Random river:
Trade and rent extraction in imperial China

Marlon Seror*

Preliminary Draft – Please do not cite

December 6, 2019

Abstract

This paper exploits exogenous changes in the course of the Yellow River in China to isolate variation in the natural distribution of economic centers across space. I compute time-varying trade costs and collect original data on population and taxation from dynastic histories and local gazetteers spanning 2,000 years. This allows me to assess the effect of exogenous variation in market access on population density and resource extraction. I find that the changes in connectedness have two opposite effects. First, they induce a very large increase in the level and concentration of economic activity in the shorter run. Second, they trigger a large increase in rent capture by local elites, which severely mitigates the concentration effect in the longer run.

*University of Bristol, marlon.seror@bristol.ac.uk.

1 Introduction

As the economy grows, economic activity becomes increasingly concentrated, mostly in just a few cities. This process can be thought of as a consequence of economies of scale and agglomeration economies, and thus as conducive to economic development (Kim, 1995; Henderson et al., 2001). But this spatial dimension of economic development may cut both ways. This paper hypothesizes that wealth concentration allows the government or local elites to extract resources more efficiently. If these resources primarily benefit other places or are not invested productively, the local economy suffers a negative externality that may limit the gains from agglomeration. Another limitation in our understanding of agglomeration economies comes from the literature's focus on short-term economic effects, keeping political structures constant. This paper studies the very long run, using data covering over 2,000 years, and thus allows political structures to react.

In this paper, I investigate whether the concentration of wealth leads to an accumulation of elites that can extract rent, and whether the concentration of elites slows down economic activity and growth in the long run. Answering this research question poses two main challenges: we need exogenous variation in the spatial concentration of wealth, and we need data on the spatial distribution of economic activity, extracted resources, and elites.

I exploit major changes in geography to isolate exogenous variation in the natural distribution of economic centers across space. I use changes in the course of the Yellow River between 305 B.C.E. and today that constitute large, durable shocks to transportation costs and thus alter the spatial distribution of economic activity. Over 4,000 years of recorded history, the Yellow River shifted its mouth by a maximum of 900 km—see Figure 1. This phenomenon may have both direct—on places that lost or gained access to the Yellow River—and indirect effects—through connections between river basins. The changes in the course of the Yellow River generate exogenous variation in trade costs and in connectedness, which I use to study the causal effect of market access on economic concentration and resource extraction.

To best exploit 2,000 years of changes in trade routes, I collect original data on population and taxation, and nest these data at the level of prefectures as of 2000.[1] These data are extracted from historical sources—dynastic histories, geographic treatises, and local gazetteers—covering over 2,000 years, which allows me

[1] Prefectures are the third level of administration, below provinces and the central government. I recast all data to consistent units, using the map of prefectures in 2000. There were 348 prefectures in 2000. The procedure followed to combine data corresponding to different administrative divisions is described in detail in the Appendix (available upon request).

to follow the evolution of economic activity and its concentration over the long run. They contain detailed information on taxation to the central government, which provides a rare picture of extracted resources across space and over a long time period. The taxation data additionally offer information on production and structural transformation, as for most of the period taxes were collected in kind. To distinguish between government taxes, which may finance public goods, and rent extraction, I complement the taxation data with the geolocated birthplaces of imperial examination graduates, which I use as a proxy for local elites. The output is a unique panel on economic concentration and rent extraction.

I find evidence of two opposite effects. First, over the shorter run improved market access generates large increases in both the level and concentration of economic activity. Second, the spatial concentration of wealth is associated with an increase in the rate of taxation and in the spatial concentration of elites. A heterogeneity analysis based on pre-existing elite concentration suggests that most of the effect is due to the greater influence of local elites. Data on social unrest further allows me to show that rents are increasingly contested over time in places that enjoy better market access.

This paper makes several contributions to the literature. The main contribution of this paper is to document the causal effect of the concentration of economic activity on rent extraction over 2,000 years of Chinese history. This is made possible by exogenous shocks and long-run data.

I put forward and test the hypothesis that economic concentration facilitates taxation, and show that the increase in taxation mostly reflects rent extraction by local elites. The paper thus contributes to a large literature—within and outside of economics—on rent-seeking (North and Weingast, 1989; Bernhardt, 1992; De Long and Shleifer, 1993; Murphy et al., 1993; Behrens and Pholo Bala, 2013). Because of its focus on imperial China and rent extraction, which has been proposed as an explanation by historians, this paper relates to the debate on the causes of the Great Divergence (Elvin, 1972, 1973; Pomeranz, 2000; Acemoglu et al., 2002; Galor and Mountford, 2006; Brandt et al., 2014) and the related issue of the prevalence of "parasite cities" in China relative to Europe (Weber, 1921; Needham, 1969; Bairoch, 1988; Bai and Jia, 2019). The paper shows indeed that the effect of improved market access on rent extraction is larger where administrative capacity is higher, which suggests that the general mechanism may have been particularly important in China, with its long-lasting and stable administrative system.

A second contribution of the paper is the use of exogenous variation in the natural distribution of economic centers. A large literature documents the role of geographic fundamentals in shaping economic activity and institutions (Acemoglu et al., 2001; Alesina et al., 2013; Mayshar et al., 2019), and the persistence of the spatial distribution of economic activity in the long run (e.g., Bleakley and Lin, 2012; Barjamovic et al., 2019). This literature relies on cross-sectional differences in geography. One issue with this approach is that such differences may affect outcomes through other channels (other geographical features, amenities, culture, etc.). In order to study the effect of the spatial concentration of wealth on rent extraction, we need variation in this concentration that (i) affects existing economic activity, (ii) has sufficiently large effects to bring about a change in the spatial equilibrium, and (iii) is orthogonal to human activity. In other words, the ideal experiment would resemble *changes in geography*. My paper approaches this ideal experiment by isolating exogenous shocks to trade costs.

A third contribution of the paper comes from its broad time coverage. This allows me to bring empirical insights to two different strands of literature—the trade literature on the impact of market access on economic activity (Hanson, 2005; Redding and Venables, 2004; Redding and Sturm, 2007; Redding, 2010; Head and Mayer, 2011; Shiue, 2002) in the long run (Fogel, 1964; Banerjee et al., 2012; Donaldson and Hornbeck, 2016; Donaldson, 2018; Barjamovic et al., 2019), and the literature on the long-run effects of an initial economic head start (Chinitz, 1961; Falck et al., 2013; Glaeser et al., 2015; Franck and Galor, 2017; Heblich et al., 2019). Additionally, the information on production available from taxation data means that this paper relates to contributions on specialization and co-agglomeration (Glaeser et al., 1992; Kim, 1995; Duranton and Puga, 2001; Faggio et al., 2017).

The remainder of this paper is organized as follows. Section 2 provides some background information on the Yellow River. Section 3 presents the data and construction of the market access measures. Section 4 explains the empirical strategy. Section 5 describes the results. Section 6 concludes.

2 Background: the Yellow River over time

The identification strategy in this paper relies on changes in the course of the Yellow River. The Yellow River has followed a variety of different routes in history, as shown in Figure 1. In this section, I (i) briefly explain the root of this phenomenon, (ii) describe the magnitude of the changes, and (iii) discuss its predictability.

2.1 Cause of the changing course

The root cause of changes in the course of the Yellow River is the uniquely high sediment load of its water in its lower basin. As it flows across the Loess plateau, the Yellow River accumulates the sediments that give it its distinctive color. This phenomenon makes it the "most hyperconcentrated sediment-laden river in the world" (Kong et al., 2017): its sediment concentration is approximately 14 times the concentration in the Mississippi and 48 times that of the Yangzi River (Lu et al., 2011; Xu et al., 2019). Huge amounts of sediments are carried by the high volume of water and can swiftly form low, invisible sand shoals, much below the opaque surface, that can suddenly raise the bottom of the river and force it out of its bed.

