The Oriental City:

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

Ying Bai† and Ruixue Jia‡

April 24, 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. In particular, our analysis shows that economic advantages driven by political factors do not persist; rather, losing (gaining) capital status is associated with worse (better) economic development as measured by population density and urbanization. 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. To explain the politico-economic link, we show that the spatial importance of a region (e.g., centrality in the transportation networks) mirrors its status in the political hierarchy, which partially accounts for the rise and fall of certain regions in the long run.

*We thank Daron Acemoglu, Konrad Burchardi, Oded Galor, Walker Hanlon, Gordon Hanson, Yi Lu, Torsten Persson, Aleh Tsyvinski, Fabarizio Zilibotti and seminar participants at CESifo, CIFAR, HKU, IIES, and Tsinghua 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 (see De Long and Shleifer 1993, Ades and Glaeser 1995, and Davis and Henderson 2003 for the implications of 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 and uncover the underlying mechanisms. In doing so, we also speak to a growing body of scholarship documenting the persistence of economic activities in the long term as a result of locational fundamentals (e.g., Davis and Weinstein 2002, Miguel and Roland 2011) or large economic shocks (Redding, Sturm, and Wolf 2011, Bleakley and Lin 2012, Kline and Moretti 2013, Michaels and Rauch 2016, Hanlon 2017). In contrast to this line of research, however, we find that economic advantages driven by political factors do not necessarily persist.

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 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

[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-county 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.

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 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 and other sources, we construct a panel dataset across 261 prefectures for 11 periods (1078, 1102, 1580, 1776, 1820, 1851, 1910, 1964, 1982, 1990, and 2000),[2] using both a difference-in-differences strategy and an IV strategy to evaluate how political status shapes economic development. We employ the first to show that gaining (losing) provincial capital status is associated with higher (lower) population density, at a magnitude of around 40-50%. One major concern with this method, however, is whether capital status changes because of omitted variables that affect population density. To partially address this concern, we show that population density trends in the capital-to-be (capital-to-lose) prefectures are no different from the rest before

[2] We use the year 2000 prefecture boundaries to build our panel dataset in our baseline analysis. In addition, we report the results of a grid-level analysis.

they gain (lose) capital status. Our findings also hold for the subsample of cities that were at some time a provincial capital when no omitted variables prevented them from becoming one.

Using hierarchical distance rank as an instrument for capital status, we derive an estimate comparable to that using difference-in-differences: the provincial capital status is associated with about a 60% 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 using alternative measures, whose outcomes also underscore the relative importance of the political and economic factors underlying our findings. First, 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. Second, the two-part structure of hierarchical distance also enables an overidentification test, which reveals that one part does not matter once we use the other part as instrument. 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 considering other important factors affecting population density during this millennium, including deadly wars and land productivity changes induced by the Columbian exchange. We also perform robustness tests using urbanization as an alternative outcome, employing grid-level data, and examining different subperiods.

What, then, explains the link between political hierarchy and economic development? In both historical and modern China, the state plays a critical role in building the infrastructure that determines market access. Indeed, empirically, we find that a region's spatial importance (e.g., centrality or gravity in the transportation networks) mirrors its status in the political hierarchy. We also show that losing (gaining) capital status decreases (increases) a prefecture's gravity, implying that political status alters market access. Lastly, to overcome the difficulty of directly measuring how resources and information flow with transportation networks, we test an auxiliary prediction on the importance of resource reallocation. That is, we expect the effect of capital status to be less important for prefectures with more natural advantages (e.g., higher agricultural productivity) because resource reallocation is less critical for their development. Our results support this conjecture.

Overall, our findings suggest that advantages driven by political status do not necessarily persist, which contributes to the literature on the long-term spatial distribution of economic activities. Yet most of these studies focus on persistent patterns in these activities, such as the minimal impact on development of large temporary shocks to a region (e.g., Davis and Weinstein 2002, Miguel and Roland 2011) 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). Our observations are different because our findings are driven by political factors, which are little studied in this research stream.

Although the importance of political factors is broadly emphasized in the literature 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 uncover important factors underlying the changes, and shed light on the mechanisms through which capital status affects development. Our study also offers a perspective on how to link this literature with the above studies on path dependence.

Our finding on the mapping from political hierarchy to spatial hierarchy is related to the literature emphasizing the importance of market access for regional development (e.g., Redding and Sturm 2008, Donaldson and Hornbeck 2015). In our context, a prefecture's market access is in part determined by its political status because the political hierarchy is an important consideration in the transportation networks constructed by the state.

Finally, our findings on what drives capital status contribute to the literature investigating the causes and consequences of national and subnational divisions (Alesina and Spolaore 1997, Gennaioli and Rainer 2007, Michalopoulos and Papaioannou 2013). Even though the long-lasting administrative hierarchy may be unique to China, the implications of our study are equally relevant for other contexts, both historic and modern. For instance, the Ottoman Empire was also organized under a systemic administrative hierarchy, which may matter for its urban landscape (Khoury 2002). Likewise, the capital cities in Latin American countries are larger because of political factors even though their status rarely changes (Galiani and Kim 2011).

The rest of the paper is organized as follows. Section 2 describes the background and provides a conceptual framework in which 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 then describes the mapping between political hierarchy and spatial hierarchy, and discusses auxiliary evidence. 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. 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).

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.

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,[3] 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.[4]

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. Like national capitals, 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

[3]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.

[4]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/).

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.[5] 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 Our first issue of interest is why capital status changes, a question that the background discussion suggests 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,

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

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 is a weight variable – such as the area of prefecture j – to capture the scale of resources. Likewise, λ is a weight value between 0 and 1 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.[6]

Even though 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).

Remarks When considering the change in provincial capitals, we take provincial redivision as given because, as administrative histories document, in most cases, the central government determined provincial boundaries first and then decided on the provincial capital (e.g., Zhang 1739).[7] On top of this, if the central government were awarding capital status only on development potential, it could do so without the provincial boundary changes. But we

[6]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 $\lambda = 0.19$.

[7]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.)

usually observe boundary changes accompany capital changes. Due to the facts, we do not endogenize N.

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 two clear patterns: First, 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. Second, the location of provincial capitals as a group evolves with national capital relocation. For instance, when Beijing (Nanjing) became the national capital at the beginning of the Yuan (Ming) Dynasty, the group of provincial capitals orbited around it in a pattern consistent with the logic of political control from the central government's perspective.

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 1078-2000 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 – 1078, 1102, 1580, 1776, 1820, 1851, 1910, 1964, 1982, 1990, and 2000 – and calculate population density based on prefecture boundaries in 2000 (see Appendix Figure A.3 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 are also able to access population data in years following three deadly conflicts: 1393, after the Mongolian conquest, 1880, after the Taiping Rebellion, and 1953, after World War II. These three wars wiped out over 20%, 15%, and 10% of the total population at that time. Although including these additional three periods does not vary our baseline results, we intentionally employ them separately as a robustness check.

Urbanization in 1580-2000 Compared with population data, urbanization data are less systematically available, being accessible for only 6 of our 11 years: 1580, 1820, 1964, 1982, 1990, and 2000. The 1580 data are estimated based on local gazetteers,[8] 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.4), 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. 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

[8]We thank Cao Shuji, a leading scholar in Chinese population history at Shanghai Jiaotong University, for providing this information.

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.[9]

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: 1078 & 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,

(2) capitals in 1078 but not in 2000,

(3) capitals in 2000 but not in 1078,

(4) not capitals in either period.

In Panel (a) of Figure 4, the x-axis indicates the standardized logged population density in 1078, while the y-axis indicates the standardized logged light density in 2000, allowing the four quadrants to be interpreted according to whether a prefecture's economic development is above average in each period. These patterns remain similar if we use logged population density instead of logged light density in year 2000 (see Appendix Figure A.5).

Our analysis of the prefecture groups (see Figure 4) reveals three important facts. First,

[9] 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.

the development proxies in 1078 and 2000 are positively correlated, which is consistent with a large body of literature documenting the persistence of economic development. Second, in addition to persistence, the figure also shows systematic changes. In particular, within group (2) (capitals in 1078 but not in 2000), a large share (50%) of the prefectures move from being above average to being below average once they lose capital status. In contrast, the positions of group (1) (capitals in both periods) are relatively stable. Within group (3) (capitals in 2000 but not in 1078), 100% of the prefectures moved from being below average to above average after gaining political capital status. Of course, even without changes in capital status, a prefecture's relative economic development can also vary over such a long period. However, as Panel (b) shows, the changes in group (2) and group (3) appear significantly more frequent than those in group (4), suggesting that (as later tested in Section 4), capital status matters for the rise and fall of economic development. Lastly, the figure reveals heterogeneity within group (2): the economic performance of a subgroup of prefectures is still above average even after losing capital status (a finding explained in Section 5.2).

4 Provincial Hierarchy and Economic Development: Empirical Results

4.1 Difference-in-Differences Analysis

Baseline Results Our baseline estimation examines the correlation between provincial capital status and population density in 1078, 1102, 1580, 1776, 1820, 1851, 1910, 1964, 1982, 1990 and 2000. We also show that the results are robust to dropping any specific year or using one year for each regime. Our difference-in-differences estimator is as follows:

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

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 the set of prefecture characteristics that includes various geographic and agricultural measures whose effects we allow 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 $\pi_m \times \gamma_t$). All standard errors are clustered at the prefecture level.

The resulting estimates indicate that provincial capital status is associated with a 43-52% 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.

Losing vs. Gaining Capital Status The major concern with interpreting the results in Table 2 as the impact of capital status is that they may reflect omitted variables that enabled a capital prefecture to experience improved development without capital status. To gauge the importance of this concern, we compare the loss of capital status with its acquisition under the assumption that if a capital prefecture could have improved its development without capital status, losing status would matter little.

