

# The roadblock effect

## War shocks, modal shifts, and population changes

**Yan Hu** \* #  
*University of Edinburgh*

[Latest version available HERE](#)

**ABSTRACT:** This paper investigates a novel “roadblock effect,” whereby temporary forced route changes catalyse the adoption of a new transport technology. The Taiping Rebellion in 19th-century China ravaged many cities, but also blocked key land routes, triggering investments facilitating steamship trade. Combining a trade model featuring modal and route choice, shipping records, and a new method to estimate historical trade costs, I show that the post-rebellion spatial variation in maritime trade was driven by blocked land routes and the feasibility of sea alternatives. This moved many trade routes to sea and shifted the population towards port cities.

Key words: path dependence, modal choice, trade costs, Taiping Rebellion  
JEL classification: F14, N95, R12, R40

\*I am deeply indebted to Diego Puga for his support and guidance. I am grateful to Guillermo Caruana, Jorge De la Roca, Mauricio Drelichman, Keith Head, James Kung, Chicheng Ma, Monica Martinez-Bravo, Joan Monras, Stephen Redding, Daniel Sturm, Matthew Turner, Maisy Wong, Noam Yuchtman, and especially Paula Bustos and Nezih Guner for their advice. I thank participants at the Meeting of the Urban Economics Association, Virtual Economic History Seminar, NEUDC, Econometrics Society European Meeting, CES North America Virtual Conference and LSE Economic History Graduate Seminar for their helpful comments and suggestions. Special thanks to Shuji Cao for sharing historical population data in China. Funding from the European Research Council under the European Union’s Horizon 2020 Programme (ERC Advanced Grant agreement 695107 - DYNURBAN) and from Spain’s Ministerio de Economía y Competitividad (Grants ECO-2016-80411-P, BES-2017-082187 and María de Maeztu Programme for Units of Excellence in R&D, MDM-2016-0684) is gratefully acknowledged. The usual disclaimer applies.

#31 Buccleuch Place, EH8 9JT, Edinburgh, UK. E-mail: yan.hu@ed.ac.uk.

## 1. Introduction

This paper uses the Taiping Rebellion in 19th-century China to study the evolution of city populations when shocks hit them, not only directly, through deaths and displacement, but also indirectly, by promoting some trade routes and modes over others. This highlights the complex interplay between large population shocks, shifting natural advantages, the persistence of agglomerations, and substantial changes in trade costs through technology and investment.

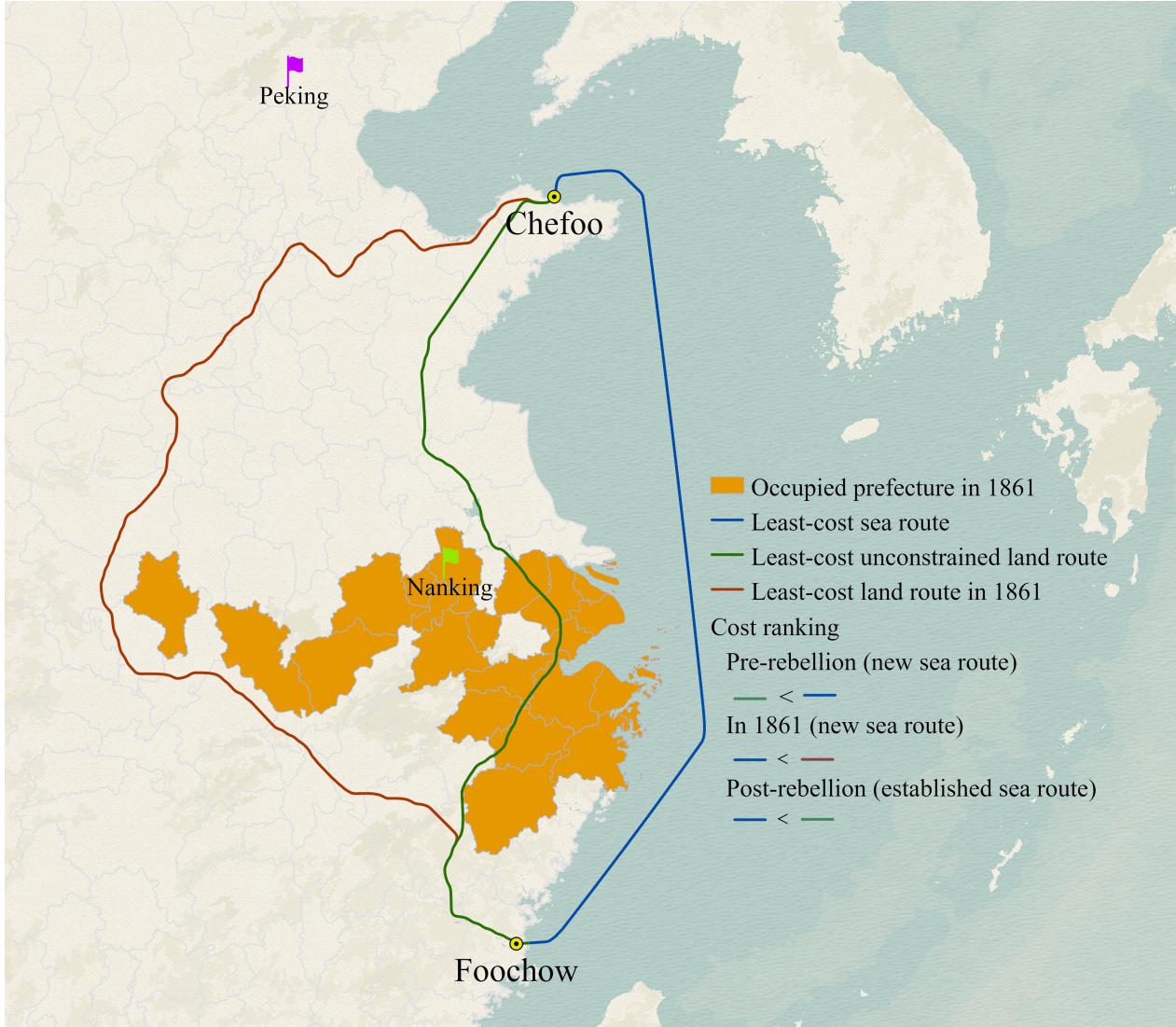
City locations and sizes reflect a combination of natural advantage and agglomeration economies. Natural endowments and technology differ across space, making some potential city locations more suitable than others. At the same time, localised increasing returns make locations that already have many firms and people more attractive to other firms and people. Comparative advantage of locations can change when new technologies emerge. However, agglomeration forces may trap populations in a location even if it is longer superior, whether because of coordination problems or sunk infrastructure costs (e.g., Bleakley and Lin, 2012; Michaels and Rauch, 2017).

Trade costs complicate the analysis of relative city sizes further because, in addition to underlying conditions (“first nature”) and the current concentration at each location (“second nature”), we have to consider the access each of them has to a larger market (Krugman, 1993). Trade costs can also change over time as new transportation technologies become available, and locations that benefit from them through increase in market access become more desirable. However, new technologies usually require large sunk investments upfront, which may deter the adoption despite their superiority. Moreover, whether or not locations adopt new technologies are often endogenous.

This study exploits the heterogeneous, and plausibly exogenous, incentives to adopt steamship transport through the route and modal changes induced by shocks of civil conflict elsewhere. The Taiping Rebellion was the largest peasant revolt in the history of China. The civil war between the rebels and the Qing government led to the death of one out of every six Chinese people (Ge, Hou and Zhang, 1999, 99). The rebellion caused direct damage and casualties in the war-affected area, concentrated mainly in the rebellion occupied region. At the same time, the indirect effect of the war, operating through changing trade routes and modes, also changed the distribution of population.

Despite a large potential for sea trade, trade activities and population in China remained largely inland before the rebellion. When the rebellion blocked some of the main inland trade routes, this forced a search for alternatives. It triggered substantial investments to facilitate sea trade, notably using steamships, which until then had a slow take-up. After the rebellion, with sunk investment already incurred for establishing sea lanes, many trade routes permanently shifted to sea transport. This in turn catalysed a shift of population towards port cities. An example of this can be seen in Figure 1. The figure shows in green the least-cost land route connecting Chefoo (located in today’s Yantai) and Foochow (Fuzhou) before the rebellion. These two coastal cities could also have been connected by sea, but my estimates show that the cost savings compared with the usual land route was not large enough to incur the sunk costs associated with establishing such a route.

Figure 1: An example of the “roadblock effect”



*Notes:* Least-cost land and sea routes between Chefoo (today's Yantai) and Foochow (Fuzhou) based on travel costs in *Panel A* of Table II. Land routes can combine courier and non-courier land routes with different levels of terrain ruggedness as well as inland waterways. The cost ranking takes into account the cost of using each mode, the cost of inter-modal transfer, and the cost of developing a new land or sea route, as estimated in *Panel B* of Table II.

These sunk costs include learning about the current and tides along the sea lane, being warned of dangerous spots, having staging-posts ready and ensuring safe anchorage on arrival. The rebellion occupied many of the prefectures traversed by the usual land route between Chefoo and Foochow. The dark shaded area in the map shows the rebellion-occupied prefectures in 1861. To transport goods between Chefoo and Foochow one would have had to take the large detour marked in red. According to my estimates presented below, the additional cost this detour entailed was large enough to instead incur the sunk cost of establishing a steamship route between these two cities. Although the unconstrained regular land route became once again usable after the rebellion, with the sea route

already in operation, steamship trade remained a less costly option. Indeed the data show that there was active sea trade between the two ports immediately after the rebellion. In contrast, regular land routes between other city pairs (e.g., Amoy and Canton) were not blocked by the rebellion and steamship trade between them did not become established. I next explain how I examine this process systematically through my analysis.

I start by providing some historical background in Section 2, then describing the data in Section 3, and characterising the impact of the rebellion on population in Section 4. After controlling for pre-existing trends and other differences across prefectures, I find that being occupied during the rebellion was associated with a permanent population loss of 54.6%, whereas being located on the coast was associated with a 31.5% population increase. This increase was even greater in coastal cities that saw a surge in domestic sea trade.

To show that population relocated to coastal locations because the rebellion shifted trade routes from land to sea, I exploit regional variations in incentives to take up sea trade. The key for my identification strategy is that, depending on their location along the regular land trade network, locations would have had their trade flows more or less affected. Maritime trade could boom broadly when the general technology of sea transport advanced with the introduction of steamships. At the same time, land trade could be directly hit in areas occupied by the rebellion, since an explicit aim was to cut off supplies from this area to the capital of the Qing government in Peking. However, regions not directly hit by the rebellion may nevertheless have been substantively affected if a regular land trade route became unusable when the rebellion hits some intermediate point on this connection. This provides them with *additional* incentives to adopt steamships compared to those regions where inland trade remained intact during the rebellion, when advances in transport technology became commonly available.<sup>1</sup> Thus, in Section 5, I proceed to show that bilateral trade flows to and from unoccupied locations were affected indirectly and differently depending on the roadblocks on the corresponding routes and the feasibility of sea trade alternatives.

To implement this strategy, I combine new data on bilateral historical trade flows with a trade model featuring modal and route choice. I collect a new data set of bilateral maritime trade in 19th-century China by digitising archives of trade reports from the China Maritime Customs Service.<sup>2</sup> Given that only sea trade after the rebellion is observed in the data, some additional theoretical structure needs to be given to the analysis to characterise the indirect effect of the rebellion through trade. Trade models typically do not incorporate modal choice, since trade data are generally not differentiated by which transport method was used to carry the cargoes. Instead, I add modal choice

---

<sup>1</sup>In other words, the identification strategy helps to “difference out” common shocks to sea transport over time, such as change of attitudes of the government towards maritime trade and opening of treaty ports to trade with Western countries. Due to the availability of data, I only compare bilateral trade between treaty ports. While all treaty ports had access to foreign trade and could have adopted the frontier technology of steamships for domestic trade due to technological diffusion, I compare whether domestic trade between any two of these ports is more likely to use maritime transport, depending on the heterogeneous increase in inland transport costs caused by the conflict.

<sup>2</sup>The China Maritime Customs Service was a bureaucratic agency set up in 1853 by foreign consuls in Shanghai to collect tariffs that went unpaid when the Qing government was overwhelmed by the Taiping Rebellion.

(by land or by sea) to an otherwise standard trade model.<sup>3</sup> Sea trade required fixed investments, and this likely held them back until the blockage of land routes made such investments a necessity. Thus, I also incorporate sunk investment costs into the model.

The model features a “roadblock effect,” by which a temporary increase in land trade costs can trigger sunk investments for sea trade, thus having permanent effects on modal choices. Across location pairs, the heterogeneous increase in land trade costs during the rebellion can be linked to the increase in the probability of sea trade immediately after the rebellion, mediated by the incentives to make sunk investments in steamship trade.

The model yields a probability of sea trade after the rebellion for each location pair, which we observe in the data, given relative costs of land to sea transport. The relative cost of using different land trade routes and their sea alternatives is a key driver of the differential take-up of steamships in different geographical areas. The literature on transport costs usually infers them from freight rates, so the lack of data on freight rates by inland transport in 19th-century China imposes a challenge.

In the second part of Section 5, I develop a new method to estimate travel costs by different transportation modes using China’s 1903 postal map. The map contains the locations of postal district headquarters and the assignment of prefectures to these. Under the assumption that the assignment was made to minimise transport costs to the postal district headquarters, a rich transport cost function can be parametrised, which includes the typical cost of different transport modes, variations in cost due to road types and the terrain, and the fixed costs associated with transitions between modes (which may be unavailable even in modern times).

Despite the limitations imposed by being able to observe sea trade only immediately after the rebellion, with the estimated transport costs before, during, and after the rebellion, the elasticity of trade costs and sunk investment costs can be pinned down jointly.

The predictions of the model closely match observed modal choices. In particular, the model predicts an average probability of sea trade between “roadblocked” location pairs of 0.696 compared with 0.375 for “non-roadblocked” location pairs immediately after the rebellion in 1867. The respective probabilities in the data are 0.693 and 0.375. The model also suggests that the “roadblock effect” was very important in practice. Under a counterfactual scenario that eliminates roadblocks between roadblocked location pairs, the model estimates that the probability of sea trade between them would have been much lower, 0.365 instead of 0.693. The model, based on parameters estimated using data in 1867, can also match the evolution of trade in subsequent periods.

The “roadblock effect” is not confounded by other explanations for the post-rebellion surge in

---

<sup>3</sup>During the rebellion, steamship technology had already been introduced in treaty ports, which also constitute the sample of analysis for trade. According to both contemporary trade reports and historians (e.g., Fan, 1985, 131-136), the new technology was superior to sailing ships and constituted the vast majority of sea trade at treaty ports. The pre-rebellion trade by sailing ships was very limited, therefore many locations leapfrogged from inland trade to steamships after the rebellion. For simplicity, the baseline model assumes that during and after the rebellion the cost of sea transport refers to that of steamships. A model with a second-step choice between sail and steam ships with idiosyncratic preferences for them generates very similar results.

maritime trade. Section 6 builds and estimates an extended model, taking into account the possibility for reduction in investment costs in steamships on sea routes that could previously have sail trade. Given that the tendencies to use sailing ships prior to the rebellion are low and are not systematically different for “roadblocked” and “non-roadblocked” pairs, the estimated “roadblock effect” remains similar. The extended model also allows decomposing the relative contribution of the rebellion to stimulating maritime trade, as well as the part resulting from technological upgrades from sail to steam ships. The quantification exercise shows that the bulk (74%) of the increase in maritime trade after the rebellion is attributable to the “roadblock effect”. Section 7 further shows that the effect of roadblocks is robust to accounting for the effects of treaty ports and foreign trade, potential changes in inland transport networks other than roadblocks during the rebellion, and that roadblocks may be more likely to occur between location pairs that are further apart.

