

NBER WORKING PAPER SERIES

THE ECONOMIC CONSEQUENCES OF POLITICAL HIERARCHY:
EVIDENCE FROM REGIME CHANGES IN CHINA, AD1000-2000

Ying Bai
Ruixue Jia

Working Paper 26652
<http://www.nber.org/papers/w26652>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020

We thank 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, Austin Wright, Xiaoxue Zhao, Junfu Zhang, Fabarizio Zilibotti and participants of various seminars for their comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2020 by Ying Bai and Ruixue Jia. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Economic Consequences of Political Hierarchy: Evidence from Regime Changes in China,
AD1000-2000

Ying Bai and Ruixue Jia

NBER Working Paper No. 26652

January 2020

JEL No. H11,N95,O11

ABSTRACT

We argue that China, with its long history, a relatively stable political system, and multiple regime changes, provides us an opportunity to investigate the political economy of administrative hierarchy. Using prefecture-level panel data and exploiting regime changes during AD1000-2000, we find that gaining and losing importance in the political hierarchy led to the rise and decline of different prefectures (measured by population density and urbanization). Moreover, political hierarchy shapes regional development via both political and market channels (reflected by public employment and transportation networks). More broadly, our study serves as new evidence on how politics shapes economic geography and offers a context to understand changes in economic activity location in the long run.

Ying Bai

Department of Economics,

The Chinese University of Hong Kong

Shatin, Hong Kong

rickybai@gmail.com

Ruixue Jia

School of Global Policy and Strategy

University of California at San Diego

9500 Gilman Drive #0519

La Jolla, CA 92093

and NBER

rxjia@ucsd.edu

A data appendix is available at <http://www.nber.org/data-appendix/w26652>

1 Introduction

In AD1000, Kaifeng was the most prosperous city in China and, with an estimated population of 1 million, arguably the largest city in the world (Mote 2003, Morris 2013). By 2015, however, its GDP ranked 129th among Chinese cities and its former glory was long forgotten. Kaifeng’s decline is closely related to its status in the political hierarchy, it having first lost political prestige as the national capital in the 13th century and then its status of provincial capital in the 20th century (Heng 1999). Kaifeng is just one of many cases whose economic status rise and fall with their position in the political hierarchy in China. In this paper, we study the economic consequences of political hierarchy by exploiting major regime changes over a millennium in China. Two findings stand out. Both gaining and losing importance in the political hierarchy matter greatly for regional development. Moreover, political hierarchy shapes regional development via both political and market channels (reflected by public employment and transportation networks).

Political hierarchy is commonly observed: most countries are divided into smaller units (provinces, states, regions and so on), with administrative centers connecting these units with the central government. Existing studies on administrative centers, however, often have to rely on cross-sectional variation as a region’s political status rarely changes.¹ China, with its long history, a relatively stable political system, and multiple regime changes that reshuffled the political landscape, provides a rare opportunity to further investigate the economic consequences of political hierarchy. By doing so, we shed light on the logic of setting up administrative centers and provide new evidence on how politics shapes economic geography. In addition, our study departs from a large literature emphasizing the persistence of economic activities and offers a context to understand changes in economic activity location in the long run.

Our empirical setting is the core regions of China (the shaded area in Figure 1, coined as China proper by historians) in the second millennium, for which rich administrative information exists for centuries. Two features of the Chinese political system are particularly important for our research questions: (i) six major distinctive dynasties/regimes existed during A.D. 1000-2000; and (ii) despite regime shifts, China’s three-tier administrative system (province-prefecture-county) that got formed around A.D. 960 remain stable over this millennium. The regime shifts led to changes in national capitals, provincial boundaries, and provincial capitals (with the latter being among the prefectures that make up provinces). Our focus is to understand the consequences of (re)locating provincial capitals – the most

¹Perhaps partly due to the limited variation, the existing literature is not large and includes DeLong and Shleifer (1993), Ades and Glaeser (1995), Davis and Henderson (2003) and Campante and Do (2014).

important regional administrative centers – on spatial location of economic activity. To do so, we construct a panel dataset across 261 prefectures for 11 periods (980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964, and 2000) for our baseline analysis, based on a large amount of historical and modern sources.

Before studying the consequences of gaining and losing provincial capital status, we discuss the underlying reasons of relocating provincial capitals. While the choices of provincial capitals certainly involved many factors whose importance could vary across regimes, there exists a general and testable logic that applies to all regimes. We argue that the location of provincial capitals reflects a trade-off between granting regional autonomy (that facilitates provincial development) and keeping central control (that is critical for political stability). This trade-off can be proxied empirically by a weighted sum of distances to the other prefectures within a province and the distance to the national capital (what we call “the hierarchical distance”). Intuitively, if regional autonomy dominates, we expect proximity to other prefectures within the province to matter more; if central control dominates, we should observe that proximity to the national center matters more. During the millennium, China underwent six dynastic regime changes that relocated national capitals and redivided provincial boundaries. As a result, the two sources of variation underlying the hierarchical distance – the location of the national center as well as provincial boundaries – are altered by regime changes. Exploiting these variations, we show that the simplest form of hierarchical distance (e.g., based on arguably exogenous geometric distances) is already powerful in explaining the change in provincial capital locations across regimes, which provides a possible instrument for provincial capital status.

We then show that the administrative hierarchy has important consequences on the spatial distribution of economic activity over the millennium, using both a difference-in-differences (DID) strategy and an instrument variable (IV) strategy (based on the hierarchical distance derived above). Specifically, gaining and losing importance in the administrative hierarchy led to the rise and decline of different prefectures (measured by population density and urbanization). The magnitude of the consequence is large. For instance, based on a difference-in-differences strategy, the change in provincial capital status is associated with a 40-50% change in population density. We discuss the empirical challenges for the DID and the IV strategies and conduct an array of checks on them such as using period-by-period analysis to examine pre-trends, dealing with uneven gaps in our data, conducting grid-level analysis, and checking the concerns of exclusion restriction for our IV analysis. Overall, the results from the two strategies corroborate each other.

We consider two broad interpretations of the link between political hierarchy and regional development: the politics hypothesis and the market hypothesis. The politics hy-

pothesis asserts that the effect can be explained by public employment (and its multiplier effect). This idea can be dated back to Max Weber, who coined the phrase “the Oriental city” and emphasized that political forces drive the rise of the administrative centers in medieval Chinese and Islamic societies (Weber 1921). In contrast, the market hypothesis assumes that politics affects development via market factors. For instance, in both history and today, we observe that the Chinese state has been critical in constructing transportation networks (that shapes market access). Empirically, using panel datasets on public offices and transportation networks, we find both channels are at work: the importance of provincial capital status is comparable to moving from being average among the 261 prefectures to top 35-15 in terms of public employment and top 60-40 in terms of spatial centrality in the transportation networks. Our finding thus suggests that it may be too narrow to assume that political hierarchy affects economic development only via the public sector.

Our study is related to a political economy literature on how politics factors affect economic geography. Compared with the existing studies mentioned above (DeLong and Shleifer 1993, Ales and Glaeser 1995, Davis and Henderson 2003, Campante and Do 2014), our setting allows us to shed light on the determinants of political status of certain localities and study both gaining and losing political importance. In this process, we uncover some new factors underlying the link between politics and development.

Second, our perspective complements the literature on the long-term spatial distribution of economic activity by documenting its changes.² A growing literature has emphasized the persistence of economic activity in the long run due to different channels (see Nunn (2009) for a review).³ Our study suggests that introducing politics to the literature can be useful for us to understand changes in economic activity. In particular, in our context, the state plays a critical role in overcoming the inertia in economic activity by providing important public goods like transportation networks.

Third, our study adds to a literature on the political economy of administrative divisions (e.g., Alesina and Spolaore 1997, Bolton and Roland 1997, Michalopoulos and Papaioannou 2013, Ko, Koyama and Sng 2016). These studies often focus on national-level variation while our research question is a within-nation one. Broadly speaking, the logic

²To be clear, 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. However, on top of some persistence, they also change systematically with political status, which has been less studied than the patterns of persistence.

³As pointed out by Nunn (2009), these channels include geography (e.g., Diamond 1998), institutions (e.g., Acemoglu, Johnson and Robinson 2001, Dell 2010), human capital (e.g., Glaeser et al. 2004), and culture (e.g., Greif and Tabellini 2017). Another literature motivated by trade theories emphasizes the role of locational fundamentals (pioneered by Davis and Weinstein 2002) or the persistent impacts of temporary advantages (exemplified by Bleakley and Lin 2012 among many others). Recently, Allen and Donaldson (2019) provide a theoretical framework to study the geography of path dependence.

of setting up administrative centers in our paper can be considered as spatial implementation of federalism discussed in an earlier theoretical literature (Blanchard and Shleifer 2001, Treisman 2007).

Finally, politics has long been recognized by scholars as an important factor in explaining development paths of different regions in China.⁴ Although rich administrative data sources exist in China for a long history thanks to its stable bureaucracy system, few of them have been employed by economists. Our study is an attempt to use both historical and modern sources to study a first-order question in China’s economic geography. Given what we have found – political hierarchy shapes both political and market factors – it is not surprising that the first-tier cities today are typically provincial capitals.

Section 2 describes the background and introduces the data. Section 3 discusses the variation in provincial capitals. Section 4 presents descriptive evidence the importance of political hierarchy. Section 5 presents the estimation results on economic consequences, and Section 6 discusses the politics hypothesis and the market hypothesis underlying the economic consequences. Section 7 concludes the paper.

2 Background and Data

2.1 Regime Changes & Capital Relocation

As mentioned in the introduction, two features of the Chinese political system are particularly important for our research design: (i) multiple distinctive dynasties/regimes existed during during A.D. 1000-2000, whose founders had no relationship to each other; and (ii) despite regime shifts, China’s three-tier administrative system (province-prefecture-county) that got formed around A.D. 960 remain stable over this millennium.⁵ The regimes shifts led to changes in national capitals, provincial boundaries, and provincial capitals, with the latter being among the prefectures that make up provinces. In each regime, however, the central government monopolizes the power of appointing, rotating, and removing officials for each tier. To reduce collusion between provincial governors and local elites, “hometown avoidance” has been employed, i.e., provincial governors are not allowed to govern their home province.

Against this background, it is reasonable to consider the administrative decisions from the perspective of the central government who cares about governing the whole country.

⁴Influential urban history works include Rowe (1993) and Chao (2006), both of which provide vivid case studies on how politics contributes to the rise of fall of several cities.

⁵In the Song dynasty, the first tier is known as circuit (*Lu* in Chinese), which is comparable to the provincial unit in the latter regimes.

Next, we summarize the major administrative changes over the millennium.

Regime Changes During the millennium studied (AD1000-2000), 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; 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).

These infrequent regime changes are hard to predict. For example, it would be difficult for the residents of the Song Dynasty capital (Kaifeng) to imagine that China could later be ruled by the Mongols, who moved the national capital to northern China (Beijing) and redivided provincial boundaries for political control. The administrative decisions made by the rulers at the beginning of each new regime – including national capital location, provincial boundaries, and provincial capitals – usually persisted until the regime's end, with only occasional changes in between. Our national and provincial capital variation is thus driven primarily by regime change. Because national capital relocation and provincial boundary redivision are the preconditions for changes in provincial capital, we describe them in order below.

