The Oriental City:

Political Hierarchy and Regional Development in China, AD1000-2000*

Ying Bai[†] and Ruixue Jia[‡]

August 21, 2018

Abstract

Because regime changes in China between AD1000 and 2000 systematically altered the relative importance of different regions in the political hierarchy, tracing the evolution of Chinese provincial capitals and economic activities during this period throws light on how political factors shape economic geography. Our analysis shows that economic advantages driven by political factors do not necessarily persist: while gaining capital status has a large and positive on economic development as measured by population density and urbanization, the economic advantage shrinks after losing capital status. This pattern is further supported by exploiting variation arising from relocation of national capitals and redivision of provincial boundaries due to regime changes as an instrument for provincial capitals. We further document that (i) public offices are malleable but their relocation only explains a small part of our finding on population density, and (ii) political hierarchy affects many second-nature factors, even the less malleable ones like human capital and transportation networks.

*We thank Daron Acemoglu, Konrad Burchardi, Oded Galor, Walker Hanlon, Gordon Hanson, Yi Lu, Torsten Persson, Aleh Tsyvinski, Thomas Vendryes, Xiaoxue Zhao, Fabarizio Zilibotti and seminar participants at Barcelona Summer Forum, BREAD Pre-conference, CESifo, China Meeting of Econometric Society, CIFAR, Henan University, HKU, HKUST, IIES, NBER, Paris School of Economics, Tsinghua, and Xiamen University for their comments.

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

[‡]School of Global Policy and Strategy, University of California, San Diego, and CIFAR, rxjia@ucsd.edu.

1 Introduction

In AD1000, the city of Kaifeng was the most prosperous city in China and, with an estimated urban population of 1 million, arguably the largest city in the world (Mote 2003, Morris 2013). By 2015, however, its GDP ranked 129^{th} 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 13^{th} century and then its status of provincial capital in the 20^{th} century (Heng 1999). The city's development path thus exemplifies "the Oriental city" model proposed by Max Weber (1921) in which politics rather than the market determines the spatial distribution of economic activities.

Since Weber, scholars in the vast political economy literature have expanded our understanding of the importance of political factors in shaping economic geography (e.g., see De Long and Shleifer 1993, Ades and Glaeser 1995, and Davis and Henderson 2003 for the implications on city size)[1]. Because most of these studies rely on cross-sectional variation, however, they face the challenge of explaining what drives the political importance of certain regions. In this paper, therefore, we address this challenge by taking a *longue duree* approach that traces the evolution of provincial (and national) capitals and economic activities in China from AD1000 to 2000 to evaluate the politico-economic link. We find that gaining capital status has a large and positive effect on economic development, which may not be surprising. A less obvious question is whether the economic advantages of capital prefectures still hold after losing capital status. To this question, we find that losing capital status has a large and negative effect, implying that economic advantages driven by political status do not necessarily persist. Moreover, it is difficult to explain the large effect of gaining and losing capital status by the mechanical effect from relocation of public offices. Instead, we emphasize that political hierarchy shapes many underlying economic factors, even human capital and transportation networks.

As the largest enduring state with a distinctive political hierarchy, China provides fertile research ground on which to investigate the link between politics and economic development. From 1000-2000, the country underwent six dynastic regime changes that brought about drastic shifts in boundaries and centers of power, with national capitals relocated five times and the method for dividing provinces amended from "suiting [i.e., following] the forms of mountains and rivers to intentionally avoiding them so that boundaries "interlocked like

[1] De Long and Shleifer (1993) use historical data on European cities to document that property rights facilitate economic growth; while Ades and Glaeser (1995) draw on cross-country data to demonstrate that national capitals are larger in autocracies than democracies. In more recent work, Davis and Henderson (2003) argue that the extent to which a country's urban resources are concentrated in one or two large cities rather than more evenly distributed is directly affected by policies and politics.

dog's teeth". Consequently, 63 out of the 261 prefectures defined by the 2000 boundaries were once a provincial capital whose status changed with a new regime. In addition, China has a long history of governmental censuses whose rich information on population, geography, infrastructure, and bureaucracy allow us to trace the changes in capital status, construct extensive prefecture-level data over time, and identify important factors in determining capital status.

The major driver of national and provincial capital relocation is regime change, which happens infrequently and is hard to predict. For example, residents of the Song Dynasty capital (Kaifeng), would never have imagined that China could later be ruled by the Mongolians, who moved the national capital to northern China (Beijing) and redivided provincial boundaries for political control. This pattern also occurs after other regime changes: the new regime chooses a national capital close to its power base and redefines provincial boundaries, thereby altering the relative location of a prefecture. As a result, a prefecture's capital status often varies across regimes.

To formalize this logic of provincial capital relocation, we assume that the central government cares about governing a province, as well as about gathering resources from the province to the center. Two parts of cost then become important: collecting resources (and information) within a province and delivering some part of them to the national center. We proxy the first part by a prefecture's distance to other prefectures within the same province and the second part by a prefecture's distance to the national capital. We then define the weighted sum of the two distances as "hierarchical distance" and show a prefecture's rank in hierarchical distance within a province to be a strong predictor of its capital status. Not only does this rank vary greatly with regime change-induced national capital relocation and provincial boundary redivision, but, as we later document empirically, these latter are typically driven by political factors unlikely to be affected by any specific prefecture's characteristics, Hence, a prefecture's hierarchical distance rank can serve as a reasonable instrumental variable (IV) for provincial capital status.

Using data from existing historical and modern censuses, we construct a panel dataset across 261 prefectures for 11 periods (980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964, and 2000),[2] using both a difference-in-differences strategy and an IV strategy to evaluate how political status shapes economic development. To build a panel dataset based on fixed boundaries, we use the year 2000 prefecture boundaries in our baseline analysis and also report the results of a grid-level analysis.

[2] There are four additional periods with censuses available: 1880, 1953, 1982, 1990. Because including them will make the gaps between periods even more uneven, we exclude them in our baseline analysis based on the principle of dropping the periods with a gap smaller than 30 years. We will show that including them does not vary our findings.

We first employ the difference-in-differences strategy to show that gaining (losing) provincial capital status is associated with higher (lower) population density, at a magnitude of 50%. Using this strategy, however we face two important empirical challenges. First, our baseline data is available infrequently and has uneven gaps. To deal with this challenge, we show that our findings are robust to employing subsamples of roughly equal gaps and dropping any specific regime. Moreover, we conduct a period-by-period analysis where the periods are defined of equal length relative to the period before capital status change and find that the effect on population occurs only after the capital status changes. Another major concern is whether whether capital status changes because of omitted variables that affect population density. To partially address this concern, we show that our findings hold for the subsample of cities that were at some time a provincial capital when no omitted variables prevented them from becoming one. In addition, the period-by-period analysis is also useful in that at least the population density trends in the capital-to be (capital-to-lose) prefectures are no different from the rest before they gain (lose) capital status.

To further deal with endogeneity challenge, we employ hierarchical distance rank as an instrument for capital status and obtain an estimate comparable to that using difference-in-differences: the provincial capital status is associated with about a 70% higher population density. A general concern with this approach, however, is that hierarchical distance might affect economic development via economic channels other than political status. We thus conduct several checks, whose outcomes also underscore the relative importance of the political and economic factors underlying our findings. First, we find that hierarchical distance is not significantly correlated with population density for the prefectures that have never been a provincial capital, suggesting that hierarchical distance *per se* does not necessarily affect development. Second, we exploit national capital relocation to derive placebo hierarchical distance ranks that reveal, for instance, that once Kaifeng lost its national capital status, the rank calculated using distance to Kaifeng lost its influence on population density. Third, we also measure a prefecture's rank in terms of distances to major economic centers (instead of the national capital) and show that our findings on hierarchical distance rank is orthogonal to these alternative ranks. All these results suggest that alternative channels are not critical for our IV approach.

Our findings remain robust to using urbanization as an alternative outcome or employing grid-level data. We also examine heterogeneous effects with respect to (first-nature) natural advantages and find that gaining/losing capital status matters less for prefectures with higher agricultural productivity, suggesting that capital status matters via some reallocation channels that matter less if a prefecture was already productive.

What, then, explains the link between political hierarchy and economic development?

Our argument is that political hierarchy shapes almost all second-nature factors in our context. It is impossible to enumerate each factor; instead, we take two steps to substantiate our argument. We start from relocation of public offices, which should be rather malleable. However, given the magnitude of our findings (e.g., 40-50% in terms of population density), it seems difficult for this factor to account for a large part of the effect. We also provide back-of-the-envelope calculation on the magnitude using cross-sectional modern data. Then, we turn to two factors that are much less malleable, namely human capital and transportation networks, for two reasons. First, we can build a panel dataset for them to examine how changes in capital status matter. Second, these factors are often argued to be persistent (e.g. Wantchekon et al. 2015, Barjamovic et al. 2018). If we find that they vary systematically with political status, it illustrates the importance of political hierarchy in shaping second-nature factors. Indeed, while we also find a positive correlation between past human capital (transportation) and present human capital (transportation), we do observe a systematic rise and fall in human capital and centrality in the transportation networks after the gaining and losing of capital status. In other words, political hierarchy is so important for a prefecture that even the less malleable economic factors change with the political status of a prefecture.

Our study contributes to the political economy literature emphasizing the impacts of political factors on cities and development (e.g., De Long and Shleifer 1993, Ades and Glaeser 1995, Davis and Henderson 2003, Galiani and Kim 2011, Campante and Do 2014). Most such studies necessarily focus on cross-sectional variation because the rarity of capital city changes in their contexts makes it difficult to study the impact of changing political status. The Chinese setting, in contrast, allows us not only to exploit multiple changes in capital status but to speak to the long-run consequences of gaining and losing political status.

Our perspective complements the literature on the long-term spatial distribution of economic activity. Most of the existing studies focus on the role of market factors and document persistent patterns in economic activity, such as the minimal impact on development of large temporary shocks to a region due to locational fundamentals (e.g., Davis and Weinstein 2002, Miguel and Roland 2011, Barjamovic et al. 2018) or the persistent impacts of temporary advantages (e.g., Redding, Sturm, and Wolf 2011, Bleakley and Lin 2012, Kline and Moretti 2013, Michaels and Rauch 2016, Hanlon 2017). In contrast with the market factors, political factors are little studied in this research stream yet play a critical role in shaping economic geography in regimes like China. To 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 vary systematically with political status.

Our study also adds to the broad literature on how nature, history, and politics shape the distribution of economic activity (e.g., Henderson et al. 2017). The long-lasing administrative hierarchy in China provides a useful context to investigate the role of politics, which is likely to be relevant in many other contexts.[3]

The rest of the paper is organized as follows. Section 2 describes the background and provides a simple algorithm to explain changes in provincial capitals. Section 3 introduces the data and descriptive patterns. Section 4 reports the empirical results for both the difference-in-differences and IV analyses. Section 5 discusses the channels. Section 6 concludes the paper.

2 Background and An Algorithm

2.1 Regime Changes & Capital Relocation

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 could even be of different ethnicity; and (ii) despite regime shifts, China's three-tier administrative system (province-prefecture-county) remained surprisingly stable over this entire millennium.[4] The regimes shifts did, however, lead to changes in national capitals, provincial boundaries, and provincial capitals, with the latter being among the prefectures that make up provinces. These features make China an ideal context in which to investigate how a region's status in the political hierarchy affects its economic development; for instance, a prefecture could be a provincial capital in one regime but lose its capital status in another and vice versa.

Regime Changes Our study focuses on the core regions of China known as "China proper" (shaded area in Figure 1) for which centuries of historical information is available. 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).

[3] It is true that China may be extreme in terms of political centralization. Whether our findings hold in other contexts may depend on the degree of political centralization. For instance, Galiani and Kim (2011) studies the capital cities and find that the size of capital cities (relative to other cities) increases with political centralization.

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

As previously emphasized, these infrequent regime changes are hard to predict, with both the Mongolians and Manchurians (who founded the Yuan and Qing dynasties, respectively) regarded as minorities by the Han of the preceding regimes. 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 capital variation is thus driven primarily by regime change.

As we will explain in detail, national capital relocation and provincial boundary redivision are the preconditions for changes in provincial capital. Thus, we describe them in order below.

National Capital Relocation Because each regime tended to locate its national capital close to its power base, 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,[5] after which Beijing also served as capital for both the Qing Dynasty and the current People's Republic, while, Nanjing was the capital of the intervening Republic.

The different preferences for national capital locations are explained by the Mongolians (founder of the Yuan) and Manchurians (founders of the Qing) originating in the north, while the Ming and Republic power bases were in the south. 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). In all cases, political considerations are usually more important than economics.[6]

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 previously mentioned principles for defining provincial boundaries: whether to follow or subsume the natural lines of mountains and rivers, which latter exemplifies a spatial "divide-and-rule" tactic to limit the power of local governments.[7] Like national capitals,

[5]Nanjing 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.

[6]In fact, Beijing is often argued to be a bad choice as national capital because it is distant from central China and has limited access to water (see the discussion in the Atlantic: https://www.theatlantic.com/international/archive/2013/02/are-the-chinese-longing-for-a-new-capital/273182/).

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

provincial boundaries were created not from economic imperatives but from political inspiration, which often represented efforts to divide various communities so "each could be dealt with separately" (Skinner 1977, Guy 2010).

The Song adhered to the first principle, generally defining provincial boundaries by natural mountains and rivers, but when the Mongolians 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, with a generally increasing pattern. This pattern is illustrated in Appendix Figure A.1, which shows the Yangtze River used as a provincial boundary by the (pre-Mongolian) Song but included within provinces by the (post-Mongolian) Ming and Qing.

