 observations from 960 to 1127. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table A.5: Parallel Trend and Dynamic Effects on Social Unrest

	$HasUnrest$	$HasUnrestEducated$	$HasUnrestUneducated$
	(1)	(2)	(3)
$Treat \times Pre(1019, 1023)$	0.010	-0.008	0.019
	(0.030)	(0.006)	(0.029)
$Treat \times Pre(1024, 1028)$	0.004	0.006	-0.003
	(0.028)	(0.013)	(0.027)
$Treat \times Pre(1029, 1033)$	0.017	0.007	0.010
	(0.032)	(0.013)	(0.030)
$Treat \times Post(1034, 1043)$	-0.004	-0.009	0.005
	(0.023)	(0.012)	(0.021)
$Treat \times Post(1044, 1048)$	-0.012	0.009	-0.021
	(0.035)	(0.020)	(0.029)
$Treat \times Post(1049, 1053)$	-0.007	-0.017**	0.010
	(0.034)	(0.008)	(0.033)
$Treat \times Post(1054, 1058)$	-0.052**	-0.017**	-0.035
	(0.025)	(0.008)	(0.024)
$Treat \times Post1059+$	-0.028**	-0.013*	-0.015
	(0.014)	(0.007)	(0.011)
Observations	10098	10098	10098
Controls × Post	Yes	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms on social unrest. The dependent variable $HasUnrest$ is a dummy variable that equals one if there is at least one instance of social unrest in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Pre(1019, 1023)$, $Pre(1024, 1028)$, $Pre(1029, 1033)$, $Post(1034, 1043)$, $Post(1044, 1048)$, $Post(1049, 1053)$, $Post(1054, 1058)$, $Post1059+$ are dummy variables indicating the terms before and after the reform year 1033. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The sample consists of 10,098 observations from 960 to 1127. Standard errors in parentheses are clustered at the prefecture level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table A.6: Placebo Test: Emperor Turnovers As Shocks

	$CommonerRatio$	
	(1)	(2)
$Treat \times$ Post Emperor	-0.023	-0.087
	(0.052)	(0.068)
Observations	691	691
Controls × Post Emperor	No	Yes
Prefecture Fixed Effect	Yes	Yes
Term Fixed Effect	Yes	Yes

Note: This table presents the placebo analyses by estimating the effects of emperor turnovers on the proportion of commoner *jinshi*. There are six emperor turnovers that do not overlap with the imperial examination reforms during our sample period. We restrict the sample to include only the terms immediately before and after each emperor turnover year. The dependent variable $CommonerRatio$ is the proportion of commoner *jinshi* in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. Post Emperor is a dummy variable that equals one if the term is immediately after an emperor turnover, and zero otherwise. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table A.7: Examination Reforms and Commoner *Jinshi*: Impact of Reform of Examination Content in 1094

	$CommonerRatio$			
	(1)	(2)	(3)	(4)
$Treat \times Between(1033, 1045)$	0.091	0.184**		
	(0.063)	(0.079)		
$Treat \times Post1045$	0.128**	0.206***	0.111**	0.179***
	(0.051)	(0.063)	(0.047)	(0.061)
$Treat \times Post1094$	-0.055	-0.029	-0.055	-0.028
	(0.035)	(0.049)	(0.035)	(0.049)
Observations	1966	1966	1966	1966
Controls × Between(1033,1045)	No	Yes	No	No
Controls × Post1045	No	Yes	No	Yes
Controls × Post1094	No	Yes	No	Yes
Prefecture Fixed Effect	Yes	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes	Yes

Note: This table presents the impact of imperial examination reforms, including the reform of examination content in 1094, on the proportion of commoner *jinshi*. The dependent variable $CommonerRatio$ is the proportion of commoner *jinshi* in a prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Between(1033, 1045)$ is an indicator that equals one if the term is after 1033 and before 1045. $Post1045$ is an indicator that equals one if the term is after 1045. $Post1094$ is an indicator that equals one if the term is after 1094. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. The sample consists of 1,966 observations from 960 to 1127. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table A.8: Anti-Corruption and Social Unrest before 1033