Specifically, we apply a first-difference strategy to our data for the 44 (19) prefectures that lost (gained) capital status once and show that change in capital status ($\Delta \text{Capital}_{i,t}$) is positively associated with change in population density (column (5)) at values similar to those derived using levels.

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 (6)). Although the negative coefficient on this control suggests that less dense prefectures grow faster, change in provincial capital status remains important.

In columns (7) and (8) of Table 2, we separate capital status loss from capital status gain and show that both have a sizable impact, with absolute magnitudes that are not significantly different (see the p-values in the last row).

This finding is particularly important for two reasons. 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 last in the long run, which seems to contradict claims that a large positive or negative shock persists over the long term (e.g., Bleakley and Lin 2012, Kline and Moretti 2013, Michaels and Rauch 2016, Hanlon 2017). As previously emphasized, however, our setting differs in the importance of political factors that have not been much studied by this line of research.

Checking Pre-Trends To check whether the baseline finding is driven by pre-trends, we employ a more flexible method that estimates the link between capital status and population densities in each separate period of losing or gaining capital status. Since the periods are not of equal length, we use two specifications, the first of which ignores period length for simplicity:

$$\ln PopDensity_{i,t} = \sum_{\tau=-3+}^{2+} \beta_\tau Capital_{i,\tau} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \quad (3)$$

where $\tau = -3+, -2, 0, 1, 2+$. 0 indicates the period of gaining (losing) capital status and -3+ indicates more than three periods before doing so, with the period before capital status change ($\tau = -1$) as the reference group. We exclude the eight prefectures that experienced multiple changes.

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.

The second specification allows the length of the period considered ($\Delta Yr_{i,t}$) to alter the impact of capital status:

$$\begin{aligned} \ln PopDensity_{i,t} = \rho_1 PRE_{i,t} + \rho_2 PRE_{i,t} \times \Delta Yr_{i,t} + \rho_3 POST_{i,t} + \rho_4 POST_{i,t} \times \Delta Yr_{i,t} \\ + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \quad (4) \end{aligned}$$

where $\Delta Yr_{i,t}$ indicates the difference between year t and the year of status change; $PRE_{i,t}$ and $POST_{i,t}$ denote the years before and after the status change, respectively; and ρ_2 and ρ_4 represent the influence of period length. In our estimations, we also separate gaining status from losing status (see Appendix Table A.4). Neither ρ_2 nor ρ_4 is significant, indicating that the difference between capital prefectures and non-capital prefectures remains stable over time.

One limitation of our data is the unavailability of high-frequency observations across the 1,000 year time span (e.g., the median of period 0 in Figure 5(a) is 132 years), which makes it impossible to match population densities with exact years. A positive aspect of this limitation, however, is our confidence that the prefectures of interest were no different from others for a considerable time before gaining or losing capital status. These results thus serve primarily to show that our baseline finding is unlikely to be driven by pre-trends.

Using Ever-Capital Prefectures The 63 prefectures that were a provincial capital at least once serve as a useful subsample for examining the importance of capital status because the variation within them comes only from gaining or losing such status, with no omitted variables preventing any one prefecture from becoming a capital. 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 Table 3, 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)).

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. We thus next turn our attention to identifying the determinants of capital status.

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 Appendix Table A.5, 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 Appendix Table A.6), 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:

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

and

$$\ln PopDen_{i,t} = \beta' \widehat{Capital}_{i,t} + \alpha_i + \gamma_t + \theta \mathbf{X}_i \times \gamma_t + \theta' \pi_m \times \gamma_t + \epsilon_{i,t}, \quad (6)$$

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

These IV estimates, whose reduced form is reported in Table 4, column (1), provide further support for our main findings. The F-stat from the first stage is 145.9 (see the lower panel of column (2)), implying that instrument weakness is unlikely to be a concern. The IV estimate using $\lambda = 0.19$ is around 0.66 (the upper panel of column (2)), and ranges from 0.57 to 0.85 when we vary λ from 0.1 to 0.9 (see Appendix Table A.7). The IV estimates are thus comparable to our difference-in-differences estimates (which range from 0.43 to 0.52) and 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 check whether our findings are specific to political capital status by exploiting placebo hierarchical distances. Then, we employ the two components of our instrument for an over-identification test. Finally, we examine whether our findings are confounded by distances to major market centers.

(i) 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 4, columns (3)-(5)).

(ii) Two Components of the Instrument Our overidentification test is enabled by the fact that the two extreme values of λ ($\lambda = 0$ and $\lambda = 1$) generate two instruments. As column (1) of Appendix Table A.8 shows, these two components, whose reduced-form estimates are given in column (2), are strongly correlated with the instrument used in our baseline. Reassuringly, in the IV estimate using the $\lambda = 1$-component while controlling for the $\lambda = 0$-component (column (3)), this latter is no longer significant once we use $\lambda = 1$-component as an instrument. A similar pattern emerges when we use the $\lambda = 0$-component as an instrument while controlling for $\lambda = 1$-component (column (4)). When we employ both components as instruments (column (5)), the p-value of the overidentification test is 0.778, confirming the patterns in columns (3)-(4). Compared with using the weighted sum in Table 4, the F-stat of the first stage is smaller, suggesting that the weighted sum is a better predictor of capital status.

(iii) Distance to Major Market Centers To check whether our findings are confounded by distances to major market centers, we calculate a prefecture's (hierarchical) distance to three types of market centers: the north China plain during the Song Dynasty and the lower Yangtze after the Song (cf. Skinner 1977), Shanghai in the east, and Guangzhou in the south. To calculate the ranks of these distances, we replace $D_{i,NationalCap}$ in equation (1) with $D_{i,Market}$. Like the placebo distances, these ranks are also correlated with our instrument and thus may be correlated with the probability of being a capital. However, once again, none explains the role of our hierarchical distance (Table 4, columns (6)-(8)).

To be clear, this finding does not mean market access is not important. In fact, as we will show in Section 5, hierarchical distance in part determines market access because it affects the construction and maintenance of the transportation networks. Instead, this finding is to show 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 1078 versus 2000, our findings remain similar if we use urbanization (instead of population density), employ grid-level data (instead of year 2000 prefectures), drop any period in the data, or consider other important changes during the millennium. 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, 1982, 1990, 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.9) 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 × 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.10), indicating that our findings are robust to alternative boundary definitions.

Examining Different Periods In Appendix Table A.11, we test for robustness by dropping periods one by one after first replicating the baseline estimate using all the periods (column (1)). More specifically, we first separately drop the data for the Song, Ming, Qing, and People's Republic, respectively (columns (2)-(5)), then retain only one period for each regime (column (6)), and then restrict the data first to 1080, 1580, and 1851 (column (7)) and then to 1080 and 2000 (column (8)) The estimates vary little across all specifications.

Other Important Changes During 1000-2000 In addition to political regime changes, the millennium is marked by three deadly wars that affected population size and the Columbian exchange that led to changes in land productivity.

Our analysis already accounts for the influence of the Columbian exchange by allowing the impact of crop suitability to vary by year. The results remain similar regardless of whether we include population data from after the Mongolian conquest, the Taiping Rebellion, or World War II (i.e., 1393, 1880, and 1953, respectively; see Appendix Table A.12).

5 From Political Hierarchy to Spatial Hierarchy

Our results document that capital status clearly matters for economic development. Why would the status in the (vertical) political hierarchy matter for (spatial) economic development? We argue that because rulers in a centralized regime like China gather and distribute information and resources through a political hierarchy, the spatial importance of a region mirrors its status in the political hierarchy. We directly test this proposition using a prefecture's position in the transportation network, whose changing importance we expect to be associated with a change in its capital status.

To evaluate the political-spatial hierarchy link, we digitize the transportation networks in the Song, Ming, and Qing dynasties and 1990 (see Figure 6), taking into account the differences between modern and historical transportation networks. After determining the correlation between political status and spatial importance, we examine how the change in capital status affects transportation networks, subsequently reporting auxiliary evidence as a further test of our mechanism.

5.1 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.

The costliness to construct and maintain historic routes led to wide variation over time and across regions as both state capacity and the relative importance of different regions changed. This variation provides us a unique opportunity to systematically investigate the link between the transportation networks and the political hierarchy.

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 – and the railroad map for the People's Republic (1990) (see Figure 6).

Descriptive Patterns Although we can easily count the number of landways, waterways, and railways in a prefecture, we also want to account for the relative importance of the different conduits in the transportation network. To this end, we employ a centrality measure termed "gravity" in the urban literature, defined for each prefecture, i, within a range, r, as follows:

$$Gravity_i^r = \sum_{j \neq i, d_{i,j} \leq r} \frac{W_j}{e^{\beta d_{i,j}}} \tag{7}$$

where $d_{i,j}$ refers to the shortest distance between i and j, and W_j is the weight variable. A smaller r or larger β indicates greater importance of the local access.

Intuitively, the gravity measure captures the degree to which a prefecture is a hub in the transportation network. As a benchmark, we choose $W = 1$, $\beta = 1$, and $r = 500$K meters, implying a discount rate of 82% and 61% for regions 200K meters and 500K meters away. We also use alternative values for robustness.

Our gravity measure confirms the importance of the provincial capitals that our maps suggest are transportation hubs (see Figure 6). It also reveals a high correlation between a prefecture's status in the political hierarchy and its status in the spatial hierarchy. For example, our estimates using the period-standardized gravity measure for national capital, provincial capitals, and other prefectures (Table 5) show a provincial capital's gravity to be around a 0.3-0.5 standard deviation higher than average. The national capitals of Kaifeng (central China) and Nanjing (central south), in contrast, enjoy gravity 2 to 3 standard deviations higher than average, while Beijing, being located in the north, had only moderate gravity initially but enjoyed an increase over time (see columns (3)-(4)).

