The Economic Consequences of Political Hierarchy:

Evidence from Regime Changes in China, 1000–2000 C.E.*

Ying Bai† Ruixue Jia‡

Date: April 28, 2021

Abstract

We study how political hierarchy shapes regional development in China, using variations driven by regime changes during the 1000–2000 period. We find that changes in the status of the provincial capital led to the rise and decline of different prefectures as measured by population and urbanization. Two other novel findings stand out: (1) the economic advantages of the provincial capitals did not persist if they lost their political status, and (2) political hierarchy shaped economic development not only through public employment but also through the development of important infrastructure, such as transportation networks. Our findings highlight the importance of politics in determining the locations of economic activities.

JEL codes: H11, N95, O11

*An earlier version of this paper was circulated under the title "The Oriental City?: Political Hierarchy and Regional Development in China, A.D. 1000–2000." We thank Rema Hanna, the anonymous referees, Treb Allen, Konrad Burchardi, Filipe Campante, Donald Davis, Edward Glaeser, Walker Hanlon, Stephan Haggard, Gordon Hanson, Jeffrey Lin, Stelios Michalopoulos, Nancy Qian, Craig McIntosh, Torsten Persson, James Rauch, Andrei Shleifer, Xiaoxue Zhao, Junfu Zhang, Fabarizio Zilibotti, and the participants in various seminars for their comments. Ying Bai acknowledges the financial support of the Research Grants Council of Hong Kong (grant no. GRF 14502118).

†Department of Economics, Chinese University of Hong Kong, ybai@cuhk.edu.hk.

‡University of California, San Diego, and London School of Economics and Political Science, ruixue.jia@gmail.com.

1 Introduction

Political hierarchy is observed in most countries, whose governance is divided into units (e.g., provinces, states, regions) with administrative centers that connect these units with the central government. Economists use cross-country or cross-regional variations to document that administrative centers (or cities) are larger than average cities (e.g., DeLong and Shleifer 1993, Ades and Glaeser 1995, Davis and Henderson 2003) and that the locations of these centers are important in terms of political accountability and stability (Campante and Do 2014, 2019). However, some related questions remain unanswered. For example, how does political hierarchy shape regional development? Should we consider administrative centers to be "parasitic cities" that achieve success primarily due to an increase in public employment but have no advantage in productivity?[1] The answers to these questions will not only inform several lines of research in the fields of political economy and urban economics but also have important policy implications, as the determination to set up administrative centers can be considered a place-based policy.

A major challenge to investigating the long-run consequences of gaining and losing political importance is the lack of a suitable empirical setting, in which there is reshuffling of administrative centers within a relatively stable political system. In addition, such gains and losses may be endogenous. For instance, it may the more successful cities that become administrative centers. In this paper, we argue that China during the second millennium offers an appealing setting in which to address these challenges. First, six major distinctive dynasties or regimes existed during the 1000–2000 C.E. period, and regime shifts during this period led to changes in national capitals, provincial boundaries, and provincial capitals (with the latter among the prefectures that comprise provinces). Second, as noted by scholars, the three-tier administrative system (province-prefecture-county) in China was formed in approximately 960 C.E. and remained stable over the following millennium.[2]

[1]This notion has long been discussed by social scientists, with such centers described as "Oriental cities," following Weber (1921), or as "parasitic cities," following Bairoch (1985)).

[2]The persistence of this administrative and bureaucratic structure is regarded by many as one of the most important

Third, thanks to this stable administrative system, rich information from several centuries is available in historical archives. We construct a dataset based on a large number of historical and modern sources that enabales us to study the consequences of political changes for development within periods comprising several decades before and after the changes. Finally, guided by historical narratives, we can address the challenge of endogeneity by providing a novel algorithm to predict the locations of administrative centers.

Our empirical setting is the core regions of China (the shaded area in Appendix A.1, termed "China proper" by historians) during the second millennium.[3] We aim to understand the effects of (re)locating provincial capitals–the most important regional administrative centers–on the spatial locations of economic activity. Accordingly, for our baseline analysis, we construct a panel dataset across 261 prefectures during 11 periods (980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964, and 2000).[4]

We first discuss the underlying reasons for relocating provincial capitals. Although the choice of a provincial capital certainly involves many factors, the process is subject to general and testable political logic. Specifically, a provincial capital plays two roles: it acts as an administrative center responsible for fiscal, judicial and welfare governance within a province, and as a node through which to transfer resources and information from a province to the central government. Thus, when determining the location of a provincial capital, the central government faces a trade-off between granting provincial autonomy and maintaining central control. This trade-off can be proxied empirically by the weighted sum of the distances to the other prefectures within a province and the distance to the national capital (what we call the "hierarchical distance"). Intuitively, if provincial autonomy dominates, we would expect proximity to other prefectures within the province to be more important, whereas if central control dominates, proximity to the national center should be more important. During the millennium considered in this study, China underwent six dynastic regime

features of the Chinese political system (e.g., Ma and Rubin 2019).

[3]Historical China is usually divided into proper and peripheral regions. We do not study the peripheral regions because they were not always under the control of the central government, and few historical data are available.

[4]The median area of a Chinese prefecture is 11,649 square kilometers, roughly comparable to that of seven median-sized counties in the United States.

changes that led to relocations of national capitals and the redivision of provincial boundaries. As a result, the two sources of variation that underlie hierarchical distance–the locations of the national center and the provincial boundaries–are altered by regime change. Utilizing these variations, we show that the simplest form of hierarchical distance (e.g., based on exogenous geometric distances) already has a powerful influence on changes in provincial capital locations across regimes. Therefore, this form of hierarchical distance may be a useful instrument for determining the status of provincial capitals.

Using a difference-in-differences (DID) strategy and an instrumental variable (IV) strategy based on the hierarchical distance, we find that administrative hierarchy has important effects on the spatial distribution of economic activity over the millennium. Specifically, changes in administrative hierarchy led to the rise and decline of different prefectures (as measured by population density and urbanization). The magnitude of this consequence is large. For instance, using our DID strategy, we find that the change in provincial capital status is associated with a 40–50% change in population density. We discuss the empirical challenges facing the DID and IV strategies and conduct a series of checks, such as period-by-period analysis to examine pre-trends, grid-level analysis to address uneven gaps in our data, and checking concerns about the exclusion restrictions applied to our IV analysis. Overall, the results obtained using one strategy corroborate those obtained using the other.

We examine the impact of gaining and losing provincial status separately. Unsurprisingly, we find that gaining political importance led to improved development. In light of the large body of literature on the persistence of economic activities (discussed below), however, it is less clear whether losing political importance leads to significant declines of provincial capital cities. In this context, we find that such a loss of status is very important: once status was lost, a provincial capital returned to a level of development similar to that of its non-capital peers within 75 years. The observed lack of persistence prompts us to ask why these economic advantages disappeared. This question is related to the underlying mechanism by which politics determines the allocation of economic activities.

Three economic forces may drive this link between politics and development: the direct effects of public employment, the spillover effects of public employment (through private employment and agglomeration), and improvements in market access (mediated by investments in infrastructure). Using public office and transportation network panel datasets, we find that these forces can work together. Importantly, our finding suggests that the consideration of provincial capitals as "parasitic cities" may be too narrow, as market access is also altered by political status. The market access finding has implications on the heterogeneous impact of capital status on population and spatial spillover, which are supported by data. In addition, we conduct a counterfactual analysis to illustrate the economic cost of setting up provincial capitals following the political logic. In a politically decentralized counterfactual scenario where provincial capitals were set at the provincial centroid, provincial aggregate market access could be 10.8% higher.

Our study adds to the political economy literature on how political factors affect economic geography. Compared with the above-noted studies, the setting of our study allows us to shed light on the determinants of political status in certain localities and to study both gains and losses of political importance. In this process, we uncover some new factors that underlie the link between politics and development. In particular, a strong centralized state plays a critical role in overcoming inertia in economic activity by providing important public goods, such as transportation networks. Dell, Lane, and Querubin (2018) provide convincing evidence of the benefits of a Chinese-style strong and centralized state through a study of regional variation in Vietnam. Although our main findings are consistent with the effectiveness of a strong centralized state, the declines of former capital prefectures and changes in transportation networks have implications on the cost side of such a political system.

Our perspective complements the literature on the long-term spatial distribution of economic activity by documenting its changes.[5] A large body of political economy literature emphasizes the long-term persistence of economic activity due to different channels (see Nunn (2009) for a review).[6]

---

[5]To clarify, we also find some persistent patterns in our key variables (e.g., population density, human capital, and transportation), implying that geography and other market factors also matter in China. Although these factors are somewhat persistent, they change systematically with political status. However, these changes are less well studied than patterns of persistence.

[6]As noted by Nunn (2009), these channels include geography, institutions, human capital, and culture. We must omit here

Another stream of literature is motivated by trade theories and emphasizes the role of locational fundamentals (e.g., Davis and Weinstein 2002, Barjamovic et al. 2019) or the persistent impact of temporary advantages (e.g., Bleakley and Lin 2012, among many others). In light of the theory of locational fundamentals, our findings could be interpreted to mean that politics can affect geography in our setting by, for example, altering canal routes and land routes. Our findings may differ from earlier work emphasizing the persistent impact of temporary advantages because in our context, the state plays an important role in overcoming inertia in the location of economic activity.[7] Our finding thus suggests that whether the state (or other stakeholders) has the incentive and capacity to invest in important infrastructure, such as transportation networks, is an important predictor of whether we should expect persistence or not This perspective can be useful when interpreting persistence versus changes in different contexts.[8]

Our findings also inform the place-making literature (e.g., Glaeser and Gottlieb 2008). In recent years, an increasing number of countries have increased their quantity of subnational administrative units, leading to an increase in the number of administrative centers (Grossman and Lewis 2014). Although such policies are intended to enhance economic development, their economic consequences remain unclear.[9] Our finding suggests that in a context in which politics plays a critical role in resource allocation, it is unclear whether the implementation of administrative centers can boost economic development in the long run after a loss of political advantage.

Finally, our study contributes to the literature on the political economy of administrative divisions (e.g., Alesina and Spolaore 1997, Bolton and Roland 1997, Michalopoulos and Papaioannou 2013,

---

the related literature produced during the decade since his review. In particular, by taking a long temporal framework, our study joins the effort to understand long-run urban development in different parts of the world (e.g., Bosker, Buringh, and van Zanden 2013, Stasavage 2014, Duben and Krause 2020). Our study design is appealing because it allows us to trace gains and losses of status.

[7]In other settings, however, market players can be incentivized to play this role if the location of economic activity has a strong effect. For instance, Rauch (1993) provides examples wherein industrial developers engage in such behavior.

[8]For instance, Huillery (2009) and Jedwab, Kerby, and Moradi (2017) observe the strong persistence of colonial investments in Africa. From our perspective, one interpretation is that African states lack the capacity to alter these investments. Additionally, our finding that urban geography is shaped by a change in transportation networks, which is driven by a political regime shift in our setting, is closely related to the findings reported by Michaels and Rauch (2018) on Britain before and after the collapse of the Roman Empire.

[9]Bluhm et al. (2020) use cross-country nightlight data to evaluate the short-term impact of administrative status. However, the extent of the long-term impact remains to be determined .

Ko, Koyama and Sng 2018). These studies often focus on national-level variation; in contrast, our research question involves within-nation variation. Broadly, we consider the logic of setting up administrative centers to be a spatial implementation of federalism, as discussed in earlier theoretical literature (Blanchard and Shleifer 2001, Treisman 2007).

2 Background and Data

2.1 Regime Changes and Capital Relocation

The Chinese political system, which is characterized by a stable three-tier (province-prefecture-county) administrative system and multiple distinctive dynasties/regimes, is particularly suitable for our research design. Regime shifts have led to changes in national capitals, provincial boundaries, and provincial capitals. In each regime, however, the central government has monopolized the powers of appointing, rotating, and removing officials on each tier. Below, we summarize the major administrative changes that occurred over the millennium under study.

Regime Changes During the millennium studied, six major regimes came into existence: the Song Dynasty (960–1279), which coexisted with the Liao (907–1125) and Jin (1115–1234) dynasties in the north, the Xixia Dynasty (1038–1227) in the northwest, and the Dali Dynasty (937–1253) in the southwest. This was followed by the Yuan (1271–1368), Ming (1368–1644), and Qing (1636–1912) dynasties, the Republic (1912–1949), and the People's Republic (1949–the present).

Such infrequent regime changes are difficult to predict. For example, it would have been difficult for residents of the Song Dynasty capital (Kaifeng) to imagine that China would later be ruled by the Mongols, who moved the national capital to northern China (Beijing) and drew provincial boundaries to achieve political control. The administrative decisions made by the rulers at the beginning of each new regime, including the locations of the national capital, provincial boundaries, and provincial capitals, usually persisted until the end of the regime, interspersed by only occasional changes. Variations in the national and provincial capital locations are, therefore, driven primarily by regime changes. Because a national capital relocation and provincial boundary redivision are the preconditions for changes in the provincial capital, we describe these in order below.

National Capital Relocation The national capital changed five times across the six regimes, reflecting the unpredictability of where a new power base could arise. Specifically, modern-day Kaifeng (in central China), Beijing (in northern China), and Nanjing (in central-southern China) were the national capitals of the Song, Yuan, and Ming dynasties, respectively,[10] after which Beijing served as the capital of the Qing Dynasty, and Nanjing as that of the Republic. Finally, the capital of the current People's Republic was once again established in Beijing.

Historians argue that one important determinant of a national capital location is the origin of the power base (e.g., Chao 2006). For instance, the Mongols (founders of the Yuan) and Manchurians (founders of the Qing) originated in the north and, hence, chose Beijing as their national capital, whereas the power bases of the Ming and Republic were in the south, leading these regimes to favor Nanjing. Although both Beijing and Nanjing were candidates for the national capital of the People's Republic, Beijing was chosen partly due to its nearness to China's political ally at that time (the Soviet Union). Although the power base is not the only determinant of a national capital location, these observations illustrate that political considerations are usually more important than economic considerations.

Provincial Boundary Redivision Provincial boundary alterations occurred after the national capital relocation. During the studied millennium, provincial boundaries were affected by a major shift between two ways of defining provincial boundaries: *following or subsuming* the natural lines of mountains and rivers. The former is known as "*suiting [i.e., following] the forms of mountains and rivers*"; the latter, which intentionally includes natural barriers within provinces, is known as "*interlocked like dog's teeth*" (e.g., Zhou 1998). The latter exemplifies a spatial "divide-and-rule" tactic to limit the power of local governments.

The Song adhered to the first principle, generally defining provincial boundaries by natural mountains and river. However, when the Mongols rose to power, their preoccupation with the

[10]Nanjing was the national capital of the Ming Dynasty until its capital was relocated to Beijing in 1421, partly because the new emperor, who had taken power through a coup, had established his power base in Beijing. The Crown Princess, however, remained in Nanjing, which became the norm for this regime.

possibility that a usurper would mobilize resources against the central government led them to adopt an extreme version of the interlocking principle, with natural mountains and rivers intentionally included within (larger and fewer) provinces. After the Yuan Dynasty, regimes adjusted the number of provinces using both principles, leading to a general increase in the number of provinces. In Appendix A.2, we use the Yangtze River as an example to illustrate how the two principles of defining provincial boundaries were implemented across regimes. As shown, the Yangtze River was used as a provincial boundary by the (pre-Mongol) Song but included within provinces by the (post-Mongol) Ming and Qing. Historians document similar changes in provincial boundaries driven by the evolution of these two principles in other regions, including the Huang River, Qin Mountain, and Taihang Mountain (Zhou 1998).

Based on these historical narratives, Sng et al. (2018) develop a congruence index to measure how strategic provincial boundaries were drawn by comparing actual provincial boundaries with hypothetical boundaries predicted by catchment or least-cost distance. This measure reveals two groups of provinces distinguished by high and low levels of strategic boundary setting. Consistent with historians' conjecture, those with higher levels of strategic setting are located around the major rivers and mountains mentioned above. We consider these variations in our analysis.

Provincial Capital Relocation The relocation of national capitals and redivision of provinces naturally affected the relative importance of prefectures, which could become rather isolated due to the new delineation and national capital, despite their central status based on the old provincial boundaries. For instance, Luzhou was the capital of Hedong province during the Song Dynasty because it connected the national capital, Kaifeng, with other prefectures in the province (see Panel (a) of Appendix A.3). During the Ming Dynasty, however, Luzhou lost its capital status because the redrawing of the provincial boundary placed it far from other prefectures in the province, even though it was still closer to the new national capital of Nanjing (Panel (b) of Appendix A.3). When the national capital relocated to Beijing (Panel (c) of Appendix A.3), Luzhou became even further isolated and, as a result, never regained its capital status.

In contrast, Changsha was the provincial capital of Jinghu South province during the Song

Dynasty and was relatively close to the national capital and the other prefectures in the province (Appendix A.3). This prefecture became isolated and lost its provincial status during the Yuan and Ming dynasties. In the Qing Dynasty, however, it regained its capital status due to provincial redivision.

For our baseline analysis, we map historical data onto the 261 prefectures that existed in the year 2000 and construct a panel dataset (see the discussion of data construction in Section 2). Of these 261 prefectures, 63 were provincial capitals at least once during the period studied (see their locations in Figure A.1): 36 lost their capital status once, 11 gained their capital status once, eight experienced multiple changes, and eight remained capitals throughout. We summarize these changes across regimes in Figure 1.

Notably, each province has only one capital in a given regime except during the Song Dynasty, when the central government limited the power of local governments by separating capitals spatially according to fiscal affairs, judicial affairs, and welfare (Mostern 2011). As a result, the majority of provinces had two provincial capitals during this dynasty: one for fiscal affairs and one for judicial affairs and welfare. This feature is helpful because it provides a broad baseline set of possible candidates for provincial capitals. We include both in our baseline analysis and show that our findings are robust to the exclusion of the Song Dynasty (and any particular regime).

2.2 Data

We construct a prefecture-level panel dataset covering the 980–2000 period. We focus on four of the six major regimes. We do not study the Yuan (1279–1368) or the Republic (1912–1949) for two reasons. First, short-lived regimes are less suitable for our study because population sizes take time to adjust after a regime change. Second, prefecture-level data are less available for these two regimes.[11] We present the summary statistics of the key variables in Table 1, and the definition and measurement of each variable in Appendix B.1 and B.2.