This paper contributes to the theoretical and empirical literature on whether the spatial distribution of economic activities is uniquely determined and therefore whether temporary shocks can lead to permanent changes in spatial structure (i.e. path dependence).<sup>4</sup> In particular, the paper is closely related to studies that incorporate changing comparative advantage due to innovations in transport technology. Bleakley and Lin (2012) document the continuing importance of historical portage sites as city locations in the United States even though their original advantages have long since become obsolete. Michaels and Rauch (2017) compare the diverging experiences of British and French cities following the collapse of the Western Roman Empire. In Britain, subject to a more intense shock, the urban network was displaced towards emerging navigable waterways. In France, where the ensuing destruction and deurbanisation was more limited, medieval towns instead re-emerged more often on their Roman-era locations. I also consider a historical episode where shifting natural advantage (due to the introduction of steamships) was combined with large but heterogeneous population shocks (due to the Taiping Rebellion). A key contribution of my framework is that, in addition to the direct channel operating through temporary population shocks, I highlight an indirect channel operating through temporary roadblocks and the incentives they provide to invest in new transport technologies.

In this respect, the paper also connects to the literature on the importance of trade costs and market access in determining the distribution of population across cities (e.g., Redding and Venables, 2004; Faber, 2014; Donaldson, 2018) and within cities (e.g., Baum-Snow, Brandt, Henderson, Turner and Zhang, 2017; Tsivanidis, 2018; Hebligh, Redding and Sturm, 2020). An important challenge when trying to evaluate the consequences of transport improvements is the endogeneity of infrastructure

---

<sup>4</sup>The empirical evidence on path dependence is mixed. Davis and Weinstein (2002); Miguel and Roland (2011) find that the long-term distribution of relative city sizes can be highly persistent after wartime bombings. On the other hand, Hanlon (2017) finds that the American Civil War had a permanent impact on British cities specialising in textile production. Redding, Sturm and Wolf (2011) show that the division of Germany led to relocation of air hub from Berlin to Frankfurt. After German reunification, air traffic did not return to Berlin due to the sunk costs of creating hubs. Allen and Donaldson (2018) develop a dynamic quantitative spatial model featuring path dependence, which can stem from reliance on past infrastructure investment.

investments. An advantage of the setting in this paper is that the Taiping Rebellion creates exogenous variation in the incentives to invest in sea transport technology.

The paper also adds to the literature on the effect of wars and conflicts on urban development (e.g., Glaeser and Shapiro, 2002; Voigtländer and Voth, 2012; Dincecco and Onorato, 2016). While the previous papers have focused on local impacts of war, this paper shows that the indirect effect on other regions via trade disruption and the ensuing incentives to adopt new technologies can be as important.

Similar to Juhász (2018), this paper looks at war-induced trade blockade and technological adoption. In contrast to the infant industry mechanism that delayed the adoption of mechanised cotton spinning in 19th-century France, the slow adoption of sea transport in China was due to persistence of population away from the coastal areas before the rebellion and the large sunk cost needed to develop the new technology.<sup>5</sup> Because adoption of transport technology changed trade costs, hence the market access, I also look at the effect of the rebellion on the long-run population distribution.<sup>7</sup> As the “roadblocked effect” depends crucially on whether the rebellion hits some intermediate point on regular land trade routes connecting locations, an additional contribution of the paper is developing a novel method to parametrise a rich transport cost function using allocation rules of locations based on transport cost minimisation.

A debate related to the persistence of city locations also pertains to the slow adoption of new technologies even when they are obviously superior to old ones. This could be due to the trade-off between the sunk costs of developing and adopting new technologies and the accumulated learning by doing of incumbent technologies (Brezis and Krugman, 1997). In fact, the introduction of a new technology sometimes revitalises innovation of an incumbent technology. Ward (1967) coined the term “sailing ship effect” for this phenomenon, following the analysis by Gilfillan (1935) of advances made in sailing ships in the second half of the 19th century in response to the introduction of steamships. According to Ward, such advances led to greater improvements in sailing ships than those made in the previous three centuries.

The “roadblock effect” that I document and investigate in this paper works in precisely the opposite direction: the temporary blockage of land trade routes during the Taiping Rebellion triggered substantial investments to facilitate the use of steamships instead. This had long-term consequences for the distribution of population across Chinese cities in addition to the direct shocks of the rebellion.

---

<sup>5</sup>According to conservative estimates of the cost of sailing ships before the rebellion, sea transport by sail was also much less costly than inland transport.

<sup>6</sup>Because I look at long-distance interregional trade in a still largely agricultural economy, the majority of trade consisted of agricultural produces, determined by regional fundamentals, such as climate and soil, which could explain why the response to blockade was to find alternative trade routes but not to improve production technology to produce the good on its own.

<sup>7</sup>An economic geography model is not developed here because there is also a direct effect of the war on population, but data on population loss from the war in each prefecture are not available.

## 2. Historical background

### 2.1 *The Taiping Rebellion and roadblock*

The Taiping Rebellion (1851-1864) was the largest peasant revolt in the history of China and the deadliest civil war in human history. About one out of every six people in China (approximately 70 million) died during the rebellion (Ge, Hou and Zhang, 1999, 99).

The rebellion started as guerrilla warfare in the mountainous area of the Kuanghsia province in southwestern China (Figure 2). Peasants rose up against the Qing government, which levied ever-increasing taxes after a series of natural disasters and an economic crisis following its defeat in the First Opium War. The uprising escalated into a civil war with the Qing government when the rebels marched northeastward and occupied Nanking, establishing around it “the Taiping Heavenly Kingdom” in the mid and lower Yangtze region. When the rebels’ attempt to besiege the imperial capital Peking was defeated, wars between the two forces receded to the area occupied by the rebellion, causing great damage and civilian deaths. Some prefectures in the occupied region lost up to 80% of their population (Cao and Li, 2000) and many places were razed to ground (Meyer-Fong, 2013, 8-9).

Before the rebellion, the area around the Yangtze River, especially its lower section, had become the economic heartland of China. The Qing government had relied on the Grand Canal to transport tax-in-kind from this region to the imperial capital. The rebels strategically seized this area as a base to overthrow the Qing dynasty (Kuhn, 1978). They occupied prefectures sitting on the Grand Canal and the Yangtze River, which strangled the flow of goods from the Yangtze region to the imperial capital through the inland routes. According to contemporary witness, the imperial capital was struck by food price spikes when the rebellion army cut off its grain supply and many residents fled (Ni, 2005, 102). With its supplies through the regular routes combining land route, the Yangtze River and the Grand Canal cut off, the Qing government was forced to rely on shipping of the tribute grain by sea via Tientsin to the imperial capital (Xia, 1995; Dai, 2012).<sup>89</sup>

This paper shows that the modal change was not unique to the transport of tribute grain nor to trade with Tientsin, but instead affected a much broader exchange of goods across the whole country. The military confrontation between the rebellion and the Qing government severely curtailed inland transport (including canals) that previously used trade routes traversing through

---

<sup>89</sup>The Yellow River shifted its course in 1855, which further paralysed the part of the Grand Canal where they intersected. While bursting of the Yellow River was a constant concern for the Qing government and the imperial court spent strenuous efforts to contain it before the rebellion, the court was forced to set aside the problem when preoccupied with the rebellion, leading to its ultimate shift (Jia, 2009; Dai, 2012). My baseline estimation takes into account the change of the Yellow River in 1855 and also shows in robust results that the “roadblock effect” was not primarily driven by it.

<sup>9</sup>The rebellion could also paralyse sea transport if the port became occupied and my analysis takes this into account. In particular, I only consider periods in which the prefectures that had ports were not occupied and therefore its transport costs to other locations by land and sea could be affected differentially.

Figure 2: The Taiping Rebellion and Imperial postal districts



*Notes:* Author's map based on the 1903 postal working map from the China Maritime Customs Postal Series and the China W dataset. Prefectures delimited in light grey and coloured if occupied by the Taiping Rebellion. China's historical postal districts delimited by dashed lines with the district headquarters marked by a green pin.

the rebellion-occupied region, which is evidenced by a number of contemporary reports from local officials to the emperor. In 1852 the governor of Kiangxi province wrote (Liao, 2010, 282):

Min Guangdong bandits [The Taiping rebels] invaded the You River [in Kiangxi province], merchants and traders were stuck.

This blocked the overland transport of the Wuyi tea in Fukien province to Canton (Liu, 2016) and both places were not under direct attack by the rebellion. In 1854 the governor of Anhwei province reported (Liao, 2010, 283):

Roads were blocked, merchants were wary, tariffs fell short.

The superintendent of Huai-an Pass in Kiangsu province reported in the same year (Liao, 2010, 283):

Kuachow [in Yangchow] and Chinkiang are still occupied by the rebels, trade routes between the north and south are not accessible.

Depending on the roadblock of their usual land routes and the feasibility of sea trade alternatives, bilateral trade flows between unoccupied locations could be affected differently, triggering investments facilitating steamship trade on some routes but not others. Before providing systematic evidence using post-rebellion patterns of trade in Section 5, an example from the *Reports on trade at the treaty ports in China* for the years 1871-1872 renders qualitative support for the “roadblock effect.” Blocked overland trade during the rebellion catalysed the adoption of steamship trade for a newly established treaty port along the Yangtze River:

The new port [Kiukiang] offered but few material advantages over the old ones for supplying Kiangsi, unless it was just about that time the trade of Ningpo was paralysed by the invasion of the Rebels into Chekiang.

In contrast, inland trade routes to Kiangsi province that were not affected by the rebellion were persistent after the war despite opening of the treaty port Kiukiang and the advantage of steamships:

These [inland] routes became so to speak the established outlet and inlet for Native products and Foreign manufactures; so that what between the power of “custom” and their convenience, they appear to have been slow in abandoning them.

Steamships could enter from sea to the Yangtze River ports unchallenged from 1861 onwards after the Second Opium War. While Kiukiang could be connected with all other treaty ports by steamship, the only active trade link of Kiukiang immediately after the rebellion by sea was with Ningpo, which also happened to be the destination to which the inland route was affected the most by the conflict according to my estimates in Section 5.

## 2.2 The failure to adopt sea trade before the rebellion

Because I will argue that the Taiping Rebellion acted as a catalyst for sea trade, using steamships in particular, this part provides information about the status of sea trade before the rebellion.

One reason for the weakness of sea trade prior to the Taiping rebellion was the concentration of population away from coastal areas. Early during its reign, the Qing dynasty (1644 to 1912) implemented a sea ban on the mainland. The purpose was to cut off supplies to the Zheng regime in Taiwan, which was founded by loyalists of the previous Ming dynasty (Gu, 1983). As a part of the policy, coastal residents were evacuated to areas 15 to 25 kilometres inland (Gu, 1983). When the Zheng family surrendered in 1683, the ban was lifted and former coastal residents were allowed to return (Huang, 1986). Thereafter, except for a short sea ban placed between 1717 and 1727, maritime trade with foreign countries was acquiesced.<sup>10</sup> Nevertheless, sea trade with foreign destinations did not restore to its full blossom in the previous dynasties and recovery of population in coastal regions was sluggish.

Traditionally, domestic trade used a combination of land transport and inland waterways (including canals). Use of sea trade was restricted to specific commodities (e.g., export of soy beans from the northeast to Shanghai) and varied substantially from year to year (Liao, 2010, 162). For the most part, trade was conducted by inland routes even though sea transport was much less costly.<sup>11</sup>

In 1842, after the defeat in the First Opium War, China was forced to open up treaty ports, in addition to Canton, to trade with western countries. Foreign steamships started to appear in large numbers on the Chinese coast.<sup>12</sup> They were faster, less reliant on wind and currents, and cheaper than the incumbent sailing ships. Despite these new changes, trade routes remained persistent. Foochow was one of four ports signed to open in 1842, but its celebrated black tea from the nearby Wuyi region continued to take a lengthy detour overland to Canton for export. It was not until the outbreak of the Taiping Rebellion, when the regular land route to Canton became blocked, that Wuyi tea was shipped directly from Foochow in 1853 (Liu, 2016).

The inertia in the use of established land and river routes could be due to a large sunk cost usually associated with establishing new trade routes. Especially for sea trade, mariners need to harness the currents and tides along the way, be warned of dangerous spots, have supplies when needed and find safe anchorage on arrival. The large investment needed for safe voyage became evident

---

<sup>10</sup>In 1758, the Qing government established the Canton system, restricting all trade with western countries to only one port, Canton in the South. Seen by some scholars as a closed-door policy, the restriction was not strictly carried out. Foreign ship arrivals in other ports were documented and their numbers increased substantially (e.g., foreign merchant ships arrivals in Canton between 1757 and 1838 increased by 16 times over the previous 72 years and all ports combined Huang (1986)). Trade with non-western countries was not subject to the Canton rule.

<sup>11</sup>According to a rough estimate by (Deng, 2009, 55-98), 25% of grain trade was carried out by sailing ship. Immediately after the rebellion in 1867, 72% roadblocked location pairs had sea trade, compared to 34% non-roadblock location pairs. The simple model developed in this paper matches these patterns of trade.

<sup>12</sup>Steamships first appeared in China sea as early as the late 1820s and early 1830s (Fan, 1985, 87-88).

when concerted efforts were made by the China Maritime Custom Service in the late 19th century, drawing on expertise of mechanics, engineers, surveyors and constructors from home and abroad to build a chain of lighthouses and a system of navigation aids to safeguard trips along the Chinese coast and the Yangtze River.<sup>13</sup> While the Qing government was responsible for constructing and maintaining roads (Deng, 2009, 39) and exerted strenuous efforts in improving the Grand Canal (Dai, 2012), its attitudes towards sea trade was rather lukewarm if not negative. Without state investment, the gain from sea trade for any individual merchant may not be large enough to compensate for a large investment upfront, although collectively and in the long term, merchants could benefit from investment facilitating sea trade, which could result in inefficient modal choice.

### 3. Data

This paper intends to identify the “roadblock effect,” whereby blocked land routes during the Taiping Rebellion catalysed adoption of sea trade and relocation of populations to the coast. This requires data on trade, historical transport costs, population and measures of roadblocks due to the rebellion. I describe each one of them in detail in the following.