Capital Status and Spatial Gravity To test whether changes in capital status lead to changes in gravity, we replace the dependent variable in our baseline estimation with logged gravity and use it with the four-period panel data used in the transportation analysis.

The results show that a provincial capital's gravity is about 0.36 standard deviation higher than that of the non-capital prefectures (Table 6, columns (1)-(2)). In addition to deriving a similar result using our first-difference strategy (column (3)), we also show that, as in our baseline results, losing capital status implies a loss of gravity (column (4)). This latter is 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 using the alternative values of r and β, as well as employing prefecture area as the weight (columns (5)-(7)).

To additionally ensure the absence of significant pre-trends, we conduct a period-by-period analysis similar to that for our baseline findings, which shows that gravity decreases (increases) only after the loss (gain) of capital status (Appendix Table A.13). This finding is consistent with the fact that the capitals were assigned in the beginning of a dynasty whereas the transportation networks were altered later. Thus, our findings on gravity are also unlikely to be driven by reverse causality.

5.2 Auxiliary Evidence from Heterogeneous Effects

Although it would be ideal to examine how transportation networks affect the flow of resources and information, measuring this dynamic would be so difficult that we instead formulate an auxiliary hypothesis; namely, that the impact of political factors is less important for prefectures that enjoy natural advantages because for them, reallocation of resources is less critical. We proxy these 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 (Table 7, 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)), meaning that the results support this second auxiliary hypothesis.

The above findings also offer an explanation for the subgroup of prefectures whose performance remained above average in 2000 even after they lost capital status (see Figure 4). In fact, the crop suitability of this subgroup is 15%-20% higher than that of the prefectures whose performance after losing capital status was below average in 2000.

These results thus provide indirect evidence of the importance of political status, which seems to matter more for prefectures with fewer natural advantages.

6 Conclusions

Although political factors have long been argued to be important for economic development, 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 mechanism we document. That is, we show that in a regime in which the political hierarchy guides the distribution of information and resources, a region's spatial importance mirrors its status in the political hierarchy. 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] Alesina, Alberto, and Enrico Spolaore (1997), "On the Number and Size of Nations," *Quarterly Journal of Economics* 112(4): 1027-1056.

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

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

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

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

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

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

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

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

[11] 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.

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

[13] Galiani, Sebastian, and Sukkoo Kim (2011), "Political Centralization and Urban Primacy: Evidence from National and Provincial Capitals in the Americas," In Understanding Long-Run Economic Growth, edited by Costa and Lamoreaux, 121-54. Chicago.

[14] Gennaioli, Nicola, and Ilia Rainer (2007), "The Modern Impact of Precolonial Centralization in Africa," *Journal of Economic Growth* 12(3): 185-234.

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

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

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

[18] 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.

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

[20] Khoury, Dina Rizk (2002), *State and Provincial Society in the Ottoman Empire: Mosul, 1540-1834*. Cambridge University Press.

[21] 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.

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

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

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

[25] Michaels, Guy, and Ferdinand Rauch (2016), "Resetting the Urban Network: 117â2012," Economic Journal, In Print.

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

[27] Mostern, Ruth (2011), *Dividing the Realm in Order to Govern: The Spatial Organization*

of the Song State (960-1276 CE). Harvard University Press.

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

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

[30] 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.

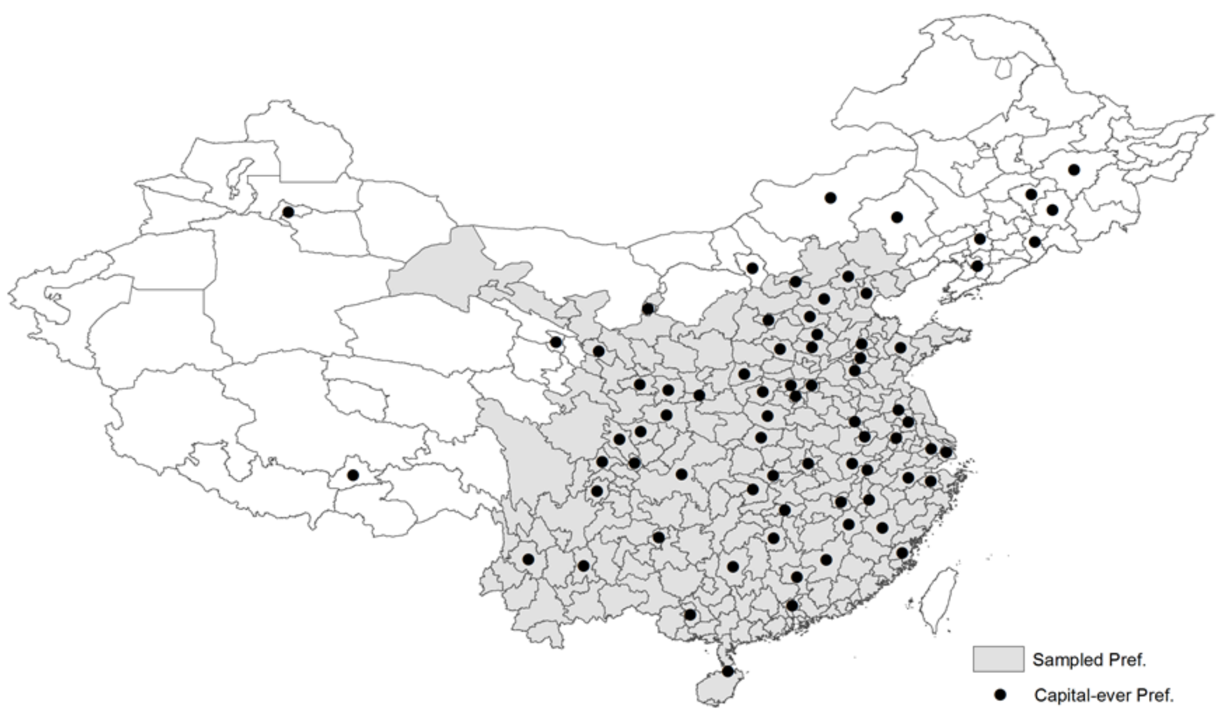
[31] Redding, Stephen J., Daniel M. Sturm, and Nikolaus Wolf (2011), "History and Industry Location: Evidence from German Airports," *Review of Economics and Statistics* 93(3): 814-831.

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

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

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

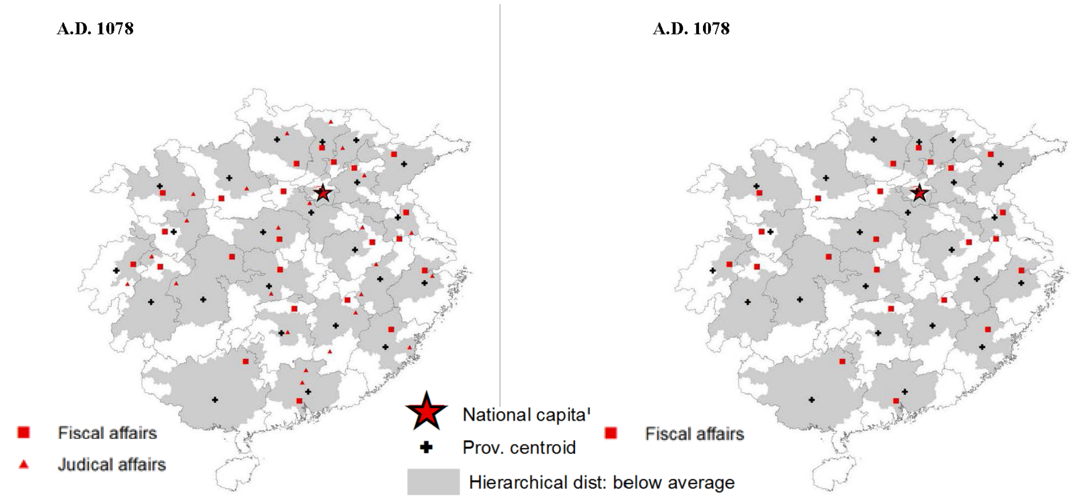
Figure 1: Capital-Ever Prefectures



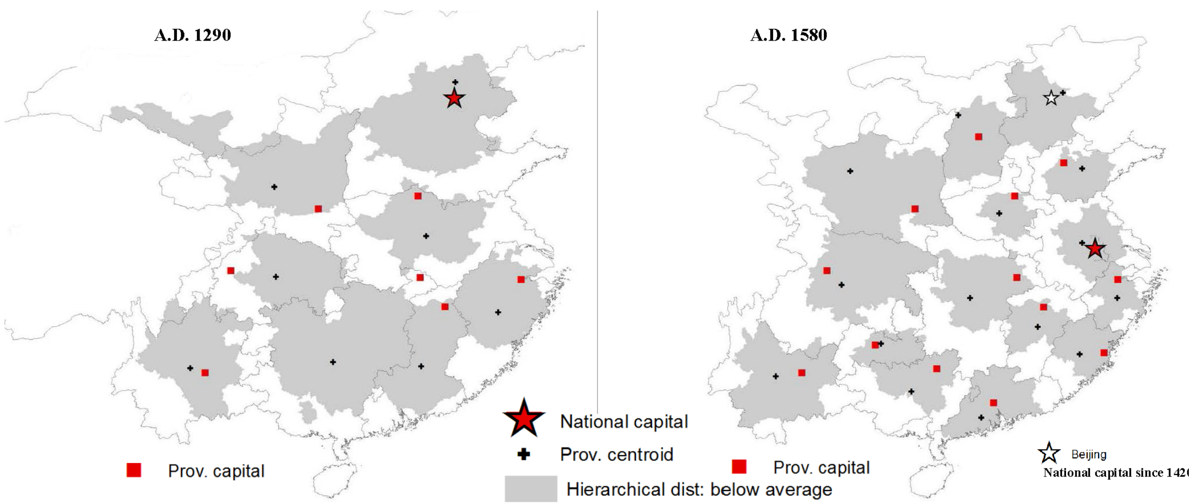
Notes: The shaded area indicate the prefectures in China proper (our sample). The dotted prefectures have ever been a provincial capital at least once during 1078-2000.