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 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 Appendix Figure 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)). When the national capital relocated to Beijing (Panel (c)), 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 (see Panel (a) of Appendix Figure A.2), became rather isolated in the Yuan and Ming dynasties and lost its provincial status (Panel (b)). It regained its capital status in the Qing Dynasty, however, because of national capital relocation and provincial redivision (Panel (c)).

For our baseline analysis, we map the historical data onto the 261 prefectures existing in the year 2000 and construct a panel dataset. 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. See Table 1 for a summary of provincial capital changes across regimes.

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, one for fiscal affairs and one for judicial affairs and others.[8] 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).

2.2 Changes in Provincial Capital: A Simple Algorithm

A Simple Algorithm We will provide a simple algorithm to predict where to locate provincial capitals, which serves two purposes: (1) to illustrate the political logic of provincial capital location, and (2) to provide an instrument for provincial capitals in our empirical analysis. Because of the second purpose, we would like to employ as few and exogenous variables as possible, conditional on the fact that the algorithm should be a strong predictor of provincial capital location.

Why does provincial capital status change? The background discussion suggests that this question can be answered from the perspective of the decision maker; namely, the central government. When deciding where to locate the provincial capital, the central government would consider multiple factors, two of which are particularly important: the governance of, and gathering resources from, a province. As a result, two factors are likely to matter: (i) the costs of gathering resources from all prefectures in a province into the provincial capital, proxied here by distance to other prefectures in the same province; and (ii) the costs of delivering a portion of these resources to the national center, proxied by distance to the national capital. Because the ruler of a new regime relocates its national capital and redraws its provincial boundaries, both types of costs could vary greatly, creating wide variation for us to explore empirically.

Expressed formally, the central government's choice of a capital for a province with prefecture $i = 1, 2, ..., N$ is a solution to minimize the following specification, which we term "hierarchical distance" (denoted by $HierDist$):

$$\underset{i}{\text{argmin}}\, HierDist_{i,t} \equiv \sum_{j=1}^{N} A_j D_{i,j,t} + \lambda \sum_{j=1}^{N} A_j D_{i,NationalCap,t} \tag{1}$$

where $D_{i,j,t}$ indicates the distance from i to another prefecture j in the same province in period t, and $D_{i,NationalCap,t}$ indicates a prefecture's distance to the national capital. A_j

[8] In a few cases, as Figure 2(a) shows, there was only one or as many as three provincial capitals.

is a weight variable – such as the area of prefecture j – to capture the scale of resources. Likewise, λ is a weight value that captures the relative importance of delivering resources to the center versus keeping resources within a province. Given one unit of resource, λ can be considered the share that must be delivered to the national capital. If $\lambda = 0$, only the within-province distance matters, so the provincial capital will be located in the provincial centroid. With an increase in λ, however, the provincial capital will be located increasingly closer to the national capital. Because we have no strong prior on the value of λ, we use the value with the most explanatory power ($\lambda = 0.19$) but also demonstrate that our results are robust to using alternative values.[9]

Although we motivate hierarchical distance from the perspective of gathering resources and distribution, one can also interpret it as the cost of distributing resources and information. Both interpretations imply the solution to equation (1).

We can extend the algorithm in several ways such as allowing λ to vary across regimes and using a more flexible weight variable A_j. However, we prefer not to do so because we will use $HierDist_{i,t}$ later as a possible instrument – for this purpose, it is useful to keep the algorithm as exogenous as possible.

Our Assumption on Provincial Boundaries When considering the change in provincial capitals, we take provincial redivision as given for both empirical and conceptual reasons.

Empirically, as administrative histories document, in most cases, the central government determined provincial boundaries first and then decided on the provincial capital.[10] Appendix Figure A.3 presents an example of the timing of changes described in the History of Ming Dynasty (Zhang 1739). As shown in panel (a), in the end of the Yuan dynasty, two regions of Zhejiang province were ruled by two warlords (Zhang Shicheng and Fang Guozhen). The new ruler of the Ming dynasty conquered the two regions and redefined the boundary of Zhejiang by incorporating the region ruled by Fang and dividing the region ruled by Zhang (see (panel (b)). After these changes, Hangzhou was designated as the new provincial capital.

Conceptually, changing the boundary of one province necessarily involves changing boundaries of nearby provinces, which would further affect relative positions of prefectures in other provinces. Therefore, it is not easy to redefine one province's boundary based on the provincial capital without leading to systematic changes.

[9] The choice is based on the magnitudes of R-squared in the following regression: $ProvCap_{i,t} = \theta HierDist_{i,t} + Prefecture_i + year_t$, where $Prefecture_i$ and $year_t$ indicate prefecture and year fixed effects. Varying the value of λ in $HierDist_{i,t}$, we find that the R-squared is the highest when $\lambda = 0.19$.

[10] Zhang's (1739) History of the Ming Dynasty, for instance, records the dates of provincial boundary and capital changes (for a brief summary, see http://www.xzqh.org/old/lishi/12ming/00.htm.)

Depiction of Hierarchical Distance In Figure 2, we graph the location of provincial capitals across major regimes (Song, Yuan, Ming, Qing, and 2000) using the case of $\lambda = 0.19$ and marking provincial centroids with crosses and regions with hierarchical distance below the provincial mean with shading. These maps reveal a clear pattern: consistent with our logic of hierarchical distance, the provincial capitals in each regime are likely to be located away from the provincial centroid and toward the national capital.

We also plot the probability of being a provincial capital by a prefecture's rank in the province in terms of hierarchical distance (see Figure 3). As panel (a) shows, for the prefectures that rank first (which vary across periods), the probability of being a provincial capital is around 0.36, whereas for those that rank second, it is around 0.26. 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 logged rank and probability of being a capital when the rank is lower than 10, one that is confirmed by the pattern in panel (b).

In sum, regime changes led to the relocation of provincial capitals based not on random decision but the logic of political control. Hence, by understanding this logic and exploiting the wide variation produced by regime change, we can identify the consequences of capital status (see Section 4).

3 Data and Descriptive Patterns

Prior to explaining the descriptive patterns that motivate our analysis (Section 3.2), we first describe our main analytic variables whose summary statistics and data sources are given in Appendix Table A.1.

3.1 Prefecture-Level Data

Population Density in 980-2000 Because population density information 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 – and calculate population density based on prefecture boundaries in 2000 (see Appendix Figure A.4 for population trends and density over time).

In our analysis, we drop these years one by one to ensure that our results are robust to different periods. We also employ subperiods of roughly equal length (e.g., one period for each regime) to ensure that our findings are not driven by the length of periods.

We are also able to access population data in years of 1880, 1953, 1982 and 1990. But as mentioned move, 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.

Urbanization in 1580-2000 Compared with population data, urbanization data are less systematically available, being accessible for only 4 of our 11 years: 1580, 1820, 1964, and 2000. The 1580 data are estimated based on local gazetteers,[11] the 1820 data are from CHGIS (2007), and the 1964-2000 information is taken from population censuses. By plotting the correlations between urbanization and population density in 1580 and 2000 (see Appendix Figure A.5), we reveal a strong link between these two measures, with correlational coefficients of 0.44 in 1580 and 0.47 in 2000. This comparison can thus be considered a validity check of our data.

Provincial Capitals and 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. 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). To construct our panel dataset, we use the boundaries for the 261 prefectures in 2000 as the baseline unit of analysis. To make sure that our results are not driven by how we define the prefecture boundary, we also conduct a 1-degree $\times$ 1-degree grid-level analysis in which the 261 prefectures are divided into 361 grids.

Additional Prefecture Characteristics We capture a prefecture's characteristics by including three additional variable sets: factors related to geography, agriculture, and regional location. The geographic variables include whether a prefecture contains a plain or major river or is on the coast, as well as its slope, elevation, longitude, and latitude. The 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. By allowing these crops' impacts to vary by year, we control for changes in land productivity.[12]

[11] We 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.

[12] See 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.

For our comparison of prefectures within a given macroregion, because the dramatic historical shift in provincial boundaries precludes a straightforward within province comparison, we use the nine physiographic macroregions defined by Skinner (1977): the north China plain, northwest, lower/mid/upper Yangtze, southeast coast, Lingnan, Yun-Gui, and Manchuria. 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 through "administrative accidents" rather than delineation of the natural boundaries of human activity.

By listing the correlations between these characteristics and ever-capital status in Appendix Table A.2, we demonstrate that having a plain and a major river matters for status because of their importance for building a city. Because the characteristics are time invariant, they are controlled for by our inclusion of prefecture fixed effects, although we also allow their impacts to vary by period.

3.2 Provincial Capitals and Development: 980 & 2000

Before analyzing our 11-period panel data, we present descriptive patterns based on the 1078 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 Panel (a) of Figure 4, the x-axis indicates the standardized logged population density in 980, while the y-axis indicates the standardized logged light density in 2000. These patterns remain similar if we use logged population density instead of logged light density in year 2000 (see Appendix Figure A.6).

Our analysis of the prefecture groups (see Figure 4) reveals systematic changes, as indicated by the four crosses. In particular, 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. Next, we will quantify these changes with different methods based on multiple-period data.

4 Provincial Hierarchy and Economic Development: Empirical Results

We study how much capital status matters for economic development using both a difference-in-differences method (Section 4.1) and an IV approach (Section 4.2), and summarize several additional checks in Section 4.3.

4.1 Difference-in-Differences Analysis

Baseline Results Our baseline estimation examines the correlation between provincial capital status and population density in 980, 1078, 1102, 1393, 1580, 1776, 1820, 1851, 1910, 1964 and 2000. We also show that the results are robust to dropping any specific year or using subperiods of roughly equal length. Our difference-in-differences estimator is as follows:

$$\ln PopDensity_{i,t} = \beta Capital_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \delta EverCap_i \times \gamma_t + \epsilon_{i,t}, \quad (2)$$

where $Capital_{i,t}$ indicates whether a prefecture i is a provincial capital in year t.

Here, we 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 year fixed effects (γ_t). $\mathbf{X}_i$ is two sets of prefecture characteristics: (i) geographical variables including whether a prefecture contains a plain or major river or is on the coast, as well as its slope, elevation, longitude, and latitude, and (ii) agricultural variables including the five types of crop suitability. We allow $\mathbf{X}_i$'s effects to vary across time by controlling for $\mathbf{X}_i \times \gamma_t$.

We also use the Skinner measures (π_m) to compare prefectures within the same macroregions, again allowing their effects to vary by year (indicated by $\pi_m \times \gamma_t$). Moreover, we allow for the prefectures that have even a provincial capital ($EverCap_i$) to have different trends from the rest by controlling for $EverCap_i \times \gamma_t$. All standard errors are clustered at the prefecture level.

The resulting estimates indicate that provincial capital status is associated with a 47-59% higher population density. Column (1) of Table 2 presents the result with all the fixed effects; column (2) controls for logged area and its interaction with γ_t; columns (3) and (4) add other geographical features and five-type crop suitability and their interactions with γ_t; finally, column (5) adds $EverCap_i \times \gamma_t$.

We also apply a first-difference strategy to our data for the 45 (19) prefectures that lost (gained) capital status once and show that change in capital status ($\Delta\text{Capital}_{i,t}$) is positively associated with a 45% change in population density (column (6)).

Next, because the change in population density is likely to depend on the initial density level, we further include lagged population density as a control (column (7)). The negative coefficient on this control indicates a convergence pattern, i.e., less dense prefectures grow faster. Nevertheless, the change in provincial capital status remains important (with a coefficient of 0.38).

In columns (8) and (9) of Table 2, we separate capital status loss from capital status gain and show that both have a sizable impact: gaining capital status is associated with a 41-45% increase in population density, whereas losing capital status is associated with a 36-46% decrease in population density.

It is worthwhile emphasizing the impact of losing capital status. First, it suggests that the omitted variable concern may not be essential; otherwise, we would observe that losing capital status matters little. Second, it indicates that the advantages brought by capital status did not necessarily last in the long run.

Dealing with Uneven Gaps We employ two methods to deal with the empirical challenge of low-frequency data with uneven gaps. In our analysis, we did not include four additional years (1880, 1953, 1982, 1990) that would make the gaps more uneven. Including them (column (2) of Table 3 gives us an estimate close to our baseline (column (1)). Now we further 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 an 2000 to construct another subsample (of a gap around 120-180 years). The estimates from these two subsamples are 0.51 and 0.57 (columns (3)-(4) of Table 3), close to our baseline estimates, suggesting that our results are not driven by the gaps. Moreover, our findings hold if we exclude any specific regime (columns (5)-(9)).

Second, we conduct a period-by-period analysis where we define periods relative to the period before the change in capital status (denoted by period0). 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. Using period0 as the reference group, we use the following specification:

$$\ln PopDensity_{i,t} = \sum_{\tau=-75+}^{300+} \beta_\tau Capital_{i,\tau} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \delta EverCap_i \times \gamma_t + \epsilon_{i,t}. \tag{3}$$

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 Appendix Table A.3. Figure 5 further visualizes the findings, where the lines connect the estimates and the dotted lines indicate the 95% confidence intervals. The patterns imply a shift in levels: gaining (and losing) capital status is associated with an increase (or decrease) in population density. However, due to the limitation of the low-frequency data, we cannot pin down the exact year when the change took place.

Using Ever-Capital Prefectures We already control for $EverCapital \times \gamma_t$ in our analysis. Nonetheless, 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. Hence, if we find a similar pattern for this subsample, then the concern over omitted variables becomes less critical. In fact, our baseline finding, using the same specifications as in our baseline estimation (see Appendix Table A.4, columns (1)-(4)), does hold for this subsample. The results also remain similar when we extend the comparison to the ever-capital prefectures and their neighbors (columns (5)-(6)).

As an additional test, we use the geographic, agricultural and regional prefecture characteristics ($\mathbf{X}_i$) to conduct propensity score matching on the ever-capital prefectures and their nearest match comparison group. The outcome is again comparable to our baseline finding (columns (7)-(8) of Appendix Table A.4).