[11] In addition, the Song Dynasty (960–1279) was divided into the Northern Song (960–1127) and the Southern Song (1127–1279). The latter controlled only the southern part of China proper and did not conduct any census. Thus, the data on the Song Dynasty cover only the Northern Song.

Panel Construction Although some prefecture boundaries remained fairly stable during the period under study, many others changed across regimes. To construct a panel dataset, one must fix the boundaries and map the original data to these fixed boundaries. We achieve this by (i) using the boundaries of the 261 prefectures in 2000 in our baseline analysis, as our readers are more familiar with these divisions; and (ii) complementing these with our grid-level analysis, in which the 261 prefectures are divided into grids with sizes that can be smaller or larger than the prefecture at baseline. We define a capital prefecture (grid) as a prefecture (grid) that contains the geocode of a provincial capital.

Population and Urbanization Because data on population density are the most comprehensive type of data with which to measure long-term economic development, our baseline estimations include population data from 11 years, based on all of the existing censuses: 980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964, and 2000. We are also able to access population data from the years 1880, 1953, 1982, and 1990. However, the gaps between periods are already uneven, which poses a challenge to the research, and including the latter group of years would make these gaps even more uneven. Therefore, we intentionally exclude the latter four years from our baseline and include them as a robustness check. Not surprisingly, the first census of each regime was usually conducted decades after the establishment of the regime.

The population data are originally available at the prefecture level based on the boundaries in each regime. We map these data to fixed prefectures in 2000 or a grid as follows. Suppose that a fixed prefecture i in 2000 comprises fractions of two prefectures (1 and 2) in 1080. We first calculate the population density in 1080 by the prefecture boundaries in 1080 and obtain $PopDensity_{1,1080}$ and $PopDensity_{2,1080}$. Second, we calculate the areas of the factions included within the fixed 2000 prefecture i to yield $Area_1$ and $Area_2$. Then, we calculate the population of prefecture i in 1080 as $PopDensity_{1,1080}$*Area$_1$ + $PopDensity_{2,1080}$*Area$_2$. Dividing the population size by the prefecture area in 2000 (i.e., $Area_1 + Area_2$), we obtain the population density of prefecture i in 1080.[12]

[12]See Appendix B.1 for a more general case, where we show that prefecture boundary changes can be considered as measurement errors in the capital status of a prefecture, which likely lead to an underestimation of the importance of capital status.

We use urbanization measures as a complement to our population density measure. Compared with population data, urbanization data are less systematically available; specifically, such data are accessible for only four of our 11 years: 1580, 1820, 1964, and 2000.[13]

Provincial Capitals and Prefecture Boundaries We use data from CHGIS (2007) on the boundaries and provincial capitals during the period from the Ming Dynasty to 2000 and digitized data on the Song Dynasty based on the Treatise of the Nine Regions from the Yuanfeng Reign (1078-1085), a Song imperial geography widely used by historians who study this dynasty (e.g., Mostern 2011). As expected, variations in the provincial boundaries and capitals can be attributed to regime change (i.e., these were set at the beginning of each political regime). Using these data, we determine the prefecture boundaries during each regime and the exact geolocation of each provincial capital.

Prefecture-level Controls We capture a prefecture's characteristics by including three sets of additional variables: factors related to geography, agriculture, and regional location. The eight geographic prefecture variables include the log area, whether it contains a plain, whether it contains a major river, whether it is located on the coast, the average slope, the elevation, the longitude, and the latitude. The five agricultural variables measure suitability for wheat, rice, fox millet, maize, and sweet potato crops (FAO GAEZ 2012). To enhance the comparability of prefectures, we compare prefectures within the well-known nine physiographic macroregions defined by Skinner (1977): the North China plain; Northwest China; the Lower, Middle, and Upper Yangtze Plains; the southeast coast; Lingnan; Yun-Gui; and Manchuria (see the map in Appendix B.2). As provincial boundaries vary over time, we also control for province fixed effects in several ways.

[13] The data from 1580 are estimated based on local gazetteers, provided by Cao Shuji, a leading scholar in Chinese population history at Shanghai Jiaotong University. The historical urbanization data from 1820 are estimated based on the population living inside and outside walled cities, obtained from CHGIS (2007), a database of place names and historical administrative units associated with Chinese dynasties, available at https://chgis.fas.harvard.edu/. The 1964–2000 data are collected from population censuses.

3 Understanding the Choice of Provincial Capitals

The selection of a provincial capital is a complex task involving many considerations. We aim to highlight important political logic that is applicable to all regimes and could be used to generate an instrument for determining provincial capital relocation. Motivated by the roles of provincial capitals, we provide a simple algorithm to characterize the selection of a provincial capital by the central government and apply this algorithm to our dataset.

3.1 A Simple Algorithm

The Roles of Provincial Capitals Given China's vast territory, it would be impossible for the central government to collect taxes and information directly from every local administrative unit. Thus, a hierarchical system is necessary to maintain political control. Provincial capitals are the most important nodes through which the central government connects with lower-ranked administrative units within a province. Conceptually, a provincial capital plays two essential roles. First, it is the administrative center of the province, where fiscal and judicial affairs and welfare issues are governed. Second, it facilitates the transfer of resources and information between the central government and all of the prefectures within the province. For instance, in the Song Dynasty, provincial governors (particularly those known as *fiscal intendants*) inventoried local resources and arranged for their transportation to the national center when appropriate (Mostern 2011). Today, the fiscal budget of a province is first determined at the level of the provincial capital and then applied to lower-ranked administrative units. Although the details vary across regimes, the two roles of the provincial capitals have persisted.

Formalization Next, we formalize the two roles of provincial capitals by exploiting a prefecture's proximity to the other prefectures within a province and to the national capital. The relative importance of these roles may have shifted across regimes. To address these changes, we also allow the relative importance of these roles to vary by regime.[14]

We use tax collection by the central government to drive the algorithm, the logic of which can

[14]Similarly, the importance of national capitals and provincial boundaries can vary across regimes. Such changes, however, should not affect our main findings because we are comparing prefectures within the same regime.

be applied to other dimensions (e.g., information collection). Assume that the gross tax revenue in a province is T. In each regime r, the aim of the central government is to maximize effective tax revenues. Here, two components of cost are important. The first is the cost of moving tax resources from different prefectures (j) in the province to the provincial capital: $T \times a \sum_{j=1}^{N} \frac{A_j}{\sum_{j=1}^{N} A_j} D_{i,j,r}$, where a indicates the marginal cost of distance; $D_{i,j,r}$ indicates the distance from i to another prefecture j in the same province in regime r; and $\frac{A_j}{\sum_{j=1}^{N} A_j}$ is a weight variable. Note that when A_j is the area of prefecture j, $\sum_{j=1}^{N} A_j$ is the total area of the province and $\sum_{j=1}^{N} \frac{A_j}{\sum_{j=1}^{N} A_j} D_{i,j,r}$ is the distance from prefecture i to the provincial centroid.[15] The second important component of cost is the cost of delivering a λ_r share of the tax to the national capital: $\lambda_r T \times aD_{i,NationalCap,r}$, where $D_{i,NationalCap,r}$ is the distance from prefecture i to the national capital.

Thus, the central government's selection of a capital for a province containing prefecture $i = 1, 2, ..., N$ aims to maximize the following:

$$T - T \times a \sum_{j=1}^{N} \frac{A_j}{\sum_{j=1}^{N} A_j} D_{i,j,r} - \lambda_r T \times aD_{i,NationalCap,r}. \tag{1}$$

This is equivalent to minimizing the following specification, which we term "hierarchical distance" (denoted by $HierDist$):

$$\sum_{j=1}^{N} A_j D_{i,j,r} + \lambda_r \sum_{j=1}^{N} A_j D_{i,NationalCap,r}. \tag{2}$$

The logic of this algorithm can be extended beyond tax collection. Broadly, one can consider that the first and second elements in $HierDist$ capture the importance of provincial autonomy and central control, respectively. λ_r reflects the tradeoff between these two considerations.

We attempt to keep the elements included in $HierDist_{i,r}$ as exogenous as possible. Therefore, we focus on geographical variables. In principle, this algorithm can be enriched by considering other factors, such as using the population to measure A_j and considering transportation networks in our calculations of D. As we show later, however, both population and transportation can be

[15]To visualize this intuition, if A_j=1, each prefecture is treated as equally sized, i.e., the relative sizes of the prefectures are ignored.

affected by political hierarchy.

Assuming that $\lambda > 0$, $HierDist_{i,r}$ implies that provincial capitals should be located farther from the provincial centroid and closer to the national capital. This prediction is confirmed by the empirical patterns. To see these patterns, we zoom in on the cases of Ming (the middle period in our data) and modern China in Figures 2(a) and 2(b) (see other regimes in Appendix C.1), where the provincial capitals and centroids are marked with red squares and crosses, respectively. As shown in both maps, the squares (provincial capitals) tend to deviate from the crosses toward the direction of the national capital.[16] Thus, our simple algorithm based on arguably exogenous features shows promise as a predictor of provincial capital status. Next, we test this argument more systematically.

3.2 Test the Role of Hierarchical Distance

Here, we test the association between $HierDist_{i,r}$ and the provincial capital status of a prefecture i in regime r. To measure $HierDist_{i,r}$ systematically, we need to assign values to λ_r. We choose the value of λ_r that delivers the highest predictive power in each regime. We find that λ_r varies from around 0.3 in the Song and Ming Dynasties to around 0.2 in the Qing Dynasty and modern China. This choice is not crucial to our analysis, and our findings are robust to alternative ways of specifying λ_r. In Appendix C.2, we demonstrate a similar pattern obtained when assuming $\lambda = 0.2$ for all regimes.[17]

We plot the probability of a prefecture's being a provincial capital by its rank in the province with respect to its hierarchical distance. As shown in Panel (a) of Figure 3, among the top-ranked prefectures (which vary across periods), the probability of being a provincial capital is approximately 0.40, whereas among the second-ranked prefectures, it is approximately 0.24. This probability decreases as the ranking level decreases: beyond the fifth-ranked level, the probability decreases to

[16] Because the administrative hierarchy was set at the beginning of the regime, the Ming considered the location of Nanjing to be more important than that of Beijing (which became the national capital after a coup). In the Song dataset, we treat Xixia and Dali as different regimes and calculate the hierarchical distance of each prefecture in these regions to the respective national capitals (Yinchuan and Dali respectively).

[17] Note that when compared with a fixed λ, allowing λ_r to change by regime does not introduce additional concerns about endogeneity, as λ_r remains consistent for all prefectures within a single regime whose impact is absorbed by year fixed effects in our analysis. Thus, we focus on λ_r instead of a fixed λ, as the former better captures variation in the tradeoffs across regimes in our baseline. We report the results using a fixed λ for robustness.

below 0.1 and approaches 0 by the tenth-ranked level. This nonlinear pattern also suggests a linear relation between the log rank ($lnRankHierDist$) and the probability of being selected as a capital below the 10th ranking level, which is confirmed by the pattern in panel (b).

Motivated by Figure 3, we evaluate how hierarchical distance influences the provincial capital location systematically using both cross-sectional and panel specifications. For each regime r, our cross-sectional analysis is conducted as follows:

$$ProvCap_{i,r} = \beta lnRankHierDist_{i,r} + \theta \mathbf{X}_i + \pi_m + \epsilon_{i,r}, \tag{3}$$

where $ProvCap_{i,r}$ indicates whether prefecture i is a provincial capital in regime r. $\mathbf{X}_i$ includes the geography- and agriculture-related variables mentioned above, and π_m indicates the Skinner macroregions.

Our panel analysis pools all regimes, enabling us to control for prefecture characteristics that do not vary or vary slowly over time (e.g., geography) and factors that affect all prefectures similarly (e.g., dynasty cycles) by including prefecture fixed effects (α_i) and regime fixed effects (γ_r). The specification is as follows:

$$ProvCap_{i,r} = \beta lnRankHierDist_{i,r} + \alpha_i + \gamma_r + \theta \mathbf{X}_i \times \gamma_r + \pi_m \times \gamma_r + \epsilon_{i,r}. \tag{4}$$

We allow the effects of $\mathbf{X}_i$ to vary across time by controlling for $\mathbf{X}_i \times \gamma_r$, and the effects of π_m effects to vary by regime (indicated by $\pi_m \times \gamma_r$). All standard errors are clustered at the prefecture level.

In each regime, we find that $lnRankHierDist_{i,r}$ is a strong predictor of provincial capital status (as shown in Appendix C.3). The explanatory power of $lnRankHierDist_{i,r}$ alone is similar to that of a large set of controls, as demonstrated by the $R-$squared value. We focus on panel specification, wherein we link the change in $HierDist_{i,r}$ with the change in capital status. Again, the change in log rank in $HierDist_{i,r}$ is a strong predictor of a change in provincial capital status. If a prefecture's rank in hierarchical distance doubles , the probability of being selected as a provincial capital decreases by eight percentage points (around 80% of the mean probability).

In our calculation of $HierDist_{i,r}$, we assume the given provincial boundaries. Although administrative histories have documented the dates of provincial capital designation (e.g., Zhang

(1739) for the Ming Dynasty) and showed that provinces were defined before a prefecture was designed as a provincial capital, we remain uncertain whether provincial boundary changes affect our findings. Relying on insights from Sng et al. (2018), we divide the provinces into two groups. The non-strategic group includes provinces whose boundaries were congruent with the hypothetical boundaries predicted by geography, and the strategic group refers to those that exhibit clear incongruence. The logic of hierarchical distance applies to both groups, suggesting that the discussed trade-off is not affected by strategic concerns about provincial boundary setting (see Appendix C.3).

4 The Economic Consequences: Descriptive Evidence

How does political hierarchy affect economic development? To answer this question, we present two sets of descriptive patterns to drive our empirical estimations.

A Case of Two Periods: 980 and 2000 Before analyzing our 11-period panel data, we present descriptive patterns based on the data from 980 and 2000. This two-period structure allows us to depict the main pattern by categorizing the prefectures into four groups:

(1) capitals in both periods, denoted by “yes-yes"

(2) capitals in 980 but not in 2000, denoted by “yes-no"

(3) capitals in 2000 but not in 980, denoted by “no-yes"

(4) not capitals in either period, denoted by “no-no"

In Figure 4(a), the x-axis indicates the standardized log population density in 980, and the y-axis indicates the standardized log nightlight density in 2000. The figure suggests that a prefecture’s economic status changes with its political status. Specifically, the crosses indicate the average economic status of different groups. As shown, an average “no-no" prefecture group is close to the mean in both periods, and an average “yes-yes" prefecture is above the mean in both periods. In contrast, an average “no-yes" prefecture is 0.3 standard deviations below the mean in 980 but one standard deviation above the mean in 2000, indicating a correlation between a gain in capital status and better economic development. An average “yes-no" prefecture is 0.5 standard deviations above the mean and comparable to a “yes-yes" prefecture in 980 (when both were provincial capitals), but decreases to near the mean, similar to a “no-no" prefecture, in 2000 after a loss of capital status.

Pooling All Changes Together: 980–2000 As in the above analysis, we can divide the prefectures into four groups every two regimes. This allows us to visualize the four groups period-by-period. To compare prefectures with similar characteristics, we first regress the log population density on all of the prefecture-level controls mentioned above and obtain the residuals. Then, we plot the average residuals for the four groups period-by-period (relative to the time of status change) in Figure 4(b).

Consistent with the pattern in Figure 4(a), the “no-no" prefecture group is close to the mean (in the macro region) in all periods, whereas the “yes-yes" prefecture group is 40% above the mean in all periods. The gaining group (i.e., “no-yes") is similar to its “no-no" peer group before the status change but increases to 48% higher than the “no-no" peer group after gaining capital status. In contrast, the losing group (i.e., the “yes-no") is around 38% higher than its “no-no" peer group before the status change, whereas the two groups are comparable after losing capital status.

5 The Economic Consequences: Estimation Results

5.1 Empirical Strategies

We use two empirical strategies to identify the economic consequences of political hierarchy: a DID and an IV strategy. Below, we present these two strategies and discuss their empirical assumptions.

The DID Strategy The DID strategy is as follows:

$$y_{i,t} = \beta Capital_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \tag{5}$$

where $Capital_{i,t}$ indicates whether a prefecture i is a provincial capital in year t. $y_{i,t}$ is the outcome variable, e.g., log population density or urbanization rate. The other variables are the same as those in equation (3), except that we can include multiple years of observations within a regime r. Therefore, we use t to indicate the year of observation.

An important assumption of this strategy is that provincial-capital prefectures are comparable to non-capital prefectures before the change in capital status. We validate this assumption in several ways. Under some specifications, we further control for $EverCapital_i \times \gamma_t$ to allow the group of prefectures that have ever been a provincial capital (the ever-capital group) to have non-parametric trends different from those of the never-capital group.

IV Strategy A second strategy is to use the hierarchical distance derived above as an instrument to predict provincial capital status. The first-stage and second-stage specifications are as follows:

$$Capital_{i,t} = \delta lnRankHierDist_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \tag{6}$$

and

$$y_{i,t} = \beta' \widehat{Capital}_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \tag{7}$$

where X_i includes all of the controls in our DID analysis.

Although $lnRankHierDist_{i,t}$ is a strong predictor of provincial status (as shown in Section 3), it must satisfy two additional conditions to be a valid instrument. (i) The change in $lnRankHierDist_{i,t}$ for a prefecture must be orthogonal to its own characteristics, and (ii) $lnRankHierDist_{i,t}$ must satisfy the exclusion restriction; that is, it must affect regional development due to the political status change rather than to omitted channels. We next examine these assumptions.

5.2 Results of Regional Development Analysis

Baseline Results We use the log population density during the 980–2000 period as the primary measure of regional development. Because the prefecture boundaries are fixed in this analysis, the coefficient also can be interpreted as the impact on population size. We further complement the analysis with urbanization data, which are available only for the latter half of the millennium studied.

Our DID estimates based on the 11-period data indicate that provincial capital status is associated with a 40–50% higher population density. Column (1) of Table 2 presents the result obtained with only the prefecture and the year fixed effects, for which we obtain an estimate of 0.62. Once we control for eight geographic variables and their interactions with γ_t, we obtain an estimate of around 0.46 (Column (2)). After adding the five-type crop suitability, regional dummies, and their interactions with γ_t to Columns (3) and (4), the estimates are around 0.47.[18]

[18] For all of the results, we report standard errors clustered at the prefecture level. Our findings are robust to Conley spatial

To put this magnitude into perspective, it is instructive to measure the variation in population density within a province. On average, the population density of the most densely populated prefecture is 158% higher than that of a median prefecture, whereas that of the least densely populated prefecture is only 60% of that of a median prefecture. Therefore, the change in density caused by a change in status can be considered as a shift from a median prefecture (normally the sixth- or seventh-ranked prefecture) to the third-ranked prefecture, or from the third-ranked prefecture to the top-ranked prefecture.