Figure 2: Hierarchical Distance, National Capitals and Provincial Capitals

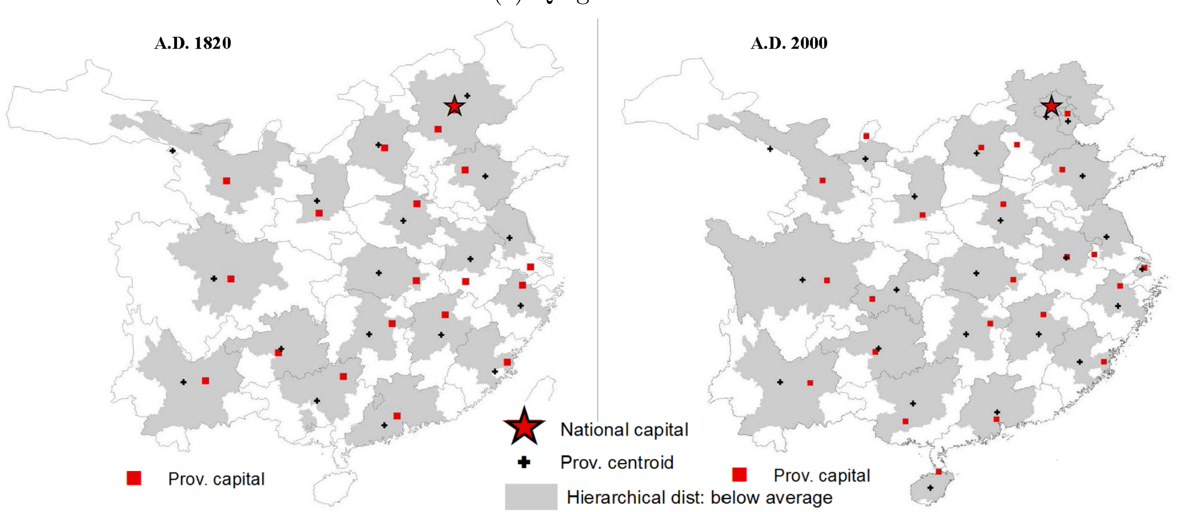
(a) Song



(b) Yuan and Ming



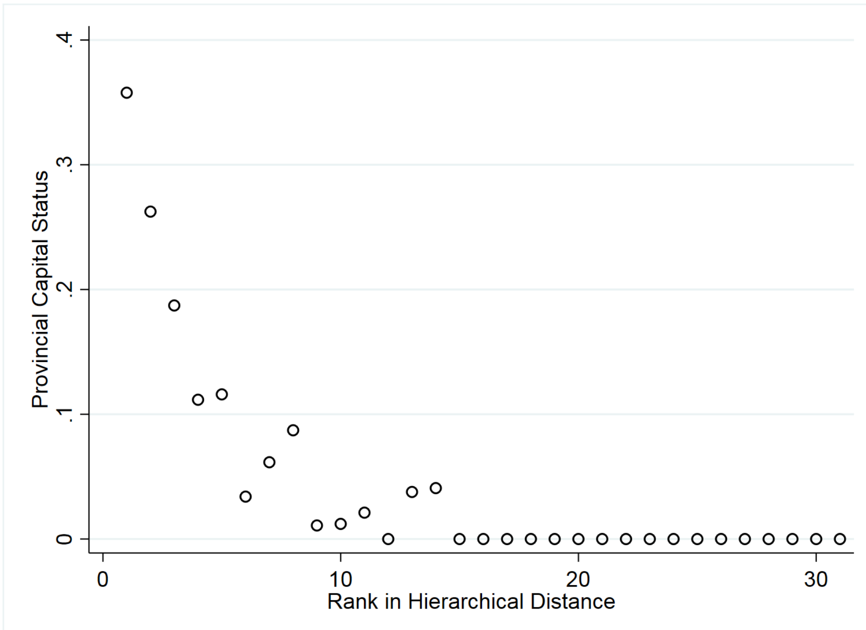
(c) Qing and 2000



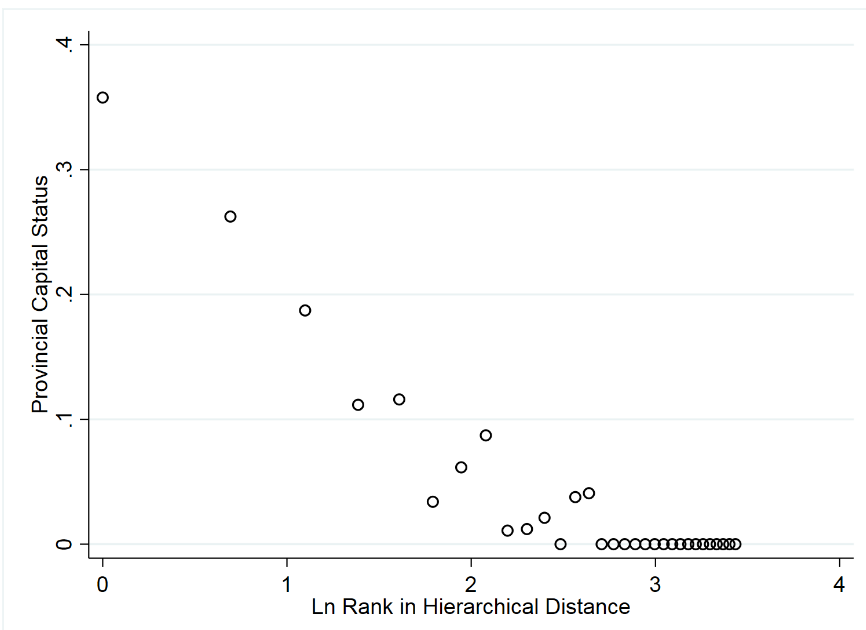
Notes: This figure shows that provincial capitals are usually located away from provincial centroid toward the shaded region with a shorter hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province). We present two maps for the Song dynasty because 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



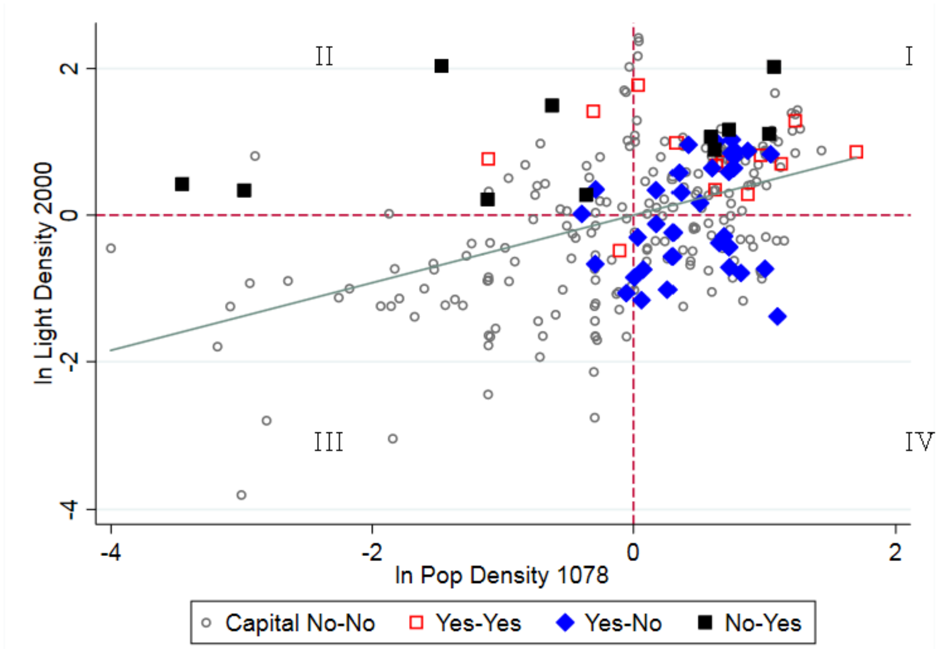
(b) Ln Rank in Hierarchical Distance vs. Prob. of Being Capital



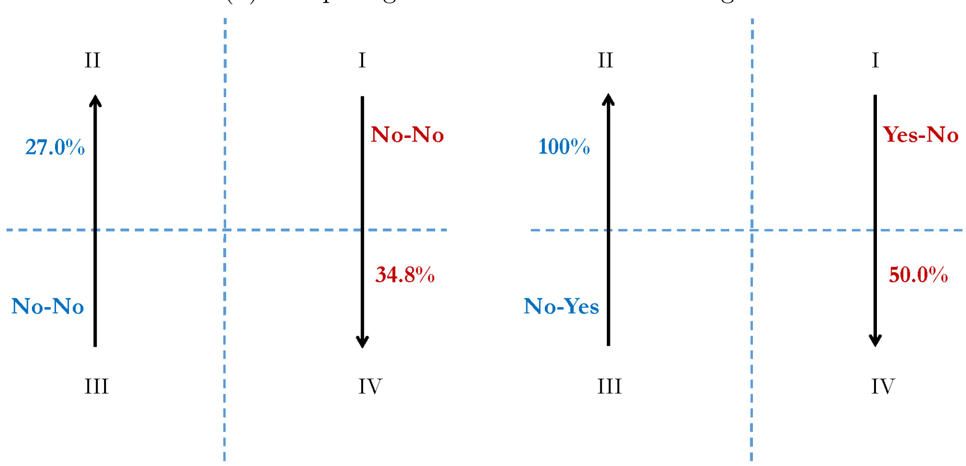
Notes: The figure shows a strong negative correlation between a prefecture's rank in hierarchical distance (weighted sum of distances to the national capital and the other prefectures within a province) within a province and its probability of being a provincial capital, which confirms the pattern in Figure 2.