Although neither the period-by-period results nor the outcomes using ever-capital prefectures fully address the endogeneity concern, they do reassure us that our baseline finding cannot simply be explained by population trends or some time-invariant omitted variables.

4.2 Rank in Hierarchical Distance as an Instrument

As explained in the background, the change in a prefecture's hierarchical distance stems from relocation of national capitals and redivision of provincial boundaries, which are unlikely to be driven by a prefecture's own characteristics. Moreover, as earlier illustrated (Figure 3), a prefecture's rank in terms of hierarchical distance within a province is strongly correlated with the probability of being a provincial capital. Thus, hierarchical distance rank seems a viable instrument for provincial capital status.

To use it, however, we must ascertain that (i) it is not driven by a prefecture's pre-change characteristics and (ii) it satisfies the exclusion restriction that it should not affect economic development via channels other than capital status.

Validity Checks To determine whether the change in hierarchical distance rank is indeed orthogonal to a prefecture's pre-change characteristics, we conduct two sets of tests. In the first, we show that ($\Delta \ln RankHierDist_{i,t}$) is not significantly correlated with past levels of logged population density ($\text{lnPopDensity}_{i,t-1}$ and $\text{lnPopDensity}_{i,t-2}$) or past changes in logged population density ($\Delta\text{lnPopDensity}_{i,t-1}$ and $\Delta\text{lnPopDensity}_{i,t-2}$) (see Table 4, columns (1)-(6)). These results indicate that the change in hierarchical distance rank is not driven by any changes 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 (see Table 5), thereby demonstrating that logged rank in hierarchical distance is a reasonable candidate as an instrument for capital status.

IV Estimates In our IV estimates, we focus on logged hierarchical distance rank because of its linear relation with the probability of being a provincial capital (Figure 3(b)). Our first-stage and second-stage specifications are as follows:

$$Capital_{i,t} = \delta lnRankHierDist_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \delta EverCap_i \times \gamma_t + \epsilon_{i,t}, \tag{4}$$

and

$$\ln PopDen_{i,t} = \beta' \widehat{Capital}_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \delta EverCap_i \times \gamma_t + \epsilon_{i,t}, \tag{5}$$

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

These IV estimates, whose reduced form is reported in column (1) of Table 6, provide further support for our main findings. The F-stat from the first stage is 133.6 (see the lower panel of column (4)), implying that instrument weakness is unlikely to be a concern. The IV estimate using $\lambda = 0.19$ is around 0.71 (the upper panel of column (4)), and ranges from 0.6

to 0.8 when we vary λ (see columns (2)-(6) Appendix Table A.5). In addition, our results are also robust to extending the linear functional form of the distances in equation (1) to a nonlinear form (i.e., $\sum_{j=1}^{N} A_j D_{i,j,t}^{\alpha}$ and $\sum_{j=1}^{N} A_j D_{i,NationalCap,t}^{\alpha}$). The IV estimates remain in a comparable range when we take $\alpha = 0.5$ or $\alpha = 2$ (columns(7)-(8) of Appendix Table A.5). Thus, these IV estimates provide further support for our analytic method.

Nevertheless, because an instrumental approach raises the concern of whether the instrument may affect channels other than capital status, we conduct three sets of tests for whether our findings are specific to political capital status. First, we examine whether hierarchical distance affects population density for the never-capital prefectures. Second, we use placebo hierarchical distances by exploiting the relocation of national capitals. Finally, we examine whether our findings are confounded by distances to major market centers.

(i) Hierarchical Distance for Never-Capital Prefectures We find that our IV results are driven by the variation within the ever-capital prefectures (Table 6, column (2)). In contract, hierarchical distance is not significantly correlated with population density for the never-capital prefectures (column (3)). This finding is reassuring: if hierarchical distance affects population density regardless of capital status, we expect to see that it also matters within the never-capital prefectures. This is not the case.

(ii) Placebo Hierarchical Distances For this test, we exploit the change in national capital status to construct a set of placebo hierarchical distance ranks. 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 (Table 6, columns (5)-(7)).

(iv) 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 (Appendix Table A.6). 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, these results indicate not only that hierarchical distance rank can serve as a useful instrument but that capital status is a critical channel through which hierarchical distance rank can affect economic development.

4.3 Additional Results

Although our primary analyses focus on changes in population density across (year 2000 defined) prefectures in 980-2000, our findings remain similar if we use urbanization (instead of population density) or employ grid-level data (instead of year 2000 prefectures). In addition, we present heterogeneous effects with respect to natural advantages in productivity, which can be considered as a sanity check of our baseline findings. We summarize these findings below and detail the results in the appendix.

Urbanization Because it is difficult to define cities over such a long period, our primary analyses focus on a prefecture's population density. In additional estimations, however, we also consider alternative outcomes using urbanization data for the latter half of the millennium (1580, 1820, 1964, 2000), which test whether our results hold after data from the first half of the millennium are excluded.

We find that provincial capital status has a similarly strong impact on urbanization (Appendix Table A.7) as on population density, being associated in these years with an 11 percentage point higher urbanization ratio, that is, 68 percent of the mean (columns (1)-(2)). When population is divided into urban and non-urban, the impact of capital status is more important for the former than the latter (columns (3)-(6)).

Grid-Level Analysis To build a panel dataset, we choose to fix prefectures using the 2000 boundary information in our baseline analysis. In addition, we can map all our data onto 1-degree $\times$ 1-degree grids by by assuming an evenly distributed population within each prefecture. We now have 361 grids instead of 261 prefectures.

The grid-level analysis again generates patterns close to our prefecture-level analysis (Appendix Table A.8), indicating that our findings are robust to alternative boundary definitions.

Heterogeneous Effects An auxiliary prediction of our baseline finding is that the impact of provincial capital status is less important for prefectures that enjoy natural advantages because for them, resources brought by capital status is less critical. Thus, we can use this hypothesis as a sanity check of our baseline findings. To test it, we proxy natural advantages by crop suitability – specifically, the maximum and mean suitability of old and new world crops – and examine how provincial capital status varies with this measure, standardized to facilitate interpretation. For old world crops (rice, wheat, and millet), we find a negative heterogeneous effect of maximum suitability (Appendix Table A.9, column (1)), with a one standard deviation increase in suitability decreasing the impact of capital status by about one third. These estimates remain similar using all crops or average suitability (columns (2)-(3)) and even using a change-on-change specification (column (5)). When we separate out gain versus loss of capital status, the patterns are consistent in either case (column (6)).

5 Underlying Factors

We have documented that capital status clearly matters for economic development proxied by population density and urbanization. What, then, explains this politico-economic link? One can conjecture that many factors alter along with political status including jobs, people and resources, which makes it impossible to enumerate each one. We thus takes an alternative approach by focusing on three factors, on a spectrum from being very malleable to being very persistent.

First, we discuss the channel of reallocating public offices (Section 5.1), which can be considered mechanical and is among the most malleable ones. Then, we turn to the other end of the spectrum by examining human capital (Section 5.2) and transportation networks (Section 5.3). We focus on these two, partly because they are less malleable, partly because we can build a panel dataset to examine the impact of capital status change.

Using this approach, we have to miss some factors. For instance, Chen et al. (2017) argue that firms in provincial capitals get cheaper loans from the government. For this type of analysis, one has to rely on cross-sectional variation in capital status. In addition, it is challenging to find a historical counterpart on access to credit. Therefore, instead of listing every possible factor, we only focus on three factors to illustrate the logic.

5.1 Relocation of Public Offices

Public offices move with provincial capitals. However, considering the size of public employment in history and today, it seems difficult for this channel to explain a large part of our findings. For instance, in the mid-19th century, the gentry and their immediate family members accounted for around 2% of the population (Chang 1955). Since the gentry class provides the talent pool for public employment, this number is likely to be an upper bound of the size of public employment. For modern China, employing various sources of data, Ang (2012) estimates that the size of public employment is around 3.1% of population in 2000.

Unfortunately there are no comparable prefecture-level data by occupation from history to today. To gauge the magnitude of this channel, we conduct a back-of-envelope analysis using cross-section census data in 2000 (reported in Table 7). First, we show that provincial capital status is associated with a 47% higher total employment (column (1)). Then, we define public employment as those employed in government agencies, the Communist Party agencies, and social organizations, plus those employed in health care, sports and social welfare. Capital status indeed matters, which doubles the size of public employment (columns (2)-(3)). However, given the small share of public employment in the total employment, it can only explains 6% of the impact on total employment (0.03/0.474).

As expected, public employment is affected by capital status. However, it is difficult for this channel alone to account for a large part of our baseline finding. Next, we turn to two other factors to illustrate the importance of capital status in shaping underlying economic factors.

5.2 Change in Human Capital

In additional to being less malleable, we would like to examine human capital for another reason: it is useful to know whether talents move with capital status. To measure human capital in history, we employ the number of presented scholars (the highest degree in the imperial examination, known as *Jinshi* in Chinese) in the Qing dynasty. For modern human capital, we use the number of individuals with high school education and above. We normalize the former by population size in 1776 (mid-Qing) and the latter by population size in 2000.

Similar to our descriptive pattern in Figure 4, we divide all the prefectures into four groups based on a prefecture capital status in the Qing and in 2000 ("No-No", "Yes-Yes", "Yes-No", and "No-Yes"). Figure 6 illustrates the pattern, where the x-axis indicates the standardized log *Jinshi* per capita in history, and the y-axis indicates the standardized log individuals with high school or above per capita in 2000. These two measures are positively correlated, indicating some persistence of human capital. However, we also observe that human capital varies systematically within gaining and losing capital status. For instance, those gaining capital status (the "No-Yes" group) are similar to the "Yes-Yes" group in 2000 in terms of modern human capital, even though they were at lower level in the Qing dynasty. In constrast, those losing capital status (the "Yes-No" group) were comparable to the "Yes-Yes" group in the Qing dynasty but performed worse in 2000, became more comparable to "No-No" group.

These patterns are confirmed by regression analysis in Table 8. Because we are using two measures, it is not straightforward to calculate change in human capital. We employ two ways to mitigate this concern. One is to use the standardized measure for both periods; the other is to categorize human capital in both periods into deciles and use the change in decile rank as the dependent variable (a higher rank indicate more human capital). As shown in Table 8, gaining (losing) capital status is associated with better (worse) human capital.

5.3 Change in Transportation Networks

Finally, we examine transportation networks. Due to the high fixed costs to build new transportation networks, one expects transportation networks to exhibit persistence (which is also true in our context). However, transportation networks also 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. Second, when a regime replaced the previous one, the ruler decided which parts of the transportation networks to be reconstructed, which usually depends on the regions' relative importance in the political hierarchy. As a result, we obtain a unique opportunity to systematically investigate changes in the transportation networks.

Historical and Modern Transportation The historical transportation networks were comprised of the Grand Canal, which connected various waterways, and a state courier system (supported by many post offices). Because the state was the largest single investor in transportation and communications facilities, the network aim was primarily political: to maintain an adequate flow of information, revenues, and personnel on which the state relied (Brook 1998). Nevertheless, 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 compare a prefecture's spatial importance in the transportation networks across periods, we employ standardized network measures for each period.

To assess how the transportation networks vary with capital status, we digitize roads and waterway maps for three historical periods (represented by specific years) – the Song (1078), Ming (1587), and Qing (1820) dynasties[13] – and the railroad map for the People's Republic (1990).

Capital Status and Spatial Centrality Although we can easily count the number of landways, waterways, and railways in a prefecture, we also want to account for the relative

[13] These maps are collected in the Historical Atlas of China (Cheng and Hsu 1980).

importance of different links in the transportation network. To this end, we employ a centrality measure, defined for each prefecture i as follows:

$$Centrality_i = \sum_{j \neq i} \frac{1}{d_{i,j}} = \sum_{j \neq i} \frac{1}{d_{i,N_i} + d_{j,N_j} + (1/\theta) d_{N_i,N_j}}, \tag{6}$$

where $d_{i,j}$ indicates shortest distance between i and j in the transportation networks.[14]

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 a 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, we will show that our results are not sensitive to the choice of θ .

Again, 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. To this end, we replace the dependent variable in our baseline estimation with logged centrality and use it with the four-period panel data used in the transportation analysis. Since centrality is an abstract measure, we report standardized coefficient in Table 9. A provincial capital's centrality is about 0.2 standard deviation higher than that of the non-capital prefectures (column (1)). Similar to our baseline finding, gaining provincial capital status increases centrality whereas losing capital status implies a loss of centrality (column (2)). The latter finding provides further evidence that our result is unlikely to be driven by an endogenous assignment of capital status to better connected locations (which is of more concern for new capitals). These findings remain robust to employing different values of θ (columns (3)-(8)).

Since the capitals were assigned in the beginning of a dynasty whereas the transportation networks were altered later, our finding on centrality is unlikely to be driven by reverse causality. With four periods of data, we do not have the same power to conduct period-by-period tests as in Figure 5. Nevertheless, we separate the two periods before capital status change and show that centrality decreases (increases) only after the loss (gain) of capital status (Appendix Table A.10).

These findings show that even transportation networks vary with the political hierarchy. Together with the earlier findings on public office and human capital, we argue that

[14] 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 the position in the transportation networks.

political capital status has a large impact on economic development because many underlying economic factors flow with it. Once a prefecture lost its political advantages, these factors also moved away, which explains our finding on losing capital status.

6 Conclusions

Although political factors have long been argued to be important for economic geography, it is difficult to provide direct evidence of this claim without understanding what drives the political importance of certain regions. In the case of China, its combination of an enduring state, a distinctive political hierarchy, and many changes in national and provincial capitals make it a particularly advantageous context for examining the impact of politics on economic development. Our analysis demonstrates that both gaining and losing capital status affect regional development.