2.2 Magnitude of the shocks

Because of the flat terrain of the Lower Yellow River basin and of the mountainous Shandong peninsula that stands in the middle, this feature of the Yellow River makes it one of the most fickle rivers on earth. Between 602 B.C.E. and today, the Lower Yellow River experienced several major course changes, shifting its mouth over about 900 km over the last 2,000 years, from present-day Tianjin in the north to the Yangzi River delta in the south—see Figure 1. The magnitude of the shocks and the fact that they happened in the cradle of the Chinese civilization mean that they had large effects on human activity.

A major source of variation in the magnitude of the changes is whether the River flows north or south of the Shandong peninsula. The Yellow River changed its course completely to the other side of the Shandong peninsula four times in recorded history. Such changes are particularly important because most major rivers in China flow west to east. Changes in the course of the Yellow River mean that northern and southern China may become connected or disconnected, depending whether China's longest rivers, the Yellow River and the Yangzi, are linked. The course shifts may thus affect connectedness and trade costs much farther than the Lower Yellow River basin.

2.3 Unpredictable courses

Both the timing and route of the course changes are difficult to model and predict. In imperial times, the lack of a clear understanding of the dynamics of the Yellow River led to heated policy debates, for instance between Li Chui and Ding Wei at the Song court (Lamouroux, 1998), and in the Republican era the government commissioned a Sino-German team of scientists to investigate the process, which led to the construction of a large outdoor model of the Yellow River in Obernach, Bavaria (Pietz, 2015). Even today, an active literature studies the Yellow River. Until recently, models could not explain how the Yellow River carries so much sediment over such distances. According to Ma et al. (2017), previous models are inadequate because they assume that the sediments coalesce into dense clusters of tall dunes that slow down the river. Ma et al. (2017) show instead that the fine sediment particles form low, dispersed dunes that preserve the energy of the flow and allows the river to move more sediment.

Anecdotal evidence illustrates the unpredictability of the changes in the course of the Yellow River. Chen et al. (2012) put forward a quote from the Ming-dynasty hydraulic engineer Wan Gong ("Floods are rapid and large. Flood stage fluctuates unpredictably.") to show that the runoff of the lower river was highly variable and known to be unpredictable. They further cite the following record from 1601 in *The History of Ming Dynasty: Gazetteer of Rivers and Channels*, which shows that sudden course shifts would take even river traders by surprise: "The water level of the Yellow River at Shangqiu County rose and the water broke through the levee at the village of Xiajiakou. The whole flow was diverted southward. The trunk channel changed into a flat sandy land, and merchant boats were stuck in the sands."[2]

Because of its economic and cultural significance (water management is a major imperial prerogative in China), central and local governments tried to control the Yellow River. This took mostly the form of levee construction, maintenance, and repair, and the construction and regular consolidation of banks. The effect of such actions was mixed, as repairing banks prevents the channel from scouring naturally, increasing the likelihood of breaches and therefore the need for repairs (Chen et al., 2012). Importantly, once a levee is breached, the new course of the river is extremely difficult to predict.

Finally, some course shifts were man-made—in 1128 and 1938. Both artificial course shifts served a military purpose: they aimed at stopping the advance of the enemy (the Jurchen in 1128 and the Japanese in 1938) by drowning its troops. In both cases, deviating the river proved very difficult, and the new course followed by the river failed to fulfill its military purpose, as the river flowed south, while the enemy was advancing from the north. In both cases, the enemy eventually took control of the entire country. I argue that these artificial course changes do not threaten identification because, even if the timing is not random, the course followed

[2]The possibility of a sudden shift inhibited water transport along the Yellow River, which was therefore more expensive than along the Yangzi (e.g., Evans, 1984). This is accounted for empirically through river-specific trade costs. Note that because large course changes are rare, the Yellow River however remained a major trade route.

by the river after a man-made breach could not be predicted.[3]

3 Data and empirical strategy

This section describes the data sources and empirical strategy. I first present the population and taxation data, as well as additional data sources. I then explain how I measure market access and construct an instrument for it based on exogenous changes in geography. I finally describe my main estimation strategy.

3.1 Population and taxation data

My main outcomes of interest are population density and taxation intensity. I collect original data on population and taxation at the prefecture level from dynastic histories and geographical treatises, and geolocate them using a combination of place lists (from the CHGIS and SongGIS databases, which contain precise geocoordinates) and textual information on location from various historical sources. This yields a unique geolocalized data set spanning more than 200 prefectures and 2,000 years, from the earliest extant population census in the world in 2 C.E. to 2010. I now describe the data and the challenges they raise.

Description I extract population and taxation data from original sources in classical Chinese.[4] The texts usually follow a similar structure: For each province, prefectures are listed and described along different dimensions—population and taxation for each prefecture, subdivisions of each prefecture, administrative changes, customs, etc. Population is typically recorded in two ways: as the total numbers of households (*hu*) and individuals (*kou*). Some sources distinguish between "master" (*zhu*) and "guest" (*ke*) households, a distinction mostly determined by asset ownership.

Taxation is provided along with population figures from the beginning of the 8$^{\text{th}}$ century. For most of Chinese imperial history, taxation was predominantly in kind. In early waves, lists of goods are available. Starting with the Northern Song dynasty (960–1127), the sources specify amounts in great detail. A wide variety of goods were involved, which allows me to create a large set of taxation variables: number of goods used for taxation, number of goods exiting or entering the list, share of manufactured goods, degree of alignment between production and comparative advantage, etc. In later periods, taxes were increasingly collected in money or in a standard metal such as silver. This allows me to study the monetization of the local

[3] I however drop in robustness checks periods in which the course of the Yellow River resulted from a man-made breach.

[4] All sources are listed in the Appendix.

economy as an additional outcome. Finally, in some years the taxation data allow me to distinguish between taxes levied in urban and rural areas. It is important to note that the data contain information on tax remittances to the *central* government. Such taxes fueled the Imperial Treasury and were typically spent on maintaining the palace, defending borders, and remunerating officials; some redistribution could take place, for instance as Treasury relief in case of famine, but they were not specifically meant to benefit the prefectures where they had been collected. As far as the collection process is concerned, it varied by locale and over time, but local officials appointed by the central government typically needed the help of the local gentry. They were attentive to preserving the rents that local elites enjoyed, through annual negotiations and occasionally by quenching rent resistance with the army (e.g., Bernhardt, 1992).

Challenges In total, I use 25 waves of population data and 13 waves of taxation data for imperial China,[5] to which I add 7 post-1949 censuses. Over such a long period (and across such a wide territory as China's), population and taxation data may not be consistently recorded. The Appendix provides details about data characteristics and harmonization, and describes the challenges that the population and taxation data raise. I present here the main challenges briefly.

First, the data are organized by prefectures, but prefecture boundaries change over time. I recast all the data to the level of prefectures as of 2000. Second, data for some prefectures are missing in some periods. The changes in administrative divisions and reweighting of the data subsequently needed makes this issue more complicated than in most empirical settings. I develop a simple methodology to deal with missing observations and assess the robustness of the results.[6] Third, population and taxation data coverage may vary according to state capacity and administration quality (Ho, 1959; Cartier and Will, 1971; Bielenstein, 1987). In some regressions, I control for the dynastic cycle, distance to the imperial capital, and distance to the border to capture the effect of state capacity on data quality.