National Capital Relocation The national capitals changed five times across the six regimes, reflecting the unpredictability of where a new power base could arise. Thus, today's Kaifeng (in central China), Beijing (in northern China), and Nanjing (in central-south China) were the national capitals for the Song, Yuan, and Ming dynasties, respectively,⁶ after which Beijing served as capital the Qing Dynasty, Nanjing for the Republic, and Beijing again for the current People's Republic.

Historians have argued that one important determinant for national capital locations is the origin of the power base. For instance, the Mongols (founders of the Yuan) and Manchurians (founders of the Qing) originated in the north and hence picked Beijing as the national capital, while the Ming and Republic power bases were in the south and favored Nanjing. Although both Beijing and Nanjing were candidates for the national capital of the People's Republic, Beijing was chosen, partly because of its nearness to China's political ally at that time (the Soviet Union). This is not to say that power base is the only factor to determine national capital location, but illustrates that political considerations are usually

⁶Nanjing was the national capital of the Ming Dynasty until its capital was relocated to Beijing in 1421, partly because the new emperor, who took power via a coup, had his power base in Beijing. The Crown Princess, however, stayed in Nanjing, which became the norm for this regime.

more important than economics.⁷

Provincial Boundary Redivision Along with national capital relocation came provincial boundary alteration, which during this millennium was affected by a major shift between the two principles for defining provincial boundaries: whether to *follow or subsume* 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 the natural barriers within provinces so that boundaries, is known as “*interlocked like dog’s teeth*”. The latter exemplifies a spatial “divide-and-rule” tactic to limit the power of local governments.⁸

The Song adhered to the first principle, generally defining provincial boundaries by natural mountains and rivers, but when the Mongols came to rule, being preoccupied with the possibility of a usurper’s mobilizing resources against the central government, they adopted the interlocking principle to an extreme, intentionally including natural mountains and rivers within (larger and fewer) provinces. The regimes following the Yuan Dynasty then adjusted the number of provinces using the two principles, with a generally increasing pattern. In Online Appendix A.1., we use the Yangtze River as an example to illustrate how the two principles of defining provincial boundaries work across regimes. 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. Historians have also documented similar changes of provincial boundaries driven by the evolution of two principles in other regions such as the Qin Mountain and the Huang River (Zhou 1998, Li 2011).

Like national capitals, provincial boundaries were created more from political inspiration than from economic imperatives, which often represented efforts to divide various communities so “each could be dealt with separately” (Skinner 1977, Guy 2010). As pointed out by Skinner, provincial boundaries arise through “administrative accidents” rather than delineation of the natural boundaries of human activity.

Provincial Capital Relocation The relocation of national capitals and redivision of provinces naturally affected the relative importance of a prefecture, which having been central based on the old provincial boundary could become rather isolated given the new delineation

⁷Kang Chao, in an influential study on urban history in China, argues that the choices of national capitals were primarily driven by political factors, which actually “wasted considerable resources” because the central governments had to support them by constructing new infrastructures to transfer resources from the rest of the country. For example, after Kaifeng became the national capital for the Song dynasty, the central government built four major canals to supply grain to the city. But the canals needed to be widened and deepened every few years. After Kaifeng’s losing capital status, these canals stopped functioning soon due to lack of maintenance (Chao 2006).

⁸The logic is similar to divide-and-rule documented in Michalopoulos and Papaioannou (2013).

and national capital. 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 Online Appendix A.2.). During the Ming dynasty, however, it lost its capital status because redrawing of the provincial boundary placed it far away from other prefectures in the province even though it was still relatively closer to the new national capital of Nanjing (Panel (b) of Online Appendix A.2.). When the national capital relocated to Beijing (Panel (c) of Online Appendix A.2.), the prefecture became even further isolated and as a result, Luzhou never regained its capital status.

In contrast, Changsha, which as provincial capital of Jinghu South province during the Song Dynasty was relatively close to the national capital and the other prefectures in the province (Online Appendix A.2.), became rather isolated in the Yuan and Ming dynasties and lost its provincial status. It regained its capital status in the Qing Dynasty, however, because of provincial redivision.

For our baseline analysis, we map the historical data onto the 261 prefectures existing in the year 2000 and construct a panel dataset (see discussion on data construction in Section 2). Of these 261 prefectures, 63 were at some time a provincial capital (see their locations in Figure 1): 36 lost their capital status once, 11 gained their capital status once, 8 have experienced multiple changes, and 8 have always been capitals. We summarize these changes across regimes in Table 1.

At this time, it should be noted that each province has always had only one capital except during the Song dynasty when the central government limited the power of local governments by spatially separating capitals according to fiscal affairs, judicial affairs, and welfare (Mostern 2011). As a result, the majority of provinces had two provincial capitals in the Song dynasty, one for fiscal affairs and one for judicial affairs and others. This feature is helpful, because we have a broad set of possible candidates for provincial capitals to start with. We include both in our baseline analysis and show that our findings are robust to dropping the Song (and any particular regime).

It should also be noted that both provincial boundaries and provincial capitals changed across regimes. This fact suggests that the administrative decision is made for the sake of national control rather than simply to reward or punish a certain city. If the ruler only wanted to reward or punish a city, he did not need to change the boundaries (which is likely to be costly). In Section 3, we will discuss the underlying logic of (re)locating provincial capitals.

2.2 Data

We construct a prefecture-level panel dataset from 980 to 2000. We focus on four of the six major regimes and do not study the Yuan (1279-1368) and the Republic (1912-1949) for two reasons: one is that the short-lived regimes are less suitable for our study because population sizes take time to adjust after a regime change. The other is that the prefecture-level data are less available for these two regimes.

One challenge is to deal with prefecture boundary changes across regimes. We first explain how we deal with this challenge. Then, we introduce the key variables in our analysis. We describe the underlying data for each variable in Online Appendix B.

Panel Construction Even though some prefecture boundaries are fairly stable, many also change across regimes. To construct a panel dataset, one needs to fix the boundaries and map the original data to these fixed boundaries. We employ two ways in our analysis: (i) using the boundaries for the 261 prefectures in 2000 in our baseline analysis since our readers are more familiar with these boundaries; and (ii) complementing our analysis grid-level analysis in which the 261 prefectures are divided into grids whose sizes can be smaller or larger than the prefecture in our baseline. If the prefecture (grid) contains the geocode of a provincial capital, we call it a capital prefecture (grid).

Next, we explain how we map the data across regimes to the same boundary, using population as an example.

Population and Urbanization Because population density is the most comprehensive data with which to measure long-term economic development, our baseline estimations employ population data for 11 years based on all the existing censuses – 980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964, and 2000. We are also able to access population data in years of 1880, 1953, 1982 and 1990. But one challenge we face is the uneven gap between periods and including these years would make the gap even more uneven. So we intentionally exclude them in our baseline and include them as a robustness check. Not surprisingly, the first census for 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 the population data to a fixed prefecture in 2000 or a grid as follows. Suppose a fixed prefecture i in 2000 comprises of fractions of two prefectures (1 and 2) in 1080. We first calculate the population density 1080 by prefecture boundary in 1080 and get $PopDensity_{1,1080}$ and $PopDensity_{2,1080}$. Second, we calculate the area of the factions included in the 2000 prefecture i , $Area_1$ and $Area_2$. Then, the population in 1080

for prefecture i is $PopDensity_{1,1080} * Area_1 + PopDensity_{2,1080} * Area_2$. Divided the population size by the prefecture area in 2000 (i.e., $Area_1 + Area_2$), we obtain the population density for prefecture i in 1080.

In Online Appendix B.1., we present a more general case with a figure to illustrate these steps and show that by such construction, prefecture boundary changes can be considered as measurement error in capital status of a prefecture, which likely leads to an underestimate of the importance of capital status.

We complement our population density measure with urbanization measures. Compared with population data, urbanization data are less systematically available, being accessible for only four of our 11 periods: 1580, 1820, 1964, and 2000. The 1580 data are estimated based on local gazetteers,⁹ the 1820 data are from CHGIS (2007), and the 1964-2000 information is taken from population censuses. These data are also originally available at the prefecture level in each regime. We map them to the fixed boundaries using the same way we deal with the population data.

Provincial Capitals and Prefecture Boundaries We use CHGIS (2007) information on the boundaries and provincial capitals from the Ming Dynasty to 2000 and digitize the information for 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 studying the Song (e.g., Mostern 2011). As expected, the variation in the provincial boundaries and capitals comes from cross-regime variation (i.e., they were set up at the beginning of each political regime).

With these data, we know the prefecture boundaries for each regime and the exact geolocation of each provincial capital. As mentioned above, we define a capital prefecture (grid) based on whether the prefecture (grid) contains the geocode of a 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. Different from population and urbanization data, these data are originally available at the grid level, which can be mapped to the fixed prefecture boundaries in our analysis. Thus, all these variables are calculated based on the fixed prefecture boundaries. For the continuous measures, we calculate the mean values for the (fixed) prefecture.

The eight geographic variables include log area of a prefecture, whether a prefecture contains a plain, whether it contains a major river, whether it is on the coast, as well as its

⁹We thank Cao Shuji, a leading scholar in Chinese population history at Shanghai Jiaotong University, for providing this information. The historical urbanization data in 1580 and 1820 are estimated based on the population living inside and outside walled cities.

average slope, elevation, longitude, and latitude. The five agricultural variables include the crop suitability of wheat, rice, fox millet, maize, and sweet potato (FAO GAEZ 2012), the first three being major old-world crops; the latter two introduced through the Columbian exchange.¹⁰

To enhance the comparability of prefectures, one may prefer within-province analysis. Because the dramatic historical shift in provincial boundaries precludes a straightforward within province comparison, we use the well-known nine physiographic macroregions defined by Skinner (1977): the north China plain, northwest, lower/mid/upper Yangtze, southeast coast, Lingnan, Yun-Gui, and Manchuria (see the map in Online Appendix B.2.). According to Skinner, these macroregions, which he defines based on major river drainage basins and other travel-constraining geomorphological features, are a better measure of markets because provincial boundaries emerged for political reasons rather than economic consideration.

Other Outcomes To understand the underlying factors for our main finding, we also examine how political hierarchy affects public employment (in the Qing and 2000) and transportation networks (in 1080, 1580, 1820 and 1990). These results help us understand what drives the politics-development link. We will discuss these measures in Section 6.

We present the summary statistics of the key variables in Table 2. To be clear, the variation in political status change stems from four major regime change (Song, Ming, Qing, and modern China). Within each regime, we can have multiple periods of population data to measure economic development of a prefecture.

3 Understanding the Choice of Provincial Capitals

The choice of provincial capitals is a complex issue affected by many factors. Some of the factors do not vary or change slowly over time. For instance, geography is likely to matter: having a river is useful as a provincial capital needs good access to water; elevation can also matter as being hilly may not be good for a capital city. It is possible that the importance of geography changes across regimes that leads to change in provincial capitals. There could also be time-varying shocks: regime changes might involve population size variation that affected capital status.

Our aim here is to focus on one general perspective that applies to all regimes and generates a possible instrument for provincial capital relocation. Specifically, we argue that

¹⁰See Nunn and Qian (2010) for a review of the exchange’s influence. Although our results are all robust to using Galor and Özaka (2016) caloric suitability as an alternative measure, we focus on crop-specific suitability for simplicity.

when locating provincial capitals, the central government in each regime faces the tradeoff between granting regional economy and keeping central control. Our argument is testable because we can empirically proxy the two considerations. Intuitively, if the ruler only cares about governing a province, the cost of within-province governance matters. Hence, it may be optimal to locate the provincial capital in the provincial centroid (to minimize the cost of collecting resources and information). However, the ruler may also care about keeping the provincial under control, which creates incentives to locate the provincial capital toward the direction of the national capital.