At the same time, given the growing literature documenting the persistence of economic activities, our finding of an association between losing capital status and worse development may appear surprising. Yet this observation is consistent with the underlying political logic we uncover. That is, we show that in a regime in which the political hierarchy guides the distribution of resources, the influence of political status goes beyond public offices, and matters for many underlying economic factors, even human capital and transportation. Hence, not only do we throw important light on these observations by showing how political factors shape economic geography, but our approach offers a perspective on how to link the political economy literature with the research on both changes and the persistence of economic activities in the long term.

References

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

[2] Ang, Yuen Yuen (2012), "Counting Cadres: A Comparative View of the Size of China's Public Employment," *China Quarterly* 211: 676-696.

[3] Barjamovic, Gojko, Thomas Chaney, Kerem A. Cosar, and Ali Hortaçsu (2017), "Trade, Merchants, and the Lost Cities of the Bronze Age," NBER Working Paper 23992.

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

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

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

[7] Chang, Chung-li (1955), *The Chinese Gentry: Studies on Their Role in Nineteenth-century Chinese Society*. University of Washington Press.

[8] Chen, Ying, J. Vernon Henderson, and Wei Cai (2017), "Political Favoritism in China?s Capital Markets and Its Effect on City Sizes," *Journal of Urban Economics* 98: 69-87.

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

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

[11] Donaldson, Dave, and Richard Hornbeck (2016), "Railroads and American Economic Growth: A 'Market Access' Approach," *Quarterly Journal of Economics* 131(2): 799-858.

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

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

[14] 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.

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

[16] Fogel, Robert W. (1964), *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins University Press

[17] Galiani, Sebastian, and Sukkoo Kim (2011), "Political Centralization and Urban Primacy: Evidence from National and Provincial Capitals in the Americas," In Understand-

ing Long-Run Economic Growth, edited by Costa and Lamoreaux, 121-54. Chicago.

[18] Galor, Oded, and Ömer Özaka (2016), "The Agricultural Origins of Time Preference," *American Economic Review* 106(10): 3064-3103.

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

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

[21] Hanlon, W. Walker (2017), "Temporary Shocks and Persistent Effects in Urban Economies: Evidence from British Cities after the US Civil War," *Review of Economics and Statistics* 99(1): 67-79.

[22] Henderson, J. Vernon, Tim Squires, Adam Storeygard, and David Weil (2017), "The Global Distribution of Economic Activity: Nature, History, and the Role of Trade," *Quarterly Journal of Economics*, 133(1): 357-406.

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

[24] Kline, Patrick, and Enrico Moretti (2013), "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *Quarterly Journal of Economics* 129(1): 275-331.

[25] Liang, Fang-zhong (1980), *Historical Statistics on Hukou, Land and Land Tax of China (Lidai Hukou, Tudi,Tianfu Tongji)*. Shanghai: Zhonghua Book Company.

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

[27] Miguel, Edward, and Gerard Roland (2011), "The Long-Run Impact of Bombing Vietnam," *Journal of Development Economics* 96(1): 1-15.

[28] Michaels, Guy, and Ferdinand Rauch (2016), "Resetting the Urban Network: 117?2012," Economic Journal, In Print.

[29] Morris, Ian (2013), *The Measure of Civilization: How Social Development Decides the Fate of Nations*. Princeton University Press.

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

[31] Mote, Frederick (2003), *Imperial China 900-1800*. Harvard University Press.

[32] Nunn, Nathan, and Nancy Qian (2010), "The Columbian Exchange: A History of Disease, Food, and Ideas," Journal of Economic Perspectives 24(2): 163-188.

[33] 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.

[34] Redding, Stephen J., Daniel M. Sturm, and Nikolaus Wolf (2011), "History and Industry

Location: Evidence from German Airports," *Review of Economics and Statistics* 93(3): 814-831.

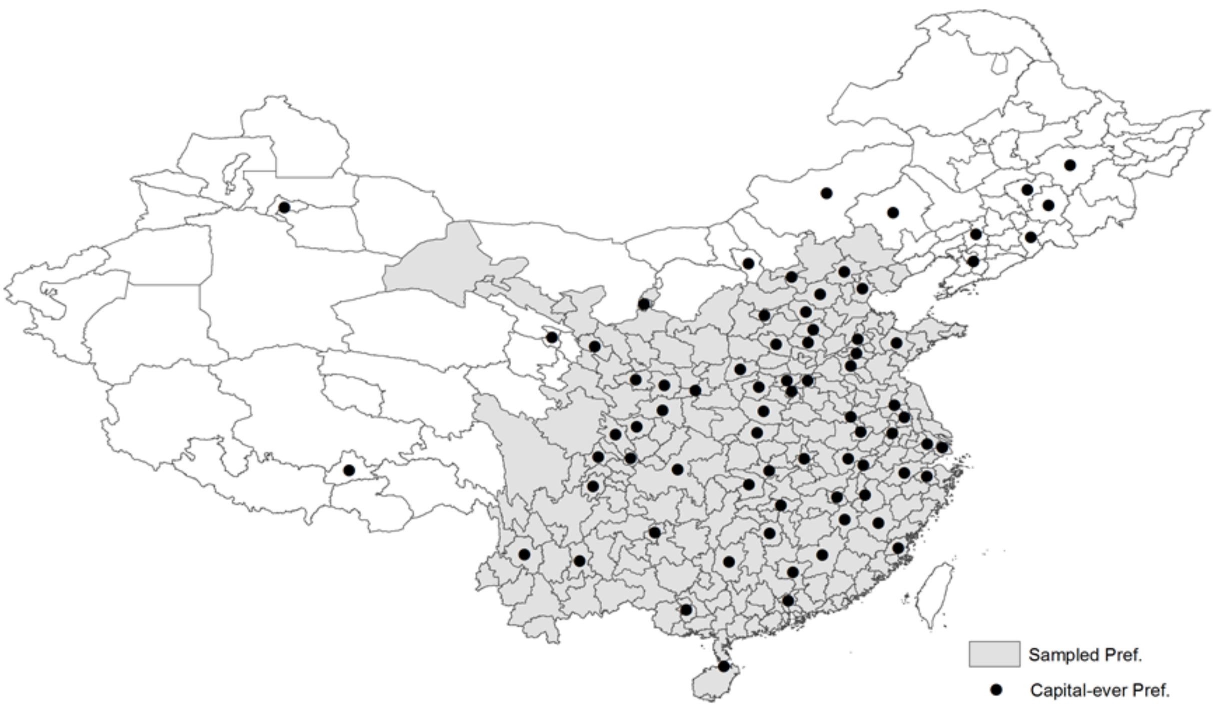
[35] Skinner, G. William, ed. (1977), *The City in Late Imperial China*, Stanford University Press

[36] Wantchekon, L., Klasnjaz, M., and Novta, N. (2015). Education and Human Capital Externalities: Evidence from Colonial Benin. Quarterly Journal of Economics, 130(2):703?757.

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

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

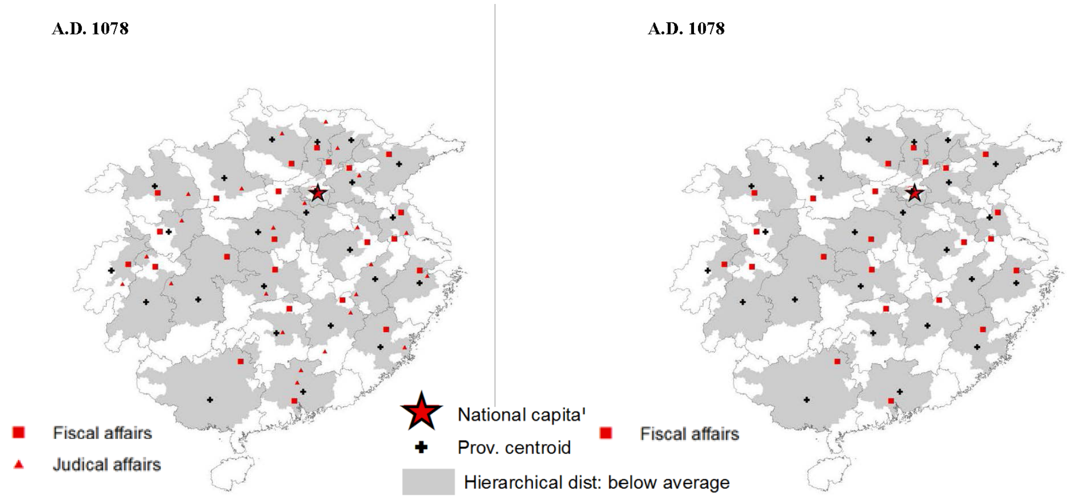
Figure 1: Capital-Ever Prefectures



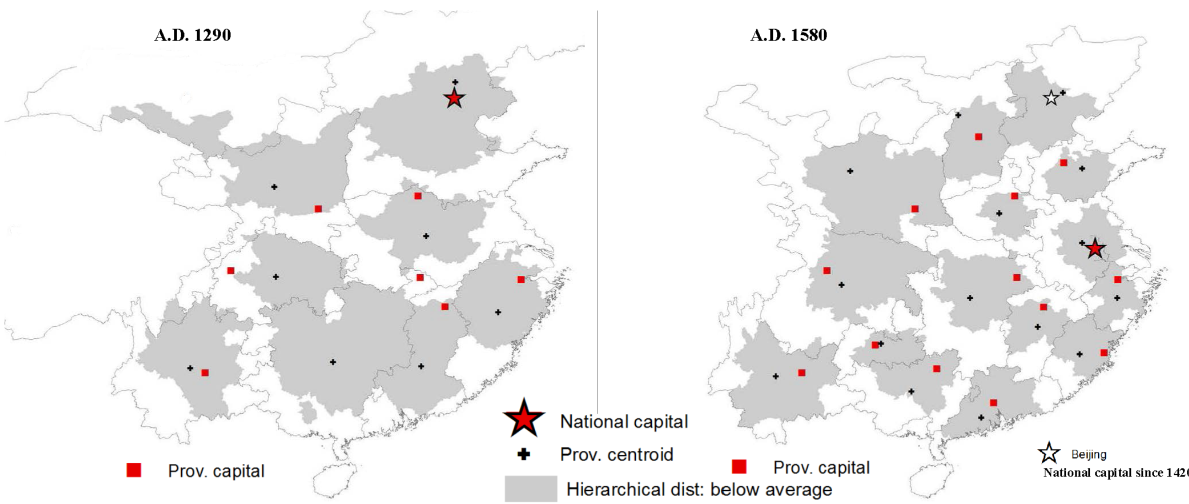
Notes: The shaded area indicate the prefectures in China proper (our sample). The dotted prefectures have ever been a provincial capital at least once during 1078-2000. This map is based on the prefecture boundaries in 2000: 63 out of 261 prefectures have ever been a provincial capital.