3.2 Market access

To construct historical market access measures between 350 B.C.E. and the present, I need to (i) determine the cost surface at different points in time based on historical geography, (ii) calculate minimum cost distances between all fixed units based on

[5]Population censuses were often used to evaluate tax revenues, so the waves usually coincide. In some rare instances, population and taxation data are slightly asynchronous; I then assume that the taxation data correspond to the closest year with population figures.

[6]This methodology is described in the Appendix.

the time-varying cost surface, and (iii) combine trade costs and population (as a proxy for economic activity) into market access measures.

Cost surface For identification, I use two measures of market access: one based on trade costs calculated *as they were*, which accounts for all rivers, roads, railroads, and canals, and the other where I isolate the exogenous component in the cost surface, keeping artificial transport infrastructure, which is likely endogenous to economic activity, fixed at baseline. These two measures rely on different specifications of the cost surface.

In both cases, I distinguish between overland (with and without roads), sea, and inland waterway transports. I assume overland (without roads) and sea trade costs to be uniform.[7] Inland waterway trade costs are allowed to differ by waterway type (natural rivers vs. canals) and by river (e.g., Yellow River vs. Wei River) to reflect differences in navigability and transportation costs. Most rivers experience only minor, very local changes in their courses (e.g., due to the natural evolution of meanders, or to floods) and are considered time-invariant. Time-invariant rivers are extracted from a map of the contemporary waterway network, from which I remove all time-varying waterways and replace them with historical courses to create time-varying waterway networks. Since rivers are cheap and therefore major trade routes, I further allow trade costs to differ for up- and downstream journeys. I rely on historical accounts to parameterize relative trade costs at different points in time.[8] I normalize the cost of sea transport to 1. For all river types, I distinguish between three scenarios—low, midpoint, and high river costs,—based on the range of costs provided by historians. I use the midpoint scenario and allow costs to differ by river in the baseline.[9]

[7]Alternatively, I can allow for different overland transportation costs by transforming Özak's (2010) human mobility index (HMI). Note, however, that the HMI does not account for inland waterways.

[8]Total trade costs include transportation costs and various barriers to trade. In the baseline strategy, I rely on historical accounts of the total costs of bringing goods from a given origin to a given destination. These accounts include transportation costs proper, the cost of labor, bribes, etc., but some barriers to trade such as time spent in transit and the cost of damages or loss might not be adequately captured. A more complete way of measuring total trade costs would be to use the difference in retail prices for goods from a known origin (Donaldson, 2018). For some periods, I use differences in retail prices to compute trade costs. In a robustness check (Data section in the Appendix), I rescale trade costs in other periods to obtain total trade costs. In the Appendix, I describe in detail the sources I use and how I estimate relative trade costs over time.

[9]I provide robustness checks relying on alternative scenarios and assuming a uniform river cost in the Appendix. Shiue and Keller (2007) use uniform (and time-invariant) costs, taking the midpoint between historical accounts. Since the Yellow River was less navigable than other natural waterways (at least in more recent periods) and most of the variation in market access comes from the Yellow River, assuming a uniform river transportation cost leads to overestimating changes in market access and thus puts a lower bound on estimated coefficients.

To calculate the *actual* trade costs, I use all the variation in the cost surface; to construct the *exogenous* component, I isolate exogenous changes, such as shifts in the course of the Lower Yellow River, and keep man-made infrastructures—canals, roads, and later railways—constant. For the Lower Yellow River, I use the maps of its historical courses since 350 B.C.E. created by Chen et al. (2012) combining geographical accounts in historical writings with NASA/JAXA Shuttle Radar Topography Mission (SRTM) data.[10] As the courses of other rivers, such as the Huai, were at times noticeably affected by the evolution of the Lower Yellow River, I additionally digitize maps from Tan (1989) to create time-varying courses for such rivers. For roads, railways, and canals, I digitize high-resolution maps from Cheng and Hsu (1983), based on records of the evolution of the transportation network in local gazetteers compiled in dynastic histories and geographical treatises. These maps distinguish between major and secondary roads, and allow me to date infrastructures.[11]

The cost surface is obtained by overlaying a 20×20-km grid on China and assigning a cost to crossing each cell, depending on the geography that characterizes that cell. I assume that a cell has a uniform trade cost determined by the lowest-cost geographical feature (land, road, railroad, river, or canal) intersecting it. For instance, if a cell is crossed by a river and overland transportation is costlier than river trade, I attribute to that cell the river trade cost.

Finally, I use the time-varying cost surfaces to calculate trade costs as the minimum cost distances d_{ijt} between all prefecture centroid pairs i, j at time t.[12] Figure 2 shows the resulting measures for Kaifeng, Henan Province, before and after the 1855 change in the course of the Yellow River.

Market access measures I measure the *actual* market access of prefecture i at time t as:

$$M_{it} = \sum_{j \neq i} d_{ijt}^{-1} N_{jt}, \tag{1}$$

[10] Accounts of earlier courses of the Yellow River are available, but they do not allow me to date course changes with precision. I am grateful to Chen et al. (2012) for sharing their maps with me.

[11] The large amounts of sediments carried by the Yellow River also mean that the coastline gradually moves out at its mouth. I rely on Cheng and Hsu's (1983) historical maps to create a time-varying map of the coastline. This map reflects both natural (due to the Yellow River) and artificial land expansions (polders).

[12] A more detailed description of the procedure is available in the Appendix.

where d_{ijt} is the minimum cost distance at time t between i and j using the observed cost surface, and N_{jt} is the population of prefecture $j \neq i$ at time t.

I next construct an instrument capturing the *exogenous* component in market access, defined as:

$$Z_{it} = \sum_{j \neq i} d_{ijt}^{*-1} N_{j0}, \tag{2}$$

where d_{ijt}^{*} is the minimum cost distance at time t between i and j based on exogenous geographical features (i.e., time-invariant or due to changes in the course of the Yellow River) and keeping man-made infrastructures fixed at baseline. N_{j0} is the population of prefecture j at baseline.

The constructed measure Z_{it} differs from the actual measure M_{it} in two respects. First, Z_{it} keeps man-made changes in transportation infrastructures fixed, which implies that variation only comes from exogenous changes in the course of the Yellow River. Second, it uses population at baseline rather than at time t. This is meant to avoid the reflection bias (Manski, 1993) that might arise from regressing population on a (population-based) measure of market access.

To mitigate spatial autocorrelation issues, I further create an alternative measure identical to M_{it}, except that d_{ijt} is replaced by the as-the-crow-flies distance. This alternative measure will enter regressions as a control.

Figure 3 illustrates the variation in market access, conditional on the as-the-crow flies measure included in the preferred specification (see Section 4). Panel (a) displays the constructed measure of market access (Z_{it}) in 1080 and Panel (b) in 1393. We can see that there is a large cross-sectional variation and that the southern shift of the Yellow River in 1128 led to a dramatic change in residual market access.

In the empirics, I average M_{it} and Z_{it} over a time window s (100 years, in the baseline specification) to deal with the random timing of changes in geography and irregular intervals in the population and taxation data. More precisely, I use:

$$\overline{M}_{it,t-s} \equiv \sum_{T=t-s}^{t} M_{iT}, \tag{3}$$

and similarly for $\overline{Z}_{it,t-s}$.[13]

[13] In the Appendix, I test the robustness of the results to (i) alternative time windows s and (ii) weighting the averages $\overline{M}_{it,t-s}$ and $\overline{Z}_{it,t-s}$ by the share of s corresponding to the exposure to each level of market access.

3.3 Other data

I use additional data sources to (i) test whether the spatial concentration of wealth facilitates rent extraction by local elites and (ii) check the robustness of the results.