Our DID analysis is challenged by the possibility that capital status might change due to omitted variables that affect development. Based on changes in the coefficient and R-squared values across specifications in Table 2, we can assess the importance of omitted variable bias using the method provided by Oster (2017). We present the bounding estimate in the last row of Table 2. As shown, omitted variable bias cannot easily explain the results in Table 2; after considering the possible biases, we determine a lower bound of β to be 0.327.[19]

Our IV strategy also confirms the importance of provincial capital status. Column (5) shows the reduced-form estimate: here, an increase in the hierarchical distance rank is strongly negatively correlated with population density. When the sample is separated into ever-capital prefectures and never-capital prefectures in Columns (6) and (7), we find that hierarchical distance has a noticeable effect on population density in the ever-capital prefectures but has little effect on the never-capital prefectures, suggesting that political status is a crucial channel through which hierarchical distance exerts its effects.

Column (8) reports the IV and first-stage estimates with an $F-$statistic of 131.9, implying that a weak instrument is not a crucial concern. Compared with the DID estimates, our IV estimate (0.79) is slightly higher, partly because it is local to those prefectures whose political status is affected by hierarchical distance. Using a fixed λ for all regimes, the estimated hierarchical distance is 0.70 (presented in Appendix D.3).

standard errors (see Appendix D.1).

[19]See the step-by-step calculation in Appendix D.2.

To address the problem of provincial boundary changes, we further include $Prov_{i,r} \times \gamma_t$, where we control for province fixed effects in all regimes and their interactions with year fixed effects. We observe a similar IV estimate in Column (9). In Appendix D.4, we present both the DID and IV results obtained by further controlling for $Prov_{i,r} \times \gamma_t$ or $(\cap Prov_{i,r}) \times \gamma_t$. As shown, none of these specifications invalidates the importance of provincial capital status. Following the discussion of provincial boundary settings in Section 3, we also separate the provinces into two groups by the level of strategy used in boundary setting (measured by Sng et al. 2018). As shown in Appendix D.5, our IV estimates and first-stage results hold within both groups, further suggesting that the role of hierarchical distance holds even if strategy is applied to provincial boundary setting.

As an additional measure of development, we use urbanization data, which are available only for the latter half of the millennium studied (1580, 1820, 1964, and 2000). Thus, the findings from this analysis can serve to determine whether our results hold after excluding data from the first half of the millennium. As shown in Table 3, the DID and IV estimates show that provincial capital status increases a prefecture's urbanization rate by 11–18 percentage points, a large increase compared with the change in the mean urbanization rate (15 percentage points). We also find that the provincial capital status has an effect on urban populations approximately five times that on rural populations. This finding further confirms that political status affects economic development in terms of not only population size but also economic activity.

Gaining vs. Losing Political Status We separate capital status loss from capital status gain by using change-on-change specifications (i.e., how a change in capital status affects a change in population density). Columns (1)–(4) in Table 4 initially confirm that change-on-change specifications deliver estimates similar to our baseline. The inclusion of lagged population density in the previous two periods does not invalidate our main finding. Furthermore, we obtain a negative coefficient on lagged log population density in the previous two periods, which suggests a convergence pattern.

As reported in Columns (5) and (6), gaining capital status is associated with an approximately 47% increase in population density, whereas losing capital status is associated with a 39% decrease in population density. These estimates decrease marginally if we further include the lagged log population density in the previous two periods (Columns (7) and (8)). By design, the DID estimates should be interpreted as changes, rather than levels. As indicated in the descriptive pattern in Figure 4(b), however, the losing group had a population density approximately 40% higher than that of its non-capital peers. Therefore, the impact of losing capital status is roughly equivalent to a transition to average prefecture status.

The impact of losing capital status is notable for two reasons. First, it suggests that the variables of concern omitted when determining capital status may not be essential; otherwise, we would observe that losing capital status has little effect. Second, it implies a lack of persistence: former provincial capitals lost a large part of their economic advantages after losing capital status. This finding suggests that the agglomeration effect of economic activity was not large enough to sustain a prefecture's economic advantages even a few decades after losing political economic status. As we discuss below, this finding is related to the ability of politics to shape both public employment and market access.

A Proxy for Migration Ideally, we would quantify the importance of migration as a determinant of population change, as this information could shed light on the importance of the displacement effect. Unfortunately, no sufficient migration flow data exist across our study period. Therefore, to gauge the importance of migration, we further collect information on surname distribution in years corresponding to four regimes (1078, 1580, 1776, and 2000).[20] Specifically, we divide the population into those with existing surnames and those with new surnames. As reported in Appendix D.6, among this sample of prefectures with available surname information, provincial capital status increases the population by 47.3%, similar to our baseline estimates. We further find that individuals with new surnames are responsible for a 11.2% increase in the overall population, while those with existing surnames are responsible for a 36.1% increase. Although this method is certainly subject from measurement error, the results suggest that both migration and demographic growth contribute

[20] The surname information for 1080 is obtained from Harvard's China Biographical Database; that for 1580 and 1776, from datasets on presented scholars (Zhu and Xie 1980); and that for 2000, from the population census of that year.

to population change.

5.3 Additional Checks

Checks of Our DID Analysis We conduct two additional checks to address concerns about omitted variables. First, we exploit data on the 63 ever-capital prefectures; in this group, no omitted variable prevented any one prefecture from becoming a capital. We find similar results when we focus on this subsample or further control for $EverCapital_i \times \gamma_t$ in our analysis (see Appendix D.7).

Second, we conduct a period-by-period analysis in which we define periods relative to the period before a change in capital status occurs (denoted by period 0). This specification allows us to check whether provincial capitals were already on a different developmental trajectory prior to a change in political status. Due to the frequency of data availability, we divide the data into 75-year increments (approximately one quarter of a regime's length). As reported in Appendix D.8, we find no significantly different trends in population density between the capital group and other prefectures prior to status changes. Figure 5 further depicts the findings; here, the lines connect the estimates and the dotted lines indicate the 95% confidence intervals. As shown, increases and decreases in population density occur only after gaining and losing capital status, respectively. Even though we cannot predict the exact number of years required for these effects to occur, the data indicate that they occur within 75 years after a status change.

We further conduct several measurement checks of our DID analysis. We show that our results are robust to using different periods of data (Appendix D.9) or using grid-level data in the analyses (Appendix D.10).

Checks of Our IV analysis For our IV analysis, we assume that the change in a prefecture's hierarchical distance is orthogonal to its own characteristics because this change stems from a regime change. This assumption can be confirmed by two sets of empirical results discussed in Appendix D.11, where we show that ($\Delta\ln RankHierDist_{i,t}$) is not significantly correlated with previous log population density levels ($\text{lnPopDensity}_{i,t-1}$ and $\text{lnPopDensity}_{i,t-2}$) or previous changes in log population density ($\Delta\text{lnPopDensity}_{i,t-1}$ and $\Delta\text{lnPopDensity}_{i,t-2}$).

We are concerned about the exclusion restriction. The finding that hierarchical distance is not significantly correlated with population density in the never-capital prefectures (Column (7) of Table 2) suggests that if hierarchical distance affects population density, regardless of capital status, we should expect to observe this relationship within the never-capital prefectures. We present three additional sets of analyses to further investigate the validity of the instrument. First, we employ ln Rank in $\sum_{j=1}^{N} A_j D_{i,j,t}$ $\times$ ln Rank in $\sum_{j=1}^{N} A_j D_{i,NatCap,t}$ as an instrument while controlling for the independent effects of the two components. When we use the interaction of these two components as the instrument, we do not find that either alone has an additional direct effect (Appendix D.12). Second, the two-component structure of hierarchical distance also allows us to create multiple placebo hierarchical distance ranks by exploiting changes in national capital status. For instance, we calculate one such placebo to Kaifeng when this location was not a capital, and similar placebos for Nanjing and Beijing before they became national capitals. Including these placebo hierarchical distance ranks does not alter our IV estimate or affect population density, implying that our findings are specific to the political status of these cities (Appendix D.13). Finally, we replace $D_{i,NationalCap,t}$ in equation (2) with $D_{i,Market,t}$, which measures a prefecture's distance to major market centers. Once again, this change cannot explain the role of hierarchical distance (Appendix D.14). Taken together, these results indicate that political status is a crucial channel through which the hierarchical distance rank can affect economic development.

6 Interpretations

We document the great impact of political hierarchy on spatial development, as measured by population density and urbanization. Next, we ask what explains this link between politics and development. As noted in our Introduction, three economic forces theoretically may contribute to this link: the direct effects of public employment, the spillover effects of public employment through private employment and agglomeration, and improvements in market access (mediated by investments in infrastructure). Here, we discuss these interpretations and use the evidence to address a question that has long attracted the interest of social scientists: can one consider these provincial capitals "parasitic cities" that gain success mainly via public employment but have little advantage in productivity?

6.1 Public Employment and Its Spillovers

Public offices naturally moved to new provincial capitals. The historical and sociological discourse often emphasizes that merchants and artisans move to political centers to serve bureaucrats (e.g., Heng 1999). Here, we measure public employment in two years, 1776 and 2000. To obtain historical information on bureaucrats, we digitized a 120-volume set of government records known as the Complete Directory of Qing Officials (Tsinghua Library Collection 2008) and use the information from 1776. A limitation of these records is that they include only individuals who hold offices and, hence, exclude a large number of employees who did not hold formal government positions. To obtain data from 2000, we use information from the population census, which refers to employment in all administrative and social organizations and in public institutions (e.g., physicians and teachers at public hospitals and schools). In contrast, the historical censuses do not include information on occupations. In our analysis, we first look at the cross-sectional correlations in both periods. Then, we use the rank of a prefecture as an outcome of our panel analysis (such that a higher value of rank indicates a smaller size of public employment) based on both the DID and IV methods.

Public offices move with provincial capitals. In Columns (1)–(4) of Table 5, we present our cross-sectional findings on log population and public employment in 1776 and 2000. In Columns (5) and (6), we show that a change in provincial capital status is associated with a change in the public employment rank. The coefficients indicate that the importance of provincial capital status is comparable to an increase in public employment from an average rank to a rank in the top 35–15 levels.

Spillover Effects Public employment can have spillover effects via private employment and agglomeration. Although we cannot directly estimate these spillover effects, it is instructive to conduct back-of-the-envelope analyses to understand how large an effect is needed to explain our main findings on population changes. For instance, using the census data and estimates by other scholars, we estimate that public employment in 2000 accounts for around 10% of the total labor force or 6–7% of the total population.[21] In Column (4), the coefficients of provincial capitals on log public employment and log population are 1.15 and 0.68, implying that public employment alone can explain 0.115 (1.15×10%) of 0.68. We thus need to apply a large multiplier (e.g., approximately 6) to the spillover effects of public employment to explain our findings on regional development. Existing estimates of the spillover effect of public employment on total private employment are less than 1.[22] This exercise suggests that although spillover effects are certainly present, they cannot easily explain all of our findings.

6.2 Market Access

The construction of new infrastructures that connect capitals is another common feature that shapes the market access of a prefecture. Historically, the construction of the Grand Canal exemplifies the role of the state in providing infrastructure that facilitates economic development. In 1992, when the central government decided to construct the national expressway network, the official objective was "to connect all provincial capitals and all cities with a population of at least 500,000" (Faber 2014).

Transportation networks have experienced great changes across regimes for two reasons. First, it is costly to maintain routes. Many land routes disappear due to a lack of maintenance; for example, several parts of the Grand Canal were in severe disrepair for a long period (Brook 1998). Second, when a regime replaces its predecessor, the ruler decides which parts of the transportation networks will be reconstructed, depending on the prefecture's relative importance in the political hierarchy. Even if some of the old routes are kept, a prefecture's centrality in the network is altered by the reconnected or newly built routes. Accordingly, the regime changes in China provide us with a rare opportunity to systematically investigate changes in transportation networks. As we show next,

[21] See the PIIE estimate: https://www.piie.com/blogs/china-economic-watch/shrinking-leviathan-state-employment-china-looms-smaller-expected. We also check the public employment shares in capital prefectures and non-capital prefectures, which were 10.3% and 5.7%, respectively, in 2000. Even if we were to use this difference to conduct a back-of-the-envelope analysis, we would still need a large multiplier for this channel alone to explain our main finding on population changes.

[22] Utilizing the relocation of the German federal government from Berlin to Bonn in the wake of the Second World War, for example, Becker, Heblich, and Sturm (2018) find that Bonn experienced a substantial increase in public employment but only modest increases in private sector employment. Using English data at the Local Authority level for 2003â2007, Faggio and Overman (2014) find that public sector employment has no identifiable effect on total private sector employment.

transportation networks are *endogenous* to political hierarchy, and this property underlies their changes over time.

Historical transportation networks comprised the Grand Canal, which connected various waterways, and a state courier system, which was supported by many post offices. Both systems were designed by the central government. A transportation network has a primarily political aim: to maintain an adequate flow of the information, revenues, and personnel on which the state relies (Brook 1998). Historically, commoners have also been allowed to use these routes (Zang 1997). Thus, by facilitating the movement of goods and people, these networks have contributed to economic development. Modern transportation is much more complicated than its historical counterpart. We focus on railroad networks because they comprise a state-owned monopoly.

To trace how transportation networks evolve over time, we digitize roads and waterway maps corresponding to three historical periods (represented by specific years)—the Song (1078), Ming (1587), and Qing (1820) dynasties—and a railroad map from the People's Republic (1990).[23] See Appendix E.1 for the maps by regime.

To account for the relative importance of different links in the transportation network, we apply a centrality measure used in the literature; this is defined for each prefecture i as follows:

$$Centrality_i = \sum_{j \neq i} \frac{1}{d_{i,j}} = \sum_{j \neq i} \frac{1}{d_{i,N_i} + d_{j,N_j} + (1/\theta) d_{N_i,N_j}}, \tag{8}$$

where $d_{i,j}$ indicates the shortest distance between i and j in a transportation network. The distances can also be weighted by population to obtain a measure similar to market access (for further discussion, see Donaldson and Hornbeck 2016). Here, we focus on the unweighted measure to highlight a region's position in a transportation network.

In practice, $d_{i,j}$ comprises three parts: d_{i,N_i} and d_{j,N_j} indicate the straight-line distances from prefecture i and prefecture j to the network, respectively (points N_i and N_j, respectively); d_{N_i,N_j} indicates the minimum distance between point N_i and point N_j within the network. Following the literature, we allow an adjustment factor of 1.5 between the shortest straight-line distance and the

[23]The historical maps are collected in the Historical Atlas of China (Cheng and Hsu 1980), which covers the major routes in different periods. The railroad map in 1990 is obtained from CHGIS (2007).

kilometers traveled and assume that the transportation cost is four times higher without the network; i.e., θ takes a value of 6. As in the literature, our results are not sensitive to the choice of θ.

A prefecture's centrality in a transportation network is somewhat persistent. However, we aim to test whether changes in capital status lead to changes in spatial centrality. Similar to our analysis of public employment, we first examine the relationship between capital status and centrality in each regime. Then, we use the rank of centrality as a dependent variable to build a panel dataset with which to study the impact of a change in capital status. Again, a higher rank value indicates lower centrality.

As shown in Columns (1)–(4) of Table 6, in each regime, provincial capital status is strongly associated with higher centrality. Columns (5) and (6) report the DID and IV estimates. Again, a change in provincial capital status is strongly correlated with a change in spatial centrality rank. These coefficients suggest that the importance of provincial capital status is comparable to moving from an average prefecture rank to a rank within the top 60–40 in terms of centrality.

One may wonder whether the selection of a capital depends on its level of access to a recently improved transportation network. Historically, capital status changes occurred in the early periods of dynasties (almost immediately after the foundation), whereas transportation networks were reconstructed after the government had accumulated more resources. Empirically, we examine the pre-trends in transportation centrality and observe no significant differences before a capital status change. Following the literature on market access, we calculate market access while considering topography and similarly find no pre-trends in terms of market access. These results are discussed in Appendix E.2.

Heterogeneous Effects and Spatial Spillovers Our findings regarding changes in transportation networks have implications in terms of heterogeneous effects and spatial spillovers. First, a change in economic development in response to a change in provincial capital status may vary by the position of the prefecture in the transportation network. Therefore, we examine the impact of provincial capital status on population by the lagged centrality of a prefecture. We find that provincial status has a larger impact on prefectures with a lower (lagged) centrality (i.e., a higher value of the rank).

In other words, politics appears to be more crucial for prefectures with lower positions in the transportation network (Panel A of Table 7).

Second, the neighboring prefectures of the provincial capitals benefit by gaining better market access, especially if they connect provincial capitals to national capitals.[24] To check this spatial spillover, we divide the neighboring prefectures into two groups: (1) those that connect the provincial capital to the national capital and (2) others. We find that both groups benefit and that the impact on the former group is twice that on the latter group (Panel B of Table 7). This result suggests that our main finding on changes in population in response to changes in capital status cannot be fully explained by a displacement effect, which would tend to predict that neighboring prefectures would lose population to the capital prefecture.

In summary, we find that political hierarchy shapes economic development not only via public employment but also via important infrastructure, such as transportation networks.

6.3 A Counterfactual Scenario of Political Decentralization

Because market access is shaped by transportation networks, our findings have implications for provincial market access. In particular, what would happen if a provincial capital were set up without following the political logic documented in Section 3? Recall that the choice of provincial capital location reflects the logic of a centralized state that considers both provincial autonomy and central control. Our political economy counterfactual of interest is a decentralized scenario affected by only provincial autonomy. Therefore, we conduct a simple analysis by comparing provincial market access in the current political system with that in a scenario in which provincial capitals are set in the provincial centroids. In this exercise, we assume that all prefectures in a province can still reach the other provinces and the only change is the position of the provincial capital.