Figure 4: Descriptive Patterns

(a) Pop. Density in 1078 vs. Light Density in 2000



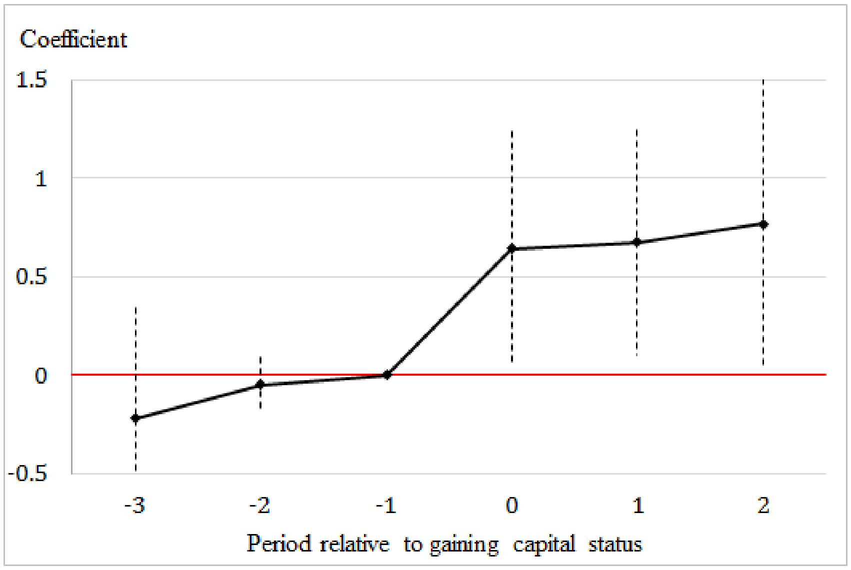
(b) Comparing With Those Without Changes



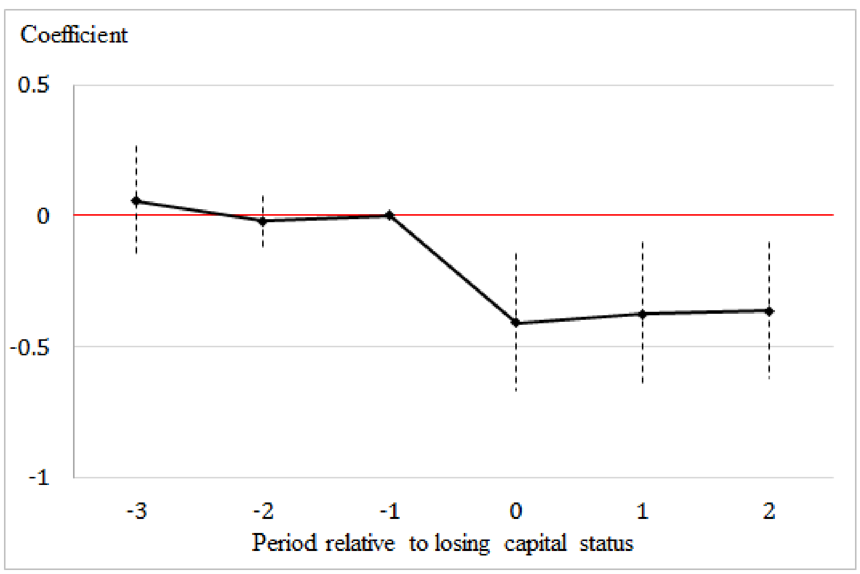
Notes: Panel (a) shows that gaining and losing capital status is associated with a systematic change in a prefecture's economic performance (measured by population density in 1078 and light density in 2000): those gaining capital status all perform better than average in 2000; a majority of those losing capital status performed worse than average in 2000.
Panel (b) uses the group of prefectures that are not provincial capitals in either period as a comparison.

Figure 5: The Impact of Capital Status on Population Density: Pre-Trends

(a) Before and After Gaining Capital Status



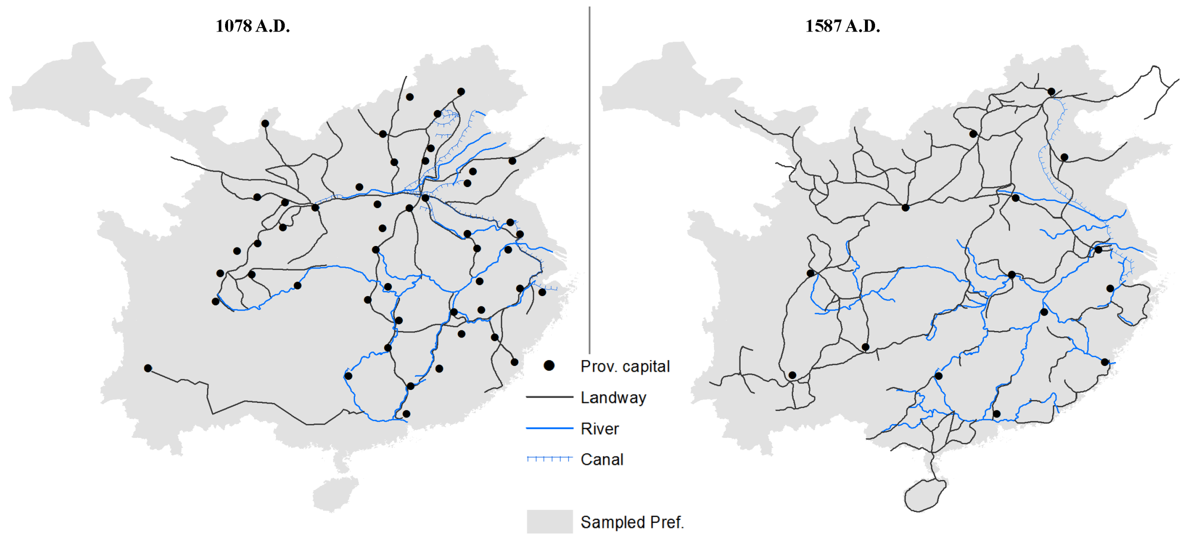
(b) Before and After Losing Capital Status



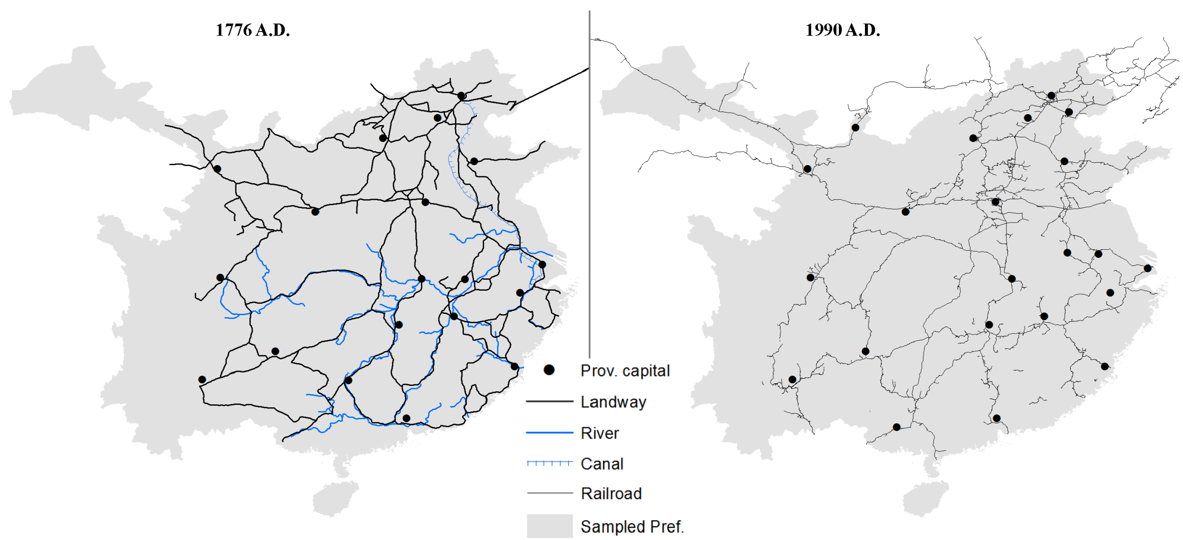
Notes: This figure visualizes the results in Appendix Table A.3. It shows that there are no systematic pre-trends before and after a prefecture gains or loses capital status. We ignore the length of periods for simplicity. See Section 4.1 for more analysis considering the length of periods.

Figure 6: 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 A.11).