First, I use two additional data sets to test the hypothesized mechanism. To assess the role of local elites in the extraction of economic resources, I complement the taxation data with the geolocated birthplaces of imperial examination graduates (*jinshi*) from all extant rosters between 618 and 1905, from the China Biographical Database. These data are available at the county level; I aggregate the number of births at the level of the 2000 prefectures. To assess the role of cities in the hypothesized mechanism, I further collect prefecture-level data on the population in urban areas to capture the concentration of population and economic activity in cities. These data are available for a subset of the study period (since the 16$^{\text{th}}$ century) and come from local gazetteers (population inside and outside of city walls), CHGIS, and 20$^{\text{th}}$-century population censuses.

Second, I use and geolocate detailed data on floods and river management activities from gazetteers and histories. These data are precisely dated; I further geolocate them and nest them at the prefecture level. These data contain information on flood events and intensity, levee construction, maintenance and repair activities, flood prevention measures such as damming, diversion, spillway, soil and water conservation, the establishment of river bureaus, and imperial Treasury relief. In robustness checks, I also use maps of administrative seats (at the county level, below the prefecture) and boundaries (at the prefecture) from CHGIS and Hartwell, respectively. Boundaries are available from the Tang dynasty; administrative seats have been geolocated from 221 B.C.E. Finally, the cost surface maps that I construct use data on topography (elevation and slope), and in some regressions I control for physiographic macroregions (Skinner, 1977) and other geographical fundamentals (latitude and longitude, access to the sea, etc.).

4 Strategy

I implement a 2SLS procedure to estimate the effect of market access on the concentration of economic activity and rent extraction. I instrument the *actual* market access instrumented by the *constructed* measure described in Section 3.

The baseline regression writes:

$$Y_{it} = \beta_0 + \beta_1 \overline{M}_{it,t-s} + \mathbf{X}_{it}\beta_{\mathbf{x}} + \delta_t + \mu_i + \varepsilon_{it}, \tag{4}$$

where $\overline{M}_{it,t-s}$ is instrumented by $\overline{Z}_{it,t-s}$, t stands for time and i for a fixed-boundary (2000) prefecture (or an arbitrary grid cell, in robustness checks), Y_{it} is the outcome of interest (population density, tax burden, elite share, etc.), $\mathbf{X}_{it}$ are time-varying characteristics of prefectures, i.e., the controls mentioned in Section 3 and the data section of the Appendix to deal with measurement and data quality issues; additional controls can be included to enhance the precision of the estimates, e.g., average cost of crossing prefecture i, sea access×year, etc. Finally, standard errors are clustered at the prefecture (2000) level.

I then introduce lags of $\overline{M}_{it,t-s}$, instrumented by the corresponding lags of $\overline{Z}_{it,t-s}$ in some specifications to compare shorter- and longer-run effects.

There are three main threats to the identification of Eq. 4 as a causal effect of market access on Y_{it}. First, the exclusion restriction would be violated if agents anticipated river course changes or could affect the course of the Yellow River. The identifying assumption under which β_1 estimates the causal effect of market access is that changes in the constructed measure of market access, which are exclusively due to changes in the course of the Yellow River, are orthogonal to agents' location decisions and rent extraction. I argue that the timing and course of the Yellow River after a change cannot be predicted, because the accumulation and movement of sediments at the bottom of the river is unobservable and challenging to model even for modern science. In some instances, a course change was artificially created for military purposes (in 1128 and 1938). I argue that even if in such cases the timing was not random, the exact course the river would adopt was impossible to predict. In robustness checks, I however deal with man-made course changes by (i) excluding periods when the river course resulted from an artificial change, and (ii) considering artificial changes in the *actual* but not the *constructed* measure of market access—in other words, I treat artificial courses as canals. The exclusion restriction would also be violated if river management systematically targeted prefectures with a higher (economic or rent) potential. In robustness checks, I control for flood events and river conversation activities, which I geolocated.

Second, because of the nature of the research question, spatial auto-correlation is a potential issue. I control for a simple, "as-the-crow-flies" measure of market access, i.e., weighting populations by the straight line distance rather than taking the cost surface into account. This reduces spatial auto-correlation: if population is highly spatially auto-correlated, then population and market access will be naturally correlated; controlling for the as-the-crow-flies measure ensures that I only use variation from exogenous changes in geography for identification.

Third, the changes in geography brought about by shifts in the course of the

Yellow River may have a variety of effects on the economy: (i) The Yellow River was at times utilized for military purposes, as a flood may drown enemy troops or its course act as a natural barrier. (ii) Rivers are not only used for trade; they are also used for irrigation. Other water resources are however available to peasants, as other rivers cross the North China Plain, so that irrigation did not critically rely on the Yellow River, contrary to the Nile, which constitutes most of Egypt's single source of water. (iii) Losing access to the Yellow River also meant that its river bed could be cultivated, increasing agricultural output; the sediments deposited by the river during floods may also enrich the soil and alter agricultural suitability. Historical sources suggest that "the benefits from the fertility of the silt brought by floods could almost offset the losses caused by floods" (Chen et al., 2012).

5 Results

5.1 Visual evidence

The identification strategy relies on large, durable changes in geography. I first provide maps to show the effect of the changes in the course of the Yellow River on the spatial equilibrium.

Figure 4 shows population density by prefecture in 1102 (Panel a) and 1784 (Panel b). Several changes in the course of the Yellow River occurred between these two dates, but the most important one was the 1128 change, which led to a shift of the river from the north to the south of the Shandong Peninsula. The two maps show that population (and economic activity) concentrates along major waterways, the Yellow River and Yangzi River. They also show that concentration largely shifted from the old to the new course of the Yellow River.

Figure 5 provides similar evidence for taxation intensity. Panel (a) maps taxes per capita in 1077 and Panel (b) in 1784.[14] Taxation intensity closely matches population concentration: taxes are higher on a per capita basis along major waterways, and the spatial distribution is dramatically affected by the change in the course of the Yellow River.

5.2 Regression-based evidence

I now move to regression analysis. I show the first stage of the 2SLS specification. I next present the effects of market access on economic concentration, structural

[14] Taxation data are not available in 1102. Note that I use taxes in strings of cash and in *dan* (a measure of volume) of wheat in Panels (a) and (b), respectively. Both maps show deciles of the prefecture distribution.

change, and taxation intensity.

Table 1 shows the results of the first stage. The dependent variable is the actual market access, and the instrument is the constructed measure that isolates exogenous changes in geography. We see that the two measures are strongly, positively associated, even controlling for the as-the-crow-flies market access measure to alleviate spatial auto-correlation.

I first find the—expected—result that market access leads to an increase in economic concentration, proxied by population density. Table 2 regresses population density (in log) on market access (standardized). Panel (a) displays OLS results, while Panel (b) instruments actual market access with the constructed measure. Panel (a) shows a large, positive correlation between market access and population density. In the preferred specification (controlling for the as-the-crow-flies market access, sea access interacted by year, and the average cost of crossing a prefecture, a 1-SD increase in market access corresponds to a 17% increase in population density. As shown in Panel (b), instrumenting confirms the positive relationship: a 1-SD increase in market access leads to a 29% increase in population density.

The increase in economic concentration translates into structural transformation, which I measure by the share of manufactured goods among the goods remitted as taxes to the central government. Table 3 shows the effect of market access on the share of manufactured goods. The effect is large: in Panel (b), we see that a 1-SD increase in market access yields a 100% increase in the share of manufactured goods. This effect is consistent with urbanization and economies of scale.

Finally, I find that improved market access increases taxation intensity. Measuring the tax burden presents two main challenges: first, the time span covered by the taxation data and China's geographical variety mean that different taxation rules and practices underpin the data; second, as taxation was mostly in kind, I need to either find goods used for taxation that are common to a large number of prefectures and to most years or to construct aggregate measures of taxation. I adopt the latter approach, first using the number of goods remitted and second computing a monetary aggregate based on Peng's (1965) historical price indices.[15] Table 4 follows the first approach and shows that a 1-SD increase in market access leads to a 29% increase in the number of goods remitted for taxation.