We define the national market access of prefecture i in province p at time t as $MA_{pi,t} = \sum_{j \neq pi} \frac{M_{j,t}}{d_{pi,j,t}}$, where the market size (M_j, t) is measured by its population . For the provincial level measure $(MA_{p,t})$, we use the average market access of all prefectures in a province p, namely $\sum_i MA_{pi,t}/N$,

[24] This exercise is similar to the research of Morten and Oliveira (2018), who study how an exogenous shock to highways due to relocating the national capital to Brasilia reduced trade and migration costs.

in which N is the total number of prefectures in the province. We then examine how the deviation of the provincial capital from the provincial centroid affects provincial market access. We find that if the provincial capital is located 100 kilometers from the provincial centroid, the provincial aggregate market access declines by 8.6% (Appendix E.3). In our data, the average deviation is 125 kilometers, suggesting that on average, provincial aggregate market access could be 10.8% higher.

7 Conclusion

Administrative hierarchy is a complex issue, and it is difficult to find a suitable empirical setting to investigate its underlying logic and evaluate its consequences. In this study, we take advantage of China's unique historical and political setting to show how political hierarchy shaped national spatial development over a millennium. The mapping of political hierarchy to economic hierarchy remains obvious today: in 25 of 27 provinces, the provincial capital is the largest city in the province.

Broadly, our study reveals the importance of politics as a determinant of economic geography. Underlying our findings is a strong centralized state that provides important public goods, such as transportation, that facilitate changes in the locations of economic activity. However, a strong centralized state also can be costly. We do not observe that economic advantages persist several decades after a prefecture loses its political status. In addition, a more decentralized approach to selecting administrative centers may increase regional market access. These findings highlight the cost side of a centralized political system.

Economists have made considerable progress in understanding the persistence of economic activities, and scholars have observed many changes in this area over the long run. Our findings suggest that introducing the role of the state to this body of work may increase understanding of both persistence and changes. Admittedly, China is a special case, as the state has had both the capacity and the incentive to redefine administrative landscapes across regimes, leading to changes in the spatial distribution of economic activities. When applying what we have learned in this setting to other contexts, it would be worthwhile to consider both the capacity and the incentives of the state.

References

[1] Ades, Alberto F., and Edward L. Glaeser (1995), "Trade and Circuses: Explaining Urban Giants," *Quarterly Journal of Economics* 110(1): 195-227.

[2] Alesina, Alberto, and Enrico Spolaore (1997), "On the Number and Size of Nations," *Quarterly Journal of Economics* 112(4): 1027-1056.

[3] Barjamovic, Gojko, Thomas Chaney, Kerem A. Cosar, and Ali Hortacsu (2019), "Trade, Merchants, and the Lost Cities of the Bronze Age," *Quarterly Journal of Economics* 13(3): 1455-1503.

[4] Bairoch, Paul (1985) [1991], *Cities and Economic Development: from the Dawn of History to the Present*. University of Chicago Press.

[5] Becker, Sascha, Stephan Heblich, and Daniel M. Sturm (2018), The Impact of Public Employment: Evidence from Bonn," Working paper.

[6] Blanchard, Olivier, and Andrei Shleifer (2001), "Federalism with and without Political Centralization: China versus Russia," *IMF Staff Papers* 48(1): 171-179.

[7] Bleakley, Hoyt, and Jeffrey Lin (2012), "Portage and Path Dependence," *Quarterly Journal of Economics* 127(2): 587-644.

[8] Bluhm, Richard, Christian Lessmann, and Paul Schaudt (2020), "The Political Geography of Cities," Working paper.

[9] Bolton, Patrick, and Gerard Roland (1997), "The Breakup of Nations: a Political Economy Analysis," *Quarterly Journal of Economics* 112(4): 1057-1090.

[10] Bosker, Maarten, Eltjo Buringh, and Jan Luiten van Zanden (2013), "From Baghdad to London: Unraveling Urban Development in Europe, the Middle East, and North Africa, 800-1800," *Review of Economics and Statistics* 95(4): 1418-1437.

[11] Brook, Tim (1998), "Communications and Commerce," *The Cambridge History of China*, 8(Pt 2), 579-707.

[12] Campante Filipe and Do Quoc-Anh (2014), "Isolated Capital Cities, Accountability, and Corruption: Evidence from US States," *American Economic Review* 104 (8): 2456-2481

[13] Campante, Filipe R., Quoc-Anh Do, and Bernardo Guimaraes (2019), "Capital Cities, Conflict, and Misgovernance," *American Economic Journal: Applied Economics* 11(3): 298-337.

[14] Chao, Kang (2006), *Essays on the History of Chinese Urban Development (Zhongguo chengshi fazhanshi lujin)*. Beijing, New Star Press.

[15] Cheng, Kuang-Yu, and Sheng-Mo Hsu (1980), *Historical Atlas of China (Zhongguo Lishi Ditu)*, Taiwan: Chinese Culture University Press.

[16] CHGIS (2007), Harvard Yenching Insitute.

[17] Davis Donald and David Weinstein (2002), “Bones, Bombs, and Break Points: The Geography of Economic Activity,” *American Economic Review* 92(5):1269-1289.

[18] Davis, James, and Vernon Henderson (2003),“Evidence on the Political Economy of the Urbanization Process,” *Journal of Urban Economics* 53(1): 98-125.

[19] De Long, J. Bradford, and Andrei Shleifer (1993), “Princes and Merchants: European City Growth before the Industrial Revolution,” *Journal of Law and Economics* 36(2): 671-702.

[20] Dell, Melissa, Nathan Lane, and Pablo Querubin (2018), “The Historical State, Local Collective action, and Economic Dvelopment in Vietnam," *Econometrica* 86(6): 2083-2121.

[21] Donaldson, Dave, and Richard Hornbeck (2016), “Railroads and American Economic Growth: A ‘Market Access’ Approach,” *Quarterly Journal of Economics* 131(2): 799-858.

[22] Duben, Christian, and Melanie Krause (2020), “The Emperor’s Geography-City Locations, Nature, and Institutional Optimization,” working paper.

[23] Faber, Benjamin (2014), “Trade Integration, Market Size, and Industrialization: Evidence from China’s National Trunk Highway System,” *Review of Economic Studies* 81(3): 1046-1070.

[24] Faggio, Giulia, and Henry Overman (2014), “The Effect of Public Sector Employment on Local Labour Markets," *Journal of Urban Economics*, 79: 91-107.

[25] FAO (2012), GAEZ (Global AgroEcological Zones), http://fao.org/Ag/AGL/agll/gaez/ index.htm.

[26] Ge, Jianxiong (2000), *History of Population in China*. Shanghai: Fudan University Press.

[27] Glaeser, Edward and Joshua D. Gottlieb (2008), “The Economics of Place-making Policies,” *Brookings Papers on Economic Activity* 39(1): 155-253.

[28] Grossman, Guy, and Janet I. Lewis(2014), “Administrative Unit Proliferation,” *American Political Science Review* 108(1): 196-217.

[29] Guy, Kent (2010), *Qing Governors and Their Provinces: the Evolution of Territorial Administration in China, 1644-1796*. University of Washington Press.

[30] Huillery, Elise (2009), “History Matters: The Long-Term Impact of Colonial Public Investments in French West Africa,” *American Economic Journal: Applied Economics*, 1(2): 176-215.

[31] Jedwab, Remi, Edward Kerby, and Alexander Moradi (2017), “History, Path Dependence and Development: Evidence from Colonial Railways, Settlers and Cities in Kenya,” *Economic Journal*, 127(603): 1467-1494.

[32] Heng, Chye Kiang (1999), *Cities of Aristocrats and Bureaucrats: the Development of Medieval Chinese Cityscapes*. NUS Press.

[33] Ko, Chiu Yu, Mark Koyama, and Tuan?Hwee Sng (2018), “Unified China and Divided Europe,” *International Economic Review* 59(1): 285-327.

[34] Liang, Fang-zhong (1980), *Historical Statistics on Hukou, Land and Land Tax of China (Lidai Hukou, Tudi,Tianfu*

Tongji). Shanghai: Zhonghua Book Company.

[35] Ma, Debin, and Jared Rubin (2019), "The Paradox of Power: Principal-agent Problems and Administrative Capacity in Imperial China (and Other Absolutist Regimes)," *Journal of Comparative Economics* 47(2): 277–294.

[36] Michalopoulos, Stelios, and Elias Papaioannou (2013), "Pre-Colonial Ethnic Institutions and Contemporary African Development,"*Econometrica* 81(1): 113-152.

[37] Morten, Melanie and Jaqueline Oliveira (2018), "The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City," Working Paper.

[38] Mostern, Ruth (2011), *Dividing the Realm in Order to Govern: The Spatial Organization of the Song State (960-1276 CE)*. Harvard University Press.

[39] Nunn, Nathan (2009), "The Importance of History for Economic Development," *Annual Review of Economics* 1(1): 65-92.

[40] Oster, Emily (2017), "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*: 1-18.

[41] Rauch, James (1993), "Does History Matter Only When It Matters Little? The Case of City-Industry Location," *Quarterly Journal of Economics* 108(3): 843-867.

[42] Skinner, G. William, ed. (1977), *The City in Late Imperial China*, Stanford University Press

[43] Sng, Tuan Hwee, Pei Zhi Chia, Chen-Chieh Feng, and Yi-Chen Wang (2018), "Are China's Provincial Boundaries Misaligned?" *Applied Geography* 98: 52-65.

[44] Stasavage, David, "Was Weber Right? The role of Urban Autonomy in Europe's Rise," *American Political Science Review* 108(2): 337-354.

[45] Treisman, Daniel (2007), *The Architecture of Government: Rethinking Political Decentralization.* Cambridge University Press.

[46] Tsinghua Library Collection (2008), *Complete Directory of Qing Officials.* Daxiang Press.

[47] Weber Max (1921)[1958], *The City*. reprint, New York: Free Press.

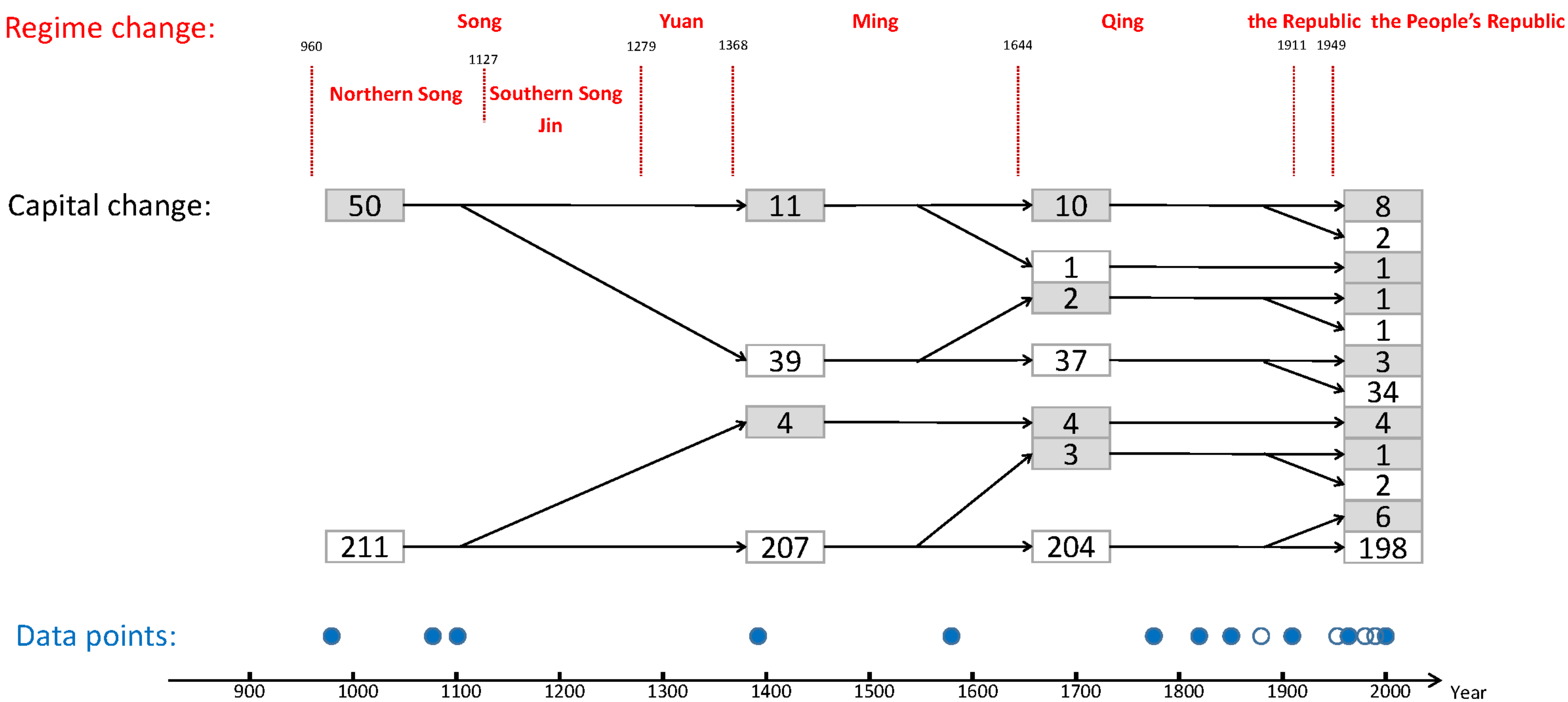
[48] Zang, Rong (1997), *Courier Route and Station in Historical China* [Zhongguo Gudai De Yizhan Yu Youchuan]. The Commercial Press.

[49] Zhang, Tingyu (1739)[1974], *The History of the Ming Dynasty (Ming Shi)*. Qianlong edition, reprint, Beijing: Zhonghua shuju.

[50] Zhou, Zhenhe (1998), *The Changes of Administrative Division in China (Zhongguo lidai xingzheng quhua de bianqian)*. Beijing: Shangwu chubanshe.

[51] Zhu, Baojiong, and Peilin Xie (Ed.) (1980), *Index of Names of Jinshi Graduates in the Ming and Qing Periods (Ming Qing jinshi timing beilu suoyin)*, Shanghai: Shanghai guji.

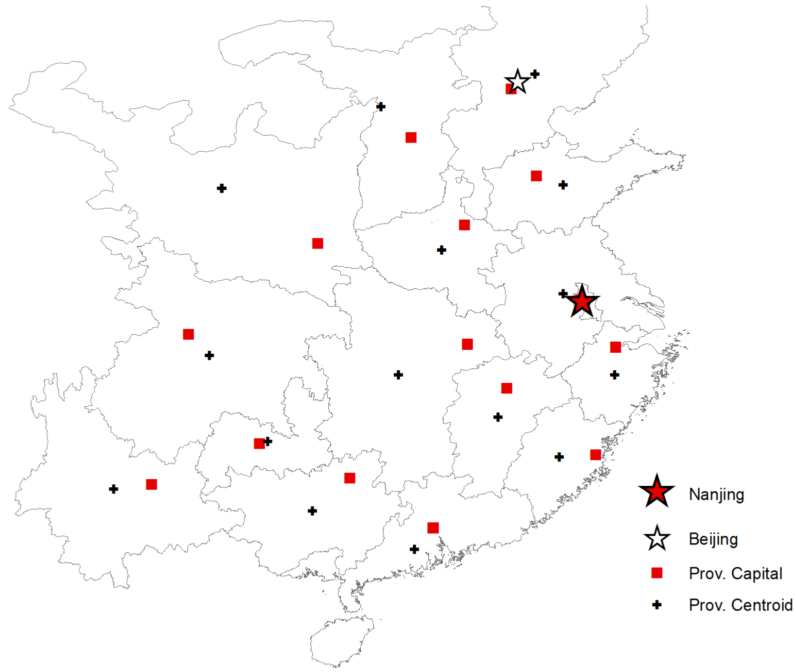
Figure 1: Changes in Provincial Capitals across Regimes



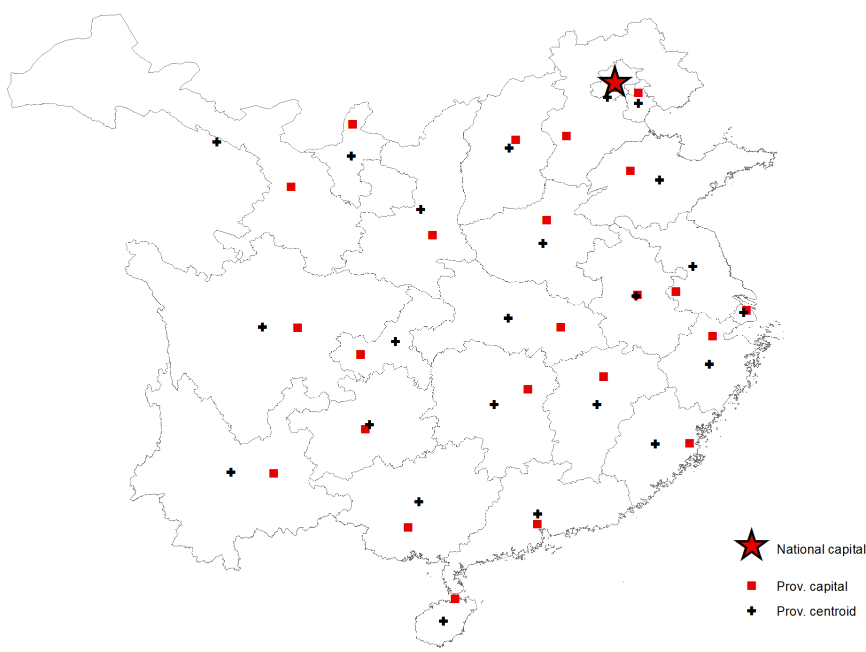
Notes: The grey boxes indicate the provincial capitals and the white box non-capital prefectures. To make the figure more readable, we omit three changes in 1964 but include them in our empirical analyses. This figure highlights the challenge of uneven gaps in the census data, which we address in several ways in the paper. For instance, in our baseline, we exclude four census years (indicated by the empty circles) to mitigate the uneven gaps.