I collect a new data set of bilateral domestic maritime trade between treaty ports by digitising archival reports of *Returns of Trade at Treaty Ports* published by China Maritime Customs Service (henceforth referred to as CMCS). CMCS was a bureaucratic agency established in 1853 by foreign consuls in Shanghai to collect tariffs that went unpaid when the Qing government became overwhelmed by the Taiping Rebellion.<sup>14</sup> CMCS supervised tariff collection of foreign exports and imports, as well as domestic maritime trade that went through treaty ports. Its statistics department published trade reports following strict statistics standards that are consistent across ports and over time. The trade data from CMCS are believed to be “the only reliable and comprehensive statistical information” to use in studying China’s historical trade (Cheng, 1956, 12). Keller, Li and Shiue (2012) made a thorough introduction of the trade data from CMCS.

The earliest data on bilateral domestic trade flow are available for 1867, when more accurate and consistent accounting practices were established, which is 3 years after the Taiping Rebellion ended. In addition, I digitised bilateral domestic trade every five year after 1867 until 1897 (i.e., 1872, 1882, 1887, 1892 and 1897), where consistent trade data are available, to look at the long-term

---

<sup>13</sup>These efforts include surveys by the British Royal Navy starting from 1842 resulting in a series of updates on *Notice to Mariners* and later construction of lighthouses, beacons and buoys along the coast of China (Bickers, 2013).

<sup>14</sup>CMCS was a Qing government agency, but its staff was a mixture of Chinese, usually in lower rank positions, and foreigners who perform managerial roles. Set up to assess taxes of foreign imports and exports, over time its role extended to collection of domestic tariffs, maintenance of harbours and lighthouses, weather inspection, payment of foreign loans and establishment of the modern postal system in China.

effects of the rebellion.<sup>15</sup> A useful feature of the CMCS data is that exporting and re-exporting are differentiated (Keller, Li and Shiue, 2012), thus I can use direct export to measure trade between treaty ports. The limitation of the CMCS data is that domestic maritime trade only covers treaty ports, which were ports opened to trade with western countries, and only the part of maritime trade carried by steamship (foreign and domestic) and foreign sailing ships (i.e., lorcas) is included. Maritime trade by Chinese sailing ships (e.g., junks) is not reported for the majority of years and ports. Inland trade is also not reported.

To deal with this data limitation, I build a trade model featuring modal and route choice, which helps to infer the missing inland trade. As the value of sea trade may be inaccurate with missing junk trade, I model whether two locations establish active trade by sea, which is less likely to be affected by this limitation. This is because steamships were superior to sail and quickly became the dominant form of sea trade when they were introduced at treaty ports (Fan, 1985, 131-136; Zhu, 2014).<sup>16</sup> According to the CMCS trade reports in 1871-1872 of Newchwang where junk trade was also recorded, sea trade between some locations still employed junk sailing ships, but in these case, steamships or lorcas were always used at the same time. For trade between other locations, only steamships or lorcas were used, either by leapfrogging from inland trade or displacing junk trade. Therefore, the information from the CMCS trade reports likely reflects accurately the use of sea trade between locations as it is unlikely that two treaty ports only had junk trade, but not any steam or lorcha trade.

Estimation of the model requires the intermediate inputs of transport costs by different modes and routes in 19th-century China. While the literature usually derives transport costs from freight rates or travel speed, this information is not available in historical China. In Section 5.2, I develop a novel method to estimate historical transport costs based on the index to the postal working map in 1903, which lists each prefecture in China under a postal district. Postal districts were subregional divisions of the Imperial post office that was established by the CMCS in 1896 to provide mail delivery service to the general public. The service expanded from postal district headquarters, usually a treaty port, to inland areas through its extensive delivery network (See Figure 2 for postal districts and their

---

<sup>15</sup>Trade was recorded in Customs Taels, a silver based monetary unit uniform across all ports and for all years except for 1867. In 1867, the southern ports reported their trade value in Mexican dollars and I convert these to the Customs Taels using the contemporary conversion rate. In 1901, CMCS took over native customs within 25 kilometres of the treaty port, at which point trade data may begin to include trade went through nearby native customs. Therefore, consistent domestic trade data up to 1897 is used.

<sup>16</sup>In Britain, steam displaced sail on short routes in 1850s and 1860s (Harley, 2015). The improvement of engine efficiency in the 1850s and the invention of higher-pressure boilers in 1860s mattered more for the competitiveness of steam over sail in long-distance international trade than intranational trade (Pascali, 2017).

headquarters<sup>17</sup>).<sup>18</sup>

The Imperial postal office developed a system of transliterating Chinese places names, referred to as postal romanisation. In this paper, postal romanised place names are used as they appear in trade reports to indicate reference to historical sites. The mapping from the postal names to their current locations is listed in Table A.1.

The inland transport cost is modelled to vary with road types (i.e., courier routes and navigable rivers) and terrain. I obtain the locations of the Qing courier route (1800-1900) from China W dataset, locations of rivers from the CHGIS V4 dataset of coded river in 1820 and terrain ruggedness from Nunn and Puga (2012).<sup>19</sup>

To measure how the rebellion affected regular inland trade routes, I collected information on prefectures occupied by the rebellion on a monthly basis from Guo (1989) and Hua (1991). By aggregating occupied prefectures within each year during the rebellion, I construct roadblocks between location pairs from 1853 to 1864.<sup>20</sup>

To look at whether roadblocks induced population changes when trade routes shifted to sea, I use the 2018 update of historical prefecture population data of Cao (2000).<sup>21</sup> The data were constructed based on government-administered census, and cross-checked with literary accounts, notes and more than 3000 local gazettes. Prefectures are the second level administrative divisions in China. They are subdivisions of provinces, followed by counties and towns. Historical prefecture population is available for the pre-rebellion period in 1680, 1776, 1820, 1851 and post-rebellion period in 1880, 1910 and 1953. I supplement with population data from the 1982 county census and from the 2010 township census and map modern population data to the historical prefecture boundaries in 1820 taken from the China Historical Geographic Information System (CHGIS V4) constructed by the Harvard Yenching Institute and Fudan Center for Historical Geography.

---

<sup>17</sup>The headquarters of each postal district was usually a treaty port with the custom commissioner as the postmaster, with the exception of two inland postal districts: Peking and Chinan.

<sup>18</sup>There were no major changes in land transportation technology in 19th-century China. Railway was still at its early stage of development in China. In 1903, only regional railways operated in the northern part of China surrounding the capital. They were rail lines connecting Chengting to Peking and Peking to Tientsin and Chinchow, and Chinchow to Chinan. Therefore, postal districts whose planning could potentially be affected by the railway network are dropped from the estimation (i.e., Peking, Tientsin, Newchwang, and Chinan postal districts).

<sup>19</sup>Rivers are coded with values between 1 to 6. Rivers of an order greater than 4 are used, which corresponds to navigable waterways. The Qing courier routes were built and maintained by the government (Deng, 2009, 39). They consisted of inland waterways and overland roads. When coded rivers overlap with courier route, I use the transport cost of river. Therefore, the remaining courier routes consist mainly of roads and small rivers.

<sup>20</sup>Although the earliest uprising took place in Kweiping, Kuangshi in 1851, the rebel advanced to the middle and lower Yangtze region and occupied the first prefecture, Chinkiang, in January, 1853. The last occupied prefectures fell to the Qing government in August, 1864.

<sup>21</sup>The previous version of the data were used by Jia (2014) to study the long-run effect of treaty ports on population and economic development in China.

## 4. The evolution of population

I begin my analysis in this section by examining the long-term evolution of the population of Chinese prefectures (cities and their immediate surroundings). This shows that prefectures directly occupied by the rebellion suffered a huge shock from which they have not recovered even after one and a half centuries. In addition, the population shifted towards coastal areas, in particular those in which domestic sea trade took off. Then, in the next section, I will show that the rapid but heterogeneous adoption of sea trade, notably using steamships, was also an effect of the rebellion through the “roadblock effect,” that explains population relocation to coastal regions during this period.

As a motivation, I estimate the following reduced-form relationship between population, timing relative to the rebellion, occupation during the rebellion, coastal location and sea trade:

$$\log \text{pop}_{it} = \psi_i + \sum_{p=\text{prov}} I_i^p \lambda_{pt} + \alpha \text{post}_t + \beta (\text{post}_t \times X_i) + \sum_{j=1910}^{2010} I_t^j \kappa_j + \sum_{j=1910}^{2010} I_t^j \gamma_j \times X_i + \epsilon_{it}, \quad (1)$$

where  $\text{pop}_{it}$  denotes the population of prefecture  $i$  in year  $t$ ,  $\psi_i$  is a prefecture fixed effect,  $I_i^p$  is an indicator for province  $p$  (taking value 1 if prefecture  $i$  belongs to province  $p$  and 0 otherwise),  $\lambda_{pt}$  is a province specific time trend,  $\text{post}_t$  is an indicator variable denoting whether year  $t$  is after the 1851–1864 Taiping Rebellion,  $I_t^j$  is an indicator for post-rebellion year  $j$  (taking value 1 if  $j = t$  and 0 otherwise),  $X_i$  is a vector of prefecture characteristics, including whether the prefecture was occupied during the rebellion, whether it is located on the coast, and a number of controls, and  $\epsilon_{it}$  is an error term. Each observation is one of the 179 prefectures from 13 provinces in the core region of 19th-century China in each data year.<sup>22</sup>

The results are presented in Table I. Robust standard errors are clustered at the prefecture level to adjust for heteroskedasticity and within-prefecture correlation over time. Column (1) looks at the different evolution of populations in those prefectures that were occupied by the rebellion and those that were not. Provincial time trends capture differential trends of population evolution across provinces. The post-rebellion indicator captures a 26% average drop in population in non-occupied prefectures during the rebellion. The post-rebellion  $\times$  occupied interaction then captures the additional drop in occupied prefectures, an extra 59% on average. The two combined gives the overall population drop in occupied prefectures, which is about 86%.

To look at population evolution in later periods, I include year fixed effects for post-rebellion years ( $\kappa_j$ ), except for immediately after the rebellion in 1880, whose effect is captured by the post dummy. Year fixed effects provide the change in population levels for later periods for all prefectures

---

<sup>22</sup>As Qing troops from all parts of the empire were transferred to suppress the Taiping Rebellion, the military vacuum spurred uprisings in Yunan, Shensi, Kansu, Szechuen and Kueichow at the same time. These were periphery regions inhabited also by non-Han minority groups, so they are dropped from the analysis. Nomadic regions of the Qing dynasty such as the northeastern provinces (Kirin, Heilungchiang and Shengching), Outer Mongolia, Inner Mongolia, Xinjiang, and Tibet are also excluded. The panel is balanced except for Taiping prefecture, whose population in 1982 is missing. Despite sharing the same name, this prefecture is not related to the rebellion.

with respect to the pre-rebellion trends after incorporating the drop immediately after the rebellion. The results show that the average levels of population for all prefectures in 1910 and 1953 increased slightly, while there was a substantial increase in 1982 and 2010. More importantly, it is of interest to know whether population evolution in occupied and non-occupied prefectures was different in later years, which is captured by the post-rebellion year dummy  $\times$  occupied interactions. The result shows slow catch-up of the level of population in occupied prefectures even after one and a half centuries (17% additional increase in the level of population in occupied prefectures in 2010 rendering 43% relative loss in 2010). This is in sharp contrast to [Davis and Weinstein \(2002\)](#), which shows swift recovery of Japanese cities to their relative sizes after the Allied bombings.

In addition to the direct impact of the rebellion on war-torn areas, I hypothesise that it could affect relative attractiveness of locations through diverting trade routes to sea. Locations that stood to benefit from sea trade would be coastal prefectures, which is investigated in column (2). The post-rebellion  $\times$  coastal interaction shows that the population of coastal prefectures increased by 29% after the rebellion on average, after controlling for the direct effect of the rebellion on the occupied region. Post-rebellion year dummies interacted with the coastal dummy are also included. They show that population of coastal prefectures relative to non-coastal prefectures decreased slightly in 1910, in contrast to the sizeable increase immediately following the rebellion.

Population growth in coastal regions immediately after the rebellion could be driven by other characteristics of coastal areas. While prefecture fixed effects account for the population differences across prefectures in levels, those characteristics may have a differential impact on population growth after a large shock to population sizes. For instance, coastal areas may have better agricultural productivity, which could attract migration after the rebellion. Therefore, in column (3), I include a variety of locational traits interacted with the post dummy. These are the initial level of human capital and a set of geographic controls including terrain ruggedness, agricultural suitability for major crops in historical China and whether the Grand Canal went through the prefecture.<sup>23</sup> The positive coefficient estimate of the interaction on post-rebellion and coastal dummies is robust to inclusion of differential impacts of these locational characteristics on populations after the rebellion, showing a 32% increase in the level of population in coastal prefectures after the rebellion.

To provide suggestive evidence that populations relocated because trade activities were diverted to the coast as the rebellion blocked regular inland trade routes, in column (4), in addition to the post-rebellion  $\times$  coastal interaction, it is further interacted with the natural logarithm of the value of domestic and foreign trade by sea relative to its average in 1867 reported by the CMCS, which consists mainly of steamship trade, to see whether port cities with a surge in steamship trade had an increase in population after the rebellion. Since trade data are only available at treaty ports, which were large ports that had trade with western countries, the number of observations drops. For

---

<sup>23</sup>Terrain ruggedness is from [Nunn and Puga \(2012\)](#). The initial level of human capital is the number of graduates from the Imperial Exam in 1840s. The measure of human capital and other geographic controls are taken from [Jia \(2014\)](#).