If our argument is valid, we expect to see relocation of provincial capitals because the new regime relocated the national capital and redivided the provincial boundary, which altered a prefecture’s access to other prefectures within the province as well as its access to the national capital. Below, we first formalize our argument with a simple algorithm. Then, show that our algorithm is a power predictor of provincial capital locations (in contrast with other factors like geography). Finally, we discuss the validity concerns of using the algorithm as an instrument.

3.1 A Simple Algorithm

Expressed formally, for the autonomy consideration, we can proxy the cost for a prefecture i to be the administrative center by its distance to all the other prefectures in the province (i.e., distance to the provincial centroid) in regime r : $\sum_{j=1}^N A_j D_{i,j,r}$, where $i = 1, 2, \dots, N$ indicate each prefecture in a province, $D_{i,j,r}$ indicates the distance from i to another prefecture j in the same province in regime r , and A_j is a weight variable – if A_j indicates the area of prefecture j , $D_{i,j,r}$ can be considered as distance from prefecture i to the provincial centroid. Similarly, we can proxy the consideration of central control by $\sum_{j=1}^N A_j D_{i,NationalCap,r}$, where $D_{i,NationalCap,r}$ indicates a prefecture’s distance to the national capital in regime r .

Thus, the central government’s choice of a capital for a province with prefecture $i = 1, 2, \dots, N$ is a solution to minimize the following specification, which we term “hierarchical distance” (denoted by *HierDist*):

$$\operatorname{argmin}_i \text{HierDist}_{i,r} \equiv \sum_{j=1}^N A_j D_{i,j,r} + \lambda_r \sum_{j=1}^N A_j D_{i,NationalCap,r} \text{ (autonomy vs. control)} \quad (1)$$

$$= (1 - \lambda_r) \sum_{j=1}^N A_j D_{i,j,r} + \lambda_r \left(\sum_{j=1}^N A_j D_{i,j,r} + \sum_{j=1}^N A_j D_{i,NationalCap,r} \right) \text{ (resource collecting)} \quad (2)$$

where λ_r is a weight value that captures the relative importance of two considerations. We

allow λ_r to change across regimes.

An alternative way to interpret the formulation is from the perspective of collecting resources and information: in equation (2), $\sum_{j=1}^N A_j D_{i,j,r}$ measures the cost of collecting resources and information within a province, and $(\sum_{j=1}^N A_j D_{i,j,r} + \sum_{j=1}^N A_j D_{i,NationalCap,r})$ measures the cost of first collecting resources within a province and then delivering it to the national center.

We attempt to keep the elements included in $HierDist_{i,r}$ as exogenous as possible. Thus, we choose to focus on geographical variables. In principle, one can enrich this algorithm by considering other factors like using population to measure A_j and considering transportation networks in calculating D . As we will show later, however, both population and transportation can be affected by political hierarchy.

As long as $\lambda > 0$, $HierDist_{i,r}$ implies that provincial capitals should be located away from the provincial centroid and toward the national capital. This prediction is confirmed by the empirical patterns. To see the pattern, we zoom in the case of Ming (the middle period in our data) and modern China in Figure 2(a) and Figure 2(b) (see other regimes in Online Appendix C.1.), where we mark the provincial capitals with red squares and provincial centroids with crosses. As shown in both maps, the squares (provincial capitals) tend to deviate from the crosses toward the national capital direction.¹¹ Thus, our simple algorithm based on arguably exogenous features appears promising as one predictor for provincial capital status. Next, we tests it more systematically.

3.2 Test the Role of Hierarchical Distance

Here, we test the association between $HierDist_{i,r}$ and 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 λ_r that delivers the highest prediction power in each regime. We find that λ_r varies from around 0.3 in the Song and Ming to around 0.2 to the Qing and modern China. Note that this choice is not critical for our analysis, and our findings are robust to alternative ways of specifying λ_r . In Online Appendix C.2., we present results using a fixed value for λ_r .

Depiction We plot the probability of being a provincial capital by a prefecture’s rank in the province in terms of its hierarchical distance. As panel (a) of Figure 3 shows, for the

¹¹Because the administrative hierarchy was set in the beginning of the regime, the location of Nanjing was more important than Beijing (which became the national capital after a coup) for the Ming.

prefectures that rank first (which vary across periods), the probability of being a provincial capital is around 0.40, whereas for those that rank second, it is around 0.24. This probability decreases as rank increases: once the rank is over 5, the probability becomes lower than 0.1; once it is above 10, the probability is close to 0. The nonlinear pattern also suggests a linear relation between log rank ($\ln RankHierDist$) and probability of being a capital when the rank is lower than 10, one that is confirmed by the pattern in panel (b). In Online Appendix C.2., we show that a similar when assuming $\lambda = 0.2$ for all regimes.

Estimations Motivated by Figure 3, we evaluate how hierarchical distance shapes provincial capital location systematically with both cross-sectional and panel specifications. Our cross-sectional analysis for each regime r is as follows:

$$ProvCap_{i,r} = \beta \ln RankHierDist_{i,r} + \theta \mathbf{X}_i + \pi_m + \epsilon_{i,r}, \quad (3)$$

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

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

$$ProvCap_{i,r} = \beta \ln RankHierDist_{i,r} + \alpha_i + \gamma_r + \theta \mathbf{X}_i \times \gamma_r + \pi_m \times \gamma_r + \epsilon_{i,r}. \quad (4)$$

We allow \mathbf{X}_i 's effects to vary across time by controlling for $\mathbf{X}_i \times \gamma_r$, and π_m 's 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 $\ln RankHierDist_{i,r}$ is strong predictor for provincial capital status (presented in Table 3A). In columns (1), (3), (5), (7), we find that this variable alone explains the variation in 8-16% of the variation in provincial capital status. In columns (2), (4), (6), (8), we add controls of geography and agriculture mentioned above as well as dummies for macroregions. Among these controls, geography does matter: the prefectures that contain rivers and plains are more likely to serve as a provincial capital. However, the explanation power of such a large set of controls together is marginally smaller than that of $\ln RankHierDist_{i,r}$ alone, as shown by the R -squared in the last two rows of Table 3A.

We obtain a similar pattern using the panel specification (presented in Table 3B). Here, the change in log rank in $HierDist_{i,r}$ is a strong predictor of provincial capital status change. If a prefecture's rank in hierarchical distance doubles, the probability of being a provincial

capital decreases by eight percentage points (around 80% of the mean probability).

3.3 Validity Concerns

We attempt to use $\ln RankHierDist_{i,r}$ as an instrument for provincial capital status in our later analysis. We face at least two important challenges in implementing this idea. The first is that provincial boundary changes offer one of the two sources of variation underlying $\ln HierDist_{i,r}$, but provincial boundary changes per se might affect a prefecture’s development; the second is that $\ln RankHierDist_{i,r}$ might affect regional development via channels beyond political status.

Provincial Boundary Changes While provincial boundary changes across regimes, we can still compare prefectures within unchanged boundaries. To this end, we calculate the intersections of provincial boundaries for all regimes ($\cap Prov_r$). In the specification controlling for these intersection fixed effects and their interactions with year fixed effects ($\cap Prov_r$) $\times \gamma_t$, we are comparing prefectures within unchanged polygons.

Exclusion Restriction Exclusion restriction is an important challenge. After all, access to provincial market alone or access to national center alone may affect economic development regardless of provincial capital status. While we cannot directly test the validity of the exclusion restriction, the two sources of variation in hierarchical distance make it possible to evaluate this concern more explicitly.

The variation in hierarchical distance stems from two sources: the relocation of national capitals and the redrawing of provincial boundaries, both of which come with regime changes. We can calculate the two sources of variation separately by examining $\ln Rank$ in $\sum_{j=1}^N A_j D_{i,j,r}$ and $\ln Rank$ in $\sum_{j=1}^N A_j D_{i,NatCap,r}$. The correlation of these two variables is weak, with a correlational coefficient of 0.14, suggesting that the two sources of variation are fairly orthogonal to each other.

This feature is useful for our analysis later when using hierarchical distance to predict provincial capital status. Conceptually, for an omitted channel to explain our findings, it has to be strongly correlated with both components, but these two components themselves are not that correlated. In Section 5.3, we will also employ $\ln Rank$ in $\sum_{j=1}^N A_j D_{i,j,r} \times \ln Rank$ in $\sum_{j=1}^N A_j D_{i,NatCap,r}$ to predict provincial capital status (while controlling for the two components), where we can explicitly control for the two elements.

In addition, we will conduct several placebo tests including examining the subgroup of prefectures that have little chance to become provincial capitals and calculating placebo hierarchical distance before and after the national capital changes.

4 The Economic Consequences: Descriptive Evidence

How does the political hierarchy matter for economic development? Here, we present two sets of descriptive patterns to motivate our empirical estimations.

A Case of Two Periods: 980 & 2000 Before analyzing our 11-period panel data, we present descriptive patterns based on the 980 and 2000 data, a two-period structure that 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, the x -axis indicates the standardized log population density in 980, while the y -axis indicates the standardized log nightlight density in 2000. The figure suggests that a prefecture’s economic status changes along with its political status. Specifically, the four crosses indicate the average economic status of the four prefecture groups. As shown, an average “no-no” prefecture group was close to mean in both periods and an average “yes-yes” prefecture was above the mean in both periods. In contrast, an average “no-yes” prefecture was 0.3 standard deviation below the mean in 980 yet became one standard deviation above the mean in 2000, indicating that gaining capital status is correlated with better economic development. An average “yes-no” prefecture was 0.5 standard deviation above mean and comparable to a “yes-yes” prefecture in 980 (when both were provincial capitals), but it became close to the mean and similar to a “no-no” prefecture in 2000 after losing capital status.

Pooling All Changes Together: 980-2000 Analogous to the analysis above, we can divide the prefectures into four groups every two regimes. This way, we can further visualize the four groups period by period. To compare prefectures with similar characteristics, we first regress log population density on all 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 5.

Consistent with the pattern in Figure 4, the “no-no” prefecture group was close to mean (in the macro region) in all periods, where the “yes-yes” prefecture group was 40% above mean in all periods. The gaining group (i.e., the “no-yes”) was similar to its “no-no” peer group before status change but become 48% higher than its “no-no” peer group after gaining capital status; in contrast, the losing group (i.e., the “yes-no”) was about 38% higher than its “no-no” peer group before status change but become comparable to the “no-no” peer group after losing capital status.

5 The Economic Consequences: Estimation Results

We document how political hierarchy shapes regional development in this section. We first present our two empirical strategies in Section 5.1. Then, we describe our main findings on regional development in Section 5.2, after which we discuss checks for each strategy in Section 5.3.

5.1 Empirical Strategies

We employ two empirical strategies to identify the economic consequences of political hierarchy: a difference-in-differences (DID) strategy and an instrument variable (IV) strategy. Below, we present these two strategies and discuss their empirical assumptions.

The Difference-in-Differences Strategy The first is a difference-in-differences strategy as follows:

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

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 rates. The other variables are the same as those in equation (3) except that we can have multiple years of observations with a regime r . Thus, we use t to indicate the year of observation here.

An important assumption of this strategy is that provincial-capital prefectures are comparable to non-capital prefectures before capital status change. We will validate this assumption in several ways in Section 5.3.