Figure 2: Illustrating the Idea Underlying Hierarchical Distance

(a) Ming



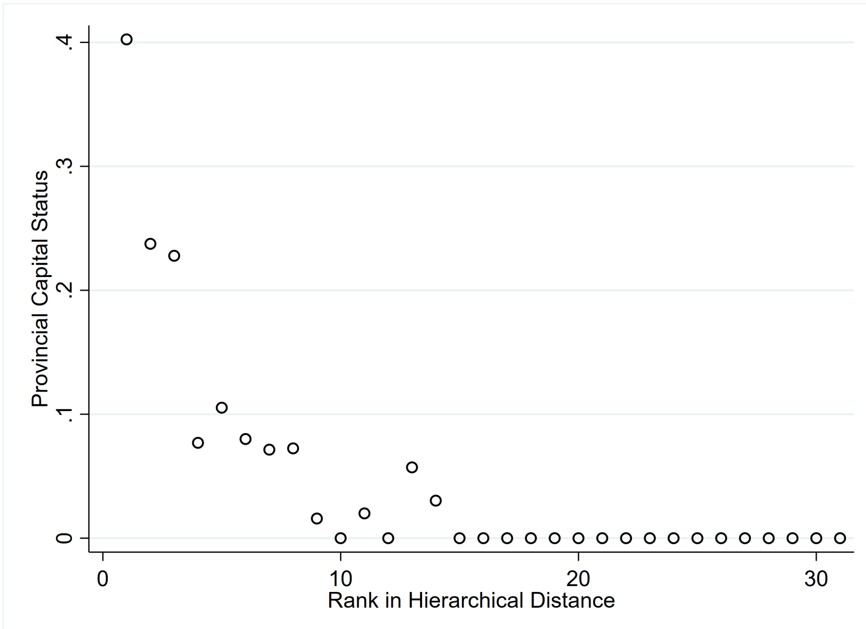
(b) Modern China



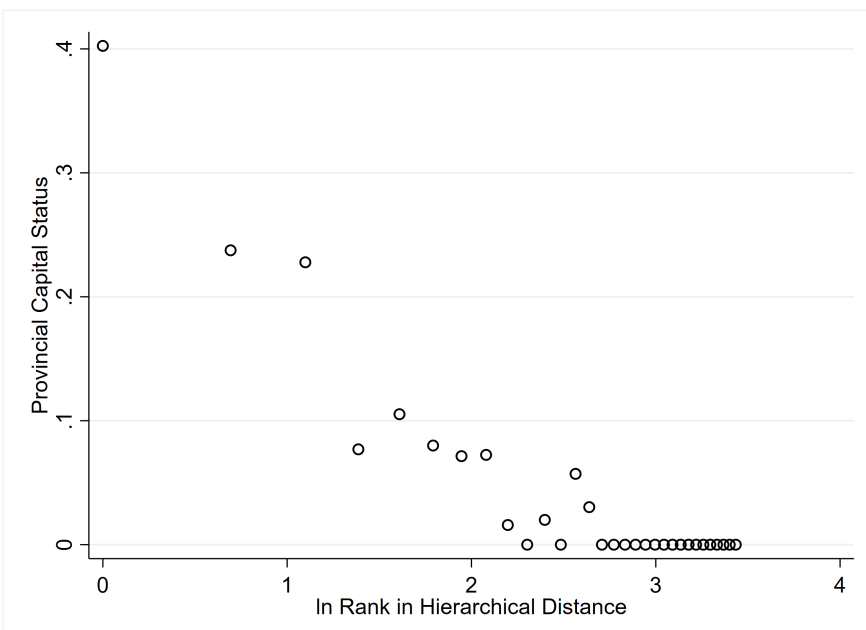
Notes: This figure illustrates the logic underlying the hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province) using the case in 1580 and 2000: provincial capitals tend to be located away from provincial centroid toward the national capital. We present the maps for other regimes in Appendix C.1.

Figure 3: Rank in Hierarchical Distance and the Probability of Being a Provincial Capital

(a) Rank in Hierarchical Distance vs. Prob. of Being Capital



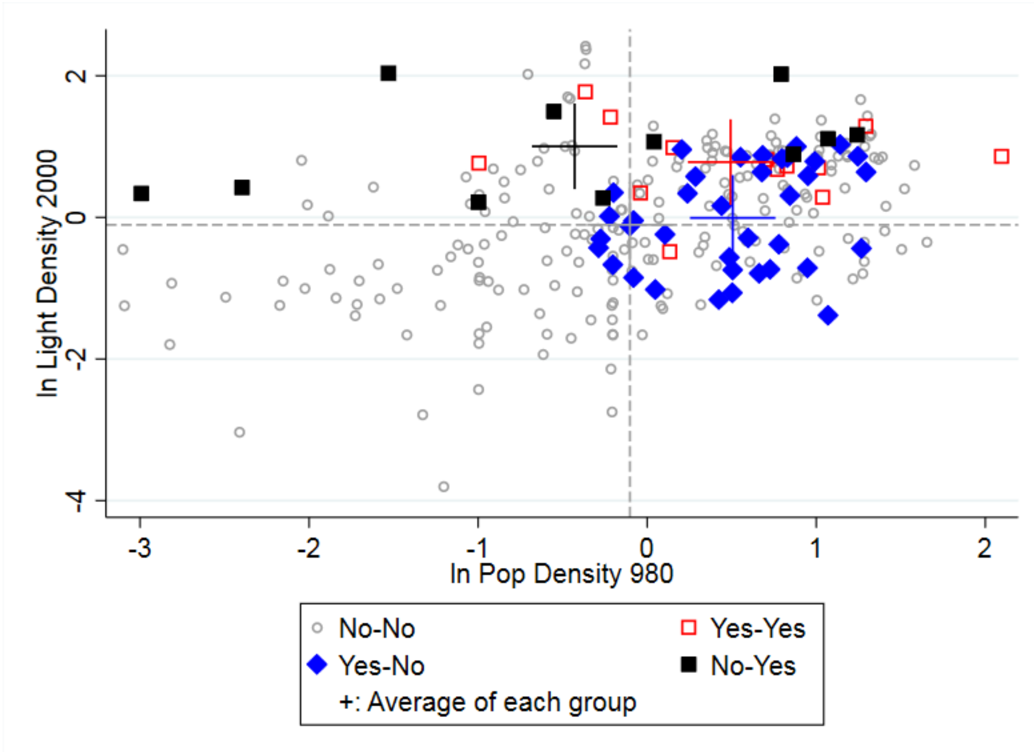
(b) ln Rank in Hierarchical Distance vs. Prob. of Being Capital



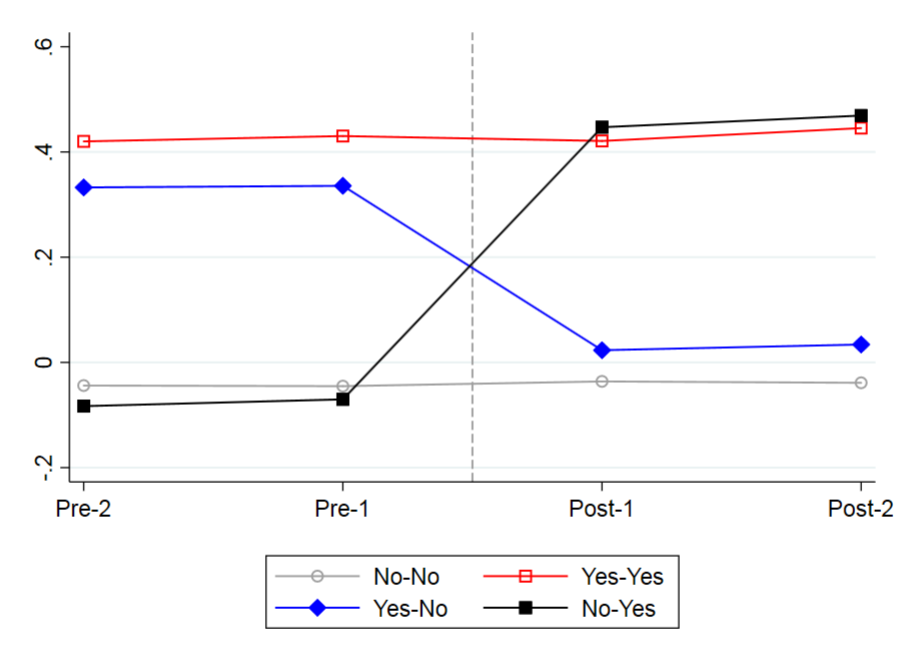
Notes: The figure shows a strong negative correlation between a prefecture's rank in hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province) within a province and its probability of being a provincial capital.

Figure 4: Descriptive Patterns

(a) 980 vs. 2000



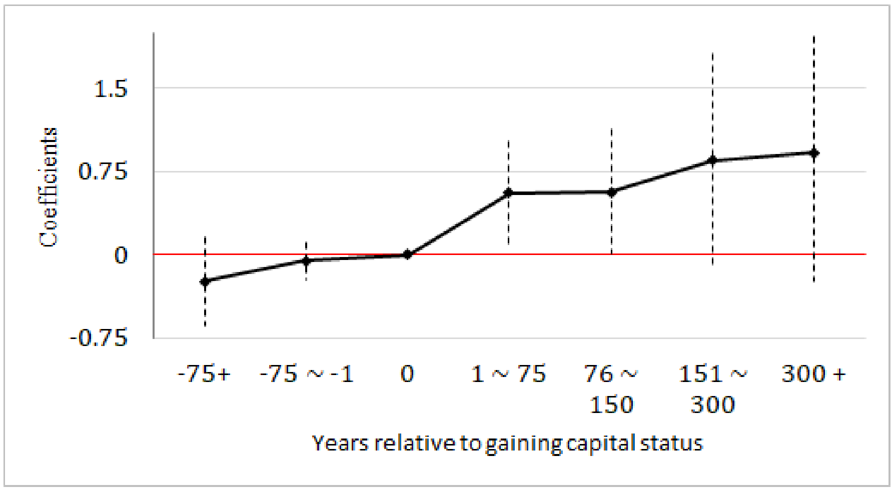
(b) Pooling all changes together 980-2000



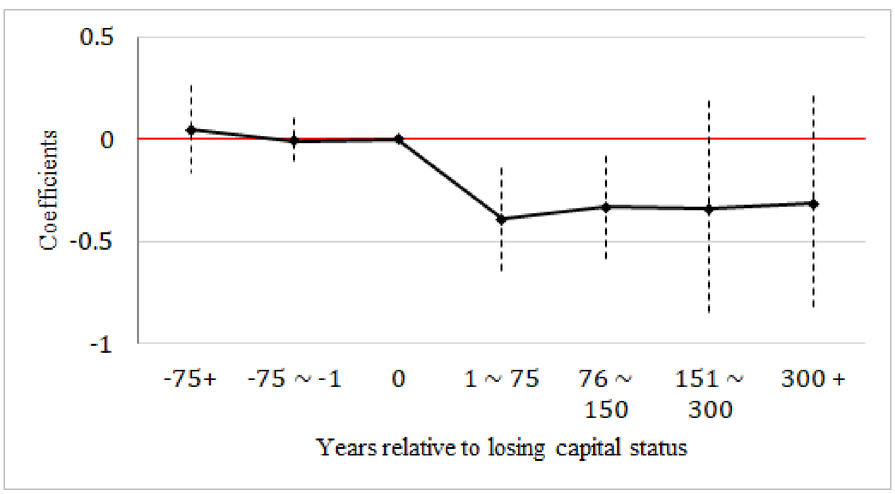
Notes: Panel (a) shows that gaining and losing capital status is associated with a systematic change in a prefecture's economic performance. The crosses indicate the averages for each group. Note that the "no-no" group averages are indicated by the big dashed cross.
In Panel (b), the four groups are defined based on the capital status every two regimes. To increase comparability, for each period, we first regress log population density on prefecture characteristics (geography variables, agricultural variables, and macroregion dummies) and obtain residuals. We plot the average residuals by group and by period.

Figure 5: The Impact of Capital Status on Population Density: Period by Period

(a) Before and After Gaining Capital Status



(b) Before and After Losing Capital Status



Notes: This figure visualizes the results in Appendix D.8. It shows that there are no systematic differences before and after a prefecture gains or loses capital status. The reference group is the period before capital status changes. Due to the low frequency of data, we cannot pin down the exact years it takes for the change to happen, but the patterns tell us that the rise and decline happen within 75 years after status change. A period of 75 years is roughly one-fourth of a regime's length.

Table 1: Summary Statistics

	Sources	Obs.	Mean	S.D.
(a) Economic development proxies				
ln Population density	1, 2, 3, 4	2,871	4.06	1.76
Urbanization ratio (%)	3, 4	1,044	15.45	15.89
(b) Time-varying provincial capital status				
Provincial capital	2, 5	2,871	0.10	0.31
(c) Hierarchical distance				
Rank in Hierarchical distance	2	2,871	8.22	5.78
ln Rank in Hierarchical distance	2	2,871	1.81	0.84
(d) Time-invariant prefecture characteristics				
ln Area	2	261	9.30	0.85
Whether a prefecture contains a plain	2	261	0.70	0.46
Whether a prefecture contains a major river	2	261	0.72	0.45
Whether it is on the coast	2	261	0.21	0.41
Slope	2	261	2.48	2.09
ln Elevation	2	261	5.52	1.60
Longitude	2	261	112.08	5.88
Latitude	2	261	30.63	5.18
Wheat suitability	6	261	3.95	1.04
Rice suitability	6	261	3.04	1.08
Fox millet suitability	6	261	3.74	1.45
Maize suitability	6	261	4.49	1.05
Sweet potato suitability	6	261	3.53	0.95
(e) Other outcomes				
Public employment				
ln Bureaucrats in 1776	7	261	3.57	0.62
ln Total employment in government/party and public institutions in 2000	8	261	9.25	0.69
Transportation networks				
Centrality in the transportation networks, θ=6, in 1078	2, 9	261	0.90	0.29
Centrality in the transportation networks, θ=6, in 1580	2, 9	261	0.99	0.22
Centrality in the transportation networks, θ=6, in 1776	2, 9	261	1.01	0.27
Centrality in the transportation networks, θ=6, in 1990	2, 9	261	0.80	0.21

Sources:
1. Liang (1980), Historical Statistics of Population, Land and Taxation in China;
2. China Historical GIS (2007), http://chgis.fas.harvard.edu/.;
3. Ge Jianxiong (2000), China Population History;
4. Population Census 1953, 1964, 1982, 1990, 2000, accessible via the library of CUHK;
5. Treatise of the Nine Regions from the Yuanfeng reign (1078-1085), https://ctext.org/wiki.pl?if=gb&res=508139;
6. FAO GAEZ (2012), http://fao.org/Ag/AGL/agll/gaez/ index.htm;
7. Tsinghua Library Collection (2008), The Complete Directory of Qing Officials;
8. Census 2000 (micro data), accessible via the library of CUHK;
9. Cheng and Hsu (1980), Historical Atlas of China.

Table 2: Economic Consequences – The Impact of Capital Status on Population Density

method	DID				Reduced form			IV	
sample	All	All	All	All	All	Ever-capital	Never-capital	All	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
								Second-stage	
Prov. capital	0.622***	0.456***	0.488***	0.470***				0.796***	0.794***
	(0.132)	(0.104)	(0.105)	(0.104)				(0.229)	(0.165)
								First-stage	
H-distance					-0.071***	-0.131***	-0.017	-0.089***	-0.101***
					(0.021)	(0.038)	(0.027)	(0.008)	(0.011)
Pref. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Geography		Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Agriculture			Y	Y	Y	Y	Y	Y	Y
Year FE * Region FE				Y	Y	Y	Y	Y	Y
Year FE * Prov$_{i,r}$ FE									Y
Observations	2,871	2,871	2,871	2,871	2,871	693	2,178	2,871	2,871
R-squared	0.757	0.850	0.865	0.882	0.878	0.924	0.881	0.880	0.957
# prefectures	261	261	261	261	261	63	198	261	261
Weak instrument test									
(F statistic)								131.9	87.6
Lower bound estimate (Oster 2017)				0.327					

Notes: This table shows that provincial capital status is associated with a higher population density. Columns (5)-(9) use ln Rank in hierarchical distance to predict provincial capital status. Column (7) shows that hierarchical distance does not affect the development of never-capital prefectures, suggesting that the political channel is critical for hierarchical distance to matter.

We add controls step by step: (i) geography variables include log area, whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude; (ii) agriculture variables refer to the suitability of rice, wheat, millet, sweet potatoes and maize; and (iii) regions refer to the 9-physiographic macroregions defined by Skinner (1977) (mapped in Figure B.2). Prov$_{i,r}$ indicates province fixed effects in all the regimes.

Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 3: The Impact of Capital Status on Urbanization

	Urbanization rate (%)			ln Urban pop			ln Rural pop		
method	DID	IV		DID	IV		DID	IV	
		First-stage	Second-stage		First-stage	Second-stage		First-stage	Second-stage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prov. capital	11.347***		18.519**	0.618***		1.983**	0.195*		0.380
	(2.549)		(9.036)	(0.141)		(0.820)	(0.103)		(0.512)
H-distance		0.055***			0.055***			0.055***	
		(0.010)			(0.010)			(0.010)	
Pref. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,044	1,044	1,044	1,044	1,044	1,044	1,044	1,044	1,044
R-squared	0.838	0.110	0.835	0.903	0.110	0.893	0.866	0.110	0.865
# prefectures	261	261	261	261	261	261	261	261	261
Weak instrument test		28.5			28.5			28.5	

Notes: This table shows that provincial capital status is associated with a higher urbanization rate.
We add controls step by step: (i) geography variables include log area, whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude; (ii) agriculture variables refer to the suitability of rice, wheat, millet, sweet potatoes and maize; and (iii) regions refer to the 9-physiographic macroregions defined by Skinner (1977) (mapped in Figure B.2). $\text{Prov}_{i,r}$ indicates province fixed effects in all the regimes. Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 4: Gaining vs. Losing Status

	Δ ln Pop Density				Δ ln Pop Density			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ prov. cap	0.431***	0.415***	0.334***	0.305***				
	(0.106)	(0.110)	(0.086)	(0.080)				
Gaining capital status					0.513**	0.465**	0.425**	0.385***
					(0.221)	(0.230)	(0.165)	(0.148)
Losing capital status					-0.390***	-0.389***	-0.287***	-0.263**
					(0.117)	(0.124)	(0.105)	(0.102)
Lag. ln pop. density			-0.428***	-0.425***			-0.428***	-0.425***
			(0.015)	(0.017)			(0.015)	(0.017)
Lag 2 ln pop. density				-0.096***				-0.096***
				(0.020)				(0.020)
Pref. FE		Y	Y	Y		Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Controls	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,610	2,610	2,610	2,610	2,610	2,610	2,610	2,610
R-squared	0.715	0.715	0.789	0.800	0.715	0.715	0.789	0.800
# prefectures	261	261	261	261	261	261	261	261

Notes: This table reports the impacts of gaining status and losing status, both of which matters.
Controls include (i) geography variables: log area, whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude; (ii) agriculture variables: the suitability of rice, wheat, millet, sweet potatoes and maize; and (iii) the 9-physiographic macroregions defined by Skinner (1977) (mapped in Figure B.2).
Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 5: Provincial Capital Status and Public Offices

	(1)	(2)	(3)	(4)	(5)	(6)
	ln Pop	ln Bureaucrats	ln Pop	ln Public Employ.	ln Rank in Public Offices	
	1776	1776	2000	2000		IV
Prov. capital	0.465***	0.477***	0.680***	1.149***	-1.090***	-2.094***
	(0.087)	(0.075)	(0.082)	(0.088)	(0.319)	(0.713)
Controls	Y	Y	Y	Y	Y	Y
Year FE * Controls					Y	Y
Pref. FE					Y	Y
Year FE					Y	Y
Observations	261	261	261	261	522	522
R-squared	0.757	0.623	0.710	0.582	0.414	0.365
# prefectures	261	261	261	261	261	261
First-stage F-statistic						25.8

Notes: This table shows that public offices are affected by capital status. Columns (1)-(4) present cross-sectional results; Columns (5)-(6) employ the panel structure by using ln Rank in public employment as the dependent variable – the negative coefficients indicate climbing up the ladder of public employment.