Table I: Prefecture population evolution 1680–2010

|                                      | Dep. var.: Log (population) |                      |                      |                      |                      |
|--------------------------------------|-----------------------------|----------------------|----------------------|----------------------|----------------------|
|                                      | (1)                         | (2)                  | (3)                  | (4)                  | (5)                  |
| Post-rebellion                       | -0.263***<br>(0.033)        | -0.323***<br>(0.037) | -0.104<br>(0.174)    | -0.319***<br>(0.038) | -0.053<br>(0.195)    |
| Post-reb. × occupied                 | -0.594***<br>(0.094)        | -0.575***<br>(0.091) | -0.546***<br>(0.102) | -0.576***<br>(0.104) | -0.557***<br>(0.118) |
| Post-reb. × coastal                  |                             | 0.286***<br>(0.052)  | 0.315***<br>(0.057)  | 0.443***<br>(0.058)  | 0.613***<br>(0.092)  |
| Post-reb. × coastal × domestic trade |                             |                      |                      | 0.198***<br>(0.064)  | 0.142*<br>(0.075)    |
| Post-reb. × coastal × foreign trade  |                             |                      |                      | 0.003<br>(0.033)     | 0.027<br>(0.041)     |
| Post-reb. × initial human capital    |                             |                      | -0.044**<br>(0.022)  |                      | -0.065**<br>(0.027)  |
| 1910                                 | 0.024<br>(0.016)            | 0.038*<br>(0.019)    | 0.020**<br>(0.009)   | 0.035*<br>(0.020)    | 0.018*<br>(0.009)    |
| 1953                                 | 0.033<br>(0.022)            | 0.046*<br>(0.027)    | 0.028<br>(0.022)     | 0.042<br>(0.027)     | 0.025<br>(0.022)     |
| 1982                                 | 0.462***<br>(0.031)         | 0.481***<br>(0.035)  | 0.445***<br>(0.028)  | 0.483***<br>(0.036)  | 0.449***<br>(0.028)  |
| 2010                                 | 0.566***<br>(0.036)         | 0.551***<br>(0.039)  | 0.523***<br>(0.033)  | 0.559***<br>(0.039)  | 0.533***<br>(0.033)  |
| 1910 × occupied                      | 0.056<br>(0.043)            | 0.055<br>(0.042)     | 0.068*<br>(0.040)    | 0.071<br>(0.047)     | 0.084*<br>(0.044)    |
| 1953 × occupied                      | 0.119*<br>(0.066)           | 0.121*<br>(0.066)    | 0.131**<br>(0.065)   | 0.152**<br>(0.072)   | 0.159**<br>(0.070)   |
| 1982 × occupied                      | 0.120*<br>(0.071)           | 0.122*<br>(0.070)    | 0.146**<br>(0.066)   | 0.127<br>(0.078)     | 0.148**<br>(0.074)   |
| 2010 × occupied                      | 0.167*<br>(0.087)           | 0.179**<br>(0.087)   | 0.194**<br>(0.083)   | 0.167*<br>(0.097)    | 0.175*<br>(0.093)    |
| 1910 × coastal                       |                             | -0.068***<br>(0.022) | -0.052***<br>(0.018) | -0.073**<br>(0.037)  | -0.058<br>(0.037)    |
| 1953 × coastal                       |                             | -0.068<br>(0.042)    | -0.051<br>(0.043)    | 0.030<br>(0.098)     | 0.046<br>(0.104)     |
| 1982 × coastal                       |                             | -0.095*<br>(0.049)   | -0.061<br>(0.044)    | -0.075<br>(0.063)    | -0.044<br>(0.060)    |
| 2010 × coastal                       |                             | 0.065<br>(0.073)     | 0.093<br>(0.070)     | 0.329***<br>(0.125)  | 0.354***<br>(0.121)  |
| Post-reb. × geographic controls      |                             |                      | Yes                  |                      | Yes                  |
| Prefecture fixed effects             | Yes                         | Yes                  | Yes                  | Yes                  | Yes                  |
| Provincial time trends               | Yes                         | Yes                  | Yes                  | Yes                  | Yes                  |
| Number of observations               | 1,610                       | 1,610                | 1,565                | 1,368                | 1,323                |
| Number of prefectures                | 179                         | 179                  | 174                  | 152                  | 147                  |
| Number of coastal prefectures        | 35                          | 35                   | 35                   | 8                    | 8                    |
| R-squared                            | 0.832                       | 0.839                | 0.867                | 0.832                | 0.866                |

Notes: In columns (4) and (5), trade val. is the natural logarithm of domestic maritime trade value relative to the average in 1867, which is only available for 8 out of 35 treaty port prefectures (and the interaction is 0 for all 144 non-coastal prefectures). In columns (3) and (5), initial human capital is proxied by the natural logarithm number of graduates from the Imperial Exam in the 1840s, which is not available for 5 prefectures. The geographic controls are an indicator for whether the Grand Canal passed through, terrain ruggedness, and categorical variables for agricultural suitability for major crops in historical China. Robust standard errors clustered by prefecture are reported in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels.

this subset of large ports, the post-rebellion  $\times$  coastal interaction indicates that their post-rebellion population increase was about 44% on average. The post-rebellion  $\times$  coastal  $\times$  domestic trade triple interaction indicates that a 1% increase in domestic sea trade reported by CMCS was associated with a 0.2% increase in coastal population after the rebellion. However, the association between coastal population and foreign sea trade was much weaker as shown by the post-rebellion  $\times$  coastal  $\times$  foreign trade triple interaction.<sup>24</sup> This suggests that the increase in coastal population during this period was not driven by opening up treaty ports and initiating trade with western countries, but instead shaped by changes in domestic trade.

On the basis of column (4), column (5) further controls for locational characteristics interacted with the post-rebellion dummy. The general picture remains the same, indicating that the positive association between population and domestic sea trade was not confounded by locational characteristics that could at the same time affect sea trade.

The evidence of a greater increase in coastal population with a surge in domestic sea trade is based on a small sample of ports and needs to be taken with a grain of salt. In [Appendix C](#), I further show that the rise in coastal population was not specific to treaty ports. The increase of population in coastal prefectures that did not have treaty ports was also substantial: a 29% on average. Furthermore, the population increase in prefectures next to treaty ports that were not coastal was smaller, suggesting that the rise of coastal population was due to sea trade rather than merely spillovers from treaty ports. [Appendix C](#) also provides evidence that the population change happened immediately after the rebellion, and not one period before or after by including year fixed effects and their interactions with rebellion occupation and coastal locations before the rebellion.<sup>25</sup> Taken together, the evidence suggests that population increase in coastal locations after the rebellion was likely to be driven by adoption of sea trade.<sup>26</sup>

Fortunately, trade data come in more detail (e.g., information about who trades with whom by sea immediately after the rebellion). By exploiting the heterogeneous incentives to adopt sea trade through roadblocks during the rebellion, the next section further advances evidence that population relocation to coastal regions found in this section can be attributed to the “roadblock effect.”

---

<sup>24</sup>While both the volume of trade and population growth could be greater in larger cities, using the value of trade per capita gives very similar results.

<sup>25</sup>While this dynamic difference in difference specification shows the exact timing of population change for occupied and coastal locations, its limitation is that we can no longer control for pre-rebellion trends.

<sup>26</sup>The results do not mean that opening up treaty ports did not contribute to population increase. As columns (5) and (6) of [Table I](#) show, the population increase in treaty ports were on average larger than other coastal locations. [Jia \(2014\)](#) exploits the timing of the opening of treaty ports between 1840 to 1910 and finds that population in treaty ports increased compared to otherwise similar prefectures. The results of this paper suggest that in addition to the treaty port effect, there was a separate effect of the rebellion on coastal population operating through modal change.

## 5. The roadblock effect

In this section, I identify the “roadblock effect” by exploiting the heterogeneous incentives for locations to adopt sea trade depending on whether the rebellion blocked their regular inland trade routes and the feasibility of sea alternatives. Given only trade by sea after the rebellion is observed in the data, the analysis needs some additional theoretical structure. Therefore I build a trade model with modal and route choice in Section 5.1. An essential input of the model is the relative transport costs of using different inland routes and their sea alternatives. As we lack information about the transport costs in 19th-century China, in Section 5.2, I develop a novel method to estimate historical transport costs. Section 5.3 shows the estimation and prediction of the model using trade data immediately after the rebellion. I use counterfactual predictions to quantify the effect of the rebellion on adoption of sea trade. I also show out-of-sample prediction on trade for the subsequent periods.

### 5.1 Model set-up

We consider a country with  $J$  cities (prefectures in historical China), indexed by  $j = 1, 2, \dots, J$ . A continuum of symmetric firms in each city produces a distinct product, which can be consumed locally as well as in other locations subject to iceberg trade costs.<sup>27</sup> If we assume that consumers have constant elasticity of substitution utility and that firms face monopolistic competition, the firm profit function for exporting merchandise from location  $j$  to  $i$  at time  $t$  can be written:<sup>28</sup>

$$\pi_{ijt} = (1 - \alpha) \left( \frac{\tau_{ijt}^m c_{jt}}{\alpha P_{it}} \right)^{1-\epsilon} Y_{it}. \quad (2)$$

The elasticity of substitution is denoted as  $\alpha$ , which falls between 0 and 1. To keep the notation concise, I introduce  $\epsilon = 1/(1 - \alpha)$ . As standard in trade models, the profit increases with income  $Y_{it}$  and the Dixit-Stiglitz price index  $P_{it}$  in the destination ( $1 - \epsilon < 1$ ) and decreases with the cost of production  $c_{jt}$ , which can be thought of as labour cost and is common to all firms in origin  $j$ .

Departing from the standard setting, I allow the iceberg trade cost  $\tau_{ijt}^m$  to vary with transport mode  $m$ . Firms can either choose land transport or sea transport, which I denote  $l, s$  respectively. The baseline model assumes that sea transport uses the superior technology of steamships between treaty ports during and after the rebellion.<sup>29</sup> There was no steamship trade between mainland ports before

---

<sup>27</sup>Recent trade models have incorporated heterogeneous firms (e.g., Helpman, Melitz and Rubinstein, 2008; Tintelnot, 2017), but since only aggregate but not firm level trade data are available in 19th-century China, there is no gain from modelling heterogeneous firms here.

<sup>28</sup>While monopolistic competition is used to characterise the market structure here, as it will become evident later, the results do not depend on this assumption. The key assumption is that firms choose the least-cost transportation mode (taking into account sunk investment cost), which is likely to hold regardless of the market structure of the commodity market, whether or not there is a separate transport sector, or the market structure of the transport sector. If firms purchase transport service from a courier and the market price of the least-cost transportation mode is not the lowest, the courier of the lowest cost mode can gain from lowering the price.

<sup>29</sup>Steamship technology was available during and after the rebellion when treaty ports became established and was the dominant mode of sea transport. According to the CMCS reports of trade in 1873, at both a river port (Chinkiang) and a seaport (Ningpo), more than 90% of sea trade was borne by steamship.

the rebellion.<sup>30</sup> In Section 6, I also consider the possibility of pre-rebellion sail trade in affecting adoption of steamship when it became available and incorporate the choice between sail and steam ships during and after the rebellion in an extended model in Section 6. Because the prevalence of sail trade was low before the rebellion, the results do not change much.

In addition to making the trade cost a function of the transport cost  $D_{ijt}^m$  that takes into account optimal route choice within each mode, I incorporate intertemporal investment for establishing new trade routes for sea transport,  $I_t^s$ , and land transport,  $I_t^l$ .<sup>31</sup> The mode-dependent iceberg trade costs are:

$$(\tau_{ijt}^l)^{1-\epsilon} = (D_{ijt}^l I_t^l)^{-\gamma} e^{\eta_{l,ijt}}, \quad (\tau_{ijt}^s)^{1-\epsilon} = (D_{ijt}^s I_t^s)^{-\gamma} e^{\eta_{s,ijt}} \quad (3)$$

for land and sea transport respectively. The investment cost is modelled here as a multiplier to the transport cost, so that we can interpret it as the cost ratio of establishing a new route over using an existing route.<sup>32</sup> The investment cost of sea transport  $I^s$  depends on whether a sea route has already been established in the previous period. Define  $S_{ijt}$  as indicator for whether sea trade happens in period  $t$ . We have:

$$I_t^s = \begin{cases} I^s, & \text{if } S_{ij,t-1} = 0; \\ 1, & \text{if } S_{ij,t-1} = 1. \end{cases} \quad (4)$$

If sea route was established in the previous period, there is no extra cost of using sea trade in addition to the transport cost. If, instead, sea trade was not used in the previous period, firms need to incur the sunk investment cost  $I^s$  for establishing a new sea route in order to use sea transport. Analogously, in order to use a new land transport route, it is assumed that firms need to incur an investment cost of  $I_t^l$ .

Trade costs are subject to shocks. The shocks to land and sea transport  $\eta_{l,ijt}$  and  $\eta_{s,ijt}$  are assumed to be log normal and independent:

$$\eta_{l,ijt} \sim \mathcal{N}(0, \sigma_l^2), \quad \eta_{s,ijt} \sim \mathcal{N}(0, \sigma_s^2). \quad (5)$$

At the beginning of each period, firms observe origin-destination specific shocks to land and sea transport and choose either land or sea transport. From the model set-up, the probability of sea trade immediately after the rebellion can be written as:

$$\Pr(S_{ij,post} = 1) = \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s} - \mathbb{1}(S_{ij,reb} = 0)i^s\right), \quad (6)$$

---

<sup>30</sup>Domestic steamship trade only existed between Hong Kong and Canton, Hong Kong and Shanghai before the rebellion. As the Qing government ceded the Hong Kong island to Britain in the 1842 Treaty of Nanking, trade with Hong Kong was not included in my sample.

<sup>31</sup> $D_{ijt}^m$  is the lowest-cost route effective distance using the terminology of Donaldson (2018) with one distinction.  $D_{ijt}^m$  here is mode-specific. I calculate the effective distance separately for sea and inland transport.

<sup>32</sup>The implicit assumption is that sunk investment is proportional to the transport cost. The assumption is reasonable given establishing a new sea route involves acquainting with tides and dangerous spots, and installing navigational aids and staging posts along the lane. These costs increase with distance.

which depends first on the post-rebellion transport cost by land relative to sea ( $D_{ij,post}^l / D_{ij,post}^s$ ). The elasticity of trade with respect to the transport cost  $\gamma$ , normalised by the standard deviation of transport shocks  $\sigma = \sqrt{\sigma_l^2 + \sigma_s^2}$ , governs how responsive usage of sea trade is to the relative cost of land transport. Second, it depends on the sunk investment cost for sea transport, which may or may not have been incurred during the rebellion. To keep the notation concise, I introduce the transformed sunk investment cost  $i^s = (\gamma/\sigma) \log I^s$ , which measures the contribution of sea transport sunk investment cost to the probability of sea trade. If sea trade was used during the rebellion, and therefore sea routes were established ( $\mathbb{1}(S_{ij,reb} = 0) = 0$ ), firms do not need to incur the sunk investment cost in the post-rebellion period, thus increasing the probability of using sea trade afterwards.

Therefore, the post-rebellion sea trade depends on whether location pairs started sea transport during the rebellion. The probability of sea trade during the rebellion period is:

$$\Pr(S_{ij,reb} = 1) = \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l - i^s\right). \quad (7)$$

Equation 7 shows how the rebellion can affect incentives to make sunk investment in sea transport through the lens of the model. First, for locality pairs whose regular inland route was blocked as the rebels occupied prefectures that it previously traversed (indicator  $\text{roadblock}_{ij} = 1$ ), finding alternative inland routes entailed an investment cost  $I^l$  (the transformed sunk investment cost  $i^l = (\gamma/\sigma) \log I^l$  measures its contribution to sea trade probability). Second, as the new route was not the unconstrained optimal one, the relative cost of land transport to sea ( $D_{ij,reb}^l / D_{ij,reb}^s$ ) increased. Both raised the relative desirability of using sea transport during the rebellion, which induce “roadblocked” location pairs to incur the sunk investment cost for steamship during the rebellion.

We can link the post-rebellion probability of sea trade with sea trade during the rebellion via the inter-temporal decision of making sunk investment in sea trade. By conditioning on trade during the rebellion, the probability of sea trade after the rebellion can be written as:

$$\begin{aligned} \Pr(S_{ij,post} = 1) &= \Pr(S_{ij,reb} = 1) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s}\right) \\ &\quad + (1 - \Pr(S_{ij,reb} = 1)) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s} - i^s\right). \end{aligned} \quad (8)$$

Combining Equations 7 and 8, the post-rebellion probability of sea trade is:

$$\begin{aligned} \Pr(S_{ij,post} = 1) &= \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l - i^s\right) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s}\right) \\ &\quad + \left(1 - \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l - i^s\right)\right) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s} - i^s\right), \end{aligned} \quad (9)$$

which is a function of (1) whether the locality pairs had roadblocks, the relative cost of land transport compared to sea during and after the rebellion, and (2) the parameters of the model ( $\gamma/\sigma$ ,  $i^s$  and  $i^l$ ).

Equation 9 shows that the model generates the “roadblock effect,” by which temporary increase in land transport cost during the rebellion could have a long-lasting effect on transport modal choices, through facilitating investment in sea transport.<sup>33</sup> It also shows that the effect of the rebellion on trade hinges crucially on whether the regular land route became unusable when the occupied region hit some intermediate point on its connection, how lengthy a detour was necessitated and the cost of sea alternatives. This requires knowledge about availability and the cost of roads, river and terrain along different inland trade routes and the sea alternatives. If we know the relative costs of theses, the parameters of the model can be estimated by maximising the likelihood of observing post-rebellion sea trade in the data. Therefore, before turning to the estimation of the model, the next section shows the intermediate step of obtaining estimates of relative transport costs.

## 5.2 Estimation of historical transport cost

The increase in inland transport costs due to roadblocks and their sea alternatives is a key determinant of the differential impacts of the rebellion on steamship take-up in different locations. This demands knowledge about the costs of using different trade routes, which depends in turn on the cost of using various transport infrastructure underlying them. While the previous literature on transport costs usually infers them from freight rates or travel speed, lack of data on them in 19th-century China imposes a challenge.

To deal with this problem, I develop a novel method of estimating historical transport costs based on China’s postal map in 1903. I allow the typical cost of different transport modes (e.g., land and waterways), variations in cost due to terrain, as well as fixed costs associated with transitions between modes (which can be unavailable even in modern times).

The index to the postal working map in 1903 lists each prefecture in China under a postal district. The postal districts were regional divisions of the Imperial post office, which expanded the mail delivery service from the postal headquarters, usually a treaty port, to inland prefectures (See Figure 2). Under the assumption that the allocation of prefectures to postal districts was made to minimise transport costs to their postal headquarters, we can infer the cost of using different transport infrastructure available along the route.<sup>34</sup>

The surface of China is divided into 0.1 degree by 0.1 degree cells ( $553 \times 826$  cells in total, a cell at 30 degree latitude covering approximately  $90 \text{ km}^2$  area) and overlaid with rivers, courier routes,

---

<sup>33</sup>The post-rebellion probability is a weighted average of probabilities  $\Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s}\right)$  and  $\Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s} - i^s\right)$ .

The weight on the higher probability  $\Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s}\right)$  becomes larger when roadblocks happen.

<sup>34</sup>The estimation excludes regions that could be potentially affected by construction of railways in late 19th century and early 20th century (which were small-scale regional railways near the imperial capital). After excluding Peking, Tientsin, Newchwang and Chinan post districts, the estimation sample is left with 28 post districts.)

and terrain (See Figure B.1). Each cell is assigned a (relative) cost to traverse based on whether it falls in a river, courier route or point of transition between land and waterway transport. Given a set of travel costs for a courier land route cell, a non-courier land route cell, a river cell, a cell involving land to river modal transition, and penalties for ruggedness, we can use the Fast Marching Method (FMM) to find the least-cost travelled postal headquarters for each prefecture.<sup>35</sup>

The optimal relative transport costs are pinned down by maximising the number of matches between actual and FMM allocations of prefectures to postal headquarters using iterations. In the best scenario, 173 out of 199 prefectures can be matched to the actual postal districts compared to 114 when using distance alone. The majority of the unmatched prefectures (24 out of 26) are very close to be matched to their actual headquarters as their neighbouring prefectures are matched to those headquarters.

The estimation based on postal districts gives the transport cost by inland waterways. To obtain the transport cost by sea, I use the information about travel time for sea transport relative to river as variable transport cost such as coal consumption and crew wage increases proportionally with travel time. According to the CMCS trade reports in 1879 and 1881, steamship trip between river ports Chinkiang and Hankow took half a week whereas a trip between seaports Ningpo and Wenchow took 26 hours. The distance ratio between the two is 1.63. Therefore, the implied cost by sea is half the cost by river.

The estimation results of transport costs are shown in *Panel A* of Table II. The cost of travelling by land without any road infrastructure or ruggedness is normalised to one so that the unit transport costs by other modes is relative to that by plain land. The cost of using courier routes was 18 % of the cost of non-courier land routes. The cost of using inland waterways was 8% of the cost by non-courier land transport and the cost by sea was 4%. An increase in terrain ruggedness by one standard deviation was associated with a 32% increase in transport cost. Regarding transition, I convert the cost into equivalence to distance travelled by waterways. The transition cost from land to water transport amounted to the cost of travelling by inland waterways for 82.5 kilometres and by sea for 165 kilometres. In other words, for a sea trip between Shanghai and Ningpo, about half the transport cost can be attributed to transshipment.

Using the estimated historical transport costs, in AppendixD I estimate the elasticity of migration with respect to transport costs in historical China. I construct novel migration flow data based on the number of migration villages founded by people from outside and information about their ancestral origin. Reassuringly, the estimated elasticities are similar to those estimated with modern data and for other parts of the world. The absolute value of the elasticity using transport costs (lowest-cost route effective distance) is also larger than the elasticity using distance alone.

---

<sup>35</sup>The FMM algorithm code of Allen and Arkolakis (2014), which is based on the "Accurate fast marching" Matlab toolbox by Dirk-Jan Kroon, is used.

### 5.3 Model estimation and prediction

With transport costs in 19th-century China in hand, we can now proceed to estimate the model and use it to predict the probability of sea trade with and without roadblocks. I also show out-of-sample predictions for trade in subsequent periods using parameters estimated using the 1867 data.

I calculate the least cost route for each locality pair before and after the rebellion, and for each year during the rebellion when the two locations are not occupied. For each trade pair during the rebellion, I identify whether their inland trade needed a detour as the occupied region hit some intermediate point of their optimal inland route. I use the average increase in land transport cost due to roadblocks to calculate the relative inland transport cost  $D_{ij,reb}^l / D_{ij,reb}^s$  in Equation 9 if roadblocks happen in multiple years and the required detours are different as occupied regions change over time.

The estimation sample includes all treaty ports that were opened before or during the rebellion, which are the same set of treaty ports that were active in 1867. Trade between two river ports is excluded because it does not involve sea transport.<sup>36</sup> Two extreme cases, Shanghai, which had active maritime trade with all other treaty ports, and Chinkiang, which had no sea trade with any other ports in 1867, are excluded from the model estimation.<sup>37</sup> Reduced-form estimation indicates that the effects are similar across samples including Shanghai and Chinkiang, excluding Shanghai, and excluding Shanghai and Chinkiang. Among the “roadblocked” pairs, sea trade is more likely to happen if the relative increase in inland transport cost is higher.

With the estimated relative transport costs of land compared to sea during and after the rebellion and their increase during the rebellion due to roadblocks, we can obtain estimates of the parameters in the model by maximising the likelihood of observing maritime trade immediately after the rebellion in the data.

The results of the baseline model are presented in *Panel B* of Table II. The elasticity of using sea transport with respect to the ratio of land to sea transport cost is positive as expected, indicating that location pairs are more likely to use maritime trade when in comparison the cost of using inland transport is high. The cost associated with initialising sea trade is substantial. The estimate implies that the cost of using a new sea route is approximately 20 times the cost of an existing one

---

<sup>36</sup>In other words, trade between seaports as well as direct trade between a sea port and a Yangtze River port that does not involve any road transit are included. In the CMCS trade report, if export from a seaport is transferred to a Yangtze port by inland route, it is recorded under transit trade instead. Steamship trade between Yangtze ports and sea ports was allowed after 1861 according to agreement after the Second Opium War (Fan, 1985, 91-93), and the estimation takes into account the timing.

<sup>37</sup>Shanghai was surrounded by occupied prefectures throughout the rebellion and land transport to Shanghai was impossible according to my estimate ( $D_{ij,reb}^l / D_{ij,reb}^s$  would be infinitely positive). Shanghai was also the only port that had sea trade with all other ports in 1867. To show that the “roadblock effect” is not entirely driven by this exceptional example, it is dropped from the estimation. Although Chinkiang was not occupied for the later part of the rebellion, sitting next to the rebellion capital, it was caught by warfare and greatly damaged, which could explain lack of trade immediately after the war in 1867. According to the CMCS trade report for the years 1871-1872, Chinkiang was “gradually recovering from the state of ruin in which they were left by the rebels” and between 1871-1872 started to export rice to the southern provinces by sea, which “was quite a new feature.”

Table II: Estimation of roadblock effect: trade cost and model prediction

| <i>Panel A:</i> Estimation results on transport costs |   |
|---|---|
| Unit cost relative to flat land                       | Punishment of ruggedness by one std.              |
| Flat land (without courier route)                     | 1.00<br>32% increase in transport cost            |
| Courier route   | 0.18<br>Transition cost from land to waterways    |
| Inland waterways                                      | 0.08<br>Equivalent to 82.5 km by inland waterways |
| Sea   | 0.04<br>Equivalent to 165 km by sea               |

| <i>Panel B:</i> Estimation results on trade model parameters      |                 |                  |
|---|-----------------|------------------|
|   | Parameter       | Value            |
| Elasticity of probability of sea trade w.r.t. land/sea cost ratio | $\gamma/\sigma$ | 0.723<br>(0.192) |
| Sunk cost of new sea route (adjusted)                             | $i^s$           | 2.170<br>(0.469) |
| Sunk cost of new land route (adjusted)                            | $i^l$           | 1.202<br>(0.390) |
| Observations  |                 | 88               |

| <i>Panel C:</i> Mean probability of steamship trade after the rebellion |              |       |                  |                   |
|---|--------------|-------|------------------|-------------------|
|   | Observations | Data  | Model prediction |                   |
|   |              |       | With rebellion   | Without rebellion |
|   | (1)          | (2)   | (3)              | (4)               |
| Roadblocked pairs   | 56           | 0.696 | 0.693<br>(0.020) | 0.365<br>(0.020)  |
| Unaffected pairs  | 32           | 0.375 | 0.379<br>(0.045) | 0.379<br>(0.045)  |

*Notes:* In *Panel A*, transition costs are measured in cost travelled by the corresponding waterway in kilometres. In *Panel B*, the estimates imply that the cost ratio of new sea route over existing route is 20 ( $I^s = \exp(i^s \times \sigma/\gamma)$ ). The implied cost ratio of new land route over existing one is 5 ( $I^l = \exp(i^l \times \sigma/\gamma)$ ). Standard errors in parentheses. In *Panel C*, the model prediction is based on parameters estimates of *Panel B*. The prediction with roadblocks in column (3) provides the mean estimated probability for city prefecture pairs for which the least costly land route was roadblocked by the rebellion, taking into account whether the required detour was costly enough to incur the sunk cost of starting to trade by sea between those locations, and whether the cost saving once that sunk cost was incurred was large enough to keep trading by sea after the rebellion. Column (2) provides the actual mean probability observed in the data. Column (4) estimates the same mean probability under a counterfactual where the route had not been roadblocked by the rebellion. Bootstrap standard errors using 500 replicates in parentheses.

( $I^m = \exp(i^m \times \sigma/\gamma)$ ). Compared to sea transport, the investment cost for land transport is modest. The cost ratio of using a new land route over an existing route is about 5.

The parameter estimates in *Panel B* can be used to predict the probability of maritime trade after the rebellion, which is shown in *Panel C*. The observations are divided into two groups:

locality pairs that experienced roadblocks during the rebellion and ones whose regular inland routes remained intact. In the data, shown in column (2), the average probability of sea trade immediately after the rebellion in 1867 was 0.696 for “roadblocked” pairs and 0.375 for unaffected pairs. Column (3) shows the prediction from the model, which matches well the observed modal choices. The average predicted probability of maritime trade is 0.693 for “roadblocked” pairs and 0.379 for “non-roadblocked” pairs. In column (4), I calculate probabilities of maritime trade under a counterfactual scenario in which the rebellion did not happen. As unaffected pairs did not have roadblocks to begin with, their counterfactual probability in column (4) is the same as in column (3). For “roadblocked” pairs, however, their counterfactual probability of maritime trade would be only 0.365 without roadblocks, which is close to the unaffected pairs. This suggests that if it were not for the rebellion, the adoption of sea trade for the two groups would have been similar. Roadblocks substantially increase the use of sea trade after the rebellion, raising the probability by about 90% from 0.365 to 0.693.

***Out-of-sample prediction*** We can use the parameters estimated using the 1867 data to predict subsequent evolution of trade. Notice that the roadblocks took place from 1853 to 1864. Therefore, one period in the model corresponds to approximately 5-10 years. I use the model parameters estimated using the 1867 data and observed trade in 1867 to make an out-of-sample prediction for trade in the subsequent period. Table III shows the results. As before, I divide observations into location pairs that had roadblocks and ones that did not and show their average probability of sea trade in data (in 1872 or 1877) in columns (2) and (3). The model prediction of trade for one period after 1867 is shown in column (4). The out-of-sample prediction is very close to the observed trade in 5 and 10 years. The probability of sea trade is slightly lower than 1867 for “roadblocked” pairs but is higher for “non-roadblocked” pairs. This pattern is borne by both the data and the model. Intuitively, the adoption of sea trade would increase for “non-roadblocked” pairs when new technology of sea transport was introduced. In the presence of shocks and inter-temporal sunk investment, the use of sea trade would increase gradually before it reaches a new equilibrium. For “roadblocked” pairs, however, because of the overwhelming effect of roadblocks in incentivising adoption of sea trade during the rebellion, the probability of sea trade did not rise further and instead suffered a small decline compared to 1867.

## 6. Model extension

In this section, I extend the model developed in Section 5.1 by considering the pre-rebellion prevalence of trade by sail ships and the post-rebellion choice between sail and steam ships. Although we do not observe pre-rebellion bilateral trade by sail, the model structure can be used to estimate the probability of sail trade before the rebellion. The extended model takes into account the possibility that some location pairs had sail trade before the rebellion and therefore may not need to incur the

Table III: Out-of-sample model prediction of sea trade

|                   | Observations | Data        |             | Model prediction |
|-------------------|--------------|-------------|-------------|------------------|
|                   |              | 1872<br>(1) | 1877<br>(2) | (4)              |
| Roadblocked pairs | 56           | 0.679       | 0.643       | 0.687<br>(0.047) |
| Unaffected pairs  | 32           | 0.406       | 0.467       | 0.457<br>(0.069) |

Notes: Columns (2) and (3) provide the mean probability observed in 1872 and 1877. Column (4) shows the out-of-sample model prediction for one period after 1867, based on parameters estimates in Table II. Bootstrap standard errors using 500 replicates in parentheses.

sunk investment of establishing a new sea route when steamship was introduced. It also allows us to break down the relative contribution of roadblocks to the increase in sea trade compared to the technological upgrade from sail to steam ships, which is discussed further in Section 7.

In addition to land transport and sea transport using steamship, it is assumed that firm  $k$  in city  $i$  can also export to  $j$  using sail ships with an iceberg trade cost:

$$(\tau_{ijt}^{sail})^{1-\epsilon} = (D_{ijt}^{sail} e^{\eta_{k,ijt}} I_t^s)^{-\gamma} e^{\eta_{s,ijt}}, \quad (10)$$

where  $D_{ijt}^{sail}$  is the transport cost of sailing ships and  $e^{\eta_{k,ijt}}$  is a shock or preference for sail trade for firm  $k$  following a normal distribution with a mean of zero and standard deviation of  $\sigma_{sail}$ . The trade costs by land and steam are the same as before:

$$(\tau_{ijt}^l)^{1-\epsilon} = (D_{ijt}^l I_t^l)^{-\gamma} e^{\eta_{l,ijt}}, \quad (\tau_{ijt}^{steam})^{1-\epsilon} = (D_{ijt}^{steam} I_t^s)^{-\gamma} e^{\eta_{s,ijt}}$$

It is further assumed that a manager of the firm who has a log utility decides transport mode and sunk investment after observing location specific shocks to land and sea transport ( $e^{\eta_{l,ijt}}$  and  $e^{\eta_{s,ijt}}$ ). Afterwards, firm-specific shocks to sailing ship ( $e^{\eta_{k,ijt}}$ ) are realised. A fraction of firms exporting from location  $i$  to  $j$  will choose sail over steam ship when both technologies of sea transport are available.<sup>38</sup> The transport cost by sea can be summarised as:

$$D_{ijt}^s = \begin{cases} D_{ijt}^{sail}, & \text{before rebellion;} \\ (D_{ijt}^{steam})^\theta (D_{ijt}^{sail})^{1-\theta}, & \text{during and after rebellion;} \end{cases} \quad (11)$$

where  $\theta = \Phi\left(\frac{\log(D_{ijt}^{sail}/D_{ijt}^{steam})}{\sigma_{sail}}\right)$  is the fraction of firms that prefer steam to sail ships given that sea transport is chosen.

<sup>38</sup>Because shocks to land and sea transport are common to trade for a specific location pair, firms exporting for the same location pair will make the same modal choice between sea and inland transport.

Therefore, the probability of sea transport in time  $t$  before the rebellion is:

$$\begin{aligned} \Pr(Sail_{ijt} = 1) &= \Pr(Sail_{ij,t-1} = 1)\Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,pre}^l}{D_{ij}^{sail}}\right) \\ &\quad + (1 - \Pr(Sail_{ij,t-1} = 1))\Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,pre}^l}{D_{ij}^{sail}} - i^s\right). \end{aligned} \quad (12)$$

If we assume that sail trade between any location pair has reached a steady state before the rebellion, the probability of sail trade  $\rho_{ij}$  between location  $i$  and  $j$  before the rebellion can be calculated by equalising  $\Pr(Sail_{ijt} = 1)$  and  $\Pr(Sail_{ij,t-1} = 1)$ . Define the steady state probability of sail trade as  $\rho_{ij}$ .<sup>39</sup> We have:

$$\rho_{ij} = \frac{\Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,pre}^l}{D_{ij}^{sail}} - i^s\right)}{1 + \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,pre}^l}{D_{ij}^{sail}} - i^s\right) - \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,pre}^l}{D_{ij}^{sail}}\right)}. \quad (13)$$

Therefore, we can obtain the post-rebellion probability of sea trade by further taking into account the pre-rebellion tendency to use sailing ships:<sup>40</sup>

$$\begin{aligned} \Pr(S_{ij,post} = 1 | sail) &= \rho \left( \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l\right) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s}\right) \right. \\ &\quad \left. + \left(1 - \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l\right)\right) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s} - i^s\right) \right) \\ &\quad + (1 - \rho) \left( \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l - i^s\right) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s}\right) \right. \\ &\quad \left. + \left(1 - \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,reb}^l}{D_{ij,reb}^s} + \text{roadblock}_{ij} \times i^l - i^s\right)\right) \Phi\left(\frac{\gamma}{\sigma} \log \frac{D_{ij,post}^l}{D_{ij,post}^s} - i^s\right) \right). \end{aligned} \quad (14)$$

As convoluted as Equation 14 may seem, it boils down to the relative transport costs of different transport modes, their change during the rebellion due to roadblocks and parameters of the model, as Equation 9. Therefore, we can estimate the parameters of the model by maximising the likelihood of observing post-rebellion sea trade. An additional input is needed, which is the relative transport

---

<sup>39</sup>Simulations show that it is very likely that the steady state of sail trade has been reached before the rebellion starting from zero trade during the sea ban. Notice that the single realisation of trade does not reach a steady state equilibrium because subject to shocks, at any time, either sea or inland transport can be chosen. However, we can know in probabilistic terms, the steady state of sail ship use for any location pair before the rebellion according to the model.

<sup>40</sup>The assumption underlying the full model is that if in the period before the rebellion, sailing ships are used, steamship trade requires no further investment during the rebellion. In contrast, the base model in Section 5.1 can be viewed as one in which using the novel technology of steamship for location pairs requires separate and new investment from sailing ships.

cost of sail to steam ship to calculate sea transport cost  $D_{ijt}^s$ , therefore I show model estimates and predictions under different transport costs by sail within reasonable ranges.

Notice that Equation 14 assumes that if bilateral sail trade took place before the rebellion, then the location pair does not need to incur the sunk investment for sea transport using steamships during the rebellion. In reality, however, it is possible that in addition to the common investment in sea trade such as learning about routes and logistics, steamships required new investment specific to the new technology. In Table A.2, I present estimation results assuming that only fraction of the investment costs for steamships are deductible if sail trade occurred previously though the results are very similar.<sup>41</sup>

I present the estimation results of the extended model in Table IV. Different columns show estimates under different cost ratios of sail to steam ships. The lower bound of the cost ratio is one, shown in column (1), which indicates that sail ships have the same variable transport cost as steam ships. The upper bound in column (5) is 2.25, which is the cost ratio of courier routes to steam ships. The estimations under intermediate values of the cost ratios 1.25, 1.5, 1.75, and 2 are shown in columns (2), (3), (4), and (5) respectively. My preferred cost ratio of sail to steam ship is 2, which corresponds to the share of grain trade by sail ships (25.6%) before the rebellion according to Deng (2009).<sup>42</sup> The estimated parameters of the extended model are shown in *Panel A*. The estimates are similar to the base model, lying within the 90% confidence interval of the estimates under the base model. The estimated elasticity of trade and transformed sunk cost of new land route are smaller, indicating that both the estimated sunk cost for new sea route and for new land route are larger in the extended model. This is reasonable because the extended model takes into account the possibility that some location pairs would have smaller or no sunk cost in establishing new sea route and their sunk investment costs are averaged with other pairs in the base model. In *Panel B*, I show the prediction of the mean probability of sea trade using the corresponding parameters in *Panel B*. “Roadblocked” pairs are predicted to have a slightly lower probability of sea trade by sail before the rebellion. Therefore, the observed post-rebellion difference in sea trade with steamship technology is unlikely a result of “roadblocked” pairs already having a higher tendency to use sea

---

<sup>41</sup>Table A.2 shows the estimation results assuming that investment costs for steamships can be reduced by 0-95%, but no more than 95% if sail trade occurred in the previous period. When this is the case, trade pairs that had previous sail trade would prefer continuing to use sail ships instead of investing in steamships. When the fraction of new investment in sea transport specific to steamships is very small, however, it would be worthwhile for location pairs that had previous sail trade to invest in steamships, although the estimated parameters and predictions of the model are very similar as shown by Table IV(which is the case that the new investment is 0) in comparison to Table A.2.

<sup>42</sup>Deng (2009) estimated that the grain trade by inland navigation and sea combined was about 80 million piculs (1 picul  $\approx$  60 kilogrammes) while the sea trade consisted of 20.5 million piculs. Using Wang and Huang (1989)'s classification of grain export and import regions during the Qing dynasty, I calculate the probability of sea trade for each grain trade pairs given their transport costs by sail and inland routes. I then use monte-carlo simulation to calculate the average share of trade by sail. The volume of trade is assumed to be proportional to the origin and destination population and inversely proportional to the transport cost (based on the simple but reasonably good approximation of gravity equation to trade, with unavailable GDP approximated by regional population). The share of trade by sail is 25.6% if the cost ratio between sail and steam is 2. For comparison, the share would be 36% and if the cost ratio is 1 and 23.6% if the cost ratio is 2.25.

Table IV: Estimation of roadblock effect: model with sail trade

|   | (1)              | (2)              | (3)              | (4)              | (5)              | (6)              |
|---|------------------|------------------|------------------|------------------|------------------|------------------|
| Sail to steam cost ratio                                    | 1                | 1.25             | 1.5              | 1.75             | 2                | 2.25             |
| <i>Panel A:</i> Estimation on model parameters              |                  |                  |                  |                  |                  |                  |
| $\gamma/\sigma$ (elasticity of trade)                       | 0.452<br>(0.103) | 0.468<br>(0.102) | 0.494<br>(0.108) | 0.514<br>(0.113) | 0.530<br>(0.117) | 0.544<br>(0.122) |
| $i^s$ (sunk cost of new sea route)                          | 1.808<br>(0.409) | 1.955<br>(0.402) | 1.994<br>(0.395) | 2.017<br>(0.391) | 2.027<br>(0.388) | 2.034<br>(0.386) |
| $i^l$ (sunk cost of new land route)                         | 1.676<br>(0.594) | 1.850<br>(0.580) | 1.751<br>(0.539) | 1.670<br>(0.512) | 1.604<br>(0.495) | 1.543<br>(0.483) |
| <i>Panel B:</i> Prediction of mean probability of sea trade |                  |                  |                  |                  |                  |                  |
| <i>Roadblocked pairs:</i>                                   |                  |                  |                  |                  |                  |                  |
| Pre-rebellion   | 0.364<br>(0.020) | 0.318<br>(0.019) | 0.294<br>(0.019) | 0.275<br>(0.018) | 0.257<br>(0.017) | 0.240<br>(0.017) |
| Post-rebellion with rebellion                               | 0.652<br>(0.016) | 0.678<br>(0.016) | 0.681<br>(0.017) | 0.683<br>(0.017) | 0.685<br>(0.017) | 0.687<br>(0.017) |
| Post-rebellion w/o rebellion                                | 0.364<br>(0.020) | 0.351<br>(0.020) | 0.356<br>(0.020) | 0.362<br>(0.020) | 0.367<br>(0.020) | 0.372<br>(0.020) |
| Observations  | 58               | 56               | 56               | 56               | 56               | 56               |
| <i>Non-roadblocked pairs:</i>                               |                  |                  |                  |                  |                  |                  |
| Pre-rebellion   | 0.419<br>(0.047) | 0.362<br>(0.044) | 0.336<br>(0.043) | 0.312<br>(0.042) | 0.290<br>(0.041) | 0.272<br>(0.040) |
| Post-rebellion  | 0.419<br>(0.047) | 0.392<br>(0.045) | 0.390<br>(0.046) | 0.388<br>(0.046) | 0.387<br>(0.046) | 0.386<br>(0.046) |
| Observations  | 30               | 32               | 32               | 32               | 32               | 32               |

Notes: *Panel A* shows model prediction with different sail to steam cost ratios. Standard errors in parentheses. *Panel B* shows the corresponding prediction based on parameter estimates of *Panel A*. Bootstrap standard errors using 500 replicates in parentheses.

trade by sail before the rebellion. The implied “roadblock effect” is also similar to the baseline model. Under the counterfactual scenario in which the rebellion did not happen, the probability of maritime trade would be only 0.367 under the sail to steam cost ratio of 2 (column (5)). This suggests that roadblocks increase the probability of sea trade by 87%, from 0.367 to 0.685.

The extended model, after taking into account the pre-rebellion sail trade, allows to decompose the relative contribution of technological change and roadblocks during the rebellion to increase in sea trade. This naturally depends on the relative cost ratio of sail to steam trade. If the cost of sail to steam ratio is low, steamships are not significantly superior to the incumbent technology, therefore the increase in sea trade due to technological improvement would be small and the increase due to roadblocks would be relatively large. On the other hand, if steamships are much more superior to sail ships, the relative contribution of technological improvement would be more substantial,

although this would be true for both “roadblocked” pairs and “non-roadblocked” pairs pairs and the identification of “roadblock effect” would be still valid exploiting their difference in sea trade. Column (5) shows that under a cost ratio of 2, which matches pre-rebellion sail trade share and is also towards the upper limit of steamship superiority, the implied contribution of roadblock effect to sea trade is still large. The predicted probability of sea trade by sail before the rebellion for “roadblocked” pairs is 0.257 on average and if roadblocks did not happen, the post-rebellion sea trade would be 0.367. This suggests that technological change contributes to only a 0.11 increase in probability of sea trade. In contrast, the majority of the increase (74%) comes from roadblocks, raising the probability of sea trade from 0.367 to 0.685. Similarly, for “non-roadblocked” pairs, their model predicted probability of sea trade increases by 0.097, from 0.29 to 0.387 with the introduction of steamship. A more thorough discussion on the role of technological change and whether it confounds the “roadblock effect” can be found in Section 7.

## 7. Alternative explanations

**Treaty ports and foreign trade** The “roadblock effect” is not confounded by opening up of treaty ports to trade with western countries. The average population growth in coastal prefectures that did not have treaty ports was also substantial: 26% immediately after the rebellion.<sup>43</sup> Regarding the results on trade, I only compare bilateral trade between treaty ports due to data availability. While all treaty ports had access to foreign trade and could adopt the frontier technology of steamships for domestic trade because of technological diffusion, the identification is to compare whether domestic trade between any two of them is more likely to adopt sea transport depending on the heterogeneous increase in their inland transport costs due to the conflict.

Another potential channel that domestic trade patterns could be affected by introduction of treaty ports is through transit trade for re-exporting domestic goods or importing foreign goods. Nevertheless, the rebellion potentially affected transit and domestic trade routes in a simultaneous manner. Before the rebellion, exporting to foreign destinations was mainly through Canton (today’s Guangzhou) and Shanghai. Due to roadblocks, a treaty port could either export the goods directly from itself when regular land trade routes to Canton became blocked, or transit goods by sea to Canton first, similar to decisions for domestic trade route between them. I also show that the magnitudes of the “roadblocked effect” are similar by excluding major ports of foreign trade.

**Technology improvement** One may wonder whether the adoption of sea trade after the rebellion was due to introduction of the new technology of steamship instead of the rebellion. My results show that even when new and superior technology became available, adoption was not guaranteed.

---

<sup>43</sup>In addition, I find that the population increase in treaty ports correlates positively with domestic sea trade value, but not foreign trade value, although this evidence is more suggestive than conclusive as the sample size for which the trade value is available is small.

Holding (new) technology constant, I show that location pairs subject to shocks to their inland route were much more likely to adopt steamship than location pairs unaffected by the rebellion. The counterfactual probability of using steamship without rebellion for “roadblocked” pairs is similar to “non-roadblocked” pairs. These provide evidence that the rebellion catalysed the adoption of the new technology.

While my identification strategy to examine the effect of the rebellion on transport modes exploits the *difference* in sea trade due to exposure to wars, it is true that the increase in the *level* of sea trade during this period can be a combination of both technological progress and roadblocks. However, quantification exercise using the extended model suggests that the contribution of technological improvement to the increase in the level of sea trade immediately after the rebellion would be relatively small. Under a cost ratio of sail to steam of 2 (which corresponds to the level of pre-rebellion grain trade by sail), column 5 of Table IV shows that the counterfactual probability of sea trade for “roadblocked” pairs without the rebellion would be 0.367, an increase of 0.11 compared to their pre-rebellion probability using sailing ships. The majority of increase in sea trade ( $0.318/0.428=74\%$ ) can be attributed to the “roadblock effect”. Similarly, for “non-roadblocked” pairs, their estimated increase in the probability of sea trade with the introduction of steamship after the rebellion is only 0.097.

In fact, even before the introduction of steamship, sea trade by sail had proved to be much more efficient than inland routes, yet was not widely adopted for domestic trade.<sup>44</sup> As discussed in Section 2, due to sea ban in the early Qing dynasty, population relocated away from the coastal region. Even after the ban was lifted, the court decided not to promote sea trade but instead made substantial efforts in maintaining inland trade routes (including the Grand Canal). The large sunk investment needed to start sea trade, even when it was allowed, could discourage individual merchants from maritime trade despite its cost-savings. The model simulation supports the possibility that in the presence of large sunk investment, sail trade will remain at low levels albeit with a lower variable cost than inland routes.

**Yellow River and the Grand Canal** Another potential mechanism that could confound the “roadblock effect” is the Yellow River shifting course in 1855, which flooded and paralysed the northern part of the Grand Canal.<sup>45</sup> While the flooding problem occurred throughout the history, the change of the river course was arguably also a result of the rebellion because the Qing government was unable to finance its maintenance as before with mounting military expenses (Jia, 2009; Dai,

---

<sup>44</sup>The transport of tribute grain largely relied on the Grand Canal before the rebellion. Because the Yellow River constantly flooded the canal, the emperor launched a trial to transport the tribute grain by sea in 1826. According to a contemporary statesman Wei Yuan, sea transport was about one-third as costly and much faster than canal shipping (Wei, 1826). However, the initiative was quickly abandoned due to political opposition.

<sup>45</sup>The Yellow River is the second-longest river in China, after the Yangtze River. The name derives from the yellow sediments discharged into its middle stream from loess deposits.

2012). Nevertheless, I show that the “roadblock effect” is not mainly driven by this alternative impact of the rebellion.

First, I take into account the potential change of the Yellow River in computing the transport costs. In particular, I assume that from 1856 onwards, the abandoned part of the Yellow River and the part of the Grand Canal that could be flooded by the new river (the part in Shantung province) were unusable. To make estimation results comparable, I use the same classification of “roadblocked” pairs as before. Hence, location pairs that could be affected by both roadblocks and the Yellow River are classified as “roadblocked” pairs whereas location pairs affected only by the Yellow River are “non-roadblocked” pairs.<sup>46</sup> Column (4) of Table A.3 shows that the prediction of the model incorporating the increase in inland transport costs due to changes in the Yellow River and the Grand Canal generates very similar results to the baseline prediction in column (1).

Second, I estimate the effect by dropping all observations that could be affected by potential changes in the river and canal, including those that could be affected by both roadblocks and the river/canal. This leaves us with a smaller sample of 62 location pairs. For this sub-sample of “roadblocked” pairs, their regular trade routes would not be affected by potential changes in the Yellow River and the Grand Canal. Still their probability of sea transport after the rebellion is 0.639 in the data, substantially and statistically larger than the “non-roadblocked” pairs. The model prediction based on parameters estimated from the full sample overestimates the post-rebellion probability for this sub-sample of “roadblocked” pairs, but not much (25 out of 36 pairs instead of 23 out of 36). On the other hand, the sub-sample of “non-roadblocked” pairs has similar probability of sea transport after excluding observations that can be affected only by potential changes in the canal and river. Both suggest that the effect of the rebellion through changes in the Yellow River, if any, was limited.

To summarise, the “roadblock effect” is robust to incorporating potential increase in transport costs due to the river/canal change as well as restricting to observations that can be affected solely by the rebellion occupation. This provides evidence that the “roadblock effect” is not mainly driven by the Yellow River shifting course during the rebellion.

**Long-distance trade** Trade route between distant location pairs may have a higher probability of crossing “roadblocked” region, therefore having a higher chance of roadblocks. At the same time, sea trade may be more likely to happen if two regions are further apart. Sea transport has a lower variable cost than inland transport methods, therefore it gains more advantage for long-distance trade, holding constant the availability of inland transport methods. Further, a fixed cost of transition

---

<sup>46</sup>This can also be justified by the observation that for locations pairs subject only to changes in the Yellow River, a small fraction of their trade routes would be affected and the frequently-travelled courier routes were nearby. Therefore, they may not need to incur the cost of finding a new trade route. In fact, the probability of maritime trade for location pairs affected only by the canal and the river was only 0.25 on average after the rebellion. On the other hand, location pairs whose optimal land route was affected by rebellion occupation required a much larger detour to distant areas (e.g. Figure 1).

between land and sea transport will also make sea transport more profitable for longer distance. To the extent that the relative transport costs and their cost of transition are taken into account correctly in the model and that they are main determinants of transport modes, this would not affect the estimated effect of roadblocks. Therefore, instead of using geographical distance or borrowing transport cost parameters from other contexts, this paper develops a new method to estimate the relative transport costs specific to China during this period in Section 5.2.

Nevertheless, one may still be concerned that there could be other reasons why location pairs further apart engage in sea trade, and the estimated “roadblock effect” is capturing them instead. To this end, I focus on a smaller sample which consists only of “roadblocked” pairs whose geographical distance as well as absolute differences in latitude and longitude are similar to “non-roadblocked” pairs. I use propensity score matching to match each “non-roadblocked” pair with a “roadblocked” pair based on log distance and their absolute differences in longitude and latitude.<sup>47</sup> Table A.5 shows that after matching, the two groups are comparable in both geographical distance and differences in longitude and latitude. I also show that the transport costs by sea and land, which are inputs of the model and not used directly for propensity score matching, are also balanced after matching.

Column (6) of Table A.3 shows that for the matched “roadblocked” pairs, their probability of sea trade after the rebellion is 0.75 in data, which is also substantially higher than the 0.375 probability of “non-roadblocked” pairs. The differences in probability in sea trade in data suggests that the roadblock effect is not driven by “roadblocked” observations that are further apart than “non-roadblocked” pairs. The model prediction based on parameters estimated using the full sample also matches well the mean probability of sea trade for this sub-sample of “roadblocked” pairs that are similar to the “non-roadblocked” pairs in terms of distance, differences in longitude and latitude, as well as relative transport costs. This further suggests that conditional on the relative transport costs and their change during the rebellion due to roadblocks, sea trade does not seem to favour more distant location pairs than less distant ones. The counterfactual probability for this subsample of “roadblocked” pairs without the rebellion would be 0.403, which indicates a 82% increase in probability of sea trade.<sup>4849</sup>

---

<sup>47</sup>Because “roadblocked” observations have both longer distance and shorter distance trade pairs while “non-roadblocked” have shorter distance trade pairs, I match each “non-roadblocked” pair with a “roadblocked” pair instead of the other way around.

<sup>48</sup>The reason that the subsample of “roadblocked” pairs that are less distant have even a slightly higher probability of sea trade, both in the data and predicted by the model, is that for long-distance trade, inland trade routes can more often use the more efficient courier routes.

<sup>49</sup>Column (6) of Table A.4 shows the prediction of the extended model incorporating the choice of sail ships. The estimated roadblock effect is similarly large for the sub-sample of “roadblocked” pairs matched to the “non-roadblocked” pairs in distance and absolute differences in longitude and latitude.

## 8. Concluding remarks

Concentration of population in a location can remain persistent even when the location is no longer advantageous due to change in technology. The introduction of new transport technology, for example steamships, may increase the relative attractiveness of coastal locations by reducing their trade costs to other locations and thus increasing their market access. However, the presence of the large sunk investment usually needed for new technology could hold it back. Failure to adopt the superior technology, however, can trap populations in sub-optimal locations for a long time.

This paper uses the Taiping Rebellion in 19th-century China to study population changes and technological adoption after a large but temporary shock to city sizes. In contrast to the large literature on the persistence of relative city sizes after war shocks, I find a permanent loss in population of about 55% in war-ravaged cities and an increase of 32% in coastal cities after the rebellion. I provide evidence that this is because, in addition to the direct impact of war through death and displacement, the rebellion affected populations indirectly through trade routes and transport technologies. Before the rebellion, the limited use of sea transport despite its large potential was due to persistence of population away from the coastal locations and lack of state investment. Even when superior technology became available, adoption of sea transport may not take place. I show that the spatial variations in maritime trade immediately after the rebellion was driven by blocked land routes and feasibility of sea alternatives. The rebellion blocked regular inland trade routes, forcing some location pairs to search for alternatives, and this triggered substantial investment to facilitate steamship trade. After the rebellion, with sunk investment already incurred, many trade routes permanently moved to sea and this catalysed a shift of population towards port cities.

I develop a simple trade model with transport modal and route choice and a novel method to estimate historical transport costs to identify this “roadblock effect.” The model estimates find a substantial cost associated with starting maritime trade. The counterfactual analysis using the model suggests that the probability of maritime trade after the rebellion would be much lower without roadblocks. The long-run evolution of trade also suggests that the initial investment in establishing sea routes could reinforce itself by facilitating further infrastructure investment in sea trade and entailing large sunk investment to return to inland routes once they become obsolete, contributing to the ultimate take-off of maritime trade. The results of the paper highlight the substantial forces of persistence in population and trade routes, the role of sunk investment in holding back superior technology, the importance of incentives for technology adoption, and population changes through trade costs and technology.

## References

- Allen, Treb and Costas Arkolakis**, “Trade and the topography of the spatial economy,” *The Quarterly Journal of Economics*, 2014, 129 (3), 1085–1140.

- and —, “Modern spatial economics: a primer,” in “World Trade Evolution,” Routledge, 2018, pp. 435–472.
- and **Dave Donaldson**, “The geography of path dependence,” *Working paper*, 2018.
- Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang**, “Roads, railroads, and decentralization of Chinese cities,” *Review of Economics and Statistics*, 2017, 99 (3), 435–448.
- Bickers, Robert**, “Infrastructural globalization: lighting the China coast, 1860s–1930s,” *The Historical Journal*, 2013, 56 (2), 431–458.
- Bleakley, Hoyt and Jeffrey Lin**, “Portage and path dependence,” *The Quarterly Journal of Economics*, 2012, 127 (2), 587–644.
- Brezs, Elise S and Paul R Krugman**, “Technology and the life cycle of cities,” *Journal of Economic Growth*, 1997, 2 (4), 369–383.
- Cao, Shuji**, *Zhongguo yi min shi [The migration history of China]*, Fujian Education Press, 1997.
- , *Zhongguo ren kou shi [The demographic history of China]*, Fudan University Press, 2000.
- and **Yushang Li**, “The effect of Taiping Rebellion on population in Zhejiang,” *Fudan Journal (Social Sciences Edition)*, 2000, 5, 5.
- Cheng, Yukuei**, *Foreign trade and industrial development of China: An historical and integrated analysis through 1948*, University Press of Washington, 1956.
- chien Wang, Yeh and Guoshu Huang**, “Shiba Shiji Zhongguo Liangshi Gongxu de Kaocha [An Investigation into the China’s Supply-Demand of Grain in the Eighteenth Century],” *Collection on Rural Economic History of Modern China*, 1989.
- Dai, Angang**, “The rise and fall of the transport of grain by water to Beijing in Qing dynasty and sailor, boat tracker,” *Historical Research in Anhui*, 2012, (6), 24–28.
- Davis, Donald R and David E Weinstein**, “Bones, bombs, and break points: the geography of economic activity,” *American Economic Review*, 2002, 92 (5), 1269–1289.
- Deng, Yibing**, *Qingdai qianqi shangpin liutong yanjiu [A study of commodity circulation in the early Qing dynasty]*, Tianjin Classic Publishing House, 2009.
- Dincecco, Mark and Massimiliano Gaetano Onorato**, “Military conflict and the rise of urban Europe,” *Journal of Economic Growth*, 2016, 21 (3), 259–282.
- Donaldson, Dave**, “Railroads of the Raj: Estimating the impact of transportation infrastructure,” *American Economic Review*, 2018, 108 (4-5), 899–934.
- Faber, Benjamin**, “Trade integration, market size, and industrialization: evidence from China’s National Trunk Highway System,” *Review of Economic Studies*, 2014, 81 (3), 1046–1070.
- Fan, Baichuan**, *Zhongguo lunchuan hangyunye de xingqi [The rise of China’s steamship industry]*, Sichuan People’s Press, 1985.

**Ge, J, Y Hou, and G Zhang,** *Population and the modernization of China since 1850*, Xuelin Press, 1999.

**Gilfillan, S Colum,** *Inventing the ship: a study of the inventions made in her history between floating log and rotorship; a self-contained but companion volume to the author's "Sociology of invention"; with 80 illustrations, bibliographies, notes and index*, Follett, 1935.

**Glaeser, Edward L and Jesse M Shapiro,** “Cities and warfare: The impact of terrorism on urban form,” *Journal of Urban Economics*, 2002, 51 (2), 205–224.

**Gu, Cheng,** “Qingchu de qianhai [The sea ban in the early Qing dynasty],” *Journal of Beijing Normal University (Social Sciences)*, 1983, (3), 60–72.

**Guo, Yisheng,** *The atlas of the Taiping Heavenly Kingdom*, SinoMaps Press, 1989.

**Hanlon, W. Walker,** “Temporary Shocks and Persistent Effects in Urban Economies: Evidence from British Cities after the U.S. Civil War,” *The Review of Economics and Statistics*, 03 2017, 99 (1), 67–79.

**Harley, Charles K,** “The shift from sailing ships to steamships, 1850–1890: a study in technological change and its diffusion,” in “Essays on a Mature Economy,” Princeton University Press, 2015, pp. 215–238.

**Heblich, Stephan, Stephen J Redding, and Daniel M Sturm,** “The making of the modern metropolis: evidence from London,” *The Quarterly Journal of Economics*, 2020, 135 (4), 2059–2133.

**Helpman, Elhanan, Marc Melitz, and Yona Rubinstein,** “Estimating trade flows: Trading partners and trading volumes,” *The Quarterly Journal of Economics*, 2008, 123 (2), 441–487.

**Hua, Qiang,** *The geography of the Taiping Rebellion*, Guangxi People Press, 1991.

**Huang, Qichen,** “Qingdai qianqi haiwai maoyi de fazhan [Development of overseas trade in the early Qing dynasty],” *Historical Research*, 1986, (4), 151–170.

**Jia, Guojing,** “Dispute of Yellow River’s new and old course after the bursting at Tongwxianang,” *Journal of Historical Science*, 2009, (12), 56–63.

**Jia, Ruixue,** “The legacies of forced freedom: China’s treaty ports,” *Review of Economics and Statistics*, 2014, 96 (4), 596–608.

**Juhász, Réka,** “Temporary protection and technology adoption: Evidence from the Napoleonic blockade,” *American Economic Review*, November 2018, 108 (11), 3339–76.

**Keller, Wolfgang, Ben Li, and Carol H Shiue,** “The evolution of domestic trade flows when foreign trade is liberalized: Evidence from the Chinese Maritime Customs Service,” *Institutions and Comparative Economic Development*, 2012, pp. 152–172.

**Krugman, Paul,** “First nature, second nature, and metropolitan location,” *Journal of Regional Science*, 1993, 33 (2), 129–144.

- Kuhn, Philip A.**, *The Taiping Rebellion*, Vol. 10 of *The Cambridge History of China*, Cambridge University Press, 1978.
- Liao, Shengfeng**, *Studies on native custom and regional economics in the Qing dynasty*, People's Press, 2010.
- Liu, Zubi**, "Qingdai Fujian Wuyi cha chukou shangdao chutan [A preliminary study on the evolution of Fujian Wuyi tea exporters' road in the Qing dynasty]," *Fujian Historical Records*, 2016, (1), 33–36.
- Meyer-Fong, Tobie**, *What remains: Coming to terms with civil war in 19th century China*, Stanford University Press, 2013.
- Michaels, Guy and Ferdinand Rauch**, "Resetting the urban network: 117–2012," *The Economic Journal*, 2017, 128 (608), 378–412.
- Miguel, Edward and Gerard Roland**, "The long-run impact of bombing Vietnam," *Journal of Development Economics*, 2011, 96 (1), 1–15.
- Ni, Yuping**, *Qingdai caoliang haiyun yu shehui bianqian [Transport of tribute grain by sea and social changes in the Qing dynasty]*, Shanghai People's Press, 2005.
- Nunn, Nathan and Diego Puga**, "Ruggedness: The blessing of bad geography in Africa," *Review of Economics and Statistics*, 2012, 94 (1), 20–36.
- Pascali, Luigi**, "The wind of change: Maritime technology, trade, and economic development," *American Economic Review*, 2017, 107 (9), 2821–54.
- Redding, Stephen and Anthony J. Venables**, "Economic geography and international inequality," *Journal of International Economics*, 2004, 62 (1), 53–82.
- Redding, Stephen J, Daniel M Sturm, and Nikolaus Wolf**, "History and industry location: evidence from German airports," *Review of Economics and Statistics*, 2011, 93 (3), 814–831.
- Tintelnot, Felix**, "Global production with export platforms," *The Quarterly Journal of Economics*, 2017, 132 (1), 157–209.
- Tsivanidis, Nick**, "The aggregate and distributional effects of urban transit infrastructure: Evidence from Bogotá's TransMilenio," *Working paper*, 2018.
- Voigtländer, Nico and Hans-Joachim Voth**, "The three horsemen of riches: Plague, war, and urbanization in early modern Europe," *Review of Economic Studies*, 2012, 80 (2), 774–811.
- Ward, WH**, "The sailing ship effect," *Physics Bulletin*, 1967, 18 (6), 169.
- Wei, Yuan**, *Daoguang bingyu haiyun ji [Memo of sea shipping in 1826 1826]*.
- Xia, Mingfang**, "The Qing government's management of the Yellow River after the Tongwaxiang burst," *The Qing History Journal*, 1995, (4), 40–51.
- Zhu, Yingui**, "The Decline of wooden ship industry and the rise of steamship industry in late Qing dynasty," *Historiography Research in Anhui*, 2014, (6), 5–17.

## AppendixA. Table appendix

Table A.1: Mapping between postal names and current locations

| Postal name<br>(Romanisation) | Current location<br>(Pinyin) | Postal name<br>(Romanisation) | Current location<br>(Pinyin) | Postal name<br>(Romanisation) | Current location<br>(Pinyin) |
|-------------------------------|------------------------------|-------------------------------|------------------------------|-------------------------------|------------------------------|
| Amoy                          | Xiamen                       | Heilungchiang                 | Heilongjiang                 | Nanking                       | Nanjing                      |
| Anhwei                        | Anhui                        | Hopeh                         | Hebei                        | Newchwang                     | Yingkou                      |
| Canton                        | Guangzhou                    | Kansu                         | Gansu                        | Ningpo                        | Ningbo                       |
| Chefoo                        | Yantai                       | Kiangsi                       | Jiangsi                      | Peking                        | Beijing                      |
| Chekiang                      | Zhejiang                     | Kiangsu                       | Jiangsu                      | Shantung                      | Shandong                     |
| Chengting                     | Shjiazhuang                  | Kiangsu                       | Jiangsu                      | Shengching                    | Liaoning                     |
| Chinan                        | Jinan                        | Kirin                         | Jilin                        | Shensi                        | Shaanxi                      |
| Chinchow                      | Jinzhou                      | Kiukiang                      | Jiujiang                     | Swatow                        | Shantou                      |
| Chinkiang                     | Zhenjiang                    | Kuachow                       | Guazhou                      | Szechuen                      | Sichuan                      |
| Foochow                       | Fuzhou                       | Kuanghsia                     | Guangxi                      | Tientsin                      | Tianjin                      |
| Fukien                        | Fujian                       | Kueichow                      | Guizhou                      | Wenchow                       | Wenzhou                      |
| Hankow                        | Hankou                       | Kweiping                      | Guiping                      | Yangchow                      | Yangzhou                     |

Table A.2: Estimation of roadblock effect: robustness results of extended model

|   | (1)              | (2)              | (3)              | (4)              | (5)              | (6)              |
|---|------------------|------------------|------------------|------------------|------------------|------------------|
| Sail to steam cost ratio                                    | 1                | 1.25             | 1.5              | 1.75             | 2                | 2.25             |
| <i>Panel A:</i> Estimation on model parameters              |                  |                  |                  |                  |                  |                  |
| $\gamma/\sigma$ (elasticity of trade)                       | 0.438<br>(0.097) | 0.482<br>(0.105) | 0.523<br>(0.112) | 0.559<br>(0.120) | 0.590<br>(0.127) | 0.619<br>(0.133) |
| $i^s$ (sunk cost of new sea route)                          | 1.845<br>(0.424) | 1.952<br>(0.406) | 1.998<br>(0.402) | 2.032<br>(0.401) | 2.057<br>(0.399) | 2.079<br>(0.398) |
| $i^l$ (sunk cost of new land route)                         | 1.707<br>(0.617) | 1.823<br>(0.586) | 1.699<br>(0.549) | 1.598<br>(0.525) | 1.515<br>(0.508) | 1.437<br>(0.497) |
| <i>Panel B:</i> Prediction of mean probability of sea trade |                  |                  |                  |                  |                  |                  |
| <i>Roadblocked pairs:</i>                                   |                  |                  |                  |                  |                  |                  |
| Pre-rebellion   | 0.364<br>(0.020) | 0.332<br>(0.020) | 0.319<br>(0.021) | 0.306<br>(0.021) | 0.292<br>(0.021) | 0.278<br>(0.020) |
| Post-rebellion with rebellion                               | 0.652<br>(0.016) | 0.676<br>(0.016) | 0.678<br>(0.017) | 0.680<br>(0.017) | 0.682<br>(0.017) | 0.684<br>(0.017) |
| Post-rebellion w/o rebellion                                | 0.364<br>(0.020) | 0.351<br>(0.020) | 0.357<br>(0.021) | 0.363<br>(0.021) | 0.370<br>(0.021) | 0.376<br>(0.021) |
| Observations  | 58               | 56               | 56               | 56               | 56               | 56               |
| <i>Non-roadblocked pairs:</i>                               |                  |                  |                  |                  |                  |                  |
| Pre-rebellion   | 0.419<br>(0.047) | 0.377<br>(0.046) | 0.360<br>(0.047) | 0.344<br>(0.047) | 0.327<br>(0.047) | 0.312<br>(0.047) |
| Post-rebellion  | 0.419<br>(0.047) | 0.392<br>(0.046) | 0.391<br>(0.047) | 0.390<br>(0.047) | 0.389<br>(0.047) | 0.389<br>(0.047) |
| Observations  | 30               | 32               | 32               | 32               | 32               | 32               |

*Notes:* The robustness estimation assumes that location pairs with previous sail trade would stick to the incumbent technology even when steamships became available whereas Table IV assumes that steamships were adopted. *Panel A* shows model prediction with different sail to steam cost ratios. Standard errors in parentheses. *Panel B* shows the corresponding prediction based on parameter estimates of *Panel A*. Bootstrap standard errors using 500 replicates in parentheses.

Table A.3: Prediction of roadblock effect: robustness results of base model

|                               | (1)<br>Baseline<br>sample | (2)<br>Excl.<br>Tientsin. | (3)<br>Excl.<br>Canton | (4)<br>Change<br>river | (5)<br>Excl.<br>river | (6)<br>Matched<br>sample |
|-------------------------------|---------------------------|---------------------------|------------------------|------------------------|-----------------------|--------------------------|
| <i>Roadblocked pairs:</i>     |                           |                           |                        |                        |                       |                          |
| Data                          | 0.696                     | 0.705                     | 0.674                  | 0.696                  | 0.639                 | 0.750                    |
| Prediction with roadblock     | 0.693<br>(0.020)          | 0.710<br>(0.022)          | 0.697<br>(0.024)       | 0.698<br>(0.020)       | 0.699<br>(0.026)      | 0.733<br>(0.026)         |
| Prediction w/o roadblock      | 0.365<br>(0.020)          | 0.383<br>(0.022)          | 0.366<br>(0.023)       | 0.370<br>(0.020)       | 0.382<br>(0.030)      | 0.403<br>(0.027)         |
| Observations                  | 56                        | 44                        | 46                     | 56                     | 36                    | 32                       |
| <i>Non-roadblocked pairs:</i> |                           |                           |                        |                        |                       |                          |
| Data                          | 0.375                     | 0.346                     | 0.417                  | 0.375                  | 0.385                 | 0.375                    |
| Prediction                    | 0.379<br>(0.045)          | 0.374<br>(0.048)          | 0.395<br>(0.051)       | 0.380<br>(0.045)       | 0.392<br>(0.043)      | 0.378<br>(0.045)         |
| Observations                  | 32                        | 26                        | 24                     | 32                     | 26                    | 32                       |

*Notes:* Estimation of parameters of the model in Section 5 with different robustness specifications. Column (1) shows the baseline estimation as in Table II for comparison. Column (2) excludes trade with Tientsin, whose trade may be directly affected by the civil war. Column (3) excludes major ports for foreign trade (Canton in addition to Shanghai). Column (4) takes into account the potential change in the Yellow River and the Grand Canal during the war when calculating transport costs. Column (5) excludes trade pairs that may be affected by changes in the Yellow River and the canal. Column (6) uses a subsample matched in geographical distance and differences in latitude and longitude. Bootstrap standard errors using 500 replicates in parentheses.

Table A.4: Prediction of roadblock effect: robustness results of extended model

|                               | (1)<br>Baseline<br>sample | (2)<br>Excl.<br>Tientsin. | (3)<br>Excl.<br>Canton | (4)<br>Change<br>river | (5)<br>Excl.<br>river | (6)<br>Matched<br>sample |
|-------------------------------|---------------------------|---------------------------|------------------------|------------------------|-----------------------|--------------------------|
| <i>Roadblocked pairs:</i>     |                           |                           |                        |                        |                       |                          |
| Data                          | 0.696                     | 0.705                     | 0.674                  | 0.696                  | 0.639                 | 0.750                    |
| Prediction with roadblock     | 0.685                     | 0.700                     | 0.688                  | 0.689                  | 0.695                 | 0.718                    |
|                               | (0.017)                   | (0.019)                   | (0.020)                | (0.017)                | (0.023)               | (0.022)                  |
| Prediction w/o roadblock      | 0.367                     | 0.386                     | 0.369                  | 0.371                  | 0.386                 | 0.408                    |
|                               | (0.020)                   | (0.024)                   | (0.023)                | (0.020)                | (0.031)               | (0.027)                  |
| Observations                  | 56                        | 44                        | 46                     | 56                     | 36                    | 32                       |
| <i>Non-roadblocked pairs:</i> |                           |                           |                        |                        |                       |                          |
| Data                          | 0.375                     | 0.346                     | 0.417                  | 0.375                  | 0.385                 | 0.375                    |
| Prediction                    | 0.387                     | 0.382                     | 0.403                  | 0.388                  | 0.399                 | 0.387                    |
|                               | (0.046)                   | (0.049)                   | (0.053)                | (0.046)                | (0.044)               | (0.047)                  |
| Observations                  | 32                        | 26                        | 24                     | 32                     | 26                    | 32                       |

*Notes:* Estimation of parameters of the model in Section 6 with different robustness specifications. Column (1) shows the baseline estimation as in Table II for comparison. Column (2) excludes trade with Tientsin, whose trade may be directly affected by the civil war. Column (3) excludes major ports for foreign trade (Canton in addition to Shanghai). Column (4) takes into account the potential change in the Yellow River and the Grand Canal during the war when calculating transport costs. Column (5) excludes trade pairs that may be affected by changes in the Yellow River and the canal. Column (6) uses a subsample matched in geographical distance and differences in latitude and longitude. Bootstrap standard errors using 500 replicates in parentheses.

Table A.5: Summary statistics of distance

|              | Full sample      |                  |                  |                  |                  | Matched sample   |                  |                  |                   |                  |
|--------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|-------------------|------------------|
|              | (1)<br>Long.     | (2)<br>Lat.      | (3)<br>Dist.     | (4)<br>Sea       | (5)<br>Land      | (6)<br>Long.     | (7)<br>Lat.      | (8)<br>Dist.     | (9)<br>Sea        | (10)<br>Land     |
| Roadblock    | 1.358<br>(0.462) | 5.419<br>(0.871) | 0.756<br>(0.110) | 0.442<br>(0.182) | 0.386<br>(0.089) | 0.458<br>(0.432) | 0.392<br>(0.856) | 0.152<br>(0.115) | -0.075<br>(0.201) | 0.043<br>(0.099) |
| Constant     | 2.925<br>(0.323) | 4.419<br>(0.537) | 6.273<br>(0.095) | 1.065<br>(0.174) | 3.071<br>(0.071) | 2.925<br>(0.325) | 4.419<br>(0.540) | 6.273<br>(0.095) | 1.065<br>(0.175)  | 3.071<br>(0.071) |
| Observations | 88               | 88               | 88               | 88               | 88               | 64               | 64               | 64               | 64                | 64               |
| R-squared    | 0.079            | 0.257            | 0.381            | 0.093            | 0.181            | 0.018            | 0.003            | 0.027            | 0.002             | 0.003            |

*Notes:* The constants show for the “non-roadblocked” location pairs, the average difference in longitude and latitude, distance, sea transport and land transport costs between trade pairs. The indicator *roadblock* shows the difference for “roadblocked” pairs relative to the “non-roadblocked” pairs. Robust standard errors in parentheses.

## AppendixB. Figure appendix