	$CommonerRatio$		$HasUnrest$	
	(1)	(2)	(3)	(4)
$Treat \times Post1009$	-0.011	-0.089	-0.005	0.009
	(0.084)	(0.091)	(0.015)	(0.017)
Observations	519	519	4284	4284
Controls × Post1009	No	Yes	No	Yes
Prefecture Fixed Effect	Yes	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes	Yes

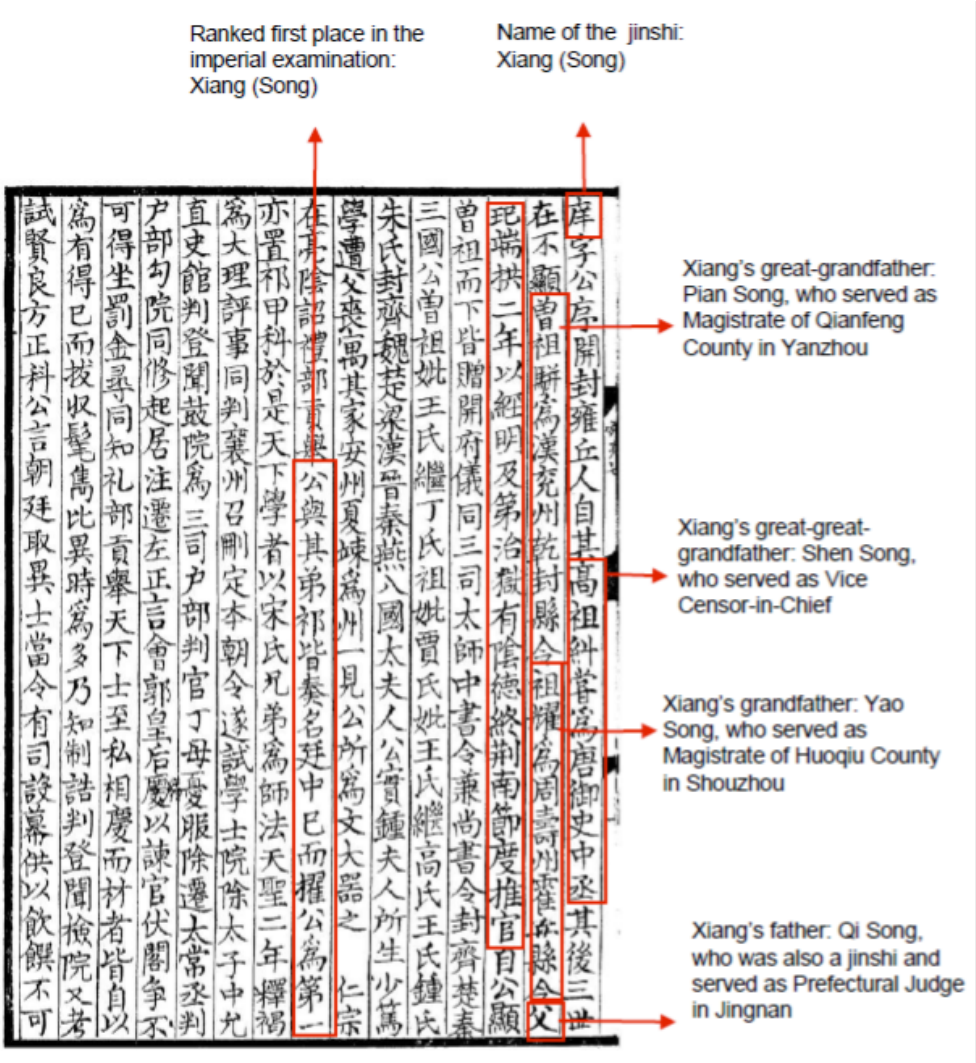
Note: This table examines the effect of a general anti-corruption event in 1009 on the proportion of commoner *jinshi* and on social unrest, estimated on the sample of terms before 1033 so that the estimates are not confounded by the later 1033 reform. Columns (1)–(2) use the prefecture-term data with dependent variable $CommonerRatio$, the proportion of commoner *jinshi* in a prefecture-term; columns (3)–(4) use the prefecture-term unrest panel with dependent variable $HasUnrest$. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Post1009$ is an indicator that equals one for terms after 1009, when a general anti-corruption event took place. Even-numbered columns add the full set of controls (population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate), each interacted with $Post1009$. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Table A.9: Examination Reforms and Local Effects: New School and Famine

	$HasNewSchool$			$Famine$		
	(1)	(2)	(3)	(4)	(5)	(6)
$Treat \times Post1033$	0.015			-0.030**		
	(0.017)			(0.012)		
$Treat \times Between(1033, 1045)$		-0.000			-0.010	
		(0.037)			(0.016)	
$Treat \times Post1045$		0.016	0.016		-0.033***	-0.031***
		(0.017)	(0.016)		(0.012)	(0.012)
Observations	10098	10098	10098	10098	10098	10098
Controls × Post1033	Yes	No	No	Yes	No	No
Controls × Between(1033,1045)	No	Yes	No	No	Yes	No
Controls × Post1045	No	Yes	Yes	No	Yes	Yes
Prefecture Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Term Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes

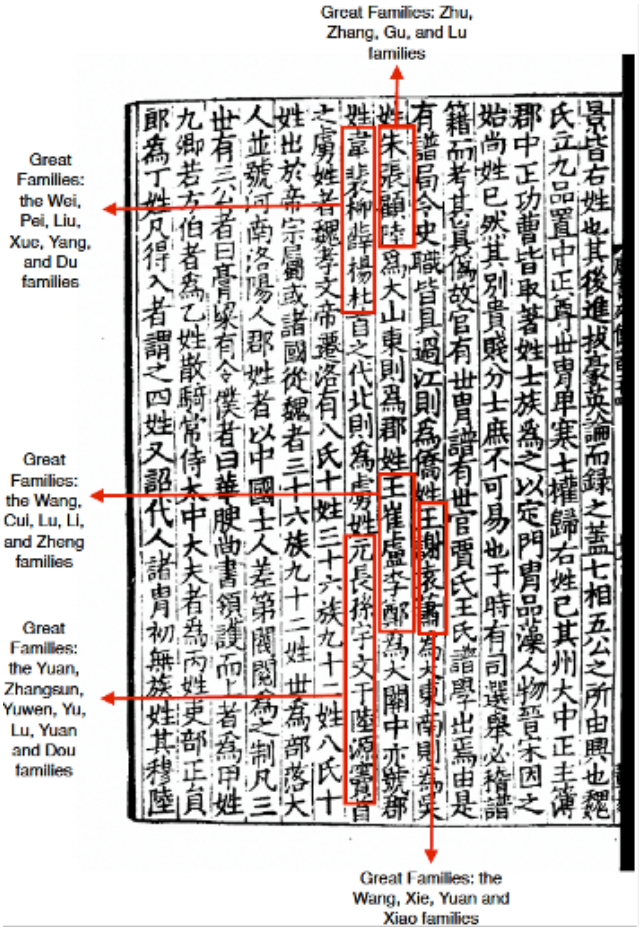
Note: This table presents the impact of imperial examination reforms on local governance outcomes. The dependent variable in Columns (1)-(3), $HasNewSchool$, is a dummy variable that equals one if there is at least one new county school built or rebuilt in the prefecture-term. In Columns (4)-(6), $Famine$ is a dummy variable that equals one if there is a famine event in the prefecture-term. $Treat$ is a dummy variable that equals one if the prefecture is a residency prefecture of a great family. $Post1033$ is an indicator that equals one if the year is after 1033. $Between(1033, 1045)$ is an indicator that equals one if the year is after 1033 and before 1045. $Post1045$ is an indicator that equals one if the year is after 1045. The controls include population, land size, distance to major river, number of Confucian philosophers, distance to publisher, distance to border, and climate. Standard errors in parentheses are clustered at the prefecture level. ***, **, and * denote significance at the 0.01, 0.05, and 0.10 levels, respectively.

Figure A.1: Records on the Family Background of a *jinshi*.

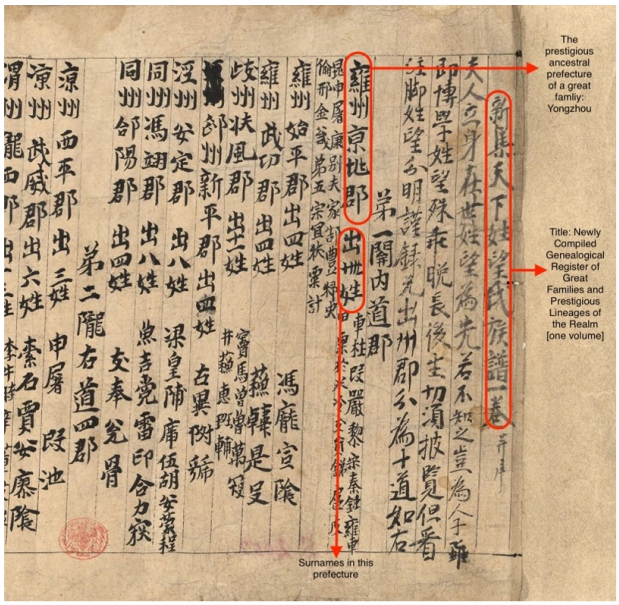


Note: This figure presents a record of a *jinshi* Xiang Song and his family background searched from the China Essential Ancient Books Database by Airusheng. In this example, the page is sourced from "Collected Inscriptions and Biographies of Illustrious Officials" (*Ming chen bei zhuan wan yan ji*) by Dagui Du in the Song dynasty (digitized in the China Essential Ancient Books Database by Airusheng). This record includes four generations of the Xiang Song family: Xiang's great-great-grandfather served as Vice Censor-in-Chief during the Tang dynasty (618-907). Xiang's great-grandfather served as the Magistrate of Qianfeng County in Yanzhou during the Later Han dynasty (947-950). Xiang's grandfather served as the magistrate of Huoqiu County in Shouzhou during the Later Zhou dynasty (951-960). Xiang's father, Qi Song, who was also a *jinshi* and served as a Prefectural Judge in Jingnan during the Song dynasty.

Figure A.2: Records on the List of Great Families.



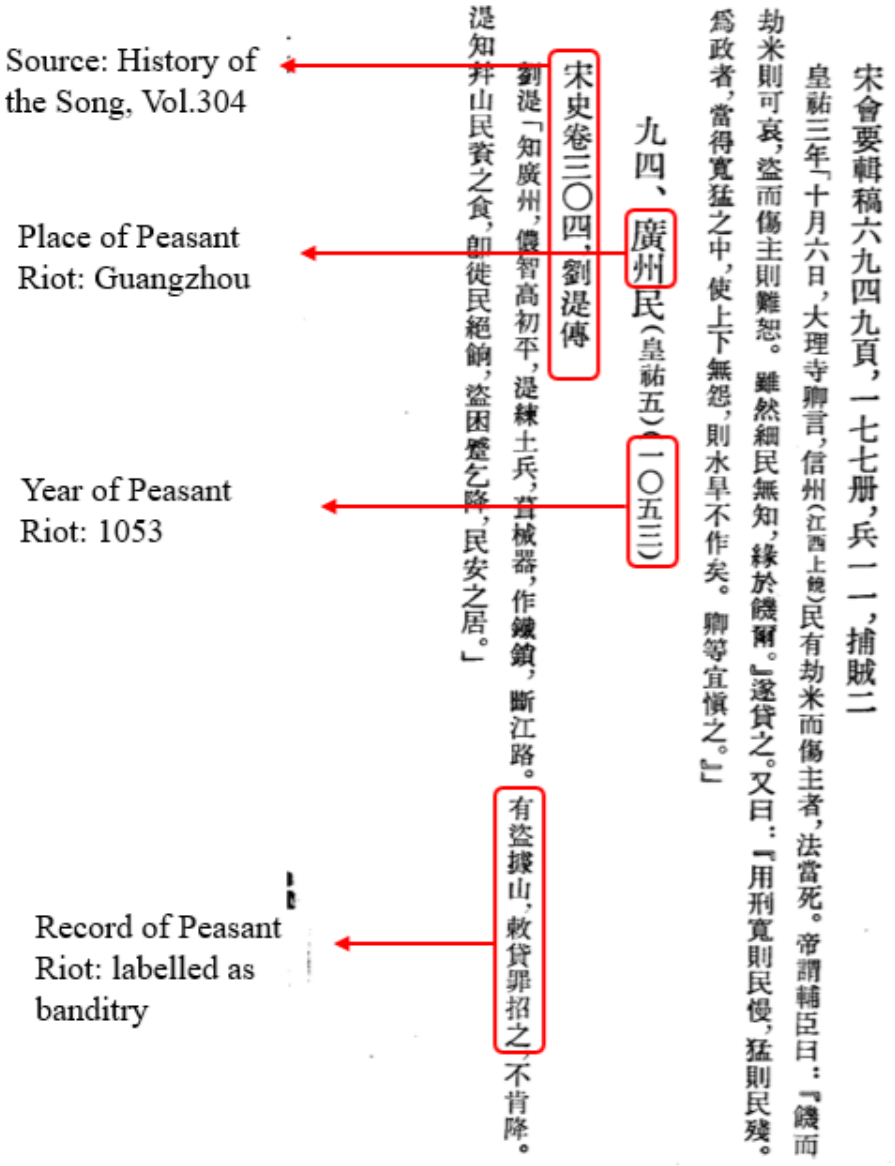
(a) Records in the *Shizu Lun* (Treatise on Great Families).



(b) Records in the Xinji tianxia xingwang shizu pu [yijuan] (Newly Compiled Genealogical Register of Great Families and Prestigious Lineages of the Realm [one volume]).

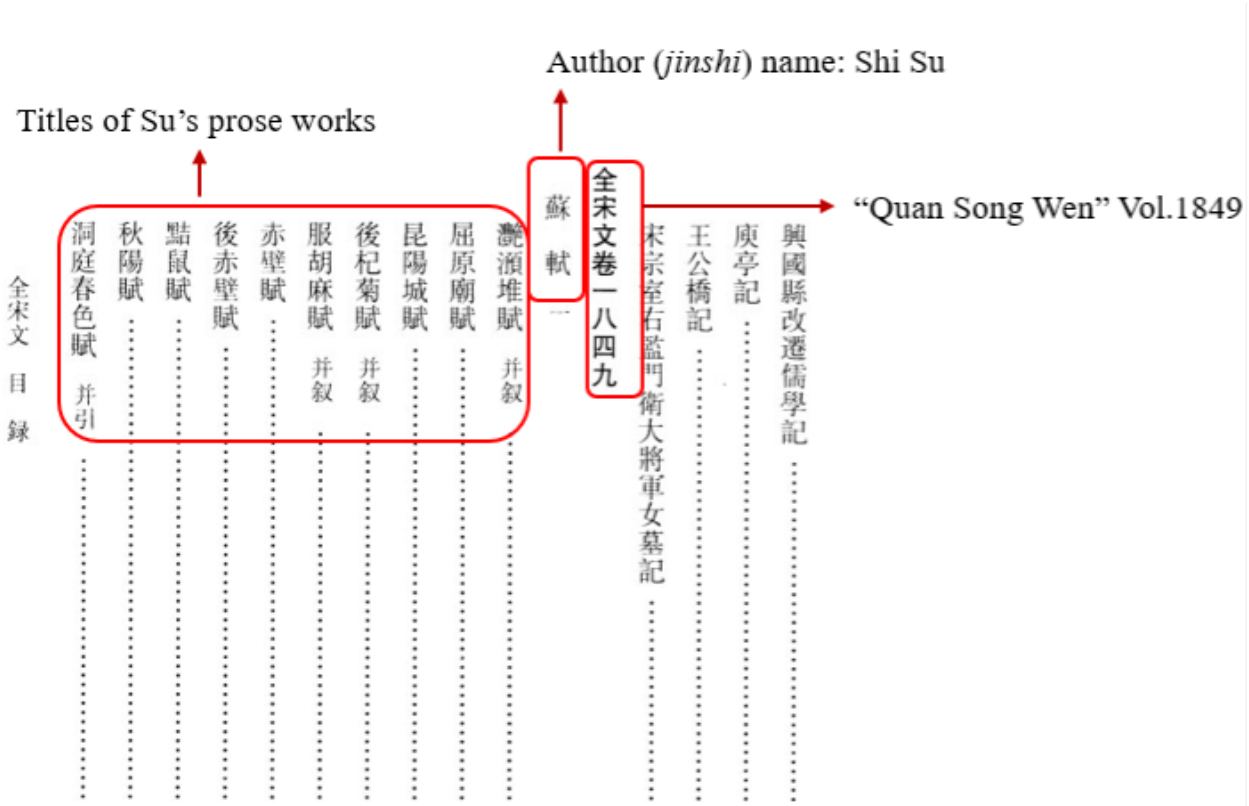
Note: The top picture presents the *Shizu Lun* (Treatise on Great Families), sourced from "New Tang Book" (Ouyang, 1976). The bottom picture presents the Newly Compiled Genealogical Register of Great Families and Prestigious Lineages of the Realm, sourced from "Dunhuang Manuscript" S.2052 (Stein Collection) dating to the Tang-Five Dynasties period. The manuscript is preserved at the British Library and was accessed via the International Dunhuang Programme. Overall, we identified 778 great families based on these two documents.

Figure A.3: Example of a record of Social Unrest



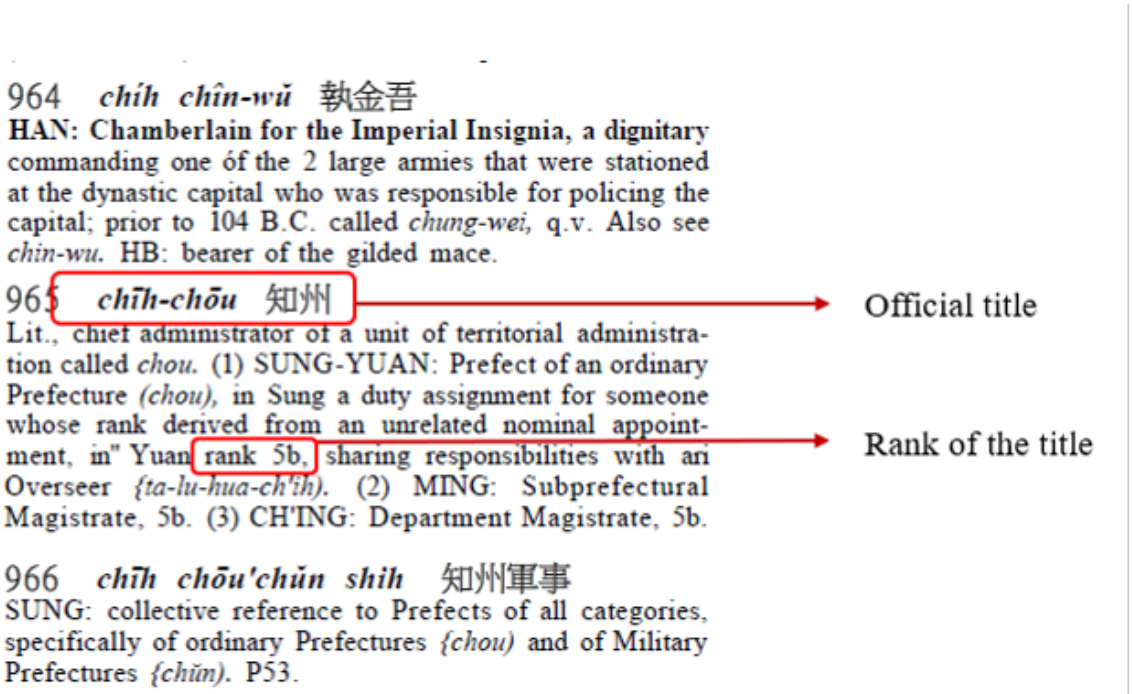
Note: This figure presents an example of a peasant riot record in He (1976), sourced from "History of the Song" (Tuotuo, 1977). The example shows the record of a peasant riot in the prefecture of Guangzhou in 1053.

Figure A.4: Example of a *jinshi*'s Prose Works



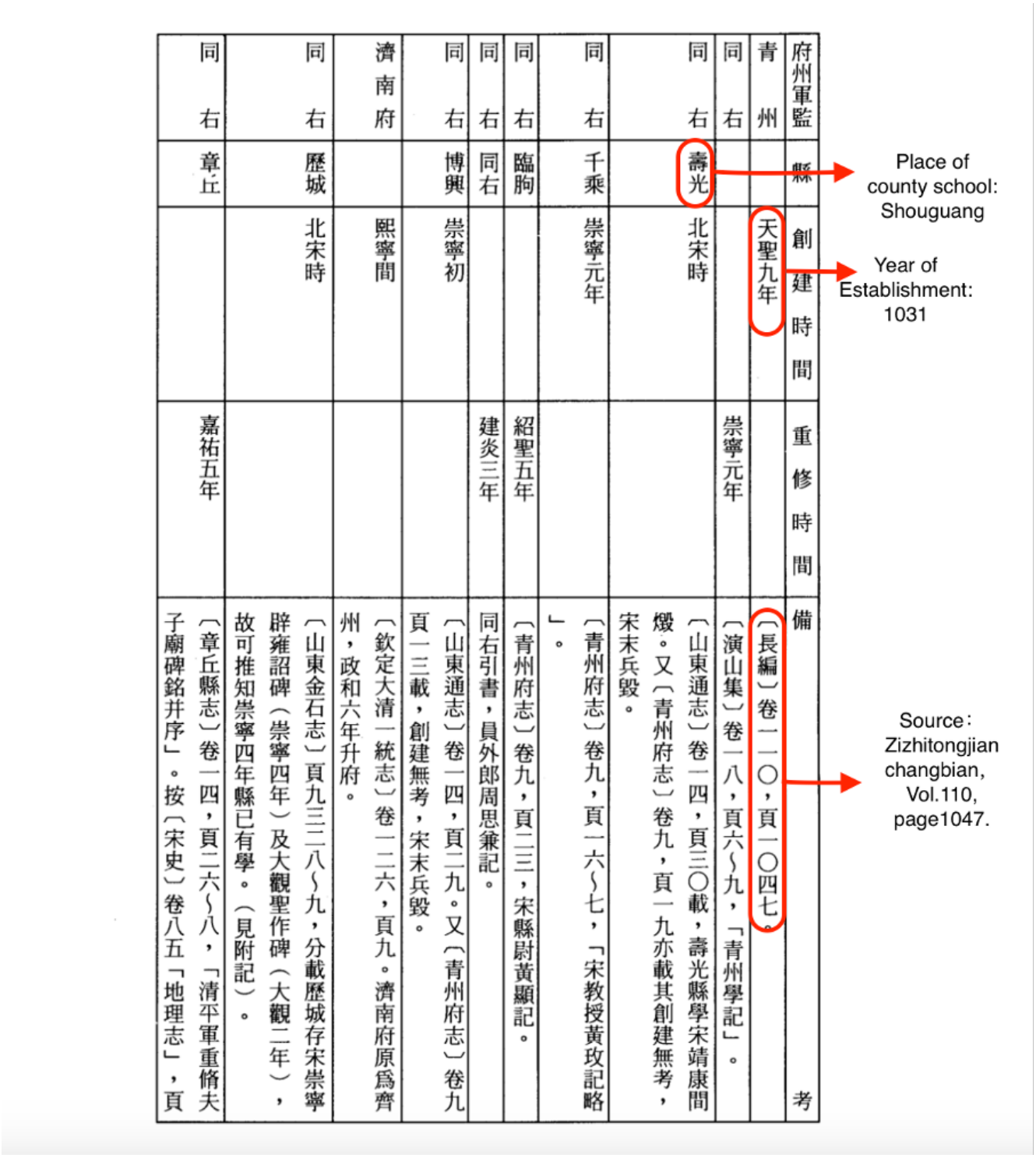
Note: This figure presents an example of the prose work by *jinshi* Shi Su collected in "Quan Song Wen" (Zeng and Liu, 1988). The figure displays a single page listing ten of his prose works; in total, Su has 4,837 prose works recorded in "Quan Song Wen".

Figure A.5: Records on the Ranks of Official Titles.



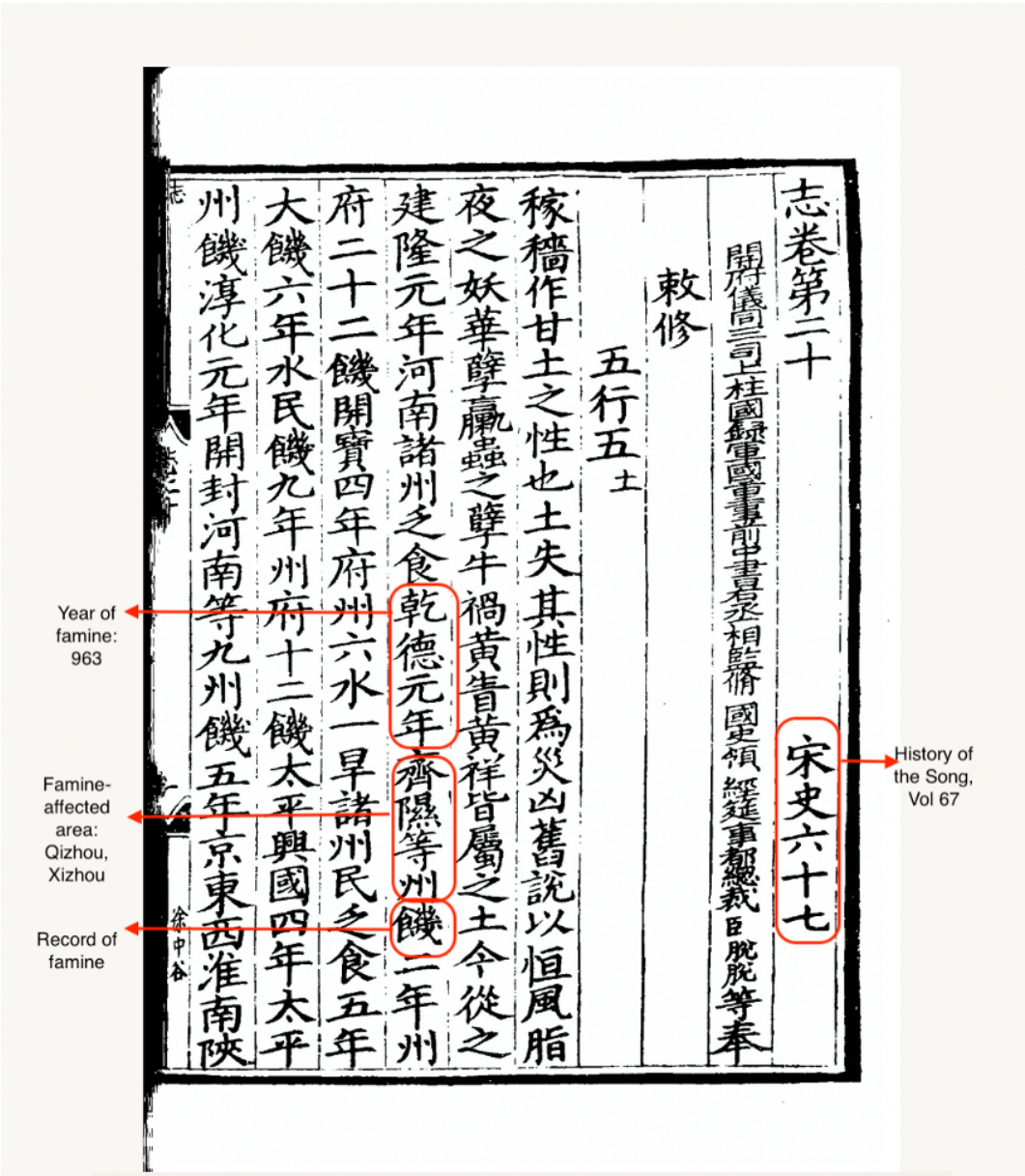
Note: This figure presents examples of the ranks of official titles from Hucker (1985): p.959. Official titles in traditional China are classified into nine grades, where grade 1 is the highest and grade 9 is the lowest. Each grade is further divided into two classes, "a" and "b", with "a" being superior. The example shows that *Zhizhou*, the governor of a prefecture, is ranked as 5b.

Figure A.6: Records on County School.



Note: This figure presents an example of a county school record in Zhou (1996), sourced from *Zizhitongjian changbian* (Li, 2004). The example shows the record of a county school established in Shouguang in the prefecture of Qingzhou in 1031.

Figure A.7: Records on Famine.



Note: This figure presents an example of a famine record documented in the *History of the Song* (Tuotuo, 1977). The example shows the record of a famine in the prefectures of Qizhou and Xizhou in 963.