[15] I however provide robustness checks in the Appendix where I use goods remitted in a subset of prefecture×years.

5.3 Robustness checks

In the Appendix, I run a wide range of robustness checks. They deal with potential concerns about (i) the exclusion restriction and (ii) data quality.

I implement seven robustness checks to establish the identifying assumption. These checks can be organized into two main categories: (i) checks of potential direct effects of course changes and (ii) checks of potential human intervention in the course of the Yellow River. As far as direct effects are concerned, I first control for flood events, flood intensity and destruction, and Imperial Treasury relief. Second, I exclude the whole Lower Yellow River basin to avoid any effect (e.g., disruption due to floods) of course changes on the outcomes of interest that does not go through market access. Third, I allow for a broader understanding of direct effects and control for access to negatively affected areas.

The role of human intervention is investigated through three checks. First, I control for levee maintenance and construction work, as well as for other river conservancy measures. Second, I exclude periods when the course of the Yellow River resulted from an artificial breach in its banks. Third, I keep man-made course changes but treat them as endogenous changes in the cost surface, i.e., I maintain the Yellow River on its pre-change course in the constructed market access measure. Fourth, I implement a series of placebo tests using proposed (but rejected or failed) artificial course changes.

Data quality raises several concerns. First, as put forward by Cartier and Will (1971), censuses and tax registries were carried more effectively in border prefectures; laws were also applied to the letter and corruption was less of an issue in those regions. I deal with this issue by controlling for a dummy equal to 1 if a prefecture lies on the border, and 0 otherwise. Second, population and taxation data are available at irregular intervals. In robustness checks, I select periods at various, approximately equal intervals.

5.4 Mechanisms

In this subsection, I first investigate the effect of market access on urbanization and test whether taxation intensity is higher in more urbanized prefectures, which we would expect if the concentration of economic transactions in one (market) place facilitates monitoring and rent extraction. I then use annual data on elite birthplaces to study the role of elite capture in the increase in taxation.

In some years, the taxation data distinguish between urban and rural amounts. Figure 6 plots the density of taxes (in cash) in 1077 for urban areas and for the rest of the prefecture (in log). In Panel (a), I plot total amounts. We can see that the urban and non-urban distributions both display a lot of dispersion; the urban curve has slightly more mass to the right, but the two distributions look broadly similar. This similarity in total amounts however hides very different per capita levels. Indeed, no prefecture in China until the 20$^{\text{th}}$ century had prefectures with an urbanization rate close to 50%. Panel (b) assumes that 10% of the population lives in cities and plots the density of taxes per capita. We see that the distribution for urban per-capita taxes now shifts much to the right of the non-urban distribution, which suggests higher taxation intensity.

[TO BE COMPLETED]

6 Conclusion

[TO BE COMPLETED]

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001, *91*, 1369–1401.

_, _, and _ , "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, Nov. 2002, *117*, 1231–1294.

Alesina, Alberto, Paola Giuliano, and Nathan Nunn, "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 2013, *128*, 469–530.

Bai, Ying and Ruixue Jia, "When History Matters Little: Political Hierarchy and Regional Development in China, AD1000–2000," Technical Report January 2019.

Bairoch, Paul, *Cities and Economic Development: from the Dawn of History to the Present*, Chicago: University of Chicago Press, 1988.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian, "On the Road: Access to Transportation Infrastructure and Economic Growth in China," 2012.

Barjamovic, Gojko, Thomas Chaney, Kerem A. Coşar, and Ali Hortaçsu, "Trade, Merchants, and the Lost Cities of the Bronze Age," *Quarterly Journal of Economics*, 2019, *134*, 1455–1503.

Behrens, Kristian and Alain Pholo Bala, "Do rent-seeking and interregional transfers contribute to urban primacy in Sub-Saharan Africa?," *Papers in Regional Science*, 2013, *92* (1), 163–195.

Bernhardt, Kathryn, *Rents, Taxes, and Peasant Resistance: The Lower Yangzi Region, 1840–1950*, Stanford, CA: Stanford University Press, 1992.

Bielenstein, Hans, "Chinese Historical Demography A.D. 2 – 1982," *Bulletin of the Museum of Far Eastern Antiquities*, 1987, *59*, 1–288.

Bleakley, Hoyt and Jeffrey Lin, "Portage and Path Dependence," *Quarterly Journal of Economics*, 2012, *127* (2), 587–644.

Brandt, Loren, Debin Ma, and Thomas G. Rawski, "From Divergence to Convergence: Reevaluating the History behind China's Economic Boom," *Journal of Economic Literature*, 2014, *52* (1), 45–123.

Cartier, Michel and Pierre-Étienne Will, "Démographie et institutions en Chine : contribution à l'analyse des recensements à l'époque impériale (2 ap. J.-C.-1750) [in French]," *Annales de Démographie Historique*, 1971, pp. 161–245.

Chen, Yunzhen, James P. M. Syvitski, Shaopu Gao, Irina Overeem, and Albert J. Kettner, "Socio-economic impacts on flooding: a 4000-year history of the Yellow River, China," *AMBIO*, 2012, *41* (7), 682–698.

Cheng, Kuang-Yu and Sheng-Mo Hsu, *Historical Atlas of China [in Chinese]* Volume 2, Chinese Culture University Press, 1983.

Chinitz, Benjamin, "Contrasts in Agglomeration: New York and Pittsburgh," *The American Economic Review*, 1961, *51* (2), 279–289.

De Long, J. Bradford and Andrei Shleifer, "Princes and Merchants: European City Growth before the Industrial Revolution," *Journal of Law and Economics*, 1993, *36* (2), 671–702.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, 2018, *108* (4–5), 899–934.

_ **and Richard Hornbeck**, "Railroads and American Economic Growth: A "Market Access" Approach," *Quarterly Journal of Economics*, 2016, *113* (2), 799–858.

Duranton, Gilles and Diego Puga, "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 2001, *91* (5), 1454–1477.

Elvin, Mark, "The high-level equilibrium trap: the causes of the decline of invention in the traditional Chinese textile industries," in W. E. Willmott, ed., *Economic Organization in Chinese Society*, Stanford, CA: Stanford University Press, 1972, pp. 137–172.

_ , *The Pattern of the Chinese Past*, Stanford, CA: Stanford University Press, 1973.

Evans, Laurence, "Junks, Rice, and Empire: Civil Logistics and the Mandate from Heaven," *Historical Reflections*, 1984, *11* (3), 271–313.

Faggio, Giulia, Olmo Silva, and William C Strange, "Heterogeneous agglomeration," *Review of Economics and Statistics*, 2017, *99* (1), 80–94.

Falck, Oliver, Christina Guenther, Stephan Heblich, and William R. Kerr, "From Russia with love: the impact of relocated firms on incumbent survival," *Journal of Economic Geography*, May 2013, *13* (3), 419–449.

Fogel, Robert, *Railroads and American economic growth: Essays in econometric history*, Baltimore, MD: Johns Hopkins University Press, 1964.

Franck, Raphaël and Oded Galor, "Industrial Development and Long-Run Prosperity," NBER Working Papers 23701, National Bureau of Economic Research, Inc August 2017.

Galor, Oded and Andrew Mountford, "Trade and the Great Divergence: The Family Connection," *American Economic Review*, 2006, *96* (2), 299–303.

Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer, "Growth in cities," *Journal of political economy*, 1992, *100* (6), 1126–1152.