The Instrumental Variable Strategy A second strategy is to use the hierarchy distance derived above as an instrument to predict provincial capital status. The first-stage and

second-stage specifications are as follows:

$$Capital_{i,t} = \delta \ln RankHierDist_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \quad (6)$$

and

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

where X_i includes all the controls in our difference-in-differences analysis.

While $\ln RankHierDist_{i,t}$ is a strong predictor of provincial status (as we have shown in Section 3), we have to be cautious about its validity as an instrument. In particular, it has to satisfy two conditions: (i) the change in $\ln RankHierDist_{i,t}$ for a prefecture is orthogonal to its own characteristics, and (ii) $\ln RankHierDist_{i,t}$ satisfies the exclusion restriction, namely that it affects regional development due to political status change, not omitted channels. In particular, we are concerned that a better rank in hierarchical distance captures some economic advantages not due to political status.

Condition (i) is not a critical challenge for us because the change in a prefecture’s hierarchical distance stems from regime change rather than any prefecture’s own characteristics. We will confirm this assumption by examining past levels and changes in population density in Section 5.3.

Condition (ii) is more subtle. As mentioned in Section 3.3, the two components underlying $\ln RankHierDist_{i,t}$ whose correlation is weak are useful. We will conduct several sets of tests by exploiting the two components of the instrument in Section 5.3.

Compared with the DID strategy, the IV strategy helps us better deal with time-variant variables. Meanwhile, the IV estimate delivers a local treatment effect. When using the IV design, our estimate is driven by the prefectures whose political status changes due to changes in their hierarchical distance.

5.2 Results on Regional Development

We use log population density during 980-2000 as the primary measure of regional development. Since we fix prefecture boundary in this analysis, the coefficient can also be interpreted as the impact on population size. In addition, we complement the analysis with urbanization data, which is only available for the latter half millennium.

Our DID estimates based on 11-period data indicate that provincial capital status is associated with a 40-50% higher population density. Column (1) of Table 4 presents the result with only prefecture and year fixed effects, where we obtain an estimate of 0.62. Once we

control for eight geographical variables and their interaction with γ_t , we obtain an estimate around 0.46 (columns (2)). After adding five-type crop suitability, regional dummies, and their interactions with γ_t in columns (3)-(4), the estimates are around 0.47.

A key challenge for our difference-in-differences analysis is that the capital status could change due to omitted variables that might affect development. Based on the changes in coefficient and R-squared across specifications in Table 4, we can assess the importance of omitted variable bias using the method provided by Altonji, Elder, and Taber (2005) and Oster (2017). We find it difficult for omitted variable bias to explain the results in Table 4: after considering possible biases, we find an lower bound of β to be 0.327 (see the step-by-step calculation in Online Appendix D.1.) We present the bounding estimate in the last row of Table 4.

Our IV strategy also confirms the importance of provincial capital status. Column (5) shows the reduced-form estimate: an increase in the rank of hierarchical distance is strongly negatively correlated with population density. When separating the sample into ever-capital prefectures and never-capital prefectures in columns (6)-(7), we find that hierarchical distance matters greatly for the population density in the ever-capital prefectures but matters little for the never-capital prefectures. This evidence suggests that political status is a critical channel for hierarchical distance to matter.

Column (8) reports the IV estimate, with a F -statistics of 131.9, implying that weak instrument is not a critical concern. Compared with the DID estimates, our IV estimate (0.79) is slightly higher, partly because this estimate is local to those whose political status is affected by their hierarchical distance. Finally, in column (9), we further include the intersections of provincial boundaries for all regimes and their interactions with year fixed effects ($\cap Prov_r \times \gamma_t$), where we are comparing prefectures within unchanged boundaries. If anything, the coefficient becomes a bit larger.

We employ urbanization data as an additional measure of development. These data are only available for the latter half of the millennium (1580, 1820, 1964, 2000). Thus, one can also consider the finding as a check on whether our results hold after excluding data from the first half of the millennium. As shown in columns (1)-(3) of Table 5, the DID and IV estimates show that provincial capital status increases a prefecture's urbanization rates by 11-18 percentage points, a large increase compared with the mean urbanization rate (15 percentage points). We further log urban population (in columns (4)-(6)) and log rural population as outcomes (in columns (7)-(9)) and find that the increase in urban population is about five times that for rural population. This finding further confirms that political status affects economic development not only in terms of population size but also economic activity.

5.3 Challenges to the Two Strategies

Both difference-in-differences and IV analyses face possible threats. Below, we summarize the conceptual issues and key results, while relegating some empirical details to Online Appendix D and E.

Omitted Variable Checks for Our DID Analysis The assessment of the omitted variable bias suggests that it is difficult to use omitted variables to explain our DID finding. Here, we present three additional checks. These results not only further mitigate the concern of omitted variables but also deliver some new insight, particularly on the importance of losing political status.

(i) Gaining vs. Losing Political Status

We separate capital status loss from capital status gain and find that both have a sizable impact (see Online Appendix D.2.): gaining capital status is associated with around 47% increase in population density, whereas losing capital status is associated with a 39% decrease in population density. By design, the difference-in-differences estimates should be interpreted as changes (rather than levels). But as indicated in the descriptive pattern in Figure 5, the losing group was about 40% higher in population density than its non-capital peers. Our estimates further suggest that the losing group became not much better than its non-capital peers after losing status. The finding on losing capital status suggests that the omitted variable concern may not be essential; otherwise, we would observe that losing capital status matters little.

We also include lagged population density in some of the specifications. We find a negative coefficient on lagged log population density, suggesting a convergence pattern. Meanwhile, including the lagged population density slightly decrease the coefficients of gaining and losing capital status.

(ii) Ever-Capital Prefectures

The 63 prefectures that were a provincial capital at least once serve as a useful subsample for examining the importance of capital status because the variation within them comes only from gaining or losing such status, with no omitted variables preventing any one prefecture from becoming a capital. We use this information in two ways. The first is to focus on this subsample; the second is to further control for $EverCapital_i \times \gamma_t$ in our analysis, which allows this subgroup to exhibit trends different from the other prefectures. Both methods deliver a pattern similar to those in Table 4 (see Online Appendix D.3.), suggesting that our baseline finding cannot simply be explained by the omitted variables affecting whether a prefecture could ever serve as a provincial capital.

(iii) Checking Pre-trends

We conduct a period-by-period analysis where we define periods relative to the period before the change in capital status (denoted by period0). This specification allows us to check whether provincial capitals were already on a different development path before their political status change. 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.

Replacing $Capital_{i,t}$ with $\sum_{\tau=-75+}^{300+} \beta_{\tau} Capital_{i,\tau}$ and using period0 as the reference group, we find no significantly different trends in population density between the capital group and other prefectures until the status changes. These results are presented in Online Appendix D.4. Figure 6 further visualizes the findings, where the lines connect the estimates and the dotted lines indicate the 95% confidence intervals. As shown, both the increase and the decrease in population density occur only after gaining and losing capital status. Even though we cannot say the exact number of years for the effects to occur, the data tell us that they occur within 75 years after status change.

Besides provides further support for our DID finding, these results also show that losing capital status leads to a great decline of regional development. In other words, while a prefecture enjoyed economic advantage due to its political status but this advantage did not necessarily persist once it moved down the ladder of the political hierarchy.

Measurement Checks for Our DID Analysis Our data come from the censuses conducted by the central government, which are of low-frequency with uneven gaps over the millennium. While our dynamic analysis in Figure 6 has partially addressed the concern of uneven gaps, we also examine the impacts by period more carefully to ensure that our findings are not driven by a particular period.

As explained above, we fix the prefecture boundaries to build a panel dataset in our earlier analysis. We conduct two sets of grid-level analysis to show that our findings are not driven by a specific way of fixing the boundary – one can also fix the prefecture boundary by any regime.

(i) Using Subperiods of the Data

First, our results hold if we include four additional years (1880, 1953, 1982, 1990) that would make the gaps more uneven. Including them delivers an estimate close to our baseline (reported in Online Appendix D.5.).

Our finding also holds if we restrict our sample to be of roughly equal length. We use the data in 980, 1393 and 1820 to construct a subsample (of a gap around 400 years)

and those in 980, 1102, 1393, 1580, 1820 and 2000 to construct another subsample (of a gap around 120-180 years). The estimates from these two subsamples are also close to our baseline estimates, suggesting that our results are not driven by the gaps (reported in Online Appendix D.5.).

Finally, our finding changes little if we exclude any specific regime. In particular, our finding also holds with and without data in modern China (also discussed in Online Appendix D.5.).

(ii) Grid-level Analysis

We conduct a 1-degree \times 1-degree grid analysis and a 2-degree \times 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 \times 1-degree grids and 94 2-degree \times 2-degree grids. These two sets of grid-level analysis again generate patterns comparable with our prefecture-level analysis (Online Appendix D.6.), indicating that our findings are robust to alternative boundary definitions.

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. This assumption can be confirmed by two sets of empirical results discussed in Online Appendix E.1.

In the first, we show that $(\Delta \ln RankHierDist_{i,t})$ is not significantly correlated with past levels of log population density ($\ln PopDensity_{i,t-1}$ and $\ln PopDensity_{i,t-2}$) or past changes in log population density ($\Delta \ln PopDensity_{i,t-1}$ and $\Delta \ln PopDensity_{i,t-2}$). Across all specifications, the correlation between $\Delta \ln RankHierDist_{i,t}$ and past population density (or its change) is close to zero. These results indicate that the change in hierarchical distance rank is not driven by any change in population density.

In the second test, because it seems unlikely that the central government would intentionally increase a prefecture’s hierarchical distance to make it lose its capital status, we separate the impacts of changed hierarchical distance on losing versus gaining status. Our seemingly unrelated regressions show that the impact of hierarchical distance rank has a similar magnitude for both cases: a one log point increase (and decrease) in hierarchical distance reduces (and enhances) the probability of being a provincial capital status by 2.7 percentage points (and 3.2 percentage points).

Exclusion Restriction of the Instrument We are concerned about the exclusion restriction. After all, better access to national capital or to the rest of the province is likely to have a direct effect on regional development. As mentioned above, the fact that the

correlation between two components of our instrument is weak is important for mitigating this concern, because it is difficult for an omitted channel to be highly correlated with both elements.

The finding that hierarchical distance is not significantly correlated with population density for the never-capital prefectures (column (7) of Table 4) is suggestive: if hierarchical distance affects population density regardless of capital status, we expect to see that it also matters within the never-capital prefectures. Next, we present three additional sets of analysis to further investigate the validity of the instrument.

(i) *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*

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 $\ln \text{Rank in } \sum_{j=1}^N A_j D_{i,j,t} \times \ln \text{Rank in } \sum_{j=1}^N A_j D_{i,NatCap,t}$ as an instrument while controlling for the independent effects of the two components. These results are presented in Online Appendix E.2.

Two findings are worth mentioning. First, both $\ln \text{Rank in } \sum_{j=1}^N A_j D_{i,j,t}$ and $\ln \text{Rank in } \sum_{j=1}^N A_j D_{i,NatCap,t}$ are indeed strongly correlated with population density; however, when using their interaction as the instrument, we do not find that either part alone has an additional direct effect. 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 4). This difference further suggests that our definition of hierarchical distance better captures the political logic of provincial capital choice by the central government.

(ii) *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. Because these placebo measures are correlated with our instrument, some of them are also correlated with the probability of being a capital. However, 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 (Online Appendix E.3.).