Figure 2: Hierarchical Distance, National Capitals and Provincial Capitals

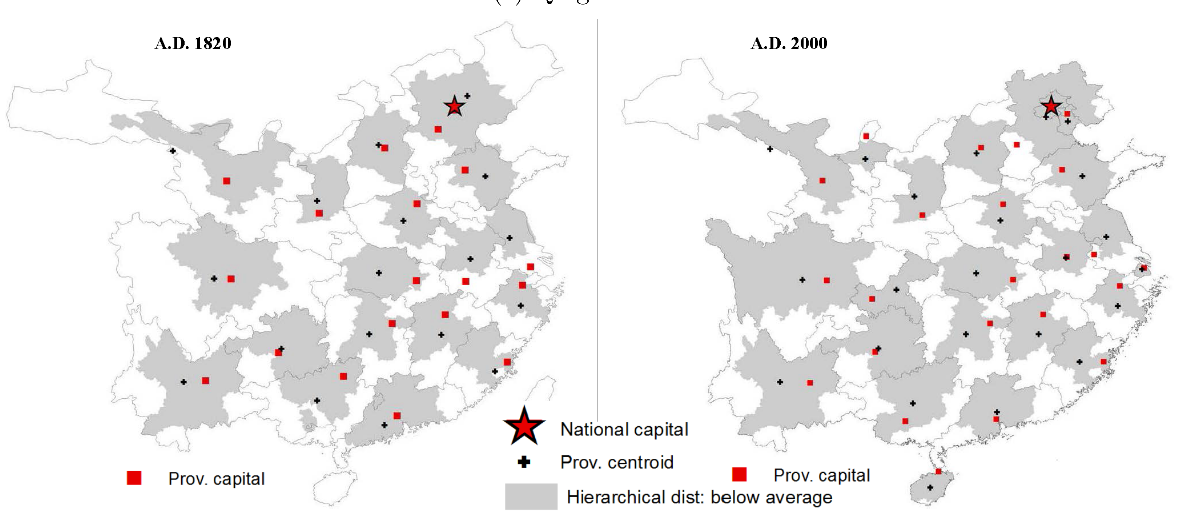
(a) Song



(b) Yuan and Ming



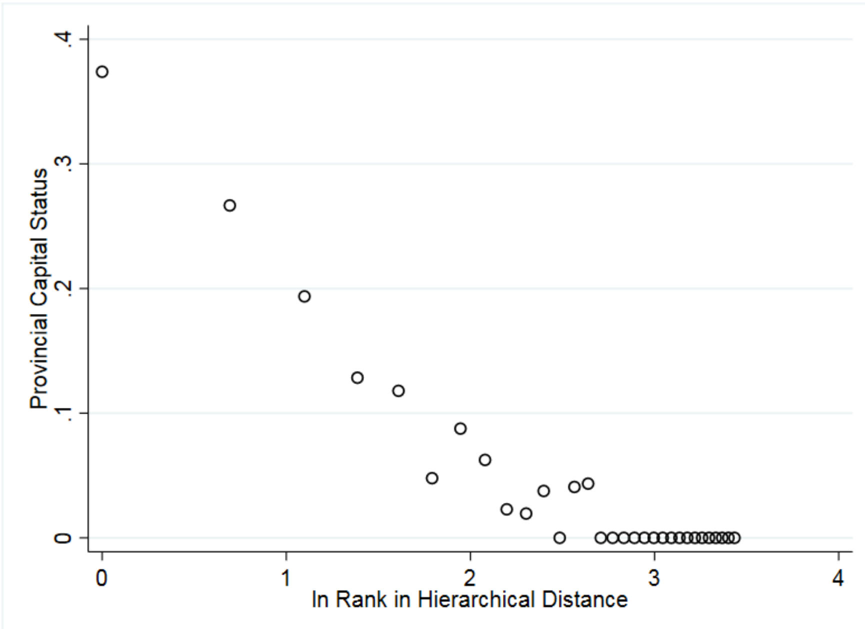
(c) Qing and 2000



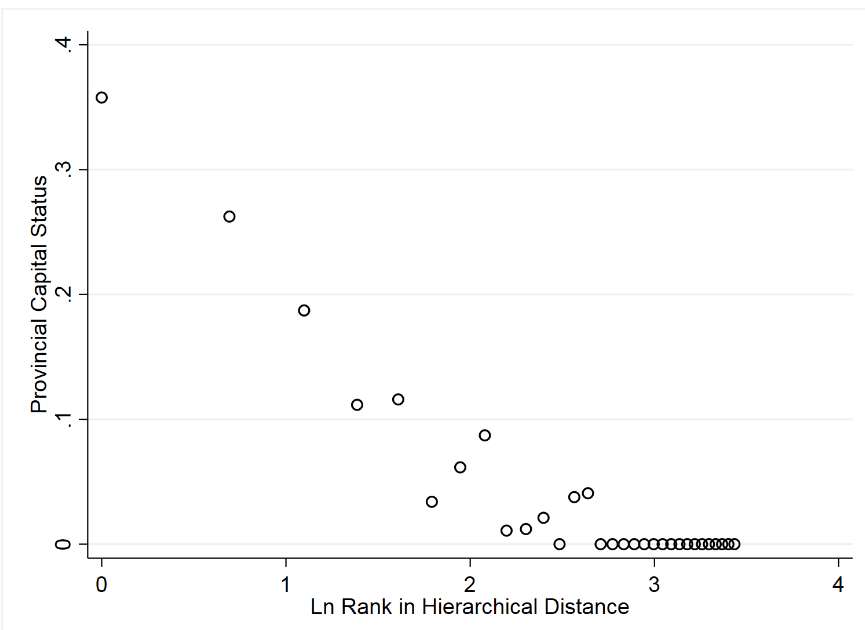
Notes: This figure shows that provincial capitals are usually located away from provincial centroid toward the shaded region with a shorter hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province). We present two maps for the Song dynasty because provinces often had two capitals, one for fiscal affairs and one for judicial affairs and others. See Section 2.1 for more discussion.

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

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

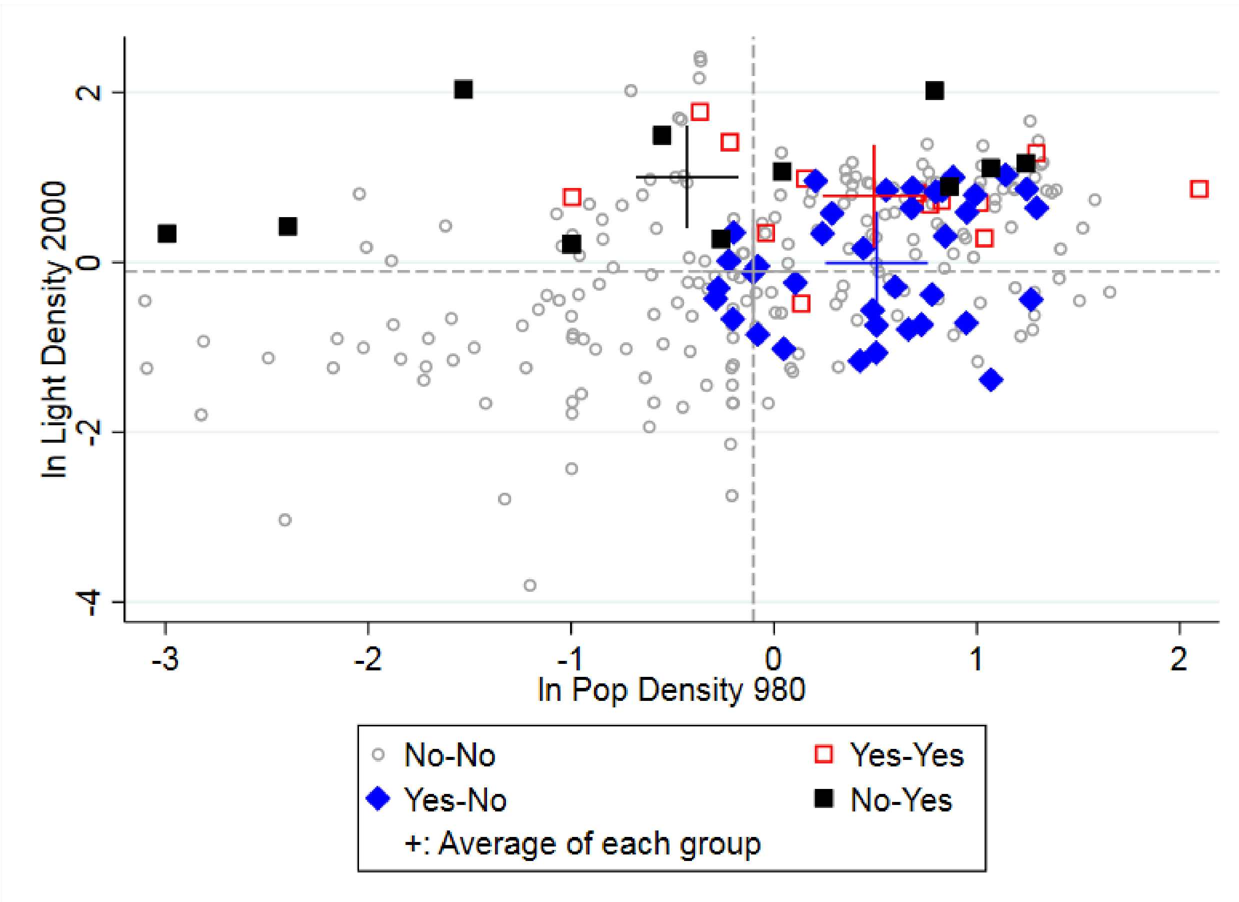


(b) ln Rank in Hierarchical Distance vs. Prob. of Being Capital



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, which confirms the pattern in Figure 2.

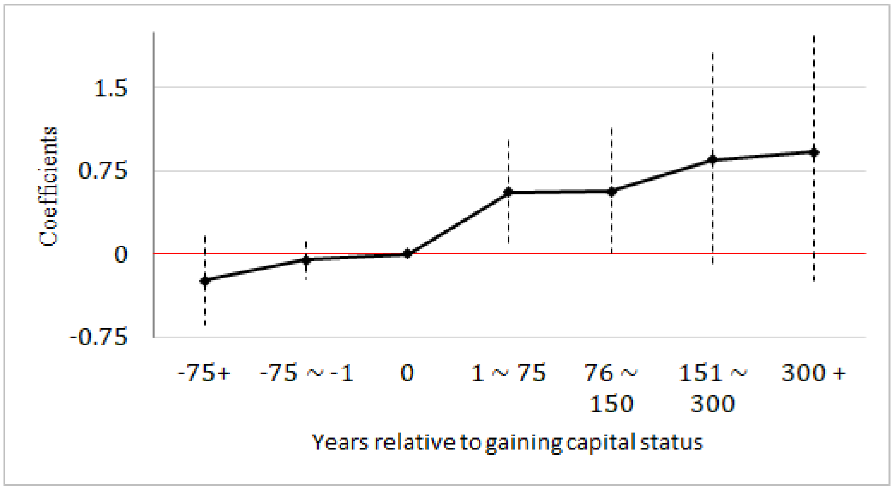
Figure 4: Descriptive Patterns



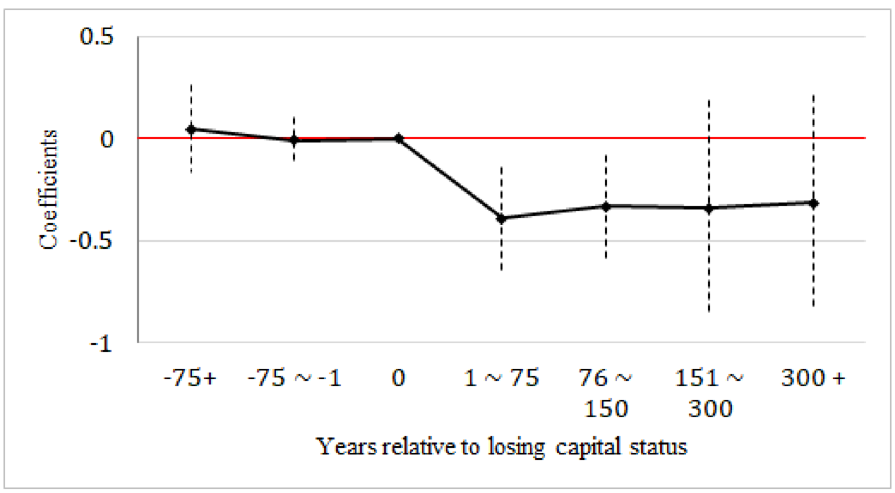
Notes: The figure shows that gaining and losing capital status is associated with a systematic change in a prefecture's economic performance (measured by population density in 980 and light density in 2000). An average "no-no" prefecture group was close to mean in both periods and an average "yes-yes" prefecture was above the mean in both periods. In contrast, an average "no-yes" prefecture was the 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. Appendix Figure A.6 presents a similar pattern using population density in 2000 instead of light density.

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

(a) Before and After Gaining Capital Status

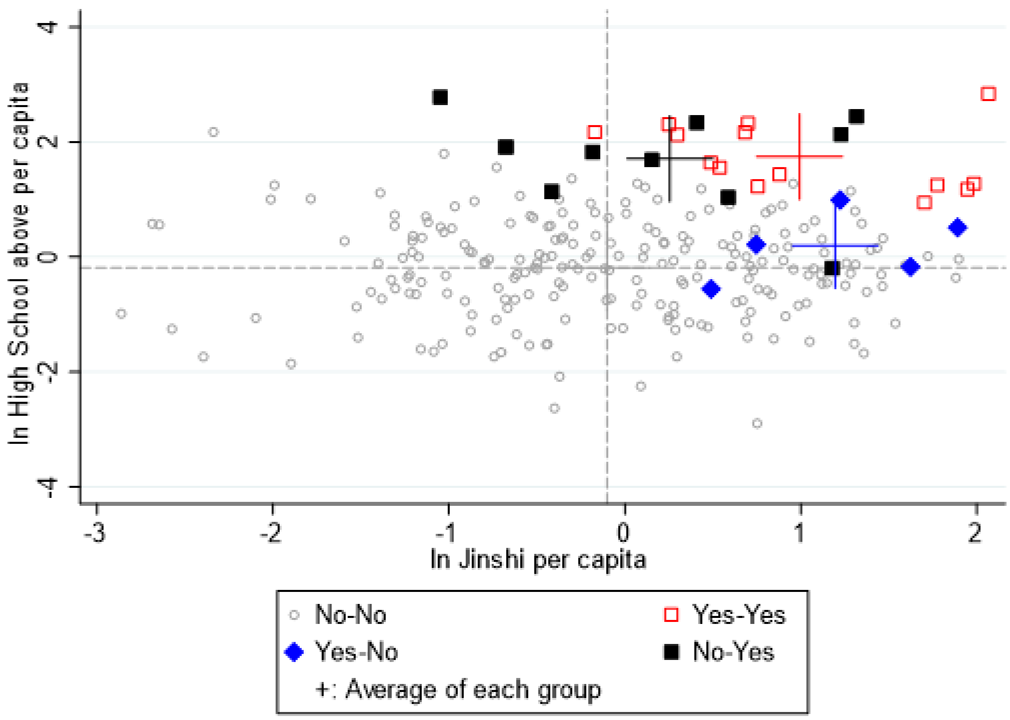


(b) Before and After Losing Capital Status



Notes: This figure visualizes the results in Appendix Table A.3. 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.