Glaeser, Edward L., Sari Pekkala Kerr, and William R. Kerr, "Entrepreneurship and urban growth: An empirical assessment with historical mines," *Review of Economics and Statistics*, 2015, *97* (2), 498–520.

Hanson, Gordon H., "Market Potential, Increasing Returns and Geographic Concentration," *Journal of International Economics*, 2005, *67*, 1–24.

Head, Keith and Thierry Mayer, "Gravity, Market Potential and Economic Development," *Journal of Economic Geography*, 2011, *11*, 281–294.

Heblich, Stephan, Marlon Seror, Hao Xu, and Yanos Zylberberg, "Industrial clusters in the long run: Evidence from Million-Rouble plants in China," *CESifo Working Paper*, 2019.

Henderson, J Vernon, Zmarak Shalizi, and Anthony J Venables, "Geography and development," *Journal of Economic Geography*, 2001, *1* (1), 81–105.

Ho, Ping-ti, *Studies on the population of China, 1368–1953*, Cambridge, MA: Harvard University Press, 1959.

Kim, Sukkoo, "Expansion of markets and the geographic distribution of economic activities: the trends in US regional manufacturing structure, 1860–1987," *Quarterly Journal of Economics*, 1995, *110* (4), 881–908.

Kong, D., C. Miao, J. Wu, A.G. Borthwick, Q. Duan, and X. Zhang, "Environmental impact assessments of the Xiaolangdi Reservoir on the most hyperconcentrated laden river, Yellow River, China," *Environmental Science and Pollution Research International*, 2017, *24* (5), 4337–4351.

Lamouroux, Christian, "From the Yellow River to the Huai: New Representations of a River Network and the Hydraulic Crisis of 1128," in Mark Elvin and Ts'ui-jung Liu, eds., *Sediments of Time: Environment and Society in Chinese History*, Cambridge: Cambridge University Press, 1998, pp. 545–584.

Lu, J., F. Qiao, X. Wang, and et al., "A numerical study of transport dynamics and seasonal variability of the Yellow River sediment in the Bohai and Yellow seas," *Estuarine, Coastal and Shelf Science*, 2011, *95*, 39–51.

Ma, Hongbo, Jeffrey A. Nittrouer, Kensuke Naito, Xudong Fu, Yuanfeng Zhang, Andrew J. Moodie, Yuanjian Wang, Baosheng Wu, and Gary Parker, "The exceptional sediment load of fine-grained dispersal systems: Example of the Yellow River, China," *Science Advances*, May 2017, *3* (5).

Manski, Charles, "dentification of endogenous social effects: the reflection problem," *Review of Economic Studies*, 1993, *60*, 531–542.

Mayshar, Joram, Omer Moav, Zvika Neeman, and Luigi Pascali, "The Origin of the State: Land Productivity and Appropriability," Technical Report October 2019.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "Why Is Rent-Seeking So Costly to Growth?," *American Economic Review*, May 1993, *83*, 409–414.

Needham, Joseph, *The Grand Titration: Science and Society in East and West*, Toronto: University of Toronto Press, 1969.

North, Douglass C. and Barry R. Weingast, "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England," *Journal of Economic History*, 1989, *49*, 803–832.
ΩÖzak

Özak, Ömer, "The Voyage of Homo-œconomicus: some economic measures of distance," Technical Report November 2010. Unpublished.

Peng, Xinwei, *A Monetary History of China [in Chinese]*, Shanghai: Shanghai Renmin Publishing House, 1965.

Pietz, David A., *The Yellow River: The Problem of Water in Modern China*, Cambridge, MA: Harvard University Press, 2015.

Pomeranz, Kenneth, *The Great Divergence: China, Europe, and the Making of the Modern World Economy*, Princeton, NJ: Princeton University Press, 2000.

Redding, Stephen, "The Empirics of New Economic Geography," *Journal of Regional Science*, 2010, *50*, 297–311.

_ **and Anthony Venables**, "Economic Geography and International Inequality," *Journal of International Economics*, 2004, *62*, 53–82.

_ **and Daniel M. Sturm**, "The Cost of Remoteness, Evidence from German Division and Reunification," *American Economic Review*, 2007, *98*, 1766–1797.

Sanderson, Eleanor and Frank Windmeijer, "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of Econometrics*, 2016, *190* (2), 212–221.

Shiue, Carol H., "Transport Costs and the Geography of Arbitrage in Eighteenth-Century China," *American Economic Review*, December 2002, *92* (5), 1189–1216.

_ **and Wolfgang Keller**, "Markets in China and Europe on the Eve of the Industrial Revolution," *American Economic Review*, September 2007, *97* (4), 127–137.

Skinner, G. William, "Regional Urbanization in Nineteenth-Century China," in G. William Skinner, ed., *The City in Late Imperial China*, Stanford, CA: Stanford University Press, 1977.

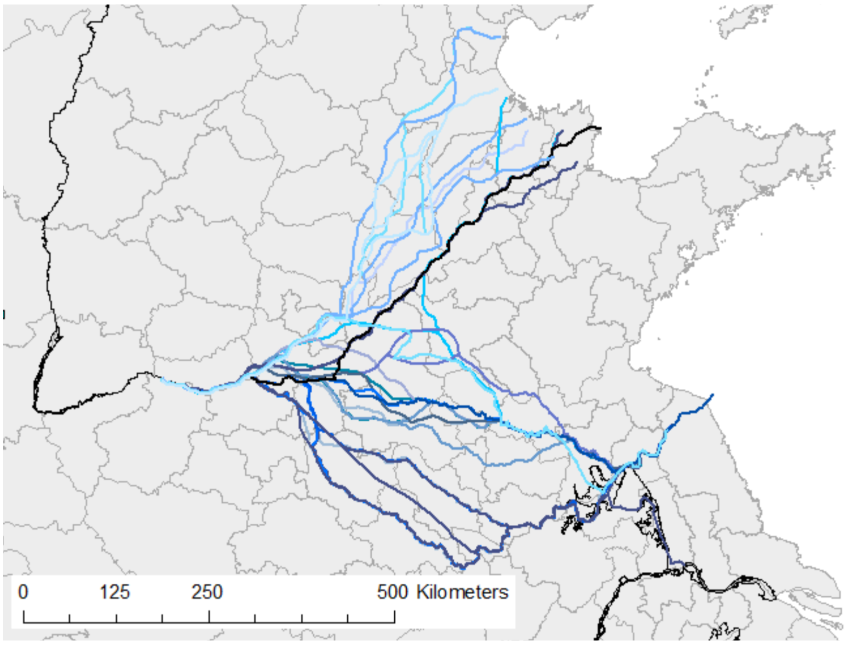
Tan, Qixiang, *Historical Atlas of China [in Chinese]*, Beijing: China Cartographic Publishing House, 1989.

Weber, Max, *The City*, New York: Free Press, 1921.

Xu, Xuegong, Zhengxiong Chen, and Zhe Feng, "From natural driving to artificial intervention: Changes of the Yellow River estuary and delta development," *Ocean and Coastal Management*, 2019, *174*, 63–70.

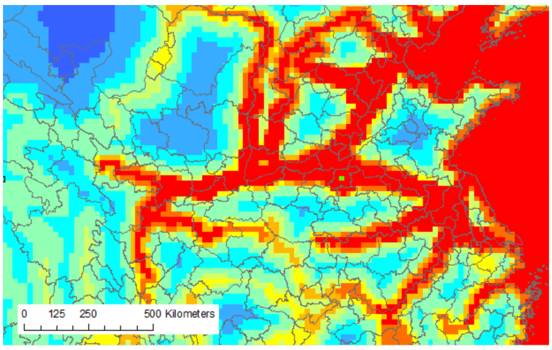
Figures and tables

Figure 1. Courses of the Yellow River since 350 B.C.E.