(iii) *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 $D_{i,NationalCap}$ in equation (1) with $D_{i,Market}$. 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 (Online Appendix E.4.). Thus, this finding shows that the hierarchical distance to the political center (instead of national economic centers) is the driver of our finding on capital status.

Taken together, even though we cannot reject the concern of exclusion restriction in theory, these results indicate that capital status is a critical channel through which hierarchical distance rank can affect economic development. Thus, the political logic of provincial capital relocation we uncover is a useful predictor of capitals status.

6 Interpretations

We have documented that political hierarchy has large impacts on spatial development measured by population density and urbanization. What explains the politics-development link? The answer is likely to be multifaceted. Instead of enumerating each factor in both history and today, we consider two broad interpretations: the politics hypothesis and the market hypothesis. Below, we explain the two hypotheses in Section 6.1 and present empirical evidence in Section 6.2.

6.1 The Politics Hypothesis and the Market Hypothesis

One interpretation for why provincial capital status matters is the role of public employment (and its spillover), which we call “*the politics hypothesis*”. This hypothesis has long been discussed by social scientists (e.g., known as “Oriental cities” following Weber (1921) or “parasite cities” following Bairoch (1985)). The historical and sociological discourse often emphasizes that merchants and artisans moved to political centers to serve the bureaucrats. One implication of this hypothesis is that such regions are usually consumption-intensive and do not necessary enjoy higher productivity. To examine this hypothesis, we will look at the link between provincial capital status and public employment.

A different interpretation is that political hierarchy also affects the allocation of production factors, which we term “*the market hypothesis*”. This interpretation is also motivated by qualitative observations in both history and today. For instance, in history, the construc-

tion 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 express way network, the official objective was “to connect all provincial capitals and all cities with a population of at least 500,000”. To examine this hypothesis, we will investigate the link between a prefecture’s provincial capital status and its position in the transportation networks.

6.2 Empirical Evidence

Change in Public Employment We measure public offices in 1776 and 2000. The measure in 2000 comes from the population census and refers to the employment in all administrative and social organizations, and that in public institutions (e.g., physicians and teachers in public hospitals and schools). The historical censuses do not have information on occupations. To get the measure for bureaucrats in history, we digitize 120-volume government records known as the Complete Directory of Qing Officials and use the information for 1776. A limitation of these records is that they only include individuals holding offices and hence miss a large amount of employees who did not hold formal government positions. 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 for our panel analysis (where a higher value of rank indicates a smaller size of public employment), based on both DID and IV methods.

Public offices move with provincial capitals. In columns (1)-(4) of Table Table 6, we present the cross-sectional findings on log population and public employment in 1776 and 2000. In column (5)-(6), we show that change in provincial capital status is associated with change in the rank of public employment. Based on these coefficients, the importance of provincial capital status is comparable to moving from being an average prefecture to top 35-15 in terms of public employment.

These estimates enable us to conduct back-of-the-envelope exercises to understand the contribution of public employment. For instance, in 2000, based on the census data and estimates by other scholars, the public employment size is estimated to be around 10% of total labor force or 6-7% of total population.¹² In column (4), we find the coefficient of provincial capitals on log public employment and log population are 1.15 and 0.68. Thus, public employment alone can explain 0.08 ($1.15 \times 7\%$) out of 0.68. We thus need a large multiplier for public employment to explain our findings on regional development.¹³ While

¹²see the PIIE estimate: <https://www.piie.com/blogs/china-economic-watch/shrinking-leviathan-state-employment-china-looms-smaller-expected>

¹³The existing estimates of the spillover effect of public employment on total private employment are less than one. Using English data at the Local Authority level for 2003-2007, Faggio and Overman (2014) find that public sector employment has no identifiable effect on total private sector employment. Exploiting the

we do not have data to estimate the multiplier, we turn to important market channels to show that politics also shapes production factors.

Change in Transportation Networks Transportation networks experience great changes across regimes, due to two sets of reasons. First, it is costly to maintain routes. Due to the lack of maintenance, many land routes disappear; several parts of the Grand Canal were ruined for a long period (Brook 1998). Second, when a regime replaced the previous one, the ruler decided which parts of the transportation networks to be reconstructed, which is likely to depend on the prefectures’ relative importance in the political hierarchy. Even if some of the old routes were kept, a prefecture’s centrality in the network got altered by the reconnected or newly-built routes. As a result, the regime change in China provides us a rare opportunity to systematically investigate changes in the transportation networks.

The historical transportation networks were comprised of the Grand Canal, which connected various waterways, and a state courier system (supported by many post offices). Both systems were designed by the central government. The aim of transportation networks was primarily political: to maintain an adequate flow of information, revenues, and personnel on which the state relied (Brook 1998). Meanwhile, commoners were also allowed to use these routes (Zang 1997). Thus, by facilitating the movement of goods and people, these networks contributed to economic development. Modern transportation is much more complicated than their historical counterpart. We choose to focus on railroad networks because they are monopolized by the state.

To trace how the transportation networks evolve over time, we digitize roads and waterway maps for three historical periods (represented by specific years) – the Song (1078), Ming (1587), and Qing (1820) dynasties– and the railroad map for the People’s Republic (1990).¹⁴

Although we can easily count the number of landways, waterways, and railways in a prefecture, we also want to account for the relative importance of different links in the transportation network. To this end, we employ a centrality measure, defined for each prefecture i as follows:

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

relocation of the German federal government from Berlin to Bonn in the wake of the Second World War, Becker, Heblich, and Sturm (2018) find that each additional public sector job destroyed around 0.2 jobs in industries and created just over one additional job in other parts of the private sector.

¹⁴The 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 comes from CHGIS (2007).

where $d_{i,j}$ indicates shortest distance between i and j in the transportation networks. One can also weight the distances by population and gets a measure similar to market potential (see more discussion on market potential and market access in Redding and Sturm (2008) and Donaldson and Hornbeck (2015)). We focus on the unweighted measure to highlight a region’s position in the transportation networks. As expected, the change in market access follows the change in spatial centrality.

In practice, $d_{i,j}$ comprises of three parts: d_{i,N_i} (and d_{j,N_j}) indicate the straight line distance from prefecture i (and prefecture j) to the network (point N_i and N_j); d_{N_i,N_j} indicates the minimum distance between point N_i and point N_j within the network. Following Fogel (1964) and Donaldson and Hornbeck (2015), we allow for an adjustment factor of 1.4 between the shortest straight-line distance and kilometers traveled and assume that the transportation cost is four times high without the network, i.e., θ takes a value of 5.6. As in this literature, our results are not sensitive to the choice of θ .

Not surprisingly, a prefecture’s centrality in the transportation networks exhibits some persistence. Our focus, however, is to test whether changes in capital status lead to changes in spatial centrality. Similar to our analysis on public employment, we first examine 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 to study the impact of capital status change. Again, a higher value of rank indicates lower centrality.

As shown in columns (1)-(4) of Table 7, in each regime, provincial capital status is strongly associated with higher centrality. Columns (5)-(6) report the DID and IV estimates: again, change in provincial capital status is strongly correlated with change in the rank of spatial centrality. Based on these coefficients, the importance of provincial capital status is comparable to moving from being an average prefecture to top 60-40 in terms of centrality.

Thus, we find that both the political channel and the market channel are at work in explaining a prefecture’s rise and fall in economic development due to its position in the political hierarchy.

7 Conclusion

Understanding administrative hierarchy is a complex issue. It is not easy to find a suitable empirical setting to investigate the underlying logic and evaluate its consequences. In this study, we take advantage of the history and politics in China. We show that the administrative hierarchy shapes China’s spatial development over the millennium. Underlying our finding, however, both political and market factors are at work.

It is widely observed that the development of Chinese cities is highly correlated with

their status in the political hierarchy, it is thus no exaggeration to say that one cannot understand the urban landscape in this country without considering its political roots.

Broadly speaking, our study serves as new evidence on how politics shapes economic geography. In the process of understanding the political economy of economic geography, we complement existing literature focusing on the persistence aspects of these factors. We hope that the dynamic variation we document can be useful for other researchers interested in understanding the evolution of these factors.