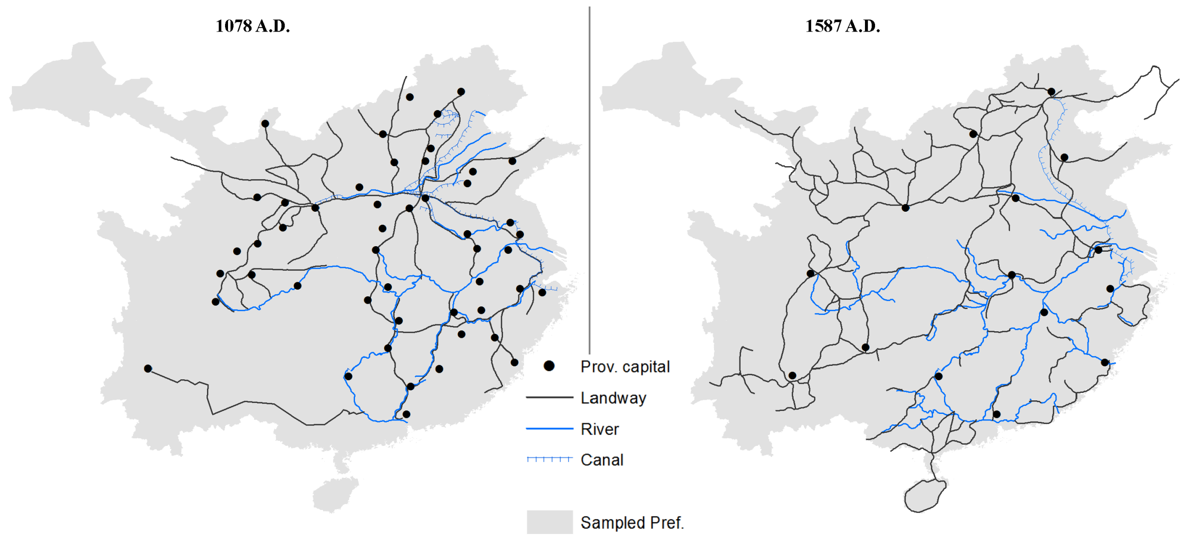
Figure 6: Human Capital in the Qing Dynasty and 2000



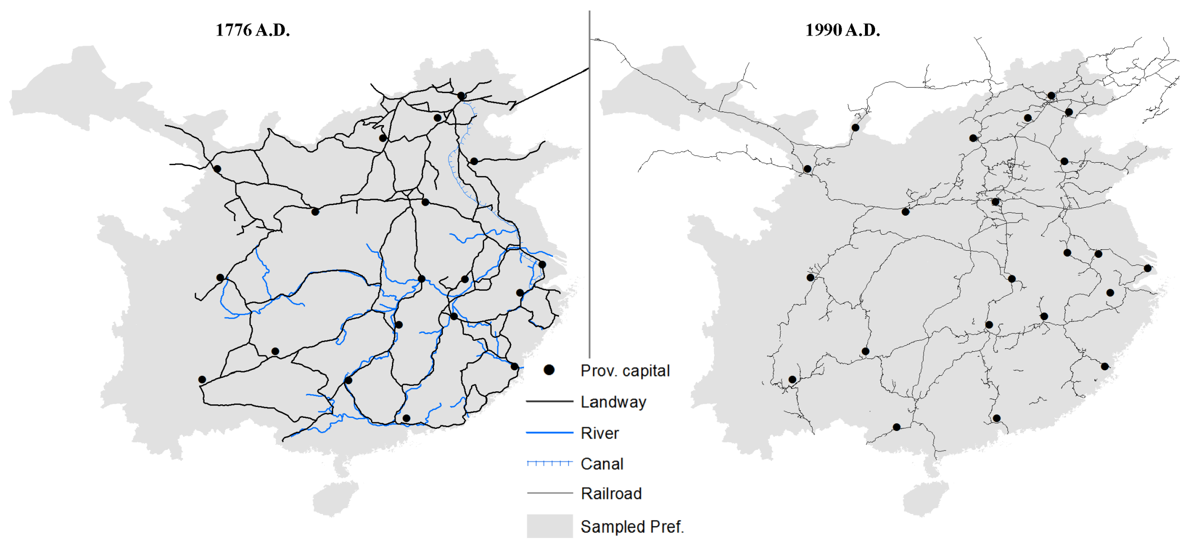
Notes: The figure shows that gaining and losing capital status is associated with a systematic change in human capital. An average "no-no" prefecture group was close to mean in both periods and an average "yes-yes" prefecture was above the mean in both periods. In contrast, an average "no-yes" prefecture was about 0.3 standard deviation higher than the mean in the Qing and became two standard deviation above the mean in 2000. An average "yes-no" prefecture was more than one standard deviation above mean and comparable to a "yes-yes" prefecture in the Qing (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 7: Transportation Networks

(a) Song and Ming



(b) Qing and 2000



Notes: This figure plots the transportation networks across regimes. In 1990, we focus on railway networks. We digitize the maps in the Song, Ming and Qing from *Historical Atlas of China* (Cheng and Hsu 1980).

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: (1) This table summarizes the change in provincial capitals across regimes. We omitted the Yuan dynasty from this table because we do not have population data in that regime. (2) Even though there are many changes from the Song to the Ming, our findings are robust to dropping the Song period (see Appendix Table **??**).

Table 2: The Impact of Capital Status on Population Density

	ln Pop Density					Δln Pop Density			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prov. Capital	0.589***	0.518***	0.472***	0.470***	0.514***				
	(0.129)	(0.110)	(0.105)	(0.104)	(0.142)				
ΔProv. Capital						0.453***	0.383***		
						(0.141)	(0.102)		
Gaining Capital Status								0.450**	0.410**
								(0.220)	(0.164)
Losing Capital Status								-0.455***	-0.358***
								(0.174)	(0.128)
lag. ln Pop Density							-0.323***		-0.323***
							(0.012)		(0.012)
Year FE * Ever-capital					Y	Y	Y	Y	Y
Year FE * Crop suitability				Y	Y	Y	Y	Y	Y
Year FE * Geography			Y	Y	Y	Y	Y	Y	Y
Year FE * ln Area		Y	Y	Y	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Pref. FE	Y	Y	Y	Y	Y				
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,871	2,871	2,871	2,871	2,871	2,610	2,610	2,610	2,610
R-squared	0.807	0.859	0.873	0.882	0.882	0.709	0.773	0.709	0.773
# prefectures	261	261	261	261	261	261	261	261	261

Notes: This table shows that provincial capital status is associated with a higher population density. Crop suitability refers to the suitability of rice, wheat, millet, sweet potatoes and maize. Geographical controls include whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude. Region refers 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 3: Dealing with Uneven Gaps (Dependent Var.: ln Pop Density)

		Baseline (1)	+small gap (2)	Gap: ca400 (3)	Gap: ca120 (4)	(5)	(6)	(7)	(8)	(9)
Prov. Capital		0.514***	0.501***	0.512***	0.575***	0.369***	0.485***	0.626***	0.473***	0.665***
		(0.142)	(0.131)	(0.154)	(0.143)	(0.123)	(0.138)	(0.154)	(0.172)	(0.173)
All baseline controls		Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations		2,871	3,915	783	1,566	3,132	3,393	2,610	2,610	1,827
R-squared		0.882	0.887	0.887	0.891	0.893	0.885	0.903	0.864	0.915
# prefectures		261	261	261	261	261	261	261	261	261
Song	980	Y	Y	Y	Y		Y	Y	Y	Y
	1080	Y	Y				Y	Y	Y	Y
	1102	Y	Y		Y		Y	Y	Y	Y
Ming	1393	Y	Y	Y	Y	Y		Y	Y	
	1580	Y	Y		Y	Y		Y	Y	
Qing	1776	Y	Y			Y	Y		Y	
	1820	Y	Y	Y	Y	Y	Y		Y	
	1851	Y	Y			Y	Y		Y	
	1880		Y			Y	Y		Y	
	1910	Y	Y			Y	Y		Y	
P R China	1953		Y			Y	Y	Y		Y
	1964	Y	Y			Y	Y	Y		Y
	1982		Y			Y	Y	Y		Y
	1990		Y			Y	Y	Y		Y
	2000	Y	Y		Y	Y	Y	Y		Y

Notes: This table shows that our results are robust to keeping subperiods of roughly equal length or dropping any specific period of data. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 4: Validity Check I: Hierarchical Distance and Pre-change Characteristics

	Δln Hierarchical Distance					
	(1)	(2)	(3)	(4)	(5)	(6)
lag. ln Pop Density	-0.024		-0.018			
	(0.024)		(0.030)			
lag2. ln Pop Density		-0.023	-0.013			
		(0.022)	(0.017)			
lag. Δln Pop Density				0.001		-0.001
				(0.017)		(0.020)
lag2. Δln Pop Density					-0.030	-0.030
					(0.037)	(0.038)
All controls	Y	Y	Y	Y	Y	Y
Observations	2,610	2,349	2,349	2,349	2,088	2,088
R-squared	0.303	0.303	0.303	0.302	0.303	0.303
#prefectures	261	261	261	261	261	261

Notes: This table shows that the change in hierarchical distance is not significantly correlated with the levels and changes in population density in the past periods. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 5: Validity Check II: Gaining vs. Losing Capital Status

		Seemingly Unrelated Regression	
	ΔCapital Status	Gaining Cap.	Losing Cap.
	(1)	(2)	(3)
Δln Hierarchical distance	-0.049***	-0.026***	0.023***
	(0.016)	(0.004)	(0.004)
All controls	Y	Y	Y
Observations	2,610	2,610	2,610
R-squared	0.376	0.192	0.607
#prefectures	261	261	261

Notes: This table shows that ln rank in hierarchical distance matters for both gaining capital status and losing capital status. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 6: Results Using Hierarchical Distance as an Instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent variable: ln Pop Density							
0		Reduced-form			IV		
Sample	All	Ever-Capital	Never-Capital	All	All	All	All
Prov. Capital				0.716***	0.922***	0.851**	0.661***
				(0.256)	(0.298)	(0.419)	(0.256)
ln Rank in Hierarchical distance	-0.059***	-0.141***	0.000				
	(0.022)	(0.038)	(0.029)				
ln Rank in H dist. KF * Post-					0.028		
					(0.028)		
ln Rank in H dist. NJ * Pre-						0.020	
						(0.039)	
ln Rank in H dist. BJ * Pre-							-0.048
							(0.044)
All controls	Y	Y	Y	Y	Y	Y	Y
		First-stage			First-stage		
Dependent var.: Prov Capital							
ln Rank in Hierarchical distance				-0.083***	-0.099***	-0.063***	-0.084***
				(0.007)	(0.010)	(0.009)	(0.007)
ln Rank in H dist. KF * Post-					0.026**		
					(0.011)		
ln Rank in H dist. NJ * Pre-						-0.034***	
						(0.010)	
ln Rank in H dist. BJ * Pre-							0.133
							(0.015)
All controls				Y	Y	Y	Y
Observations	2,871	693	2,178	2,871	2,871	2,871	2,871
R-squared	0.879	0.924	0.881	0.881	0.880	0.881	0.882
# Prefectures	261	63	198	261	261	261	261
F-Stat (Weak instrument test)				133.6	99.4	50.2	132.3

Notes: This table presents the estimates using logged rank in hierarchical distance as an instrument. Columns (2)-(3) show that rank in hierarchical distance matters only in ever-capital prefectures but not never-capital prefectures, implying that capital status is critical for our IV finding. Columns (5)-(7) show that placebo hierarchical distances calculated based on placebo national capitals do not affect our IV estimates. These results include all the controls in our baseline specification (equation (2)). ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 7: Provincial Capital Status and Public Offices in 2000

	(A) Total employed	(B) Employed pop in Governments Agencies, Party agencies and Social Organizations	(C) Employed pop in Health Care, Sports and Social Welfare	(D) = (A)-(B)-(C)
Effect on ln total employed through: = (I) * (II)	0.474	2.26%*0.870=0.020	0.96%*0.962=0.010	96.78%*0.455=0.440
(I) Percentage in total employment	100	2.26	0.96	96.78
	(1)	(2)	(3)	(4)
(II) Effect of prov. Capital on	ln (A)	ln (B)	ln (C)	ln (D)
Prov. Capital	0.474***	0.870***	0.962***	0.455***
	(0.101)	(0.091)	(0.102)	(0.101)
All controls	Y	Y	Y	Y
Observations	261	261	261	261
R-squared	0.712	0.742	0.764	0.709

Notes: This table shows that public offices are affected by capital status but this channel only explains a small part of the impact on total employment. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 8: Provincial Capital Status and Human Capital in the Qing Dynasty and 2000

	(1)	(2)	(3)	(4)	(5)	(6)
	Standardized			Decile in		
	ln Jinshi per capita, Qing	ln high school and above per capita, 2000	(2)-(1)	Jinshi per capita, Qing	High school and above per capita, 2000	(5)-(4)
Prov. Capital	0.764***	1.262***		2.134***	3.765***	
	(0.128)	(0.210)		(0.372)	(0.363)	
Gaining Capital Status			0.954***			2.057**
			(0.287)			(0.935)
Losing Capital Status			-0.696**			-1.535**
			(0.284)			(0.717)
All controls	Y	Y	Y	Y	Y	Y
Observations	261	261	261	261	261	261
R-squared	0.507	0.586	0.540	0.437	0.602	0.464

Notes: This table shows that provincial capital status affects human capital. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 9: Provincial Capital Status and Spatial Centrality in 1080, 1580, 1820, and 1990

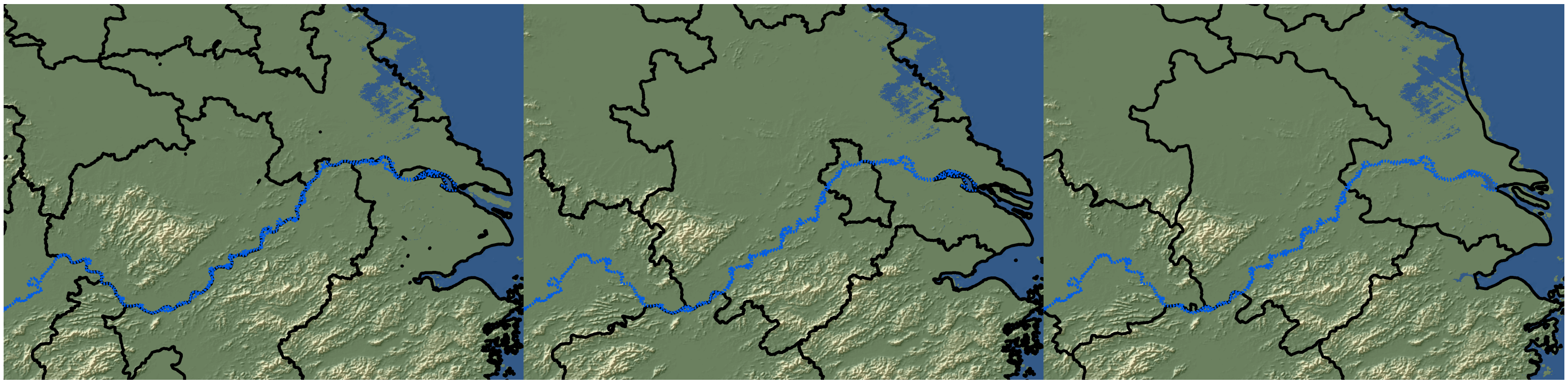
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\theta = 5.6$		$\theta = 2.8$		$\theta = 14$		$\theta = 28$	
	centrality	Δcentrality	centrality	Δcentrality	centrality	Δcentrality	centrality	Δcentrality
Prov. Capital	0.186***		0.141***		0.214***		0.227***	
	(0.048)		(0.033)		(0.062)		(0.070)	
Gaining Capital Status		0.263***		0.196***		0.290**		0.294**
		(0.094)		(0.071)		(0.118)		(0.132)
Losing Capital Status		-0.165**		-0.137***		-0.182**		-0.191**
		(0.067)		(0.053)		(0.078)		(0.085)
All controls	Y	Y	Y	Y	Y	Y	Y	Y
Observations	1,044	783	1,044	783	1,044	783	1,044	783
R-squared	0.699	0.727	0.752	0.747	0.603	0.661	0.549	0.612
# prefectures	261	261	261	261	261	261	261	261

Notes: This table shows that provincial capital status affects the spatial centrality of a prefecture. θ captures the relative cost of traveling without and with the transportation networks. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Online Appendix

Figure A.1: An Example of Boundary Changes

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



Notes: This figure presents an example of the evolution of provincial boundaries from following the natural mountains and rivers (known as "*suiting the forms of mountains and rivers*") to intentionally avoiding them (coined as "*interlocking like dog's teeth*"). Yangtze River was used as part of provincial boundaries in the Song but got included within provinces in the Ming and the Qing.