Table 2: The Impact of Capital Status on Population Density

	ln Pop Density				Δ ln Pop Density			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prov. Capital	0.518***	0.484***	0.432***	0.433***				
	(0.125)	(0.106)	(0.102)	(0.101)				
Δ Prov. Capital					0.414***	0.321***		
					(0.106)	(0.082)		
Gaining Capital Status							0.459**	0.476***
							(0.220)	(0.156)
Losing Capital Status							-0.392***	-0.242**
							(0.116)	(0.094)
L. ln Pop Density						-0.344***		-0.345***
						(0.017)		(0.017)
Year FE * Crop suitability				Y	Y	Y	Y	Y
Year FE * Geography			Y	Y	Y	Y	Y	Y
Year FE * ln Area		Y	Y	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Pref. FE	Y	Y	Y	Y				
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	2,871	2,871	2,871	2,871	2,610	2,610	2,610	2,610
R-squared	0.771	0.841	0.856	0.865	0.656	0.738	0.656	0.739
#Prefectures	261	261	261	261	261	261	261	261
Magnitude: gaining vs. losing capital (p-value):							0.790	0.203

Notes: This table shows that provincial capital status is associated with a higher population density. The impacts of gaining and losing capital status have a similar magnitude. 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: The Impact of Capital Status on Population Density: Ever-Capital Prefectures

	Capital-ever Prefectures				Capital-ever Prefs + Comparison Group			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
					Neighbors<100 km		Propensity Score Matching	
Prov. Capital	0.470***	0.426***	0.280**	0.295***	0.467***	0.341***	0.512***	0.375***
	(0.151)	(0.132)	(0.105)	(0.110)	(0.126)	(0.104)	(0.129)	(0.127)
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	0.778	0.880	0.904	0.919	0.798	0.914	0.781	0.878
#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 4: Results Using Hierarchical Distance as an Instrument

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Var.: ln Pop Density								
	Reduced-form	IV	IV	IV	IV	IV	IV	IV
Prov. Capital		0.663***	0.624***	0.688***	0.638**	0.631***	0.726***	0.681***
		(0.201)	(0.190)	(0.261)	(0.249)	(0.220)	(0.215)	(0.203)
ln Rank in Hierarchical distance	-0.069***							
	(0.022)							
ln Rank in H dist. KF * Post-			-0.006					
			(0.024)					
ln Rank in H dist. NJ * Pre-				0.012				
				(0.047)				
ln Rank in H dist. BJ * Pre-					-0.013			
					(0.046)			
ln Rank in H dist. to major econ region						-0.017		
						(0.024)		
ln Rank in H dist. to the East (Shanghai)							0.036	
							(0.026)	
ln Rank in H dist. to the South (Guangzhou)								0.012
								(0.022)
All controls	Y	Y	Y	Y	Y	Y	Y	Y
First-stage: Provincial Capital								
ln Rank in Hierarchical distance		-0.104***	-0.153***	-0.085***	-0.088***	-0.097***	-0.100***	-0.105***
		(0.009)	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
ln Rank in H dist. KF * Post-			0.076***					
			(0.013)					
ln Rank in H dist. NJ * Pre-				-0.084***				
				(0.014)				
ln Rank in H dist. BJ * Pre-					-0.085***			
					(0.014)			
ln Rank in H dist. to major econ region						-0.037***		
						(0.009)		
ln Rank in H dist. to the East (Shanghai)							-0.024**	
							(0.010)	
ln Rank in H dist. to the South (Guangzhou)								0.002
								(0.009)
All controls		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	0.861	0.864	0.864	0.863	0.864	0.864	0.863	0.863
# Prefectures	261	261	261	261	261	261	261	261
F-Stat (Weak instrument test)		145.9	164.6	87.5	96.2	122.0	128.5	143.2

Notes: This table presents the estimates using logged rank in hierarchical distance as an instrument. Columns (3)-(8) show that it is difficult to non-political factors to explain the role of hierarchical distance in economic development. ***: significant at 1%, **: significant at 5%, *: significant at 10%.

Table 5: Descriptive Pattern: Gravity along the Political Hierarchy
($\beta = 1, r = 500, W = 1$)

	Standardized gravity			
	Song	Ming	Qing	1990
National Capital	3.30 (Kaifeng)	2.08 (Nanjing)	0.38 (Beijing)	0.59 (Beijing)
Provincial capitals	0.58	0.34	0.55	0.40
	(0.74)	(0.76)	(0.85)	(0.97)
Non-capital prefectures	-0.15	-0.03	-0.04	-0.04
	(0.98)	(1.00)	(1.00)	(1.00)

Notes: This table presents the mean and standard deviations (in paraphrases) of gravity along the political hierarchy: provincial capitals have a higher gravity than non-capital prefectures.

Table 6: Provincial Capital Status and Gravity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Gravity	Gravity	$\Delta Gravity$		Gravity (no restriction of r)		
	$\beta = 1, r = 500, W = 1$				$\beta = 1$	$\beta = 1$	$\beta = 10$
					W=1,	W=area	W=1
Prov. Capital	0.413***	0.361***			0.312***	0.315***	0.293***
	(0.086)	(0.089)			(0.093)	(0.094)	(0.103)
$\Delta Prov.Capital$			0.334***				
			(0.106)				
Gaining Capital Status				0.403**			
				(0.176)			
Losing Capital Status				-0.302**			
				(0.137)			
Year FE * Crop suit.		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	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y	Y
Pref. FE, Year FE	Y	Y			Y	Y	Y
Observations	1,044	1,044	783	783	1,044	1,044	1,044
R-squared	0.178	0.293	0.302	0.302	0.239	0.259	0.210
#Prefectures	261	261	261	261	261	261	261

Notes: This table shows that provincial capital status affects the spatial centrality of a prefecture. 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 7: Heterogeneous Effects w.r.t. Natural Advantages

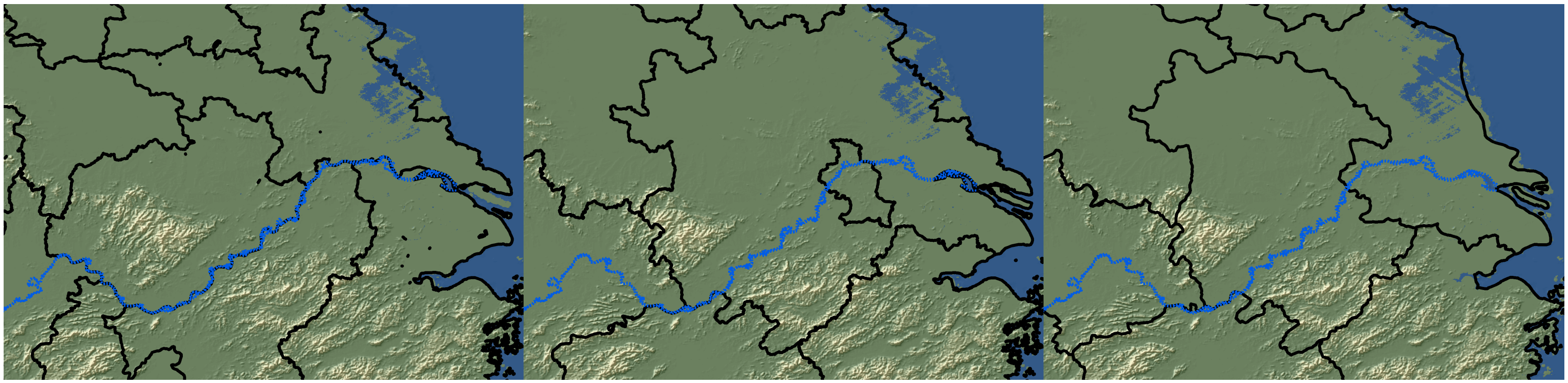
	ln Pop Density				Δ ln Pop Density	
	(1)	(2)	(3)	(4)	(5)	(6)
Prov. Capital	0.454***	0.444***	0.457***	0.443***		
	(0.103)	(0.102)	(0.102)	(0.101)		
Prov. Capital * max suitability (old world)	-0.169**					
	(0.083)					
Prov. Capital * max suitability (old+new)		-0.139*				
		(0.084)				
Prov. Capital * 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%.

Online Appendix

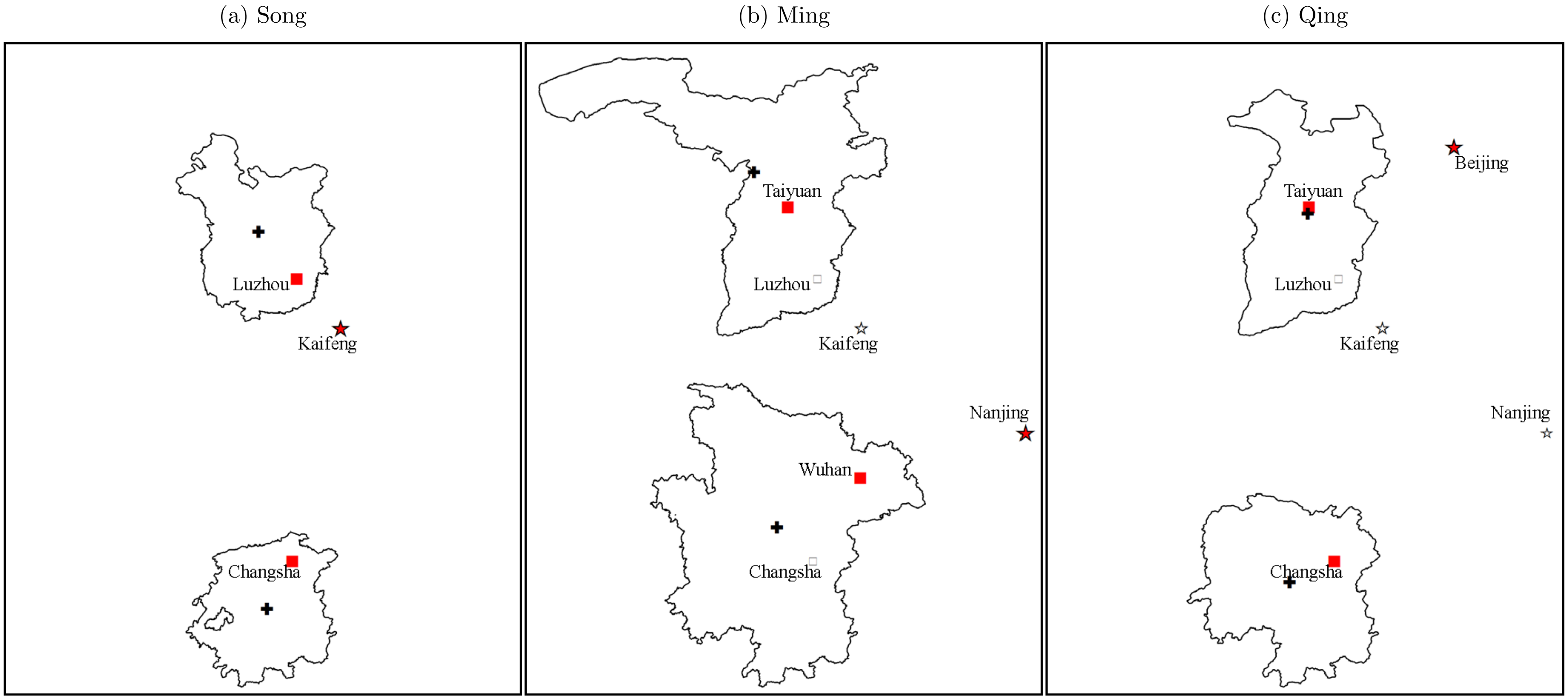
Figure A.1: An Example of Boundary Changes