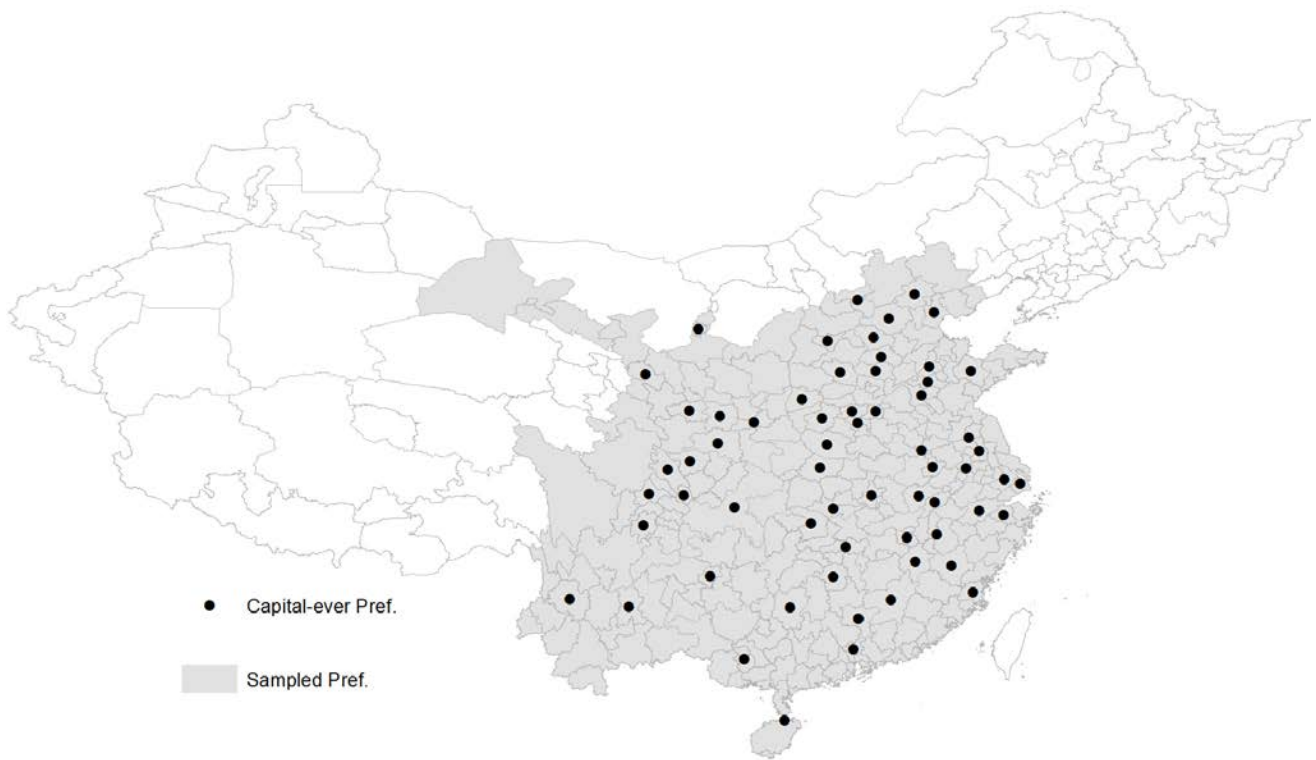
References

- [1] Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001), “The Colonial Origins of Comparative Development: An Empirical Investigation,” *American Economic Review* 91 (5): 1369-1401.
- [2] Ades, Alberto F., and Edward L. Glaeser (1995), “Trade and Circuses: Explaining Urban Giants,” *Quarterly Journal of Economics* 110(1): 195-227.
- [3] Allen, Treb, and Dave Donaldson (2019), “The Geography of Path Dependence,” Working Paper.
- [4] Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005), “Selection on Observed and Unobserved variables: Assessing the Effectiveness of Catholic Schools,” *Journal of Political Economy* 113 (1): 151-184.
- [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] Bolton, Patrick, and Gerard Roland (1997), “The Breakup of Nations: a Political Economy Analysis,” *Quarterly Journal of Economics* 112(4): 1057-1090.
- [9] Brook, Tim (1998), “Communications and Commerce,” *The Cambridge History of China*, 8(Pt 2), 579-707.
- [10] Campante Filipe and Do Quoc-Anh (2014), “Isolated Capital Cities, Accountability, and Corruption: Evidence from US States,” *American Economic Review* 104 (8): 2456-2481
- [11] Chang, Chung-li (1955), *The Chinese Gentry: Studies on Their Role in Nineteenth-century Chinese Society*. University of Washington Press.
- [12] Chao, Kang (2006), *Essays on the History of Chinese Urban Development (Zhongguo chengshi fazhanshi lujin)*. Beijing, New Star Press.
- [13] Cheng, Kuang-Yu, and Sheng-Mo Hsu (1980), *Historical Atlas of China (Zhongguo Lishi Ditu)*, Taiwan: Chinese Culture University Press.
- [14] CHGIS (2007), Harvard Yenching Insitute.
- [15] Dell, Melissa (2010), “The Persistent Effects of Peru’s Mining Mita,” *Econometrica* 78(6): 1863-1903.

- [16] Diamond, Jared M (1998), *Guns, Germs and Steel: a Short History of Everybody for the Last 13,000 Years*. Random House.
- [17] Donaldson, Dave, and Richard Hornbeck (2016), "Railroads and American Economic Growth: A 'Market Access' Approach," *Quarterly Journal of Economics* 131(2): 799-858.
- [18] Davis Donald and David Weinstein (2002), "Bones, Bombs, and Break Points: The Geography of Economic Activity," *American Economic Review* 92(5):1269-1289.
- [19] Davis, James, and Vernon Henderson (2003), "Evidence on the Political Economy of the Urbanization Process," *Journal of Urban Economics* 53(1): 98-125.
- [20] 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.
- [21] Faggio, Giulia, and Henry Overman (2014) "The Effect of Public Sector Employment on Local Labour Markets," *Journal of Urban Economics*, 79: 91-107.
- [22] FAO (2012), GAEZ (Global AgroEcological Zones), <http://fao.org/Ag/AGL/agll/gaez/index.htm>.
- [23] Galor, Oded, and Ömer Özaka (2016), "The Agricultural Origins of Time Preference," *American Economic Review* 106(10): 3064-3103.
- [24] Ge, Jianxiong (2000), *History of Population in China*. Shanghai: Fudan University Press.
- [25] Glaeser, Edward L., Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer (2004), "Do Institutions Cause Growth?," *Journal of Economic Growth* 9(3): 271-303.
- [26] Greif, Avner, and Guido Tabellini (2017), "The Clan and the Corporation: Sustaining Cooperation in China and Europe," *Journal of Comparative Economics* 45(1): 1-35.
- [27] Guy, Kent (2010), *Qing Governors and Their Provinces: the Evolution of Territorial Administration in China, 1644-1796*. University of Washington Press.
- [28] Heng, Chye Kiang (1999), *Cities of Aristocrats and Bureaucrats: the Development of Medieval Chinese Cityscapes*. NUS Press.
- [29] Li, Xiaojie (2011), *Historical Boundaries and Territorial Administration in China (Jiangyu yu zhengqu)*. Nanjing: Jiangsu chubanshe.
- [30] Liang, Fang-zhong (1980), *Historical Statistics on Hukou, Land and Land Tax of China (Lidai Hukou, Tudi, Tianfu Tongji)*. Shanghai: Zhonghua Book Company.
- [31] Michalopoulos, Stelios, and Elias Papaioannou (2013), "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica* 81(1): 113-152.
- [32] Morris, Ian (2013), *The Measure of Civilization: How Social Development Decides the Fate of Nations*. Princeton University Press.
- [33] Mostern, Ruth (2011), *Dividing the Realm in Order to Govern: The Spatial Organization of the Song State (960-1276 CE)*. Harvard University Press.
- [34] Nunn, Nathan, and Nancy Qian (2010), "The Columbian Exchange: A History of Disease, Food, and Ideas," *Journal of Economic Perspectives* 24(2): 163-188.
- [35] Nunn, Nathan (2009), "The Importance of History for Economic Development," *Annual Review of Economics* 1(1): 65-92.
- [36] Oster, Emily (2017) "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*: 1-18.
- [37] Redding, Stephen J., and Daniel M. Sturm (2008), "The Costs of Remoteness: Evidence from German Division and Reunification," *American Economic Review* 98(5): 1766-1797.

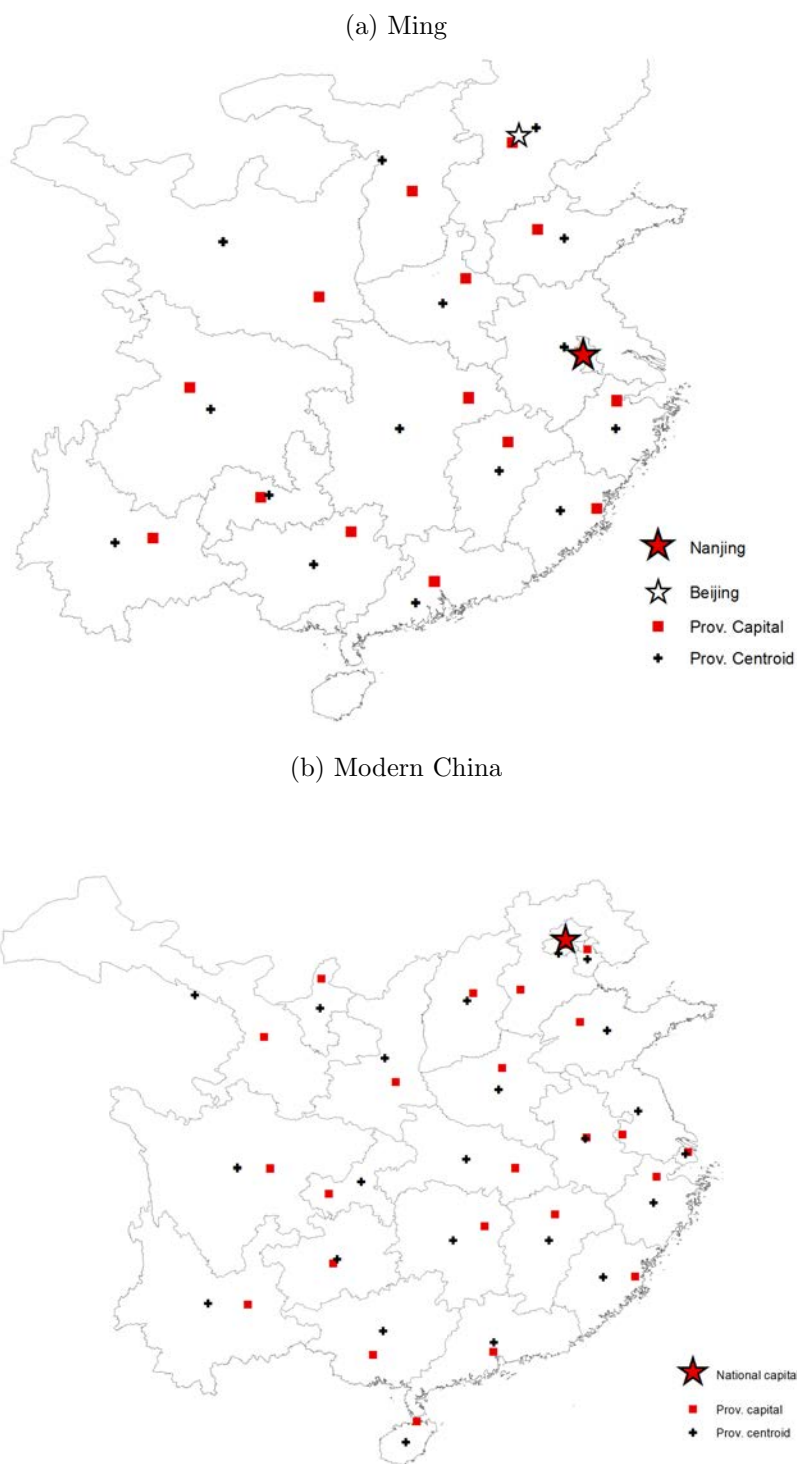
- [38] Rowe William (1993), “Introduction: City and Region in the Lower Yangzi”, in *Cities of Jiangnan in Late Imperial China*, edited by Linda Cooke Johnson, SUNY Press.
- [39] Skinner, G. William, ed. (1977), *The City in Late Imperial China*, Stanford University Press
- [40] Treisman, Daniel (2007), *The Architecture of Government: Rethinking Political Decentralization*. Cambridge University Press.
- [41] Weber Max (1921)[1958], *The City*. reprint, New York: Free Press.
- [42] Zang, Rong (1997), *Courier Route and Station in Historical China* [Zhongguo Gudai De Yizhan Yu Youchuan]. The Commercial Press.
- [43] Zhang, Tingyu (1739)[1974], *The History of the Ming Dynasty (Ming Shi)*. Qianlong edition, reprint, Beijing: Zhonghua shuju.
- [44] Zhou, Zhenhe (1998), *The Changes of Administrative Division in China (Zhongguo lidai xingzheng quhua de bianqian)*. Beijing: Shangwu chubanshe.

Figure 1: Ever-Capital Prefectures



Notes: The shaded area indicate 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.