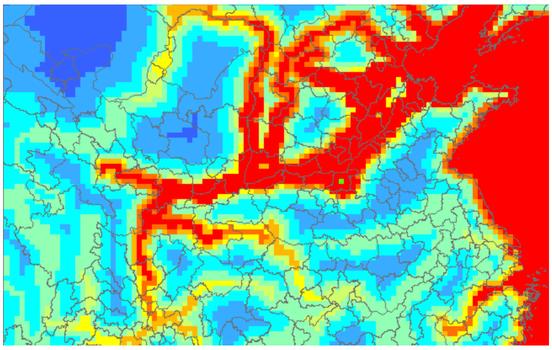


Notes: This map represents major changes in the course of the Yellow River between 305 B.C.E. and the present. It focuses on the North China Plain, between Tianjin and the Yangzi River delta. The black line corresponds to the current courses of the Yellow River and Yangzi River. Blue lines correspond to major past courses of the Yellow River (the lighter the color, the older the course). Grey lines materialize prefecture boundaries (as of 2000).

Figure 2. Least cost distances from Kaifeng, before and after the 1855 Yellow River diversion.



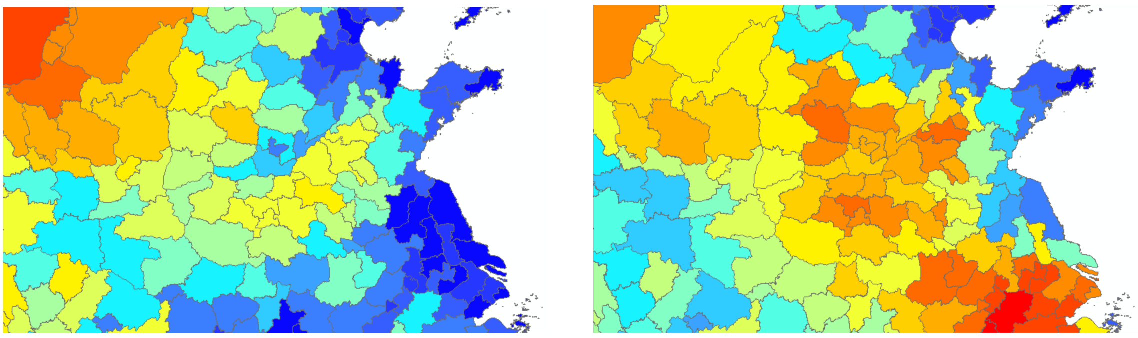
(a) Travel costs from Kaifeng, 1644–1855



(b) Travel costs from Kaifeng, 1855–1938

Notes: This map represents the minimum cost distances from Kaifeng (d_{ijt}), before (a) and after (b) the 1855 change in the course of the Yellow River. It focuses on the North China Plain, between Tianjin and the Yangzi River delta. The baseline scenario is used to define relative trade costs—see Section 3 for details. A darker (lighter) color indicates a lower (higher) travel cost. Grey lines materialize prefecture boundaries (as of 2000). Kaifeng is materialized by a green square.

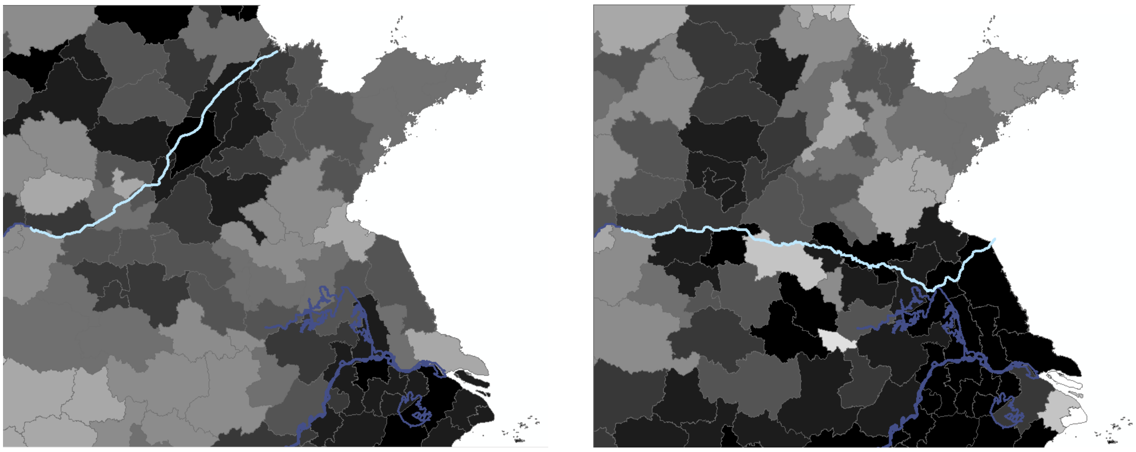
Figure 3. Variation in market access.



(a) Constructed market access, 1080 (b) Constructed market access, 1393

Notes: This map represents the residual of the constructed measure of market access (Z_{it}), in 1080 (a) and in 1393 (b) by quantiles (20), from a regression on the as-the-crow-flies market access. It focuses on the North China Plain, between Tianjin and the Yangzi River delta. The baseline scenario is used to define relative trade costs—see Section 3 for details. A warmer (colder) color indicates a higher (lower) residual. Grey lines materialize prefecture boundaries (as of 2000).

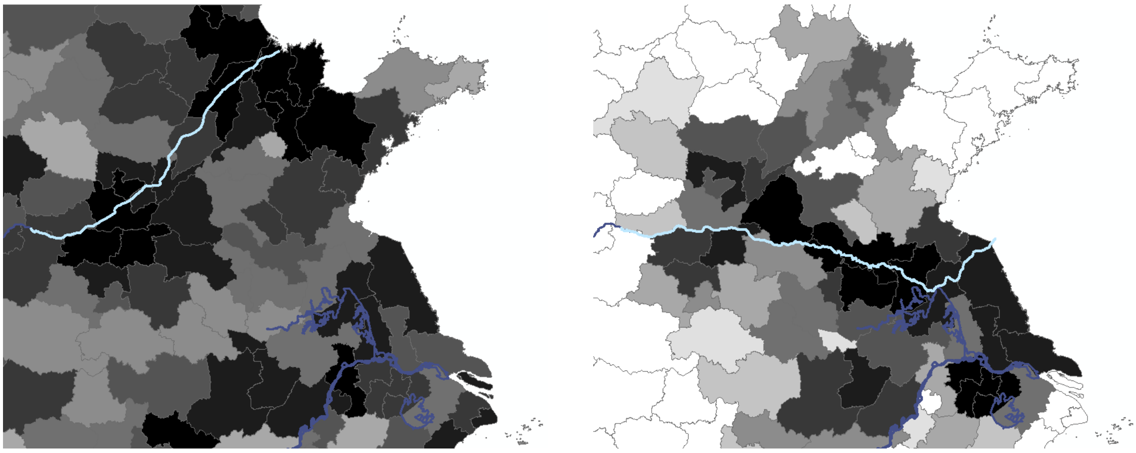
Figure 4. Variation in economic activity.



(a) Population density, 1102 (b) Population density, 1784

Notes: This map represents population density in 1102 (a) and 1784 (b), by deciles. It focuses on the North China Plain, between Tianjin and the Yangzi River delta. The lower Yellow River is shown in light blue. The Yangzi river system is shown in navy blue. A darker (lighter) shade indicates a higher (lower) value. Grey lines materialize prefecture boundaries (as of 2000).