Controls include (i) geography variables: log area, whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude; (ii) agriculture variables: the suitability of rice, wheat, millet, sweet potatoes and maize; and (iii) the 9-physiographic macroregions defined by Skinner (1977) (mapped in Figure B.2).

Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 6: Provincial Capital Status and Spatial Centrality

	(1)	(2)	(3)	(4)	(5)	(6)
	Centrality				ln rank in centrality	
	1078	1580	1776	1990		IV
Prov. capital	0.401***	0.575***	0.486***	0.513***	-0.149**	-0.669**
	(0.076)	(0.099)	(0.098)	(0.150)	(0.066)	(0.322)
Controls	Y	Y	Y	Y	Y	Y
Year FE * Controls					Y	Y
Pref. FE					Y	Y
Year FE					Y	Y
Observations	261	261	261	261	1,044	1,044
R-squared	0.778	0.685	0.764	0.624	0.401	0.354
# prefectures	261	261	261	261	261	261
First-stage F-statistic						35.0

Notes: This table shows that provincial capital status affects the spatial centrality of a prefecture. Columns (1)-(4) present cross-sectional results across four regimes; Columns (5)-(6) employ the panel structure by using ln Rank in centrality as the dependent variable – the negative coefficients indicate climbing up the ladder of centrality.

Controls include (i) geography variables: log area, whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude; (ii) agriculture variables: the suitability of rice, wheat, millet, sweet potatoes and maize; and (iii) the 9-physiographic macroregions defined by Skinner (1977) (mapped in Figure B.2).

Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 7: Heterogeneity and Spatial Spillover

A. Heterogeneity by centrality	(1)	(2)	(3)
	ln Pop. den	Δln Pop. den	Δln Pop. den
Prov. capital	0.627***		
	(0.141)		
Prov. capital * ln Rank in centrality (lagged)	0.108**		
	(0.046)		
ΔProv. capital		0.548***	
		(0.147)	
ΔProv. capital * ln Rank in centrality (lagged)		0.239**	
		(0.097)	
Gaining capital status			0.726**
			(0.324)
Losing capital status			-0.491***
			(0.170)
Gaining capital status * ln Rank in centrality (lagged)			0.226*
			(0.116)
Losing capital status * ln Rank in centrality (lagged)			-0.262*
			(0.146)
ln Rank in centrality (lagged)	-0.064**	0.089*	0.091*
	(0.029)	(0.050)	(0.051)
All controls	Y	Y	Y
Observations	1,012	1,012	1,012
B. Spatial spillover	ln Pop. den	ln Pop. den	Δln Pop. den
Prov. capital	0.473***	0.469***	
	(0.106)	(0.105)	
Neighbor	0.166*		
	(0.086)		
Neighbor * Connecting natl. capital		0.265**	
		(0.123)	
Neighbor * Others		0.144*	
		(0.087)	
Gaining capital status			0.560**
			(0.221)
Losing capital status			-0.391***
			(0.115)
Gaining neighbor status (connecting natl. capital)			0.362**
			(0.150)
Losing neighbor status (connecting natl. capital)			-0.281**
			(0.117)
Gaining neighbor status (others)			0.177*
			(0.101)
Losing neighbor status (others)			-0.120
			(0.094)
All controls	Y	Y	Y
Observations	2,871	2,871	2,610

Notes: A higher value of rank indicates *lower* centrality. All controls in the previous tables are included. Standard errors are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Appendix

Table of Contents

A Additional Information on the Background

A.1 Ever-Capital Prefectures

Figure A.1: Ever-Capital Prefectures

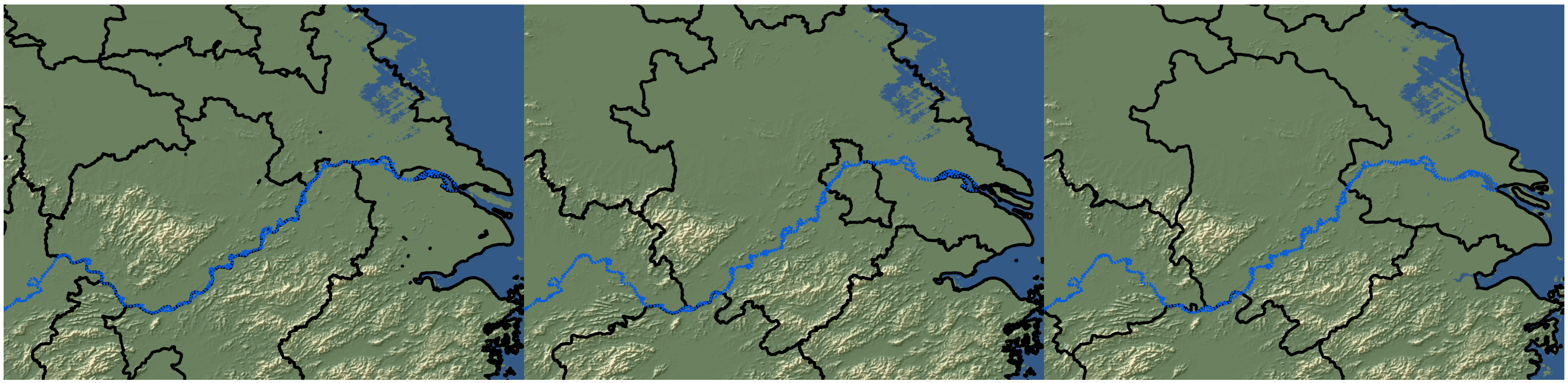
Notes: The shaded area indicates the prefectures in China proper (our sample). The dotted prefectures have ever been a provincial capital at least once during 1078-2000. This map is based on the prefecture boundaries in 2000: 63 out of 261 prefectures have ever been a provincial capital.

A.2 Two Principles to Define Provincial Boundary: An Example

Figure A.2 presents an example where changes in provincial boundaries were driven by the evolution from "*following the natural mountains and rivers*" to "*interlocking like dog's teeth*". As shown, the Yangtze River used as a provincial boundary by the (pre-Mongol) Song but included within provinces by the (post-Mongol) Ming and Qing.

Figure A.2: Two Principles of Provincial Boundary Revision

(a) The Yangtze River: Song　**(b)** The Yangtze River: Ming　**(c)** The Yangtze River: Qing



Similar patterns exist regarding provinces around the Huang River, Qin Mountain, and Taihang Mountain. We compare these provinces of more strategic boundary settings with the rest.

A.3 Examples of Provincial Capital Relocation

Figure A.3 presents two examples of provincial capital relocation. The cross indicates the provincial centroid, the hollow/solid star indicates the past/current national capital, and the hollow/solid square indicates the past/current provincial capital. Luzhou and Changsha were capitals in the Song. Both lost their capital status in the Ming. Changsha regained the capital status in the Qing but Luzhou didn't. These patterns are driven by the relocation of national capitals and redivision of provinces across regimes, which can be captured by our algorithm.

Figure A.3: Examples of Provincial Capital Relocation

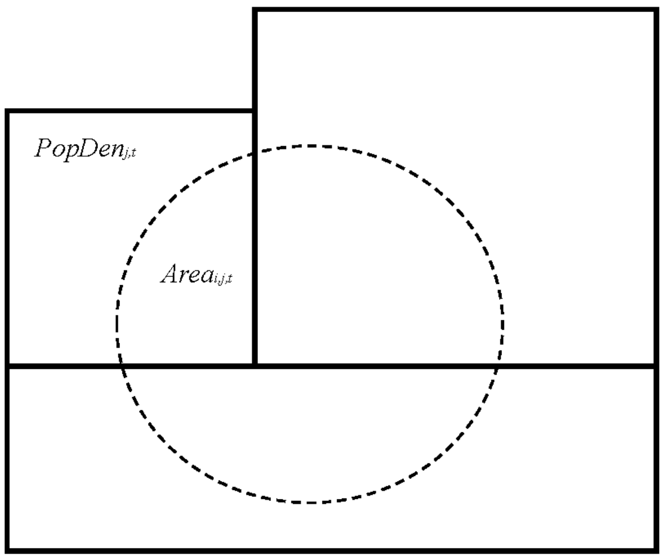
(a) Song **(b)** Ming **(c)** Qing

B Data Construction

B.1 Data on Population and Urbanization

We collect the information on prefecture population in 980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964, and 2000, according to the historical boundaries in each census. To construct a panel dataset, we map the original data to the prefecture boundary in 2000. The idea could be illustrated by Figure B.1.

Figure B.1: Illustration of Population Mapping



Here, the circle (dash line) represents prefecture i in 2000, which is overlapped with J historical prefectures in year t (represented by the rectangular). Prefecture i in 2000 is then divided into J polygons by the historical boundary in year t. The area of each polygon is denoted as $Area_{i,j,t}$. We first calculate the population density in historical prefecture j in year t (each rectangular), denoted by $PopDen_{j,t}$ ($j = 1, 2, ..., J$). Second, we assume the population is evenly distributed in each rectangular and hence calculate the population in each small polygon as $Area_{i,j,t} \times PopDen_{j,t}$. The population density of prefecture i in year t could be thus calculated as:

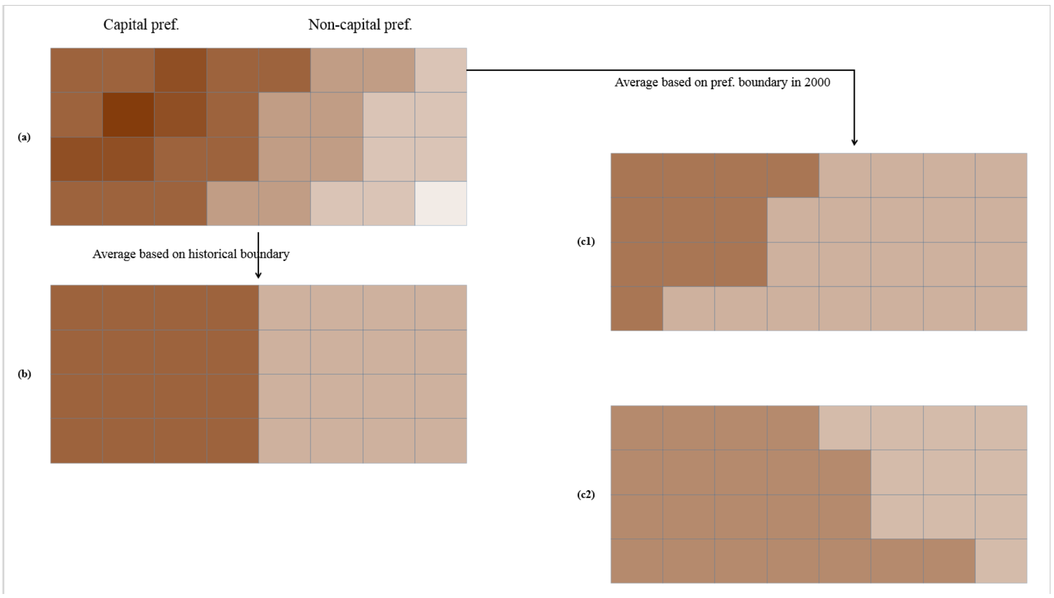
$$PopDen_{i,t} = \frac{\sum_{j=1}^{J} Area_{i,j,t} \times PopDen_{j,t}}{Area_i}$$

To check whether our findings are sensitive with mapping historical original data into the boundary in 2000, we use the same method convert the data into 1-degree X 1-degree grid or 2-degree X 2-degree grid level.

We collect information on urban population in 1580, 1820, 1964 and 2000, and map the original data into the prefecture in 2000 using the same method applied to population data.

If such construction causes measurement error in population density, the measure error may lead to a downward bias of the estimated effect of capital status. We illustrate the reason using Figure B.1.

Figure B.1: Possible Measurement Error



Graph (a) includes two historical prefectures with equal size. The left sixteen grids represent a capital prefecture, while the right sixteen grids represent a non-capital prefecture. The darkness represents population density.[25] Due to the lack of information on population density in each grid, we can only calculate the population density of each historical prefecture (sixteen grids), as graph (b) shows. To construct a panel, we follow the method introduced above to map original data (graph (b)) into the prefecture in 2000 (graph (c1) or (c2)). There are two possible cases. The first is that the capital prefecture was much larger in the history as graph (c1) shows. The non-capital prefecture will include some parts of capital prefecture, and the average population density of the non-capital prefecture should be over-estimated (much darker in graph (c1) than (b)). The second is that the capital prefecture was much smaller in the history as graph (c2) shows. The capital prefecture then includes some grids of non-capital prefecture, and the average population density of the capital prefecture should be under-estimated (much lighter in graph (c2) than (b)). In both cases, the difference in population densities between the capital-prefecture and the non-capital prefectures (in graph (c1) or (c2)) is much smaller than the true difference (in graph (b)).

Therefore, no matter whether the prefecture becomes smaller or bigger, simple OLS regressions will underestimate the effect of capital status on population density.

[25] In this figure, we assume that capital-prefectures have a higher population density. However, the logic about measurement error leading to an underestimate is not sensitive to this assumption – it holds even if we make the opposite assumption.

B.2 Other Variables

Provincial capital We collect information on the geolocation of provincial capitals in each regime. If a prefecture (or a grid) contains a provincial capital, we term it a capital prefecture (grid).

Hierarchical distance We use the geographical coordinates of prefecture centroid to calculate hierarchical distance, which comprises of two components. The first is the great circle distance between prefecture i to a peer prefecture j in the same province in regime r (denoted as $D_{i,j,r}$). The second is the great circle distance between prefecture i to the national capital in regime r (denoted as $D_{i,NationalCap,r}$). The hierarchical distance is then calculated by $\sum_{j=1}^{N} A_j D_{i,j,r} + \lambda_r \sum_{j=1}^{N} A_j D_{i,NationalCap,r}$, in which A_j represents the area of the peer prefecture j. We employ λ_r in two ways: (1) allowing λ_r to vary and searching for the optimal λ_r which has the highest prediction power on capital status; and (2) fixing it to be a specific value.

Geography To construct the dummy indicating whether the prefecture contains a plain, we first use CHGIS V4 DEM (Digital Elevation Model) to calculate the average slope for all 0.25-degree X 0.25-degree grids. If the average slope is less than 1, the grid will be defined as a plain. We then match the grid to the prefecture boundary in 2000, and construct a dummy indicating whether the prefecture contains at least one plain. By matching CHGIS V4 DEM with prefecture boundary in 2000, we could calculate the average elevation and slope within each prefecture. The other geographic variables, such as dummy indicating whether a prefecture contains major river or is on the coast, longitude, latitude, and area are directly generated from CHGIS V4 (2007) using ArcGIS software.

Crop suitability Based on the suitability index from the Food and Agriculture Organization's 2012 Global Agro-Ecological Zones database, which ranges from 1 ("not suitable") to 8 ("very high") in each 0.5-degree x 0.5-degree grid cell, we measure prefecture-level crop suitability as the average for all cells located in each prefecture with a primary focus on the suitability for wheat, rice, fox millet, maize, and sweet potato.

Macroregions By analyzing the urban and the associated local and regional system hierarchies, Skinner (1977) divides traditional China (today's China excluding Inner Asian territories, or China proper) into nine physiographic macro-regions presented in Figure B.2. Specifically, they include Northeast China, North China, Northwest China, Upper Yangzi, Middle Yangzi, Lower Yangzi, Southeast Coast, Lingnan, and Yungui (see the map below). We construct a set of dummy variables indicating in which region the prefecture is located.

Figure B.2: Macroregions in Skinner (1977)

Public employment The data in 1776 comes from the Complete Directory of Qing Officials. A limitation of these records is that they only include individuals holding offices. The data in 2000 comes from the population census. We include the population employed in administrations and public institutions as public employment. The former includes those working in government agencies, Party agencies and social organizations; the latter refers to those employed in health care, sports and social welfare, education and culture, and scientific research and polytechnic services.