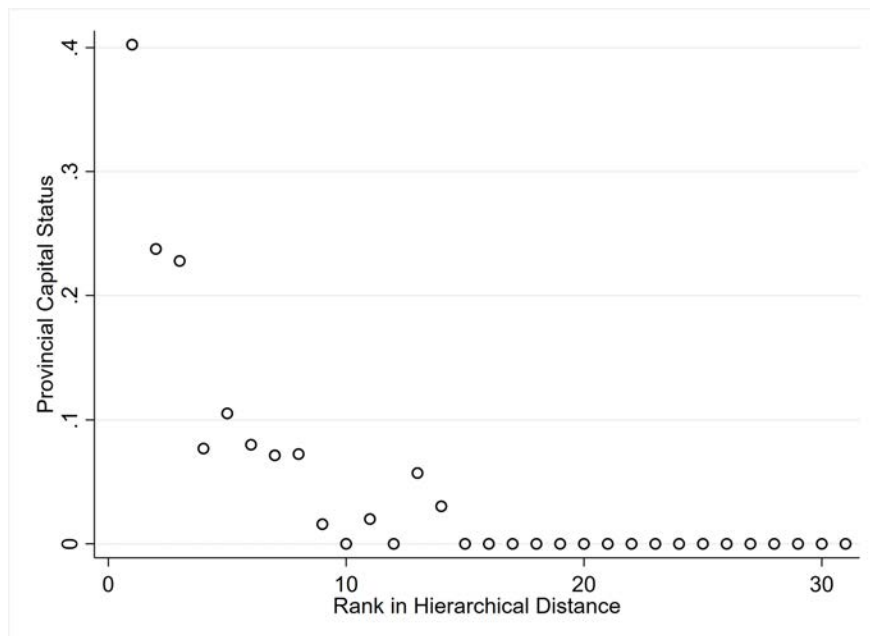
Figure 2: Illustrating the Idea Underlying Hierarchical Distance



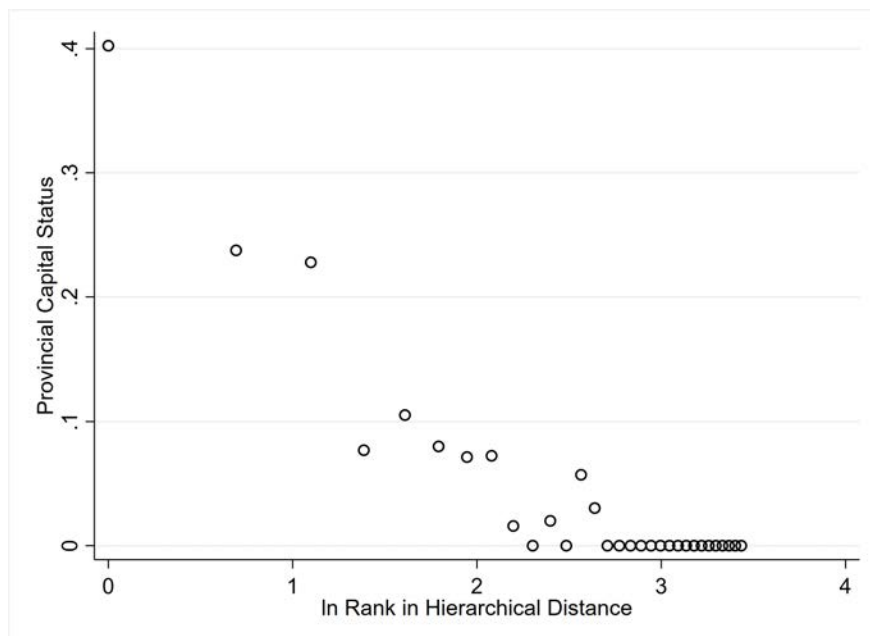
Notes: This figure illustrates the logic underlying the hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province) using the case in 1580 and 2000: provincial capitals tend to be located away from provincial centroid toward the national capital.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

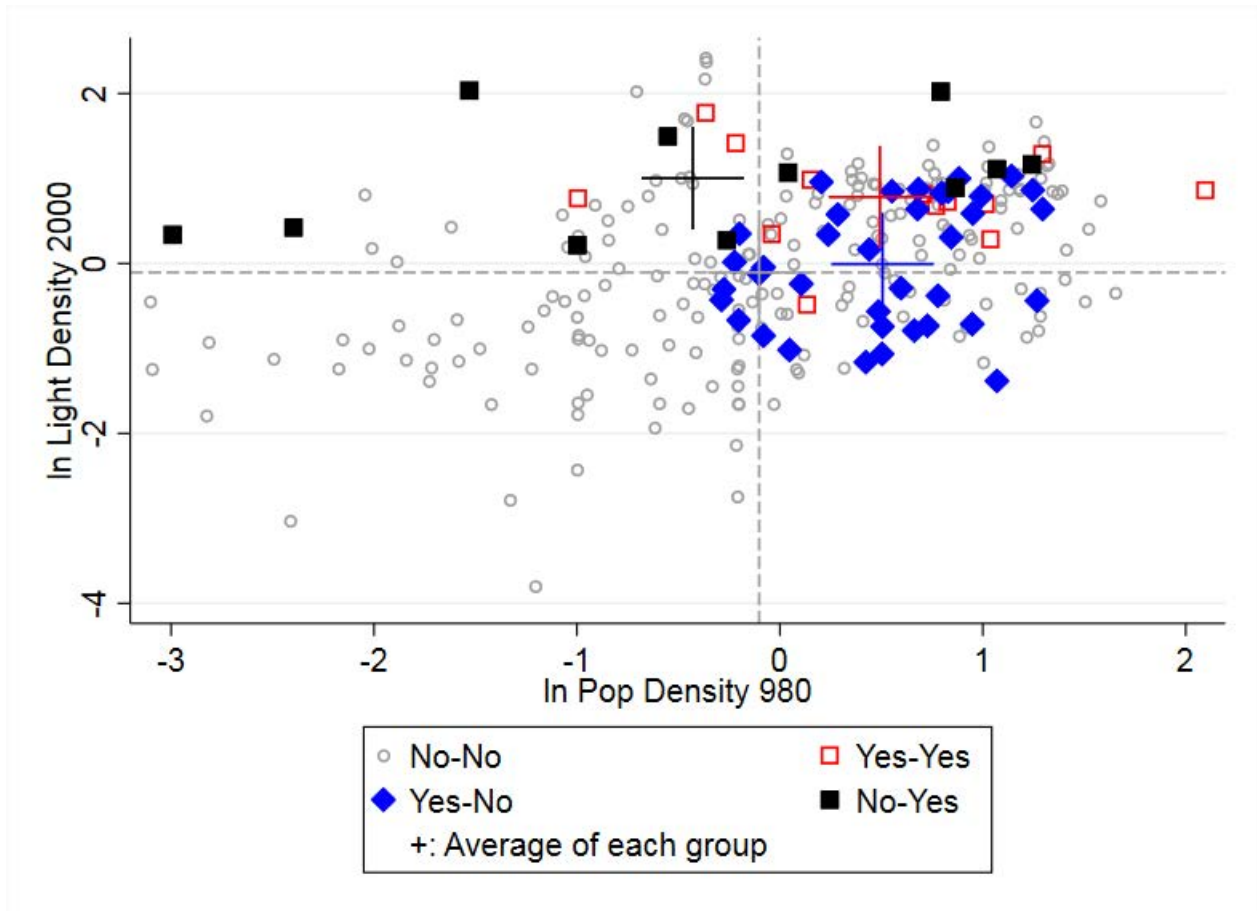


(b) ln Rank in Hierarchical Distance vs. Prob. of Being Capital



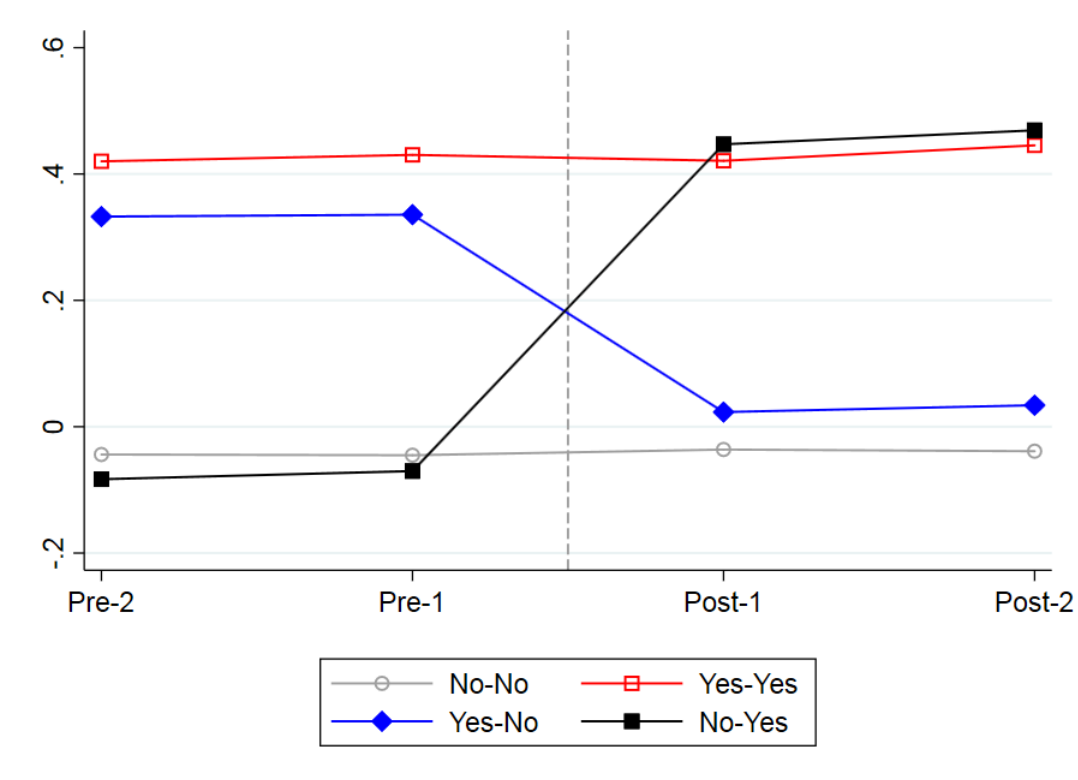
Notes: The figure shows a strong negative correlation between a prefecture's rank in hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province) within a province and its probability of being a provincial capital.

Figure 4: Descriptive Patterns I: 980 vs. 2000



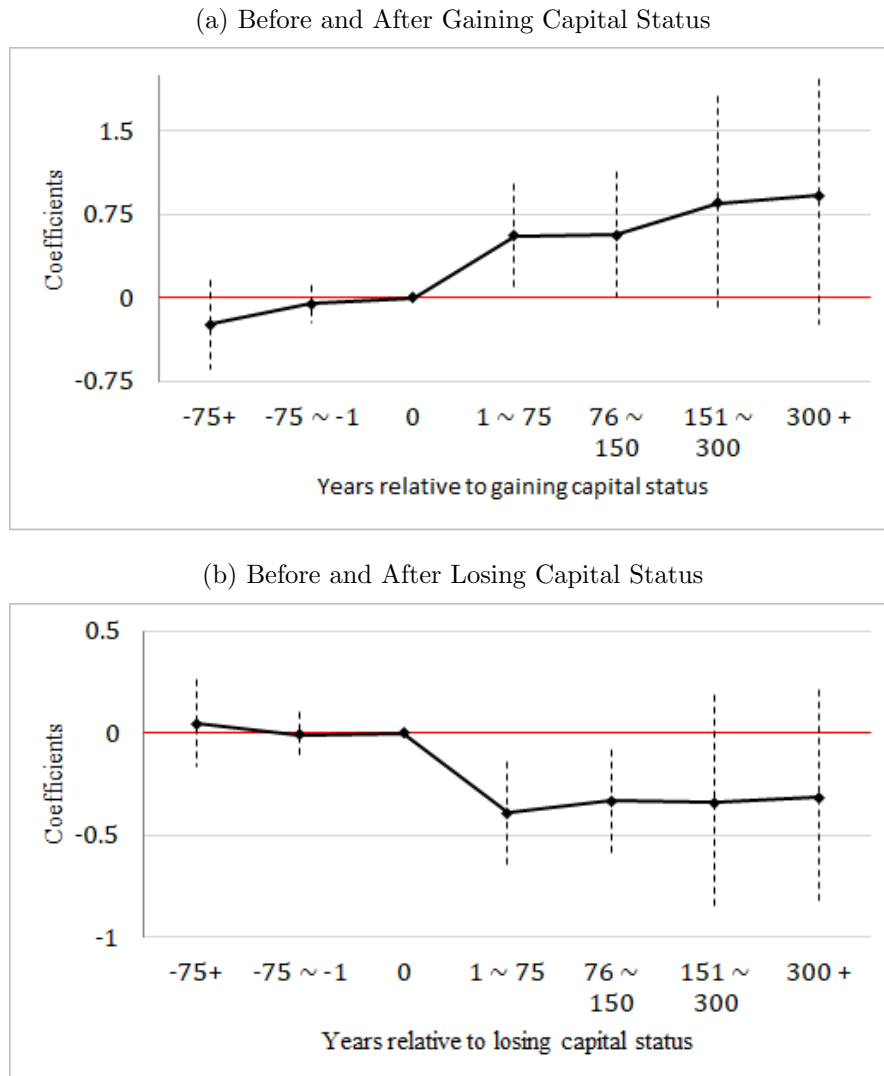
Notes: The figure shows that gaining and losing capital status is associated with a systematic change in a prefecture's economic performance (measured by population density in 980 and light density in 2000). An average "no-yes" prefecture was below the mean in 980 yet became above the mean in 2000, indicating that gaining capital status is correlated with better economic development. An average "yes-no" prefecture was above mean and comparable to a "yes-yes" prefecture in 980 (when both were provincial capitals), but it became close to the mean and similar to a "no-no" prefecture in 2000 after losing capital status.