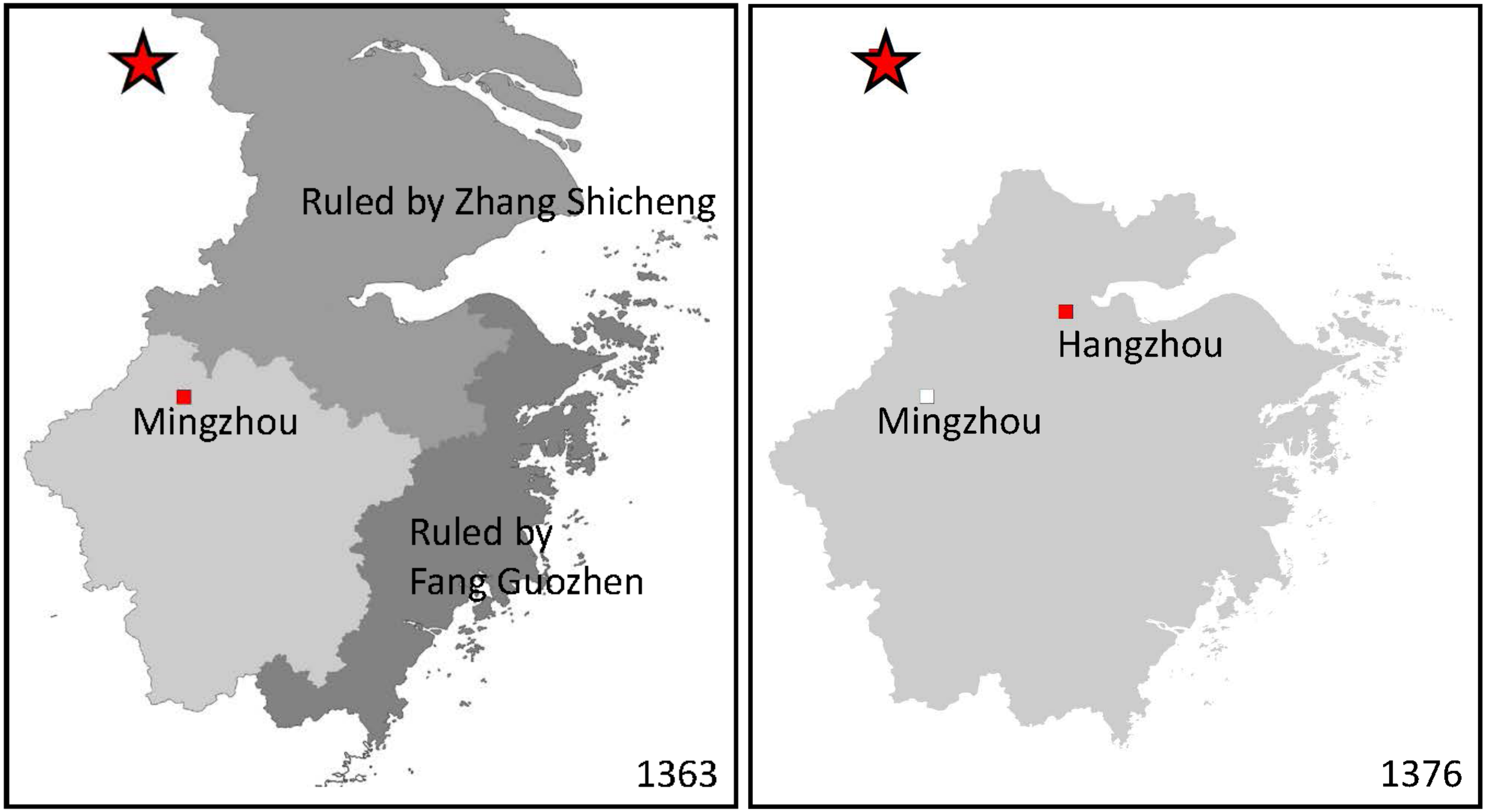
Figure A.2: Examples of Provincial Capital Relocation

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

Notes: This figure presents two examples of provincial capital relocation. The cross indicates the provincial centroid, the hollow/solid star indicates the past/current national capital, and the hollow/solid square indicates the past/current provincial capital. Luzhou and Changsha were capitals in the Song. Both lost their capital status in the Ming. Changsha regained the capital status in the Qing but Luzhou didn't. These patterns are driven by the relocation of national capitals and redivision of provinces across regimes, which can be captured by the algorithm in equation (1).

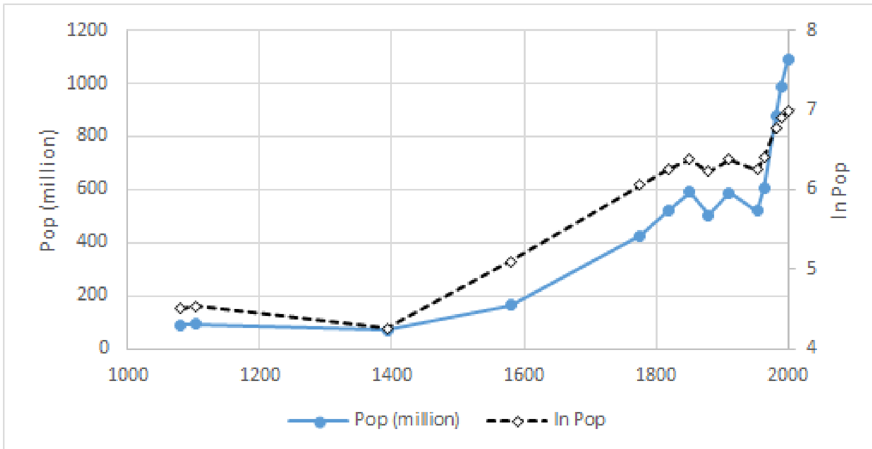
Figure A.3: An Example on the Timing of Changes in Provincial Boundary and Capital

(a) Provincial Boundary and Capital in 1363 (b) Provincial Boundary and Capital in 1376



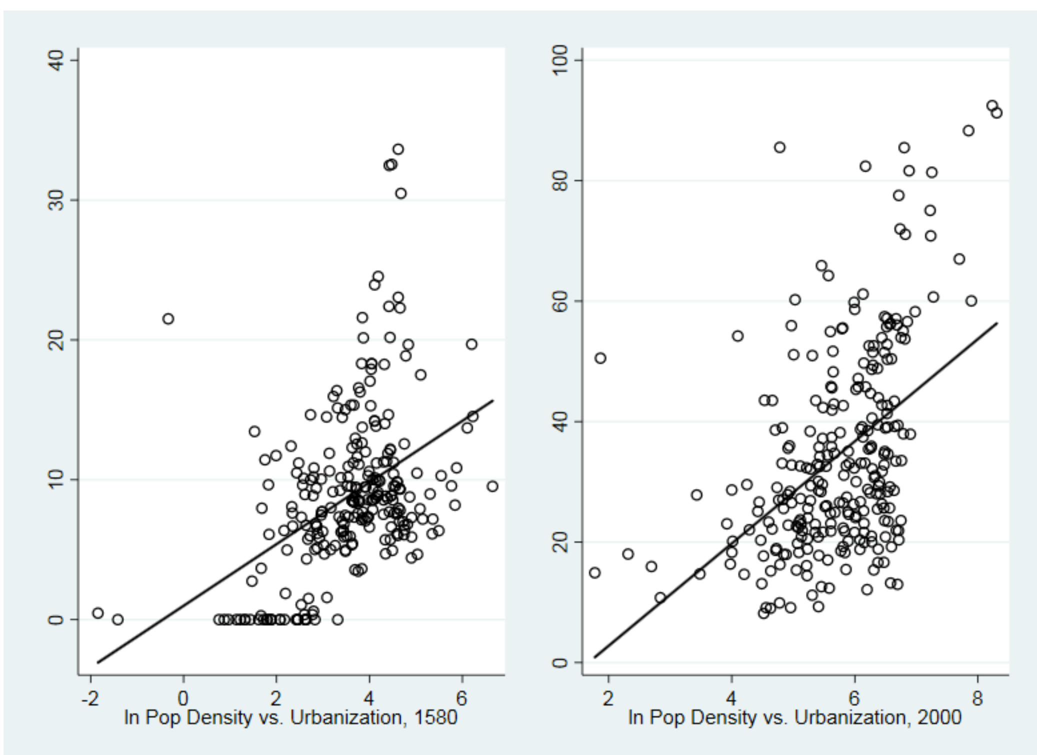
Notes: This figure presents an example on the timing of changes in provincial boundary and capital recorded in the History of Ming (Zhang 1739). In the end of the Yuan dynasty, two regions of Zhejiang province were ruled by two warlords (Zhang Shicheng and Fang Guozhen), as shown in panel (a). The new ruler of the Ming dynasty conquered the two regions and redefined the boundary of Zhejiang by incorporating the region ruled by Fang and dividing the region ruled by Zhang (see (panel (b)). After these changes, Hangzhou was designated as the new provincial capital.

Figure A.4: The Trends in Population



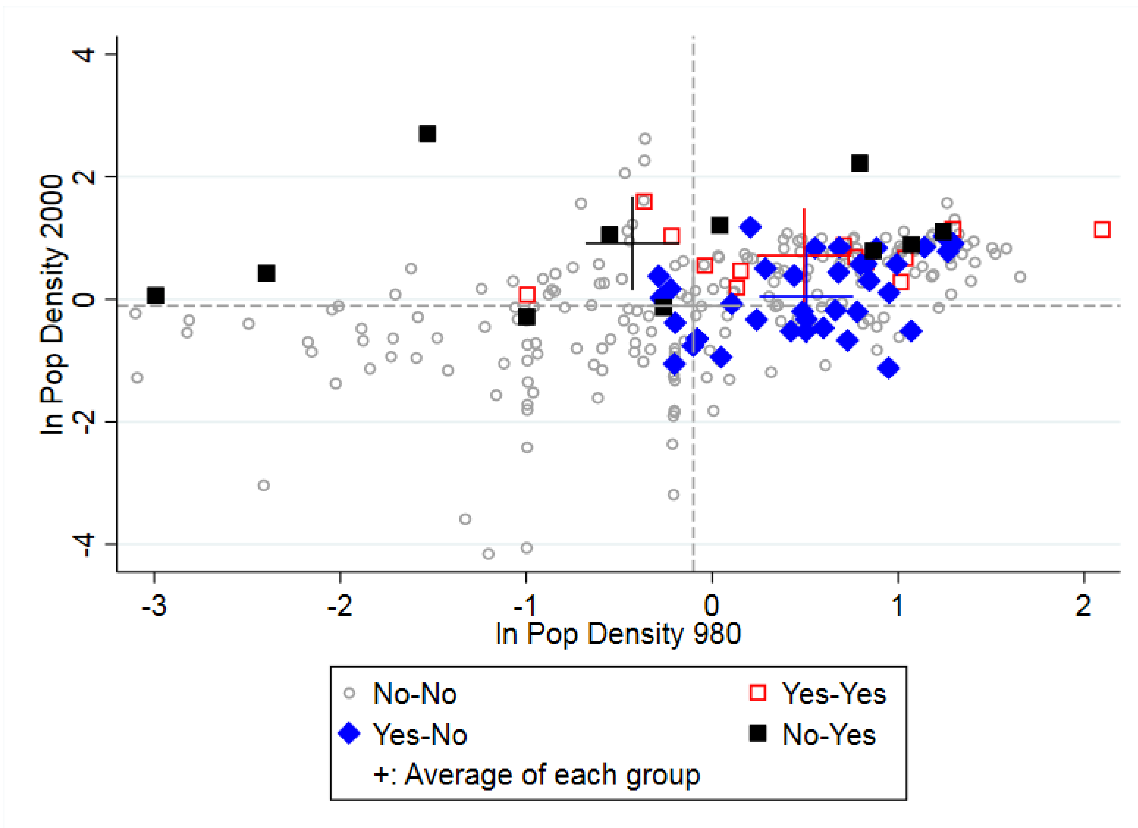
Notes: The figure plots the trends in population in our data.