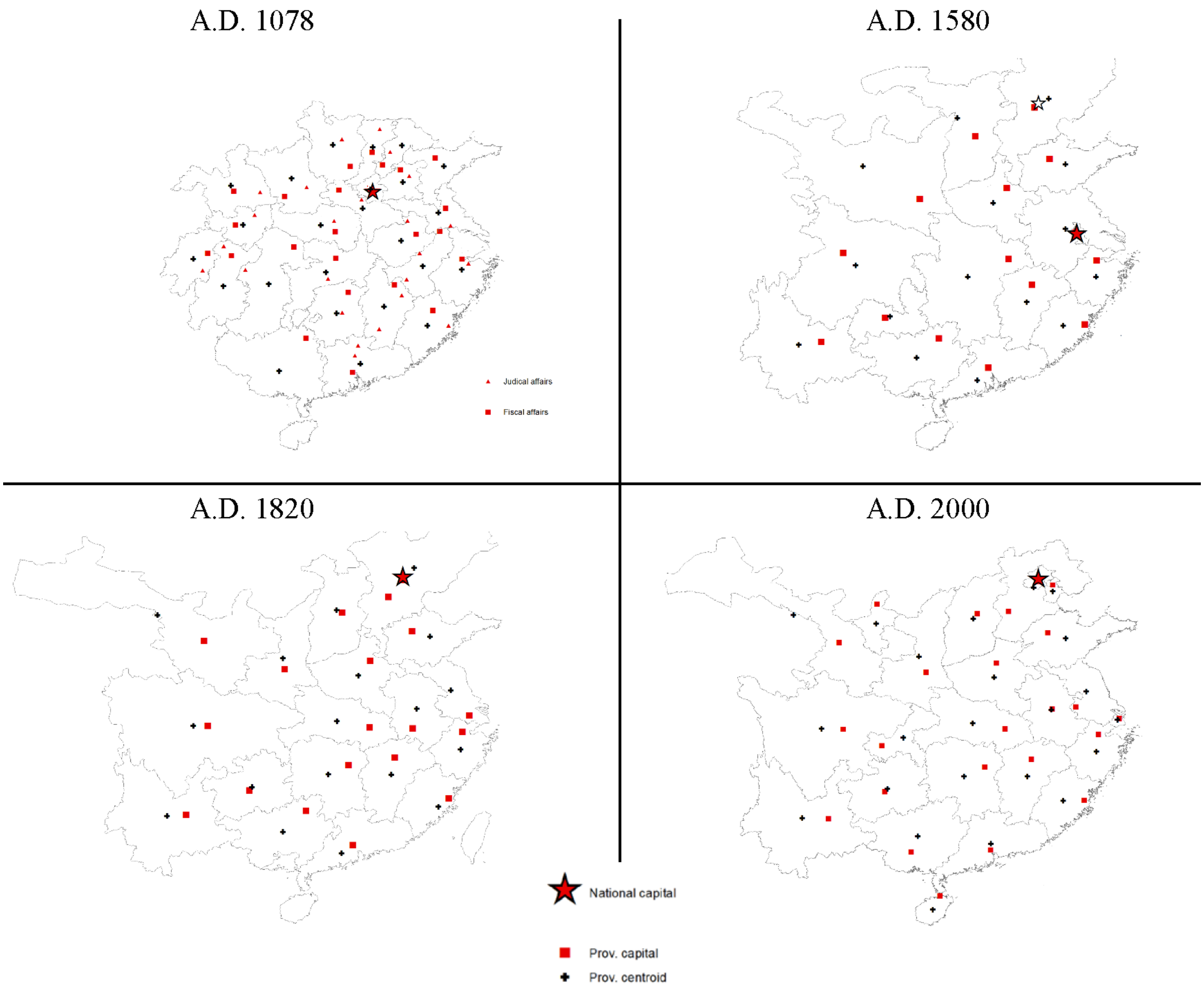
Transportation networks We digitize historical maps on the roads (both water and land way) in 1078, 1580 and 1820, as Figure 7 shows, and calculate the centrality measures as discussed in Section 6. For modern transportation networks, we focus on railways in 1990 because we are interested in those monopolized by the state.

C Additional Results on Hierarchical Distance

C.1 Maps Across Regimes

As long as $\lambda > 0$, the definition of hierarchical distance implies that provincial capitals should be located away from the provincial centroid and toward the national capital. This prediction is confirmed by the pattern in each regime. Figure C.1 plots the locations of provincial capitals (indicated by red squares), national capitals (indicated by red stars), and the provincial centroids (indicated by black crosses) regime by regime.

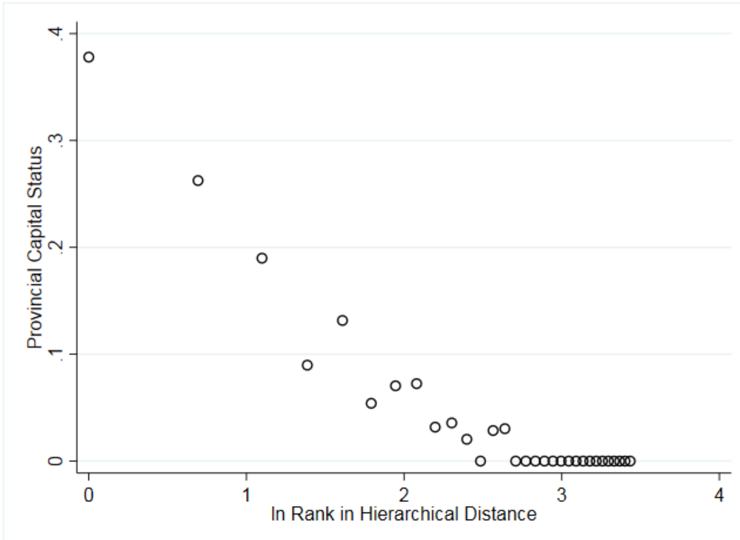
Figure C.1: Provincial Capital Locations by Regime



C.2 Fixing the Value of λ

In our main analysis, we use the λ_r that gives the largest prediction power for each regime r. This assumption is not essential for our analysis. The main patterns look similar if we vary λ from 0.1 to 0.9. Figure C.2 plots the case when λ=0.2 as an example, which is similar to the results using varied λ in our main text.

Figure C.2: ln Rank in Hierarchical Distance and the Probability of Being a Provincial Capital



C.3 Testing the Role of Hierarchical Distance

Hierarchical distance is a strong predictor of provincial capital status, both in terms of cross-sectional patterns by regime (Table C.3A) and within-prefecture changes (Table C.3B).

Table C.3: A. Hierarchical Distance and Provincial Capital Status: Regime by Regime
Dependent var.: Prov. capital =1/0

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regime	Song		Ming		Qing		P.R.C. (1964)	
λ =	0.29		0.33		0.20		0.17	
H-distance	-0.198***	-0.176***	-0.079***	-0.071***	-0.090***	-0.087***	-0.129***	-0.125***
	(0.031)	(0.042)	(0.023)	(0.025)	(0.026)	(0.027)	(0.028)	(0.030)
River		0.151***		0.077***		0.092***		0.058
		(0.052)		(0.027)		(0.028)		(0.037)
Plain		0.119*		0.062		0.058		0.051
		(0.063)		(0.040)		(0.040)		(0.041)
Other geography variables		Y		Y		Y		Y
Agriculture variables		Y		Y		Y		Y
Region FE		Y		Y		Y		Y
Observations	261	261	261	261	261	261	261	261
R-squared	0.160	0.226	0.088	0.141	0.079	0.138	0.148	0.184
R-squared w/o H-distance		0.140		0.085		0.077		0.070

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table C.3: B. Hierarchical Distance and Provincial Capital Status: Panel Analysis
Dependent var.: Prov. capital =1/0

	(1)	(2)	(3)	(4)	(5)	(6)
					Prov. Boundary Setting	
					Not strategic	Strategic
H-distance	-0.094***	-0.088***	-0.086***	-0.081***	-0.083***	-0.109***
	(0.019)	(0.018)	(0.018)	(0.019)	(0.024)	(0.031)
Prefecture FE, Year FE	Y	Y	Y	Y	Y	Y
Year FE*Geography variables		Y	Y	Y	Y	Y
Year FE*Agriculture variables			Y	Y	Y	Y
Year FE*Region FE				Y	Y	Y
Observations	1,044	1,044	1,044	1,044	716	328
R-squared	0.137	0.167	0.180	0.199	0.225	0.311

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D Additional Results for the DID and IV Analysis

D.1 Conley Spatial Standard Errors

Our results are robust to using Conley spatial standard errors (reported in the squares), where we use the distance cutoff of 1000 KM.

Table D.1: Conley Spatial Standard Errors

	(1)	(2)	(3)	(4)
	ln Pop.	ln Pop.	ln Pop.	ln Pop.
Prov. capital	0.622***	0.456***	0.488***	0.470***
	(0.132)	(0.104)	(0.105)	(0.104)
	[0.105]	[0.076]	[0.074]	[0.072]
Pref. FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Year FE * Geography		Y	Y	Y
Year FE * Crop suit.			Y	Y
Year FE * Region FE				Y
Observations	2,871	2,871	2,871	2,871
R-squared	0.757	0.859	0.873	0.882
Number of prefectures	261	261	261	261

D.2 Accessing the Importance of Omitted Variable Bias

We evaluate our baseline's robustness to omitted variable bias using the method proposed by Oster (2017), which assumes that if the relation between treatment and unobservables can be fully captured from the relation between treatment and observables, then the omitted variable bias will be proportional to the coefficient movements scaled by the change in R-squared when controls are included. Specifically, Oster (2017) proposes the following equation to calculate an approximation of the bias-adjusted treatment effect:

$$\beta^* \approx \tilde{\beta} - \delta \frac{R_{\max} - \tilde{R}}{\tilde{R} - \dot{R}} (\tilde{\beta} - \dot{\beta})$$

in which β^* represents the bias-adjusted effect of capital status on population density, while $\dot{\beta}$ and $\dot{R}$ ($\tilde{\beta}$ and $\tilde{R}$) denote the coefficients of capital status and R-squared with the fewest (most) controls.

In Table 2, we obtain $\dot{\beta} = 0.622$ and $\dot{R} = 0.757$ when we regress logged population density on capital status with only controlling prefecture and year fixed effects. As we include more controls, the coefficient of provincial capital decreases while R-squared increases. Specifically, $\tilde{\beta} = 0.470$ and $\tilde{R} = 0.882$ when we controlled the most controls as Column (4) of Table 2 shows. In the equation above, δ measures the relative importance of unobservables to observables, and R_{max} is the R-squared if the regression controls for both these latter. Using the above equation with only the upper bounds of δ and R_{max} enables derivation of a set of lower bounds for the bias-adjusted effect of capital status (β^*). Following Oster (2017), we take 1 as the upper bound of δ and assume that $R_{max} = 1$. Then, based on the two assumptions, our results show that the lower bound of bias-adjusted effect of capital on logged population density is 0.327.[26] The robustness of this approximation can be evaluated using two standards (Oster (2017)): First, because the capital effect moves toward zero when more variables are controlled for, we check whether the lower bound of β is smaller than zero. The answer is clearly negative. Second, we consider whether the lower bound of β is within the +/- 2.8 standard error of the capital coefficient with the most controls. The answer is positive, suggesting that the size of the estimate from the regression with the most controls is similar to the bias-adjusted estimate.[27] These results imply a lower likelihood that the estimated capital effect is fully driven by unobservables.

D.3 IV Results from Fixed λ

In our main results, we allow λ to change over time to better capture regime heterogeneity. Our IV estimates are similar if we fix λ to be 0.2, as shown in Table D.3.

[26]The bias-adjusted effect is 0.470-(0.622-0.470)*(1-0.882)/(0.882-0.757)=0.327

[27]The +/- 2.8 standard error of the coefficient of capital status based on Column (4) of Table 2 is [0.178, 0.762].

Table D.3: IV Results from Fixed λ

Method	Reduced form			IV	
	(1)	(2)	(3)	(4)	(5)
Sample	All	Ever-capital	Never-capital	All	All
				Second-stage	
Prov. capital				0.700***	0.671***
				(0.234)	(0168)
				First-stage	
H-distance	-0.064***	-0.141***	0.000	-0.091***	-0.099***
	(0.022)	(0.038)	(0.029)	(0.008)	(0.011)
Pref. FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Year FE * Geography	Y	Y	Y	Y	Y
Year FE * Crop suit.	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y
Year FE * $Prov_{i,r}$ FE					Y
Observations	2,871	693	2,178	2,871	2,871
R-squared	0.878	0.924	0.881	0.881	0.959
# prefectures	261	63	198	261	261
Weak instrument test (F statistic)				125.8	80.8

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.4 Various Provincial Fixed Effects

Our instrument exploits provincial boundary changes. To ensure our findings are not driven by these boundary changes, we include provincial fixed effects in two ways, including (1) $Prov_{i,r} \times \gamma_t$ where $Prov_{i,r}$ is the province fixed effects in all the regimes, and (2) $(\cap Prov_{i,r}) \times \gamma_t$ where we are comparing prefectures within unchanged polygons.

As shown in Table D.4, none of the specifications invalidate the importance of provincial capital status. Thus, while provincial boundary change adds complexity to our study, it is not an important driver for our main finding.

Table D.4: Dealing with Provincial Boundary Changes

	(1)	(2)	(3)	(4)
			IV Estimates	
Prov. capital	0.291***	0.279***	0.794***	0.990***
	(0.070)	(0.073)	(0.165)	(0.163)
			First-stage	
H-distance			-0.101***	-0.107***
			(0.011)	(0.012)
Pref. FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Year FE * Geography	Y	Y	Y	Y
Year FE * Crop suit.	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y
Year FE * Prov. FE in Song	Y		Y	
Year FE * Prov. FE in Ming	Y		Y	
Year FE * Prov. FE in Qing	Y		Y	
Year FE * Prov. FE in PRC	Y		Y	
Year FE * Prov-intersect FE		Y		Y
Observations	2,871	2,871	2,871	2,871
R-squared	0.961	0.967	0.957	0.960
# prefectures	261	261	261	261
Weak instrument test (F statistic)			87.6	80.4

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.5 Strategic Provincial Boundary Setting

Based on Sng et al. (2018), we divide provinces into two groups. The first group refers to the provinces whose boundaries were congruent with hypothetical boundaries based on catchment or least cost distance, suggesting that strategic setting is less a concern. The second is those whose congruence were low. Consistent with historical narratives, the second group concerns regions around the Yangtze River, the Huang River, the Qin Mountain and the Taihang Mountain. We present our IV estimates for both groups in Table . As shown, neither group negates our finding, suggesting that our idea about hierarchical distance works even if provincial boundary setting involved strategic concerns.

Table D.5: Dividing Provinces by How Strategic Their Boundary Is

Provincial boundary setting	(1) Less strategic		(2) More strategic	
	First-stage	IV Estimates	First-stage	IV Estimates
Prov. capital		0.689**		0.958***
		(0.271)		(0.235)
H-distance	-0.097***		-0.111***	
	(0.009)		(0.017)	
Pref. FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Year FE * Geography	Y	Y	Y	Y
Year FE * Crop suit.	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y
Observations	1,969	1,969	902	902
# prefectures	179	179	82	82
Weak instrument test (F-stat)	110.4		43.9	

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.6 A Proxy for Migration

As a method to gauge the importance of migration flows, we examine the population change by existing and new surnames in Table D.6, the latter of which is a proxy for migration inflows.

Table D.6: Population Change by Surnames

	(1)	(2)	(3)	(4)
Sample:	All	Subsample with surname information		
Normalized by average pop.	Pop. change	Pop. change	Pop. change: existing surname	Pop. change: new surname
Δ Capital status	0.504**	0.473**	0.361*	0.112**
	(0.241)	(0.236)	(0.206)	(0.046)
Pref. FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Year FE * Geography	Y	Y	Y	Y
Year FE * Crop suit.	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y
Observations	783	660	660	660
R-squared	0.598	0.625	0.463	0.710
# prefectures	261	220	220	220

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.7 Ever-Capital Prefectures

We also consider the role of ever-capital prefectures in two ways. First, we find that provincial capital status is associated with a higher population density within the subgroup of prefectures that have ever been a provincial capital (Columns (1) and (2) of Table D.7).

Second, our baseline finding also holds if we further control for all the ever-capital prefectures to exhibit different trends by controlling for ever-capital status status and its interaction with year dummies (Columns (3)–(8)).

Table D.7: The Impact of Capital Status on Population Density, Ever-Capital Prefectures

	Ever-Capital Pref. ln Pop Density		All Prefectures ln Pop Density		All Prefectures Δln Pop Density			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prov. capital	0.655***	0.311**	0.655***	0.514***				
	(0.198)	(0.124)	(0.195)	(0.142)				
Δ Prov. capital					0.448**	0.453***		
					(0.178)	(0.141)		
Gaining capital status							0.218	0.450**
							(0.275)	(0.220)
Losing capital status							-0.649***	-0.455***
							(0.211)	(0.174)
Pref. FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Controls		Y		Y		Y		Y
Year FE * Ever-capital			Y	Y	Y	Y	Y	Y
Observations	693	693	2,871	2,871	2,610	2,610	2,610	2,610
R-squared	0.738	0.926	0.758	0.882	0.411	0.716	0.412	0.716
# prefectures	63	63	261	261	261	261	261	261

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.8 Period-by-Period Results

Due to the frequency of data availability, we have more periods after capital status change than those before change. Thus, we divide the pre-periods before period0 into 1–75 years before 0, and 75+ years before 0; the post-periods into 1–75 years after 0, and 76–150 years after 0, 151–300 years after 0 and 300+ years after 0. Roughly speaking, a period of 75 years is about one-fourth of a regime's length. As shown in Table D.8, both the increase and the decrease in population density occur only after gaining and losing capital status.

Table D.8: Period-by-Period Estimates (Dependent Var.: ln Pop Density)

	(1)	(2)	(3)
Pre-Gaining: -75+	-0.329		-0.239
	(0.205)		(0.202)
Pre-Gaining: -75 ~-1	-0.064		-0.050
	(0.081)		(0.087)
Post-Gaining: 1 ~75	0.589**		0.563**
	(0.237)		(0.237)
Post-Gaining: 76 ~150	0.601**		0.569**
	(0.284)		(0.286)
Post-Gaining: 151 ~300	1.010**		0.857*
	(0.470)		(0.484)
Post-Gaining: 300+	1.099*		0.921
	(0.563)		(0.591)
Pre-Losing: -75+		0.124	0.046
		(0.110)	(0.110)
Pre- Losing: -75 ~-1		0.001	-0.009
		(0.053)	(0.057)
Post- Losing: 1 ~75		-0.512***	-0.391***
		(0.144)	(0.128)
Post- Losing: 76 ~150		-0.470***	-0.333***
		(0.142)	(0.128)
Post- Losing: 151 ~300		-0.497*	-0.340
		(0.269)	(0.267)
Post- Losing: 300+		-0.611**	-0.317
		(0.273)	(0.268)
Pref. FE, Year FE	Y	Y	Y
Year FE * Controls	Y	Y	Y
Year FE * Ever-capital	Y	Y	Y
Observations	2,783	2,783	2,783
R-squared	0.882	0.881	0.883
# Prefectures	253	253	253

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.9 Using Subperiods of the Data

Our finding is robust to using different subperiods of the data. In Columns (1) and (2) of Table D.9, we show that including the data in 1880, 1953, 1982, 1990 does not change our main finding, even though doing so make the gaps across periods more uneven. In Columns (3) and (4), we find that our results are robust to keeping subperiods of roughly equal length; Columns (5)–(9) show that our results are also robust to dropping any specific regime.