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



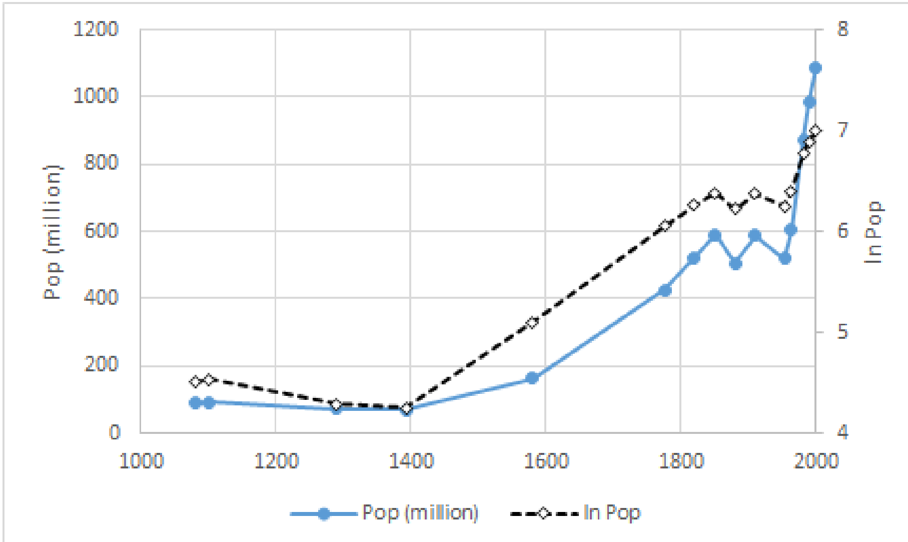
Notes: This figure presents an example of the evolution of provincial boundaries from following the natural mountains and rivers (known as "*suiting the forms of mountains and rivers*") to intentionally avoiding them (coined as "*interlocking like dog's teeth*"). Yangtze River was used as part of provincial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

Figure A.2: Examples of Capital Changes



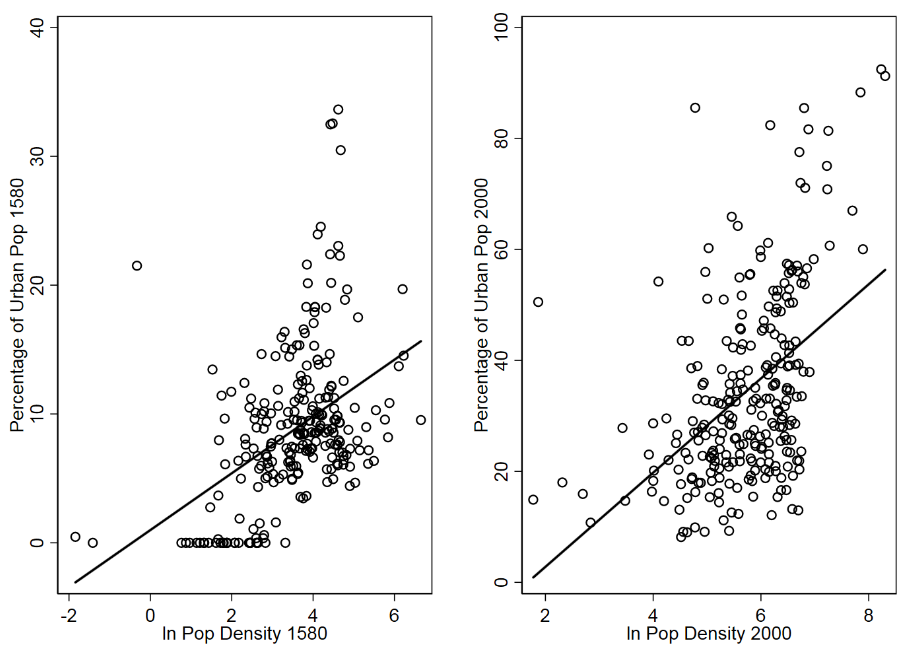
Notes: This figure presents two examples of capital relocation. The cross indicates the provincial centroid, the hollow/solid star indicates the past/current national capital, and the hollow/solid square indicates the past/current provincial capital. Luzhou and Changsha were capitals in the Song. Both 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.

Figure A.3: The Trends in Population



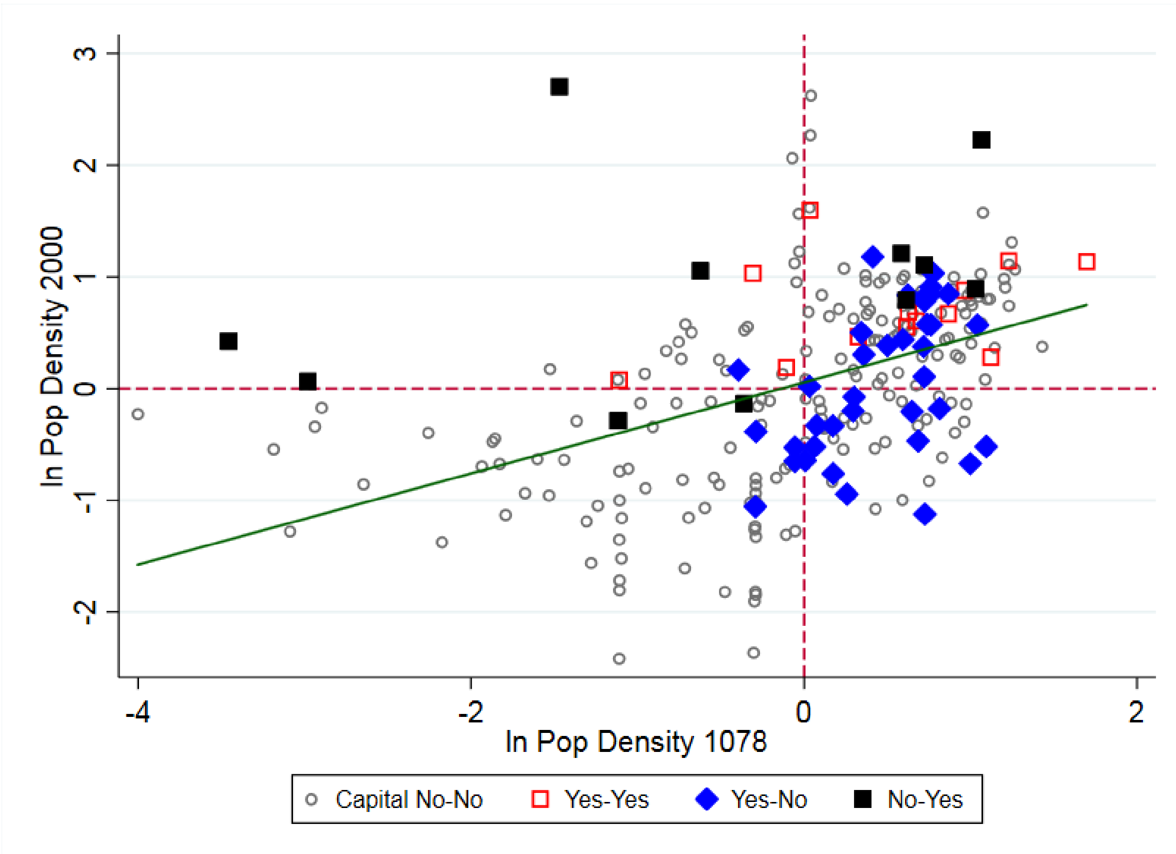
Notes: The figure plots the trends in population in our data.