Figure A.5: Pop Density vs. Urbanization in 1580 and 2000



Notes: This figure shows that population density is strongly correlated with urbanization rates in both 1580 and 2000. The correlational coefficients are 0.440 in 1580 and 0.471 in 2000.

Figure A.6: Pop Density in 980 vs. Pop Density in 2000



Notes: This figure shows a similar pattern as in Figure 4. Here, we use logged population density instead of logged light density in 2000 (Figure 4).

Table A.1: Summary Statistics & Data Sources

	Sources	Obs.	Mean	S.D.
ln Population density	1, 2, 3, 4	2,871	4.66	1.54
Urbanization ratio (%)	3, 4	1,566	15.67	14.54
Provincial capital	2, 5	2,871	0.10	0.30
Capital-ever dummy	2, 5	261	0.24	0.43
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
ln Area	2	261	9.30	0.85
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
ln Rank in Hierarchical distance to national capital ($\lambda = 0.19$)	2	2,871	1.81	0.83
ln Rank in Hierarchical distance to national capital ($\lambda = 0$)	2	2,871	1.81	0.82
ln Rank in Hierarchical distance to national capital ($\lambda = 1$)	2	2,871	1.81	0.83
ln Centrality, $\theta = 2.8$	2, 7	1,044	-0.64	0.25
ln Centrality, $\theta = 5.6$	2, 7	1,044	-0.16	0.29
ln Centrality, $\theta = 14$	2, 7	1,044	0.40	0.36
ln Centrality, $\theta = 28$	2, 7	1,044	0.73	0.42
Distance to provincial capital based on transportation (100 km)	7	1,038	1.15	0.92
Great circle distance to provincial capital (100 km)	7	1,038	1.97	1.44
Distance to neighbor provincial capital based on transportation (100 km)	7	1,044	1.54	0.92

Notes: This table presents the summary statistics for the major variables in our analysis.
Sources:
1. Liang (1980);
2. CHGIS (2007);
3. Ge (2000);
4. Population Census 1953, 1964, 1982, 1990, 2000;
5. Treatise of the Nine Regions from the Yuanfeng reign (1078-1085);
6. FAO GAEZ (2012);
7. Cheng and Hsu (1980)

Table A.2: Correlation b/w Prefecture Characteristics and Ever-Capital Status
Dependent Var.: Ever-Capital=1/0

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	OLS	OLS	Probit
Plain	0.228***	0.238***	0.235***	0.233***	0.231***	0.248***	0.224***
	(0.046)	(0.056)	(0.060)	(0.060)	(0.060)	(0.061)	(0.048)
Main River	0.152***	0.113**	0.149**	0.162**	0.163**	0.169**	0.154***
	(0.050)	(0.056)	(0.066)	(0.068)	(0.068)	(0.066)	(0.056)
ln Area			-0.062	-0.055	-0.054	-0.099	-0.090
			(0.087)	(0.088)	(0.088)	(0.098)	(0.075)
Coastal			0.003	0.002	0.002	-0.003	-0.010
			(0.023)	(0.023)	(0.023)	(0.026)	(0.027)
Slope			0.034	0.032	0.032	0.028	0.030
			(0.039)	(0.039)	(0.039)	(0.046)	(0.042)
ln Elevation			0.014	0.007	0.006	0.014	0.015
			(0.012)	(0.014)	(0.014)	(0.013)	(0.014)
Latitude			0.005	0.010	0.011	-0.010	-0.010
			(0.015)	(0.016)	(0.016)	(0.017)	(0.016)
Longitude			0.003	-0.030	-0.033	0.003	0.005
			(0.039)	(0.047)	(0.046)	(0.039)	(0.040)
ln Calories: Old World Crops				0.032			
				(0.028)			
ln Calories: All Crops					0.035		
					(0.027)		
Suitability: wheat						-0.011	-0.007
						(0.047)	(0.043)
Suitability: rice						-0.065	-0.069
						(0.051)	(0.044)
Suitability: maize						0.084**	0.087*
						(0.041)	(0.047)
Suitability: sweet potato						-0.072	-0.067
						(0.054)	(0.048)
Suitability: millet						0.023	0.022
						(0.041)	(0.038)
Region FE		Y	Y	Y	Y	Y	Y
Observations	261	261	261	261	261	261	261
R-squared	0.082	0.100	0.111	0.115	0.116	0.134	

Notes: This table presents the correlations between a prefecture's time-invariant characteristics and its ever-capital status. Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table A.3: The Impact of Gaining and Losing Provincial Capital Status: Period-by-Period Results (Dependent Var.: ln Pop Density)

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

Notes: Using the period before the capital status changes as the reference group, this table shows that there are no systematic differences in population density before the status change. These results include all the controls in our baseline specification (equation (2)). The eight prefectures with multiples changes are excluded from this analysis. Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table A.4: The Impact of Capital Status on Population Density: Ever-Capital Prefectures

	Ever-Capital Prefs				Ever-Capital Prefs + Comparison Group			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Neighbors within 100km		Propensity score matching	
Propensity matching								
Prov. Capital	0.544***	0.459***	0.291**	0.311**	0.535***	0.368***	0.588***	0.403***
	(0.164)	(0.146)	(0.122)	(0.124)	(0.128)	(0.111)	(0.136)	(0.140)
Capital-ever * Year FE						Y		Y
Year FE * Crop suitability				Y		Y		Y
Year FE * Geography			Y	Y		Y		Y
Year FE * ln Area		Y	Y	Y		Y		Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Pref. FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	693	693	693	693	1,397	1,397	1,221	1,221
R-squared	0.803	0.888	0.913	0.926	0.836	0.925	0.813	0.890
# Prefectures	63	63	63	63	127	127	111	111

Notes: This table shows that provincial capital status is associated with a higher population density within the subgroup of prefectures that have ever been a provincial capital (columns (1)-(4)). The finding also holds if we add a comparison group of their neighbors defined by distance or by the nearest score using propensity score matching (columns (5)-(8)). Crop suitability refers to the suitability of rice, wheat, millet, sweet potatoes and maize. Geographical controls include whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude. Region refers 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 A.5: IV Estimates: Alternative Functional Forms of Hierarchical Distance

	$\alpha = 1$						$\alpha = 0.5$	$\alpha = 2$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\lambda =$	0.19 (optimal)	0.1	0.2	0.3	0.4	0.5	0.25 (optimal)	0.10 (optimal)
IV Estimates: ln Pop Density								
Prov. Capital	0.716***	0.611**	0.829***	0.684**	0.761**	0.849**	0.819***	0.699**
	(0.256)	(0.289)	(0.262)	(0.266)	(0.299)	(0.330)	(0.246)	(0.294)
All controls	Y	Y	Y	Y	Y	Y	Y	Y
First-stage: Provincial Capital								
ln Rank in H dist.	-0.083***	-0.073***	-0.080***	-0.077***	-0.066***	-0.060***	-0.086***	-0.070***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)
All controls	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,871	2,871	2,871	2,871	2,871	2,871	2,871	2,871
R-squared	0.881	0.882	0.881	0.882	0.881	0.881	0.881	0.882
# Prefectures	261	261	261	261	261	261	261	261
F-Stat (Weak instrument)	133.6	127.4	122.3	96.3	78.8	74.2	146.3	99.3

Notes: This table shows that that our IV results are robust to different functional forms of hierarchical distance rank. These results include all the controls in our baseline specification (equation (2)). Standard errors are presented in the paraphrases. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table A.6: IV Estimates: Considering Distance to Market Centers

	(1)	(2)	(3)	(4)
IV Estimates: ln Pop Density				
Prov. Capital	0.716***	0.626***	0.728***	0.704***
	(0.256)	(0.229)	(0.227)	(0.221)
ln Rank in H dist. to major econ region		-0.020		
		(0.023)		
ln Rank in H dist. to the East (Shanghai)			0.033	
			(0.026)	
ln Rank in H dist. to the South (Guangzhou)				0.024
				(0.023)
All controls	Y	Y	Y	Y
First-stage: Provincial Capital				
ln Rank in Hierarchical distance	-0.083***	-0.093****	-0.095***	-0.096***
	(0.007)	(0.008)	(0.008)	(0.008)
ln Rank in H dist. to major econ region		-0.024***		
		(0.008)		
ln Rank in H dist. to the East (Shanghai)			-0.019**	
			(0.009)	
ln Rank in H dist. to the South (Guangzhou)				-0.013
				(0.008)
All controls	Y	Y	Y	Y
Observations	2,871	2,871	2,871	2,871
R-squared	0.881	0.865	0.864	0.865
# Prefectures	261	261	261	261
F-Stat (Weak instrument test)	133.6	134.5	137.5	144.2

Notes: This table shows that our IV estimates are not affected by distance to national market centers. These results include all the controls in our baseline specification (equation (2)). Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table A.7: The Impact of Capital Status on Urbanization (1580-2000)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Var.	Urbanization Ratio		ln Urban Pop		ln Rural Pop	
Prov. Capital	11.011***	10.934***	0.636***	0.595***	0.197**	0.186**
	(2.100)	(2.287)	(0.166)	(0.151)	(0.081)	(0.094)
Year FE * Crop suit.		Y		Y		Y
Year FE *Geography		Y		Y		Y
Year FE * Ln Area	Y	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y
Pref. FE, Year FE	Y	Y	Y	Y	Y	Y
Observations	1,531	1,531	1,531	1,531	1,531	1,531
R-squared	0.782	0.814	0.879	0.892	0.845	0.872
# Prefectures	261	261	261	261	261	261

Notes: This table shows that our findings hold when using urbanization as an alternative outcome. Crop suitability refers to the suitability of rice, wheat, millet, sweet potatoes and maize. Geographical controls include whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude. Region refers 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 A.8: Grid-level Analysis (1-degree × 1-degree)

	ln Pop Density				Δln Pop Density	
	(1)	(2)	(3)	(4)	(5)	(6)
Prov. Capital	0.381***	0.379***	0.353***	0.338***		
	(0.079)	(0.080)	(0.070)	(0.065)		
ΔProv. Capital					0.291***	
					(0.061)	
Gaining Capital Status						0.246***
						(0.091)
Losing Capital Status						-0.324***
						(0.075)
Year FE * Crop suit.				Y	Y	Y
Year FE *Geography			Y	Y	Y	Y
Year FE * Ln Area		Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y
Pref. FE, Year FE	Y	Y	Y	Y	Y	Y
Observations	3,971	3,971	3,971	3,971	3,610	3,610
R-squared	0.804	0.804	0.833	0.845	0.602	0.602
#Prefectures	361	361	361	361	361	361

Notes: This table shows that our findings hold when mapping the data to the grid level. Crop suitability refers to the suitability of rice, wheat, millet, sweet potatoes and maize. Geographical controls include whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude. Region refers 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 A.9: Heterogeneous Effects w.r.t. Natural Advantages

	ln Pop Density				Δ ln Pop Density	
	(1)	(2)	(3)	(4)	(5)	(6)
Prov. Capital	0.454***	0.444***	0.457***	0.443***		
	(0.103)	(0.102)	(0.102)	(0.101)		
Prov. Capital * max suitability (old world)	-0.169**					
	(0.083)					
Prov. Capital * max suitability (old+new)		-0.139*				
		(0.084)				
Prov. Capital * avg. suitability (old world)			-0.169**			
			(0.079)			
Prov. Capital * avg. suitability (old+new)				-0.150*		
				(0.078)		
Δ Prov. Capital					0.436***	
					(0.110)	
Δ Prov. Capital * max suitability (old world)					-0.186*	
					(0.096)	
Gaining Capital Status						0.444**
						(0.210)
Losing Capital Status						-0.431***
						(0.127)
Gaining Status * max suitability (old world)						-0.193*
						(0.110)
Losing Status * max suitability (old world)						0.179
						(0.149)
Year FE * Crop Suitability	Y	Y	Y	Y	Y	Y
Year FE * Geography	Y	Y	Y	Y	Y	Y
Year FE * ln Area	Y	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y		
Year FE	Y	Y	Y	Y	Y	Y
Observations	2,871	2,871	2,871	2,871	2,610	2,610
R-squared	0.865	0.865	0.865	0.865	0.658	0.658
# Prefectures	261	261	261	261	261	261

Notes: This table shows that provincial capital status matters less for those with higher agricultural productivity. The finding also provides evidence for the heterogeneity pattern in Figure 4. Crop suitability refers to the suitability of rice, wheat, millet, sweet potatoes and maize. Geographical controls include whether a prefecture contains a plain, a major river, whether it is on the coast, as well as its slope, elevation, longitude, and latitude. Region refers 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 A.10: Centrality – Separating the Two Periods before Provincial Capital Status Changes

	Centrality		
	(1)	(2)	(3)
Gaining Pre-2	0.097		0.103
	(0.132)		(0.131)
Gaining Post	0.309**		0.298*
	(0.148)		(0.152)
Losing Pre-2		-0.005	-0.004
		(0.102)	(0.101)
Losing Post		-0.174***	-0.168***
		(0.057)	(0.057)
All controls	Y	Y	Y
Observations	1,016	1,016	1,016
R-squared	0.696	0.696	0.699
#Prefecture	254	254	254

Notes: This table shows that there are no systematic differences in centrality before the capital status change – we only have four periods of data on transportation networks though. These results include all the controls in our baseline specification (equation (2)). The seven prefectures with multiple changes in these four periods are excluded from this analysis. Standard errors presented in the paraphrases are clustered at the prefecture level. ***: significant at 1%, **: significant at 5%, *: significant at 10%.