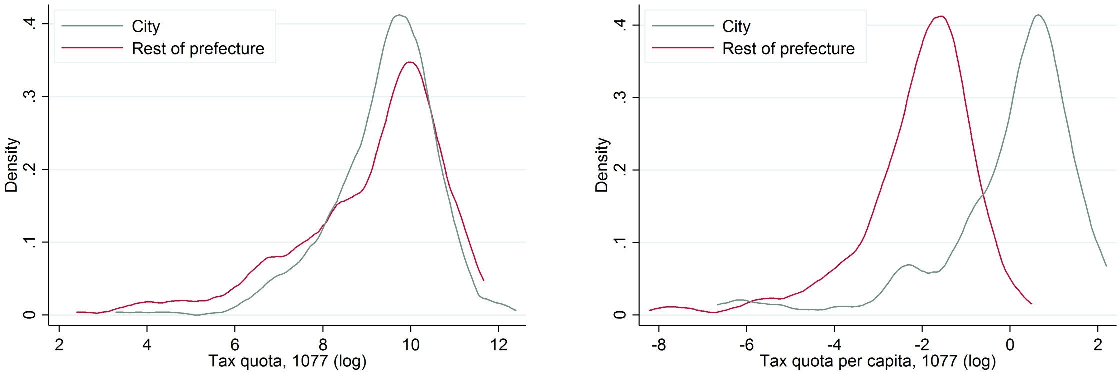
Figure 5. Variation in taxation intensity.



(a) Tax per capita (strings of cash), 1077 (b) Tax per capita (*dan* of wheat), 1784

Notes: This map represents the amount of tax per capita in 1077 (a) and in 1784 (b), by deciles. It focuses on the North China Plain, between Tianjin and the Yangzi River delta. Taxes are measured in strings of cash (*guan*) in 1077 and in *dan* (a measure of volume) of wheat in 1784. The lower Yellow River is shown in light blue. The Yangzi river system is shown in navy blue. A darker (lighter) shade indicates a higher (lower) value. Prefectures for which wheat was not used for taxation are in white. Grey lines materialize prefecture boundaries (as of 2000).

Figure 6. Taxation intensity in urban and rural areas.



(a) Urban vs. non-urban tax levels (b) Urban vs. non-urban tax burden

Notes: These two figures represent the distribution of tax quotas (left panel) and tax quotas per capita (right panel) in 1077. The green line shows the distribution for cities; the red line shows the distribution of the rest of the prefecture. Tax quotas and tax quotas per capita are expressed in log and trimmed at 1%. Panel (b) assumes that 10% of the population lives in cities.

Table 1. First stage.

Actual market access (M_{it})	(1)	(2)	(3)
Constructed market access (Z_{it})	0.982***	0.725***	0.801***
	(0.241)	(0.167)	(0.172)
Observations	2,620	2,620	2,620
Prefecture fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
As-the-crow-flies market access	Yes	Yes	Yes
Sea access×year	No	Yes	Yes
Cost surface	No	No	Yes

Notes: This table shows the first stage of the main specification. The sample is restricted to observations with non-missing population density in 1080, 1393, 1491, and 1578. An observation is a grid cell×year. Both the left- and right-hand side variables are standardized; they follow the baseline specification of the cost surface—see Section 3 for details. All regressions control for as-the-crow-flies market access. Column 2 introduces a dummy equal to 1 if the cell has access to the sea, interacted with the year fixed effects. Column 3 additionally controls for the cost of crossing the cell (cost surface used to compute M_{it}). Standard errors are clustered at the prefecture (2000) level.

Table 2. Effect of market access on economic concentration.

	Population density (log)		
	(1)	(2)	(3)
Panel A: OLS specification			
Market access	0.148*	0.169**	0.167**
	(0.085)	(0.082)	(0.082)
Observations	2,620	2,620	2,620
Panel B: IV specification			
Market access	0.238*	0.332**	0.284*
	(0.137)	(0.158)	(0.159)
Observations	2,620	2,620	2,620
Prefecture fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
As-the-crow-flies market access	Yes	Yes	Yes
Sea access×year	No	Yes	Yes
Cost surface	No	No	Yes
F-Stat. (first stage)	16.67	18.77	21.69

Notes: This table shows OLS (Panel A) and IV (Panel B) regressions of (log) population density on market access (standardized). The sample covers the following years: 1080, 1393, 1491, and 1578. An observation is a grid cell×year. In the IV specification, the actual market access (M_{it}) is instrumented by the exogenous market access (Z_{it}). Both follow the baseline specification of the cost surface—see Section 3 for details. All regressions control for as-the-crow-flies market access. Column 2 introduces a dummy equal to 1 if the cell has access to the sea, interacted with the year fixed effects. Column 3 additionally controls for cost of crossing the cell (cost surface used to compute M_{it}). Standard errors are clustered at the prefecture (2000) level. The F-statistic displayed is the one recommended by Sanderson and Windmeijer (2016).

Table 3. Effect of market access on structural transformation.

	Share manufacturing		
	(1)	(2)	(3)
Panel A: OLS specification			
Market access	0.311***	0.346***	0.346***
	(0.085)	(0.099)	(0.100)
Observations	1,919	1,919	1,919
Panel B: IV specification			
Market access	0.769***	1.002***	1.092***
	(0.209)	(0.239)	(0.252)
Observations	1,919	1,919	1,919
Prefecture fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
As-the-crow-flies market access	Yes	Yes	Yes
Sea access×year	No	Yes	Yes
Cost surface	No	No	Yes
F-Stat. (first stage)	20.10	25.11	30.20

Notes: This table shows OLS (Panel A) and IV (Panel B) regressions of the share of manufactured goods or handicrafts (among the goods remitted as taxes by the prefecture) on market access. The sample covers the following years: 1080, 1393, 1502, and 1578. An observation is a grid cell×year. In the IV specification, the actual market access (M_{it}) is instrumented by the exogenous market access (Z_{it}). Both the dependent variable and market access measures are standardized. Market access measures follow the baseline specification of the cost surface—see Section 3 for details. All regressions control for as-the-crow-flies market access. Column 2 introduces a dummy equal to 1 if the cell has access to the sea, interacted with the year fixed effects. Column 3 additionally controls for cost of crossing the cell (cost surface used to compute M_{it}). Standard errors are clustered at the prefecture (2000) level. The F-statistic displayed is the one recommended by Sanderson and Windmeijer (2016).

Table 4. Effect of market access on taxation intensity.

	Number of goods taxed		
	(1)	(2)	(3)
Panel A: OLS specification			
Market access	0.237***	0.274***	0.275***
	(0.071)	(0.081)	(0.080)
Observations	1,919		
Panel B: IV specification			
Market access	0.242*	0.267	0.293*
	(0.135)	(0.173)	(0.173)
Observations	1,919	1,919	1,919
Prefecture fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
As-the-crow-flies market access	Yes	Yes	Yes
Sea access×year	No	Yes	Yes
Cost surface	No	No	Yes
F-Stat. (first stage)	20.10	25.11	30.20

Notes: This table shows OLS (Panel A) and IV (Panel B) regressions of the number of goods remitted as taxes by the prefecture on market access. The sample covers the following years: 1080, 1393, 1502, and 1578. An observation is a grid cell×year. In the IV specification, the actual market access (M_{it}) is instrumented by the exogenous market access (Z_{it}). Both the dependent variable and market access measures are standardized. Market access measures follow the baseline specification of the cost surface—see Section 3 for details. All regressions control for as-the-crow-flies market access. Column 2 introduces a dummy equal to 1 if the cell has access to the sea, interacted with the year fixed effects. Column 3 additionally controls for cost of crossing the cell (cost surface used to compute M_{it}). Standard errors are clustered at the prefecture (2000) level. The F-statistic displayed is the one recommended by Sanderson and Windmeijer (2016).