Figure 5: Descriptive Patterns II: Pooling all changes together 980-2000



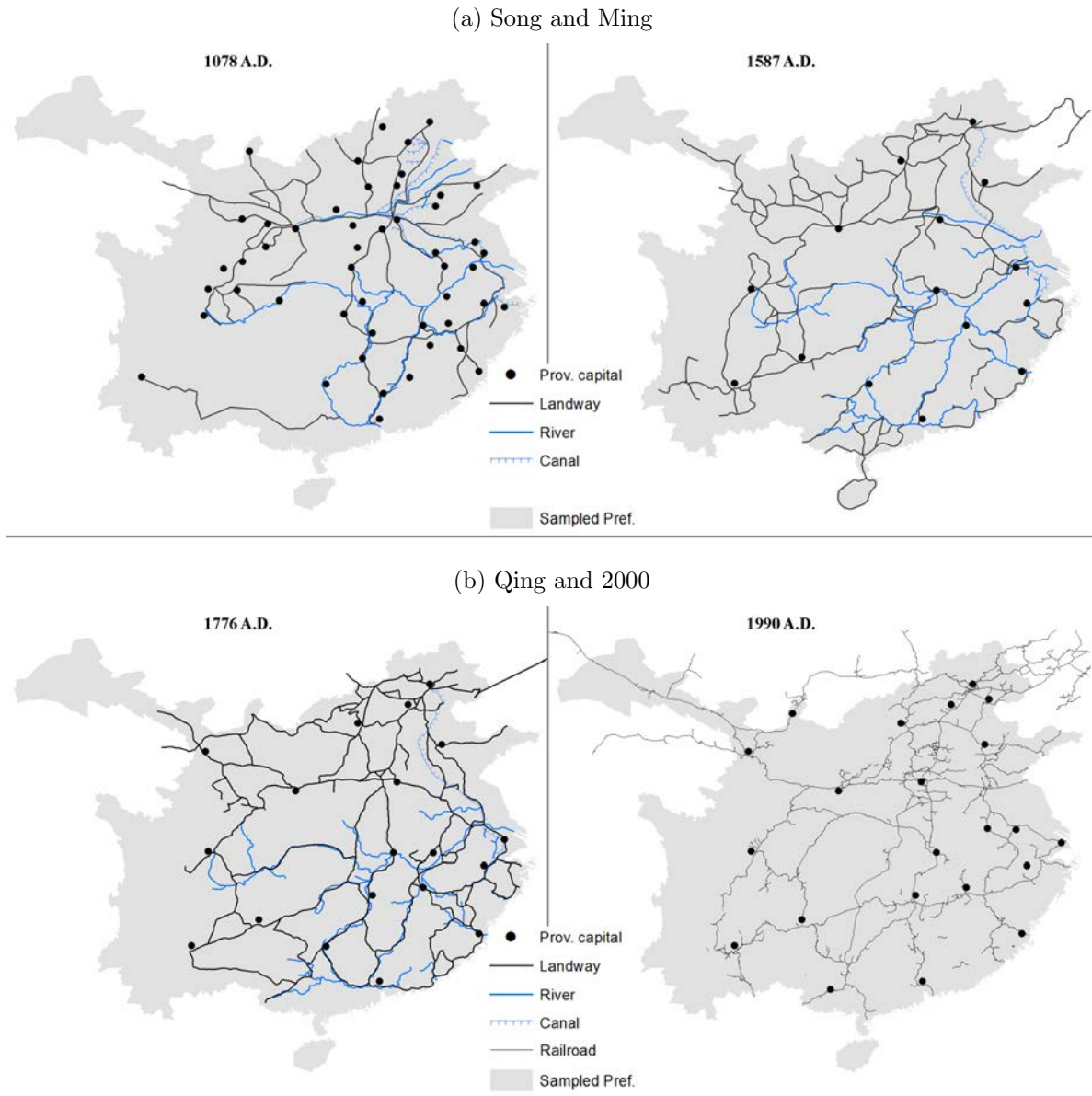
Notes: The figure shows that gaining and losing status indeed matters for population density. The four groups are defined based on the capital status every two regimes. To increase comparability, for each period, we first regress log population density on prefecture characteristics (geography variables, agricultural variables, log area, and macroregion dummies) and obtain residuals. We plot the average residuals by group and by period. After losing status, a prefecture became comparable to their non-capital peer prefectures.

Figure 6: The Impact of Capital Status on Population Density: Period by Period



Notes: This figure visualizes the results in Online Appendix D.4. 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.

Figure 7: Transportation Networks



Notes: This figure 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.

Table 1: Changes in Provincial Capitals, 1000-2000

			Ming		Qing		P.R.China (1964)		P.R.China (2000)	
			Capital	Non-capital	Capital	Non-capital	Capital	Non-capital	Capital	Non-capital
			15	246	19	242	21	240	24	237
Song	Capital	50	11	39	12	38	12	38	13	37
	Non-capital	212	4	207	7	204	9	202	11	200
Ming	Capital	15		14	1	13	2	13	2	
	Non-capital	246		5	241	8	238	11	235	
Qing	Capital	19				14	5	14	5	
	Non-capital	241				7	235	10	232	
P. R. China (1964)	Capital	21						21	0	
	Non-capital	240						3	237	

Notes: This table summarizes the change in provincial capitals across regimes. We omitted the Yuan dynasty and the Republic because they are relatively short-lived where data are less systematically available.

Table 2: 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 geography and agriculture controls				
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 2000	8	261	12.11	0.72
ln Total employment in government/party and social organization in 2000	8	261	8.29	0.67
ln Total employment in public institutions in 2000	8	261	8.76	0.71
Transportation networks				
Centrality in the transportation networks, $\theta=5.6$, in 1078	2, 9	261	11.68	2.94
Centrality in the transportation networks, $\theta=5.6$, in 1580	2, 9	261	12.98	2.31
Centrality in the transportation networks, $\theta=5.6$, in 1776	2, 9	261	13.18	2.75
Centrality in the transportation networks, $\theta=5.6$, in 1990	2, 9	261	10.86	2.24

Sources:

1. Liang (1980), Historical Statistics of Population, Land and Taxation in China.;
2. China Historical GIS (2007);
3. Ge Jianxiong (2000), China Population History;
4. Population Census 1953, 1964, 1982, 1990, 2000;
5. Treatise of the Nine Regions from the Yuanfeng reign (1078-1085);
6. FAO GAEZ (2012), <http://fao.org/Ag/AGL/agll/gaez/index.htm>.;
7. The Complete Directory of Qing Officials;
8. Census 2000 (micro data);
9. Cheng and Hsu (1980), Historical Atlas of China.

Table 3A: Hierarchical Distance and Provincial Capital Status: Regime by Regime

Regime	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\lambda =$	Song		Ming		Qing		P.R.C. (1964)	
	0.29		0.33		0.20		0.17	
H-distance	-0.198*** (0.031)	-0.176*** (0.042)	-0.079*** (0.023)	-0.071*** (0.025)	-0.090*** (0.026)	-0.087*** (0.027)	-0.129*** (0.028)	-0.125*** (0.030)
River		0.151*** (0.052)		0.077*** (0.027)		0.092*** (0.028)		0.058 (0.037)
Plain		0.119* (0.063)		0.062 (0.040)		0.058 (0.040)		0.051 (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: This table shows that hierarchical distance is a strong predictor of provincial capital status for each regime.

(i) Other geography variables include log area, 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) the 9-physiographic macroregions are defined by Skinner (1977).

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

Table 3B: Hierarchical Distance and Provincial Capital Status: Panel Analysis

	(1)	(2)	(3)	(4)
H-Distance	-0.094*** (0.019)	-0.088*** (0.018)	-0.086*** (0.018)	-0.081*** (0.019)
Prefecture FE, Year FE	Y	Y	Y	Y
Year FE*Geography variables		Y	Y	Y
Year FE*Agriculture variables			Y	Y
Year FE*Region FE				Y
Observations	1,044	1,044	1,044	1,044
R-squared	0.137	0.167	0.180	0.199

Notes: This table shows that hierarchical distance is a strong predictor of provincial capital changes.

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).

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

Table 4: Economic Consequences – The Impact of Capital Status on Population Density

method sample	DID				Reduced form			IV	
	All (1)	All (2)	All (3)	All (4)	All (5)	Ever-capital (6)	Never-capital (7)	All (8)	All (9)
Prov. Capital	0.622*** (0.132)	0.456*** (0.104)	0.488*** (0.105)	0.470*** (0.104)				Second-stage 0.796*** (0.229) 0.990*** (0.163)	
H-distance					-0.071*** (0.021)	-0.131*** (0.038)	-0.017 (0.027)	First-stage -0.089*** (0.008) -0.107*** (0.012)	
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-intersect 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	0.960
# prefectures	261	261	261	261	261	63	198	261	261
Weak instrument test (F statistic)								131.9	80.4
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).

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

Table 5: Economic Consequences – The Impact of Capital Status on Urbanization

method	Urbanization rate (%)			ln Urban Pop			ln Rural Pop		
	DID	IV		DID	IV		DID	IV	
	(1)	First-stage (2)	Second-stage (3)	(4)	First-stage (5)	Second-stage (6)	(7)	First-stage (8)	Second-stage (9)
Prov. Capital	11.347*** (2.549)		18.519** (9.036)	0.618*** (0.141)		1.983** (0.820)	0.195* (0.103)		0.380 (0.512)
H-distance		0.055*** (0.010)			0.055*** (0.010)			0.055*** (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. Columns (4)-(9) show that the impact on urban population is about five times that on rural population.

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).

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

Table 6: The Politics Hypothesis – 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*** (0.087)	0.477*** (0.075)	0.680*** (0.082)	1.149*** (0.088)	-1.090*** (0.319)	-2.094*** (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).

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

Table 7: The Market Hypothesis – Provincial Capital Status and Spatial Centrality

	(1)	(2)	(3)	(4)	(5)	(6)
		Centrality			ln rank in centrality	
	1078	1580	1776	1990		IV
Prov. Capital	0.423*** (0.080)	0.621*** (0.121)	0.567*** (0.109)	0.547*** (0.155)	-0.188*** (0.065)	-0.774** (0.351)
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	0.757	0.623	0.710	0.582	0.319	0.262
# 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). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.