Figure A.4: Pop Density vs. Urbanization in 1580 and 2000



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

Figure A.5: Pop Density in 1078 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
Mongolian conquest	1, 3	260	0.88	1.96
Taiping rebellion	1, 3	261	0.25	0.56
WWII	1, 3, 4	261	0.13	0.86
Hierarchical distance to national capital	2	2,871	-0.01	0.14
Within-prov. distance	2	2,859	-0.02	0.19
Gravity ($\beta = 1$, $r = 500$)	2, 7	1,044	39.71	26.96
Gravity ($\beta = 1$)	2, 7	1,044	55.23	36.64
Gravity ($\beta = 1$, weighted by area)	2, 7	1,044	1.05	0.88
Gravity ($\beta = 10$)	2, 7	1,044	8.44	7.14
Connect	7	1,044	0.58	0.49
# of Ways	7	783	1.66	1.85
# of Land ways	7	783	1.00	1.37
# of Water Ways	7	783	0.66	1.13

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) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) OLS	(7) 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: Pre-trends Check I: Period-by-Period Results
Dependent Var.: ln Pop Density

	(1)	(2)	(3)
Pre-Gaining: -3+	-0.223		-0.218
	(0.278)		(0.284)
Pre-Gaining: -2	-0.049		-0.048
	(0.070)		(0.071)
Post-Gaining: 0	0.666**		0.644**
	(0.294)		(0.301)
Post-Gaining: 1	0.696**		0.675**
	(0.285)		(0.291)
Post-Gaining: 2+	0.792**		0.767**
	(0.366)		(0.375)
Pre-Losing: -3+		0.067	0.056
		(0.111)	(0.107)
Pre-Losing: -2		-0.020	-0.022
		(0.049)	(0.049)
Post-Losing: 0		-0.421***	-0.409***
		(0.133)	(0.133)
Post-Losing: 1		-0.390***	-0.375***
		(0.137)	(0.138)
Post-Losing: 2+		-0.393***	-0.362***
		(0.130)	(0.131)
All controls	Y	Y	Y
Observations	2,783	2,783	2,783
R-squared	0.865	0.862	0.866
#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 pre-trends 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: Pre-trends Check II: Considering the Length of Periods
Dependent Var.: ln Pop Density

	(1)	(2)	(3)
Pre-Gaining	0.005		-0.002
	(0.128)		(0.127)
Pre-Gaining * Time trend	0.773		0.717
	(0.556)		(0.558)
Post-Gaining	0.509**		0.506**
	(0.218)		(0.221)
Post-Gaining * Time trend	0.615		0.562
	(0.879)		(0.901)
Pre-Losing		-0.038	-0.039
		(0.051)	(0.051)
Pre- Losing * Time trend		-0.405	-0.351
		(0.284)	(0.293)
Post- Losing		-0.359**	-0.364**
		(0.157)	(0.158)
Post- Losing * Time trend		-0.055	0.004
		(0.209)	(0.214)
All controls	Y	Y	Y
Observations	2,783	2,783	2,783
R-squared	0.865	0.862	0.867
#Prefectures	253	253	253

Notes: This table shows that considering the length of periods does not alter our main findings. 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.5: Validity Check I: Hierarchical Distance and Pre-change Characteristics

	Δ ln Rank in Hierarchical Distance					
	(1)	(2)	(3)	(4)	(5)	(6)
l. ln Pop Density	-0.031		-0.059			
	(0.032)		(0.061)			
l2. ln Pop Density		0.008	0.036			
		(0.020)	(0.034)			
l. Δ ln Pop Density				-0.044		-0.036
				(0.041)		(0.038)
l2. Δ ln Pop Density					0.011	0.002
					(0.011)	(0.011)
All controls	Y	Y	Y	Y	Y	Y
Observations	2,610	2,349	2,349	2,349	2,088	2,088
R-squared	0.301	0.300	0.302	0.301	0.197	0.199
#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 A.6: Validity Check II: Gaining vs. Losing Capital Status

		Seemingly Unrelated Regression	
	ΔCapital Status	Gaining	Losing
Δln Rank in Hierarchical Distance	-0.057***	0.031***	-0.026***
	(0.019)	(0.005)	(0.004)
All controls	Y	Y	Y
Observations	2,610	2,610	2,610
R-squared	0.183	0.298	0.196
#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 A.7: Alternative Functional Forms of Hierarchical Distance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\lambda =$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
IV Estimates									
Prov. Capital	0.578***	0.745***	0.600***	0.585**	0.583**	0.640**	0.786**	0.846***	0.741***
	(0.215)	(0.209)	(0.205)	(0.236)	(0.257)	(0.261)	(0.317)	(0.290)	(0.284)
All controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
First-stage									
ln Rank in H dist.	-0.097***	-0.101***	-0.099***	-0.084***	-0.077***	-0.075***	-0.062***	-0.068***	-0.069***
	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
All controls	Y	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	2,871
R-squared	0.864	0.863	0.864	0.864	0.864	0.864	0.862	0.861	0.863
# Prefectures	261	261	261	261	261	261	261	261	261
F-Stat (Weak instrument)	125.4	135.6	138.7	103.8	86.9	84.3	57.1	69.3	71.3

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.8: Using the Two Components of Our Instrument for an Over-Identification Test

	(1)	(2)	(3)	(4)	(5)
Dependent var.:	ln Rank in Hierarchical distance	ln Pop Density			
	($\lambda = 0.19$)	Reduced-form	IV	IV	IV
Prov. Capital			0.761**	0.907***	0.844***
			(0.358)	(0.305)	(0.207)
ln Rank in H dist. ($\lambda = 1$)	0.280***	-0.042**		0.008	
	(0.010)	(0.021)		(0.029)	
ln Rank in H dist. ($\lambda = 0$)	0.769***	-0.066***	-0.011		
	(0.011)	(0.023)	(0.037)		
All controls	Y	Y	Y	Y	Y
First-stage: Provincial Capital					
ln Rank in H dist. ($\lambda = 1$)			-0.056***	-0.056***	-0.056***
			(0.008)	(0.008)	(0.008)
ln Rank in H dist. ($\lambda = 0$)			-0.073***	-0.073***	-0.073***
			(0.009)	(0.009)	(0.009)
All controls			Y	Y	Y
Observations	2,871	2,871	2,871	2,871	2,871
R-squared	0.819	0.861	0.862	0.860	0.861
# Prefectures	261	261	261	261	261
F-Stat (Weak instrument test)			46.1	64.9	69.7
Over-identification test (P-value)					0.778

Notes: We use ln rank in hierarchical distance to capital with $\lambda = 1$ and that with $\lambda = 0$ as instruments. This table shows that the two instruments have no significant direct effect on population density beyond the channel of 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 A.9: 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.10: Grid-level Analysis (1-degree × 1-degree)

	ln Pop Density				Δ ln Pop Density	
	(1)	(2)	(3)	(4)	(5)	(6)
Prov. Capital	0.328***	0.328***	0.304***	0.295***		
	(0.065)	(0.065)	(0.059)	(0.054)		
Δ Prov. Capital					0.293***	
					(0.053)	
Gaining Capital status						0.225***
						(0.083)
Losing Capital status						-0.344***
						(0.063)
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
Grid FE, Year FE	Y	Y	Y	Y		
Observations	3,971	3,971	3,971	3,971	3,610	3,610
R-squared	0.797	0.797	0.825	0.836	0.526	0.526
# grids	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.11: Dropping Data Period by Period

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Baseline							
Prov Capital		0.429***	0.366***	0.415***	0.519***	0.397***	0.469***	0.410***	0.547***
		(0.102)	(0.128)	(0.109)	(0.116)	(0.127)	(0.100)	(0.119)	(0.134)
All controls		Y	Y	Y	Y	Y	Y	Y	Y
Observations		2,871	2,349	2,610	1,827	1,827	1,044	783	522
R-squared		0.865	0.856	0.866	0.889	0.836	0.875	0.822	0.913
#Prefectures		261	261	261	261	261	261	261	261
Song	1080	Y		Y	Y	Y	Y	Y	Y
	1102	Y		Y	Y	Y			
Ming	1580	Y	Y		Y	Y	Y	Y	
Qing	1776	Y	Y	Y		Y			
	1820	Y	Y	Y		Y			
	1851	Y	Y	Y		Y	Y	Y	
	1910	Y	Y	Y		Y			
P R China	1964	Y	Y	Y	Y				
	1982	Y	Y	Y	Y				
	1990	Y	Y	Y	Y				
	2000	Y	Y	Y	Y		Y		Y

Notes: This table shows that our results are robust to 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 A.12: Including Population Data in 1393, 1880, and 1953

	ln Pop Density				Δ ln Pop Density			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Prov. Capital	0.502***	0.489***	0.432***	0.440***				
	(0.119)	(0.101)	(0.097)	(0.098)				
Δ Prov. Capital					0.420***	0.339***		
					(0.105)	(0.086)		
Gaining Capital Status							0.457**	0.471***
							(0.218)	(0.164)
Losing Capital Status							-0.401***	-0.272***
							(0.117)	(0.100)
L. ln Pop Density						-0.292***		-0.293***
						(0.017)		(0.017)
Year FE * Crop suitability				Y	Y	Y	Y	Y
Year FE * Geography			Y	Y	Y	Y	Y	Y
Year FE * ln Area		Y	Y	Y	Y	Y	Y	Y
Year FE * Region FE	Y	Y	Y	Y	Y	Y	Y	Y
Pref. FE	Y	Y	Y	Y				
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Observations	3,654	3,654	3,654	3,654	3,393	3,393	3,393	3,393
R-squared	0.782	0.849	0.866	0.874	0.718	0.789	0.718	0.789
# prefectures	261	261	261	261	261	261	261	261
Magnitude: gaining vs. losing capital (p-value):							0.825	0.305

Notes: This table shows that our results are robust to including the three periods after deadly wars. 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.13: Checking Pre-trends for Gravity

	Standardized Gravity		
	(1)	(2)	(3)
Gaining Pre-2	0.038		0.056
	(0.151)		(0.154)
Gaining Post-1	0.610**		0.593**
	(0.252)		(0.258)
Gaining Post-2	0.495***		0.461**
	(0.184)		(0.182)
Losing Pre-2		0.128	0.126
		(0.212)	(0.216)
Losing Post-1		-0.329**	-0.318**
		(0.156)	(0.156)
Losing Post-2		-0.390***	-0.374***
		(0.123)	(0.123)
All controls	Y	Y	Y
Observations	1,016	1,016	1,016
R-squared	0.274	0.276	0.283
#Prefecture	254	254	254

Notes: This table shows that there are no systematic pre-trends in gravity before the capital status change. 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%.