Table D.9: Reducing Uneven Gaps and Using Different Subperiods (Dependent Var.: ln Pop Density)

		Baseline	+smaller gaps	Gap: ca400	Gap: ca120	Excluding: Song	Ming	Qing	PR China	Ming/Qing
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prov. Capital		0.514***	0.501***	0.512***	0.575***	0.369***	0.485***	0.626***	0.473***	0.665***
		(0.142)	(0.131)	(0.154)	(0.143)	(0.123)	(0.138)	(0.154)	(0.172)	(0.173)
Song	980	Y	Y	Y	Y		Y	Y	Y	Y
	1078	Y	Y				Y	Y	Y	Y
	1102	Y	Y		Y		Y	Y	Y	Y
Ming	1393	Y	Y	Y	Y	Y		Y	Y	
	1580	Y	Y		Y	Y		Y	Y	
Qing	1776	Y	Y			Y	Y		Y	
	1820	Y	Y	Y	Y	Y	Y		Y	
	1851	Y	Y			Y	Y		Y	
	1880		Y			Y	Y		Y	
	1910	Y	Y			Y	Y		Y	
P R China	1953		Y			Y	Y	Y		Y
	1964	Y	Y			Y	Y	Y		Y
	1982		Y			Y	Y	Y		Y
	1990		Y			Y	Y	Y		Y
	2000	Y	Y		Y	Y	Y	Y		Y
Pref. FE		Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE		Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE*Controls		Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE*Ever-capital		Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations		2,871	3,915	783	1,566	3,132	3,393	2,610	2,610	1,827
R-squared		0.882	0.887	0.887	0.891	0.893	0.885	0.903	0.864	0.915
# prefectures		261	261	261	261	261	261	261	261	261

D.10 Grid-level Analysis

We conduct a 1-degree × 1-degree grid analysis and a 2-degree × 2-degree grid analysis. The size of a prefecture in our main analysis is between the size of the two grids: instead of 261 prefectures, we have 361 1-degree × 1-degree grids and 94 2-degree × 2-degree grids. These two sets of grid-level analysis again generate patterns comparable with our prefecture-level analysis (presented in Table D.10) indicating that our findings are robust to alternative boundary definitions.

Table D.10: The Impact of Capital Status on Population Density: Grid-level Analysis

	1-degree × 1-degree				2-degree × 2-degree			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prov. capital	0.323***	0.288***	0.270***	0.338***	0.270**	0.197**	0.166*	0.323***
	(0.077)	(0.066)	(0.065)	(0.065)	(0.103)	(0.094)	(0.098)	(0.100)
Grid FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Geography			Y	Y			Y	Y
Year FE * Agriculture			Y	Y			Y	Y
Year FE * Region FE				Y				Y
Observations	3,971	3,971	3,971	3,971	1,034	1,034	1,034	1,034
R-squared	0.755	0.792	0.806	0.845	0.821	0.877	0.889	0.920
# grid	361	361	361	361	94	94	94	94

Notes: Standard errors presented in the paraphrases are clustered at the grid level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.11 Validity Checks of the Instrument

For our IV analysis, we assume that the change of a prefecture's hierarchical distance is orthogonal to its own characteristics, because the change in a prefecture's hierarchical distance stems from regime change.

To check whether this is the case, we first show that the change in hierarchical distance is not significantly correlated with the levels and changes in population density in the past periods (presented in part A of Table D.11).

In addition, we find that ln rank in hierarchical distance matters for both gaining capital status and losing capital status (presented in part B of Table D.11). Since it seems unlikely that the central government would intentionally increase a prefecture's hierarchical distance to make it lose its capital status, this finding is reassuring for the validity of hierarchical distance.

Table D.11: A. Validity Check I – Hierarchical Distance and Pre-change Characteristics

	Δln Hierarchical distance					
	(1)	(2)	(3)	(4)	(5)	(6)
lag. ln Pop density	-0.029		-0.030			
	(0.024)		(0.031)			
lag2. ln Pop density		-0.021	-0.004			
		(0.022)	(0.018)			
lag. Δln Pop density				-0.010		-0.014
				(0.018)		(0.021)
lag2. Δln Pop density					-0.032	-0.034
					(0.038)	(0.039)
Prefecture FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Year FE * Controls	Y	Y	Y	Y	Y	Y
Observations	2,610	2,349	2,349	2,349	2,088	2,088
R-squared	0.278	0.278	0.279	0.278	0.278	0.279
# prefectures	261	261	261	261	261	261

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table D.11: B. Validity Check II – Gaining vs. Losing Capital Status

		Seemingly Unrelated Regression	
	ΔCapital status	Gaining status	Losing status
	(1)	(2)	(3)
Δln Hierarchical distance	-0.059***	-0.027***	0.032***
	(0.016)	(0.004)	(0.005)
Prefecture FE	Y	Y	Y
Year FE	Y	Y	Y
Year FE * Controls	Y	Y	Y
Observations	2,610	2,610	2,610
R-squared	0.187	0.195	0.300
# prefectures	261	261	261

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

D.12 ln Rank in $\sum_{j=1}^{N} A_j D_{i,j,t}$ × ln Rank in $\sum_{j=1}^{N} A_j D_{i,NatCap,t}$ as an instrument

Since our main concern is whether being closer to the national capital or the other prefectures affects development through channels beyond provincial capital status, we further employ ln Rank in $\sum_{j=1}^{N} A_j D_{i,j,t} \times \ln$ Rank in $\sum_{j=1}^{N} A_j D_{i,NatCap,t}$ as an instrument while controlling for the independent effects of the two components. These results are presented in Table D.12.

Table D.12: Exclusion Restriction? Alternative Specification of the Instrument

	Reduced form			IV	
	All	Ever-capital	Never-capital	First-stage	Second-stage
	(1)	(2)	(3)	(4)	(5)
Prov. capital					0.850***
					(0.324)
ln Rank in dist to national cap.*	0.057**	0.176***	0.006	0.067***	
ln Rank in dist to prov centroid	(0.023)	(0.047)	(0.029)	(0.008)	
ln Rank in dist to national cap.	-0.197***	-0.489***	-0.049	-0.187***	-0.038
	(0.052)	(0.089)	(0.070)	(0.019)	(0.025)
ln Rank in dist to prov centroid	-0.161***	-0.391***	-0.039	-0.191***	0.001
	(0.047)	(0.090)	(0.061)	(0.167)	(0.032)
Pref. FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Year FE * Controls	Y	Y	Y	Y	Y
Observations	2,871	693	2,178	2,871	2,871
R-squared	0.879	0.929	0.882	0.216	0.879
# prefectures	261	63	198	261	261
Weak instrument test				67.6	

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

As shown in Column (4), since a higher value of rank in distance to national capital (or provincial centroid) indicates longer distance, each component decreases the probability of being a provincial capital. However, the marginal effect of each component declines with the other component. Intuitively, for a prefecture that is already far from the provincial centroid, the change in distance to the national capital matters less. This is why we obtain a positive interaction effect.

Two findings are worth emphasizing. First, both ln Rank in $\sum_{j=1}^{N} A_j D_{i,j,t}$ and ln Rank in $\sum_{j=1}^{N} A_j D_{i,NatCap,t}$ are indeed strongly correlated with population density (Column (1)); however, when using their interaction as the instrument, we do not find that either part alone has an additional direct effect (Column (5)). Once again, this finding suggests that the variation in the interaction mitigates the concern of exclusion restriction. Second, compared with the estimate using hierarchical distance, the IV estimate from using the interaction is slightly higher, but the first-stage is smaller (with a F-statistics of 67.6 in contrast with 131.9 in Table 2). This difference further suggests that our definition of hierarchical distance better captures the political logic of provincial capital choice by the central government.

D.13 Placebo Hierarchical Distances

The two-component structure of hierarchical distance also allows to create multiple placebo hierarchical distance ranks by exploiting the changes in national capital status. For instance, we calculate one such placebo to Kaifeng when Kaifeng was not a capital and similar ones for Nanjing and Beijing before they became national capitals.

Table D.13: Considering Placebo Hierarchical Distance

	(1)	(2)	(3)
IV Estimates: ln Pop density			
Prov. capital	0.901***	0.981**	0.779***
	(0.206)	(0.418)	(0.229)
ln Rank in H dist. Kaifeng * Post-	0.017		
	(0.024)		
ln Rank in H dist. Nanjing * Pre-		0.030	
		(0.044)	
ln Rank in H dist. Beijing * Pre-			-0.028
			(0.045)
Pref. FE	Y	Y	Y
Year FE	Y	Y	Y
Year FE * Controls	Y	Y	Y
Observations	2,871	2,871	2,871
R-squared	0.878	0.877	0.880
# Prefectures	261	261	261
F-Stat (Weak instrument test)	168.2	40.9	132.5

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

As shown in Table D.13, We find that including these placebo hierarchical distance ranks does not alter our IV estimate. Nor does it affect population density, implying that our findings are specific to these cities' political status.

D.14 Distance to Major Market Centers

To check whether our findings are confounded by distances to major market centers, we calculate a prefecture's (hierarchical) distance to three types of market centers: the north China plain during the Song Dynasty and the lower Yangtze after the Song (cf. Skinner 1977), Shanghai in the east, and Guangzhou in the south. To calculate the ranks of these distances, we replace $D_{i,NatCap,t}$ with $D_{i,Market,t}$. Like the placebo distances, these ranks are also correlated with our instrument and thus may be correlated with the probability of being a capital. However, once again, none explains the role of our hierarchical distance (see Table D.14), which confirms the importance of hierarchical distance to the political center.

Table D.14: Considering Distance to Major Market Centers

	(1)	(2)	(3)
IV Estimates: ln Pop density			
Prov. capital	0.716***	0.858***	0.814***
	(0.256)	(0.248)	(0.227)
ln Rank in H dist. to major econ region	-0.029		
	(0.026)		
ln Rank in H dist. to the East (Shanghai)		0.024	
		(0.027)	
ln Rank in H dist. to the South (Guangzhou)			0.017
			(0.023)
Pref. FE	Y	Y	Y
Year FE	Y	Y	Y
Year FE * Controls	Y	Y	Y
Observations	2,871	2,871	2,871
R-squared	0.881	0.879	0.880
# Prefectures	261	261	261
F-Stat (Weak instrument test)	105.2	113.1	134.9

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

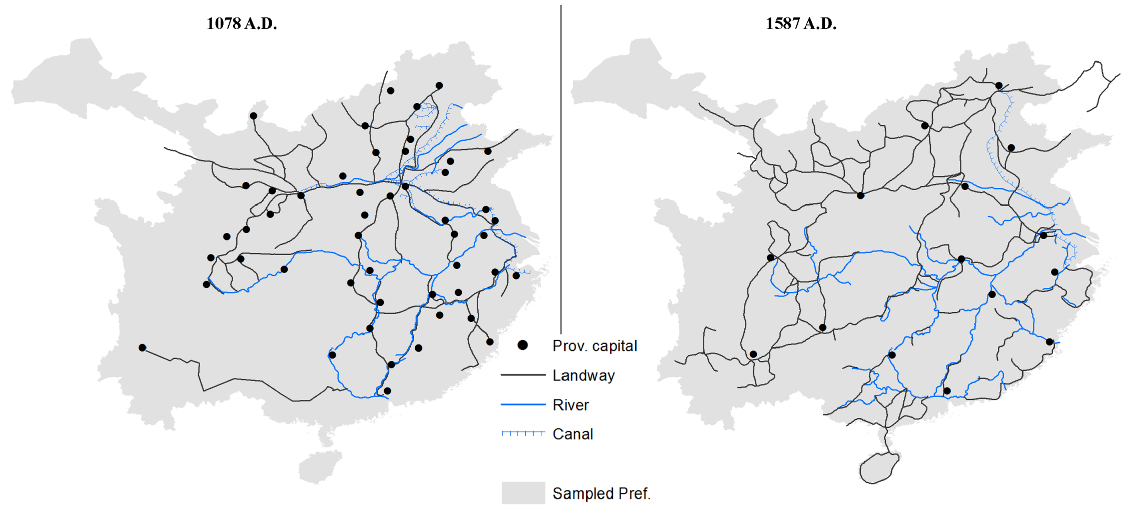
E Additional Results on Interpretations

E.1 Transportation Networks

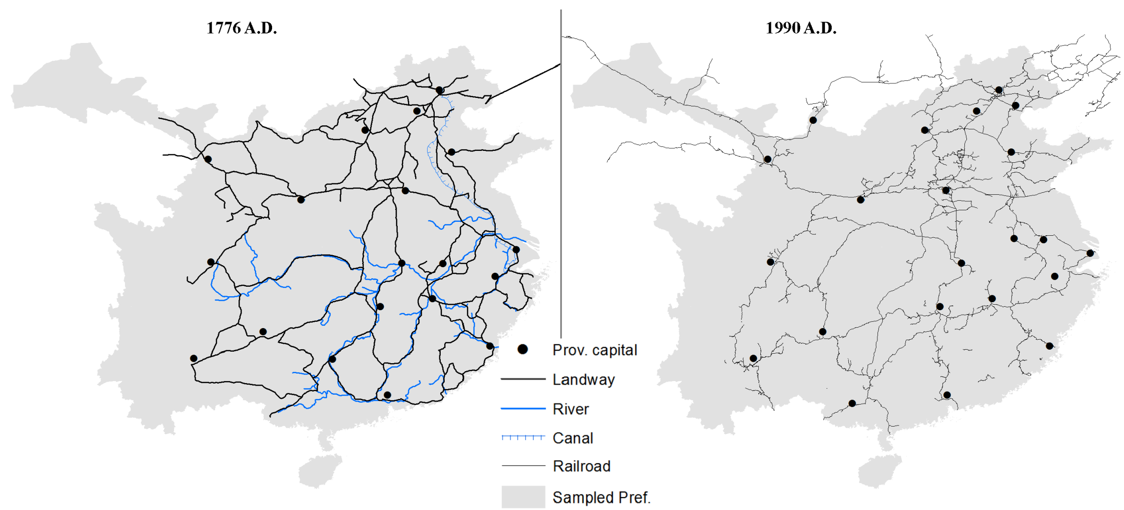
Figure E.1 plots the transportation networks across regimes. In 1990, we focus on railway networks (CHGIS 2007). We digitize the maps in the Song, Ming and Qing from *Historical Atlas of China* (Cheng and Hsu 1980), which covers major routes across periods.

Figure E.1: Transportation Networks

(a) Song and Ming



(b) Qing and 2000



E.2 Pre-trends in Centrality and Market Access

We examine the pre-trends in transportation centrality and find that no significant differences in centrality existed before capital status change (Columns (1) and (2) of Table E.2). In addition, we define market access of prefecture at time t as $ln \sum_{j \neq i} \frac{MA_{jt}}{(\tau d_{ijt}^{\delta})}$. The market size (MA_{jt}) is measured by the prefecture's area in Columns (3) and (4) and its population in Columns (5)–(8). d_{ijt} indicates shortest distance between prefecture i and j in the transportation network at time t. In Columns (7) and (8), we take topography into account by using the ruggedness to weight the cost of distance.

Table E.2: Pre-trends in Centrality and Market Access

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ln Centrality		ln Market access Area weighted		ln Market access Pop weighted		ln Market access Pop and topography weighted	
Prov. capital	0.062***		0.438***		0.369***		0.612**	
	(0.019)		(0.138)		(0.135)		(0.239)	
Before gaining status: t-2		0.006		-0.013		-0.048		-0.127
		(0.054)		(0.330)		(0.377)		(0.462)
After gaining status		0.090		0.617		0.528		0.636
		(0.059)		(0.385)		(0.397)		(0.733)
Before losing status: t-2		-0.016		-0.057		-0.097		0.182
		(0.034)		(0.267)		(0.270)		(0.279)
After losing status		-0.055***		-0.382***		-0.316**		-0.582**
		(0.019)		(0.139)		(0.130)		(0.231)
Pref. FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Geography	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Crop suit.	Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,016	1,016	1,016	1,016	1,016	1,016	1,016	1,016
R-squared	0.692	0.692	0.779	0.779	0.495	0.496	0.595	0.595
# prefectures	254	254	254	254	254	254	254	254

Notes: Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

E.3 A Counterfactual of Political Decentralization

A political economy counterfactual we are interested in is a decentralized scenario where only provincial autonomy matters. Therefore, we conduct a simple analysis by comparing provincial market access in the current political system with that in a scenario where provincial capitals were set in the provincial centroids. In this exercise, we assume that all prefectures in a province can still reach other provinces, and the only change is the position of the provincial capital.

We define national market access of prefecture i in province p at time t as $MA_{pi,t} = \sum_{j \neq pi} \frac{M_{j,t}}{d_{pi,j,t}}$, where the market size (M_j, t) could be measured by its population. We use the average market access of all prefectures in a province p, namely $\sum_i MA_{pi,t}/N$ in which N is the total number of prefectures in the province, as its provincial level measure $(MA_{p,t})$, and then examine how the impact of a provincial capital's deviation from the provincial centroid as follows:

$$\ln MA_{p,t} = \rho_1 Dist_{p,t}^{Centroid} + \rho_2 Dist_{p,t}^{NationalCapital} + \theta X_{p,t} + \gamma_t + \epsilon_{p,t}, \tag{9}$$

where $Dist_{p,t}^{Centroid}$ denotes the distance from the provincial capital in province p to the provincial centroid. Since deviation toward national capital can potentially increase provincial market access, we control for the distance from the provincial capital to the national capital $(Dist_{p,t}^{NationalCapital})$. $X_{p,t}$ refer to all geographic controls in our main results.

Table E.3: How the Location of Provincial Capitals Affects Provincial Aggregate Market Access

	(1)	(2)	(3)	(4)	(5)	(6)
Market access measure	Weighted by pop. in 1078			Weighted by pop. t		
Deviation from prov. centroid	-0.106***	-0.106***	-0.086***	-0.109***	-0.109***	-0.083**
	(0.033)	(0.033)	(0.031)	(0.034)	(0.034)	(0.032)
Deviation from natl. capita		-0.024***	-0.008		-0.028***	-0.011
		(0.006)	(0.007)		(0.006)	(0.007)
Year FE	Y	Y	Y	Y	Y	Y
Geographic controls			Y			Y
Observations	81	81	81	81	81	81
R-squared	0.430	0.430	0.681	0.946	0.946	0.970

Notes: Robust Standard errors are presented in the paraphrases. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

As shown in E.3, if the provincial capital is 100 KM away from the provincial centroid, provincial aggregate market access declines by around 8.6% (Column (3)). In our data, the average deviation is 125 KM. Thus, average provincial aggregate market access could be 10.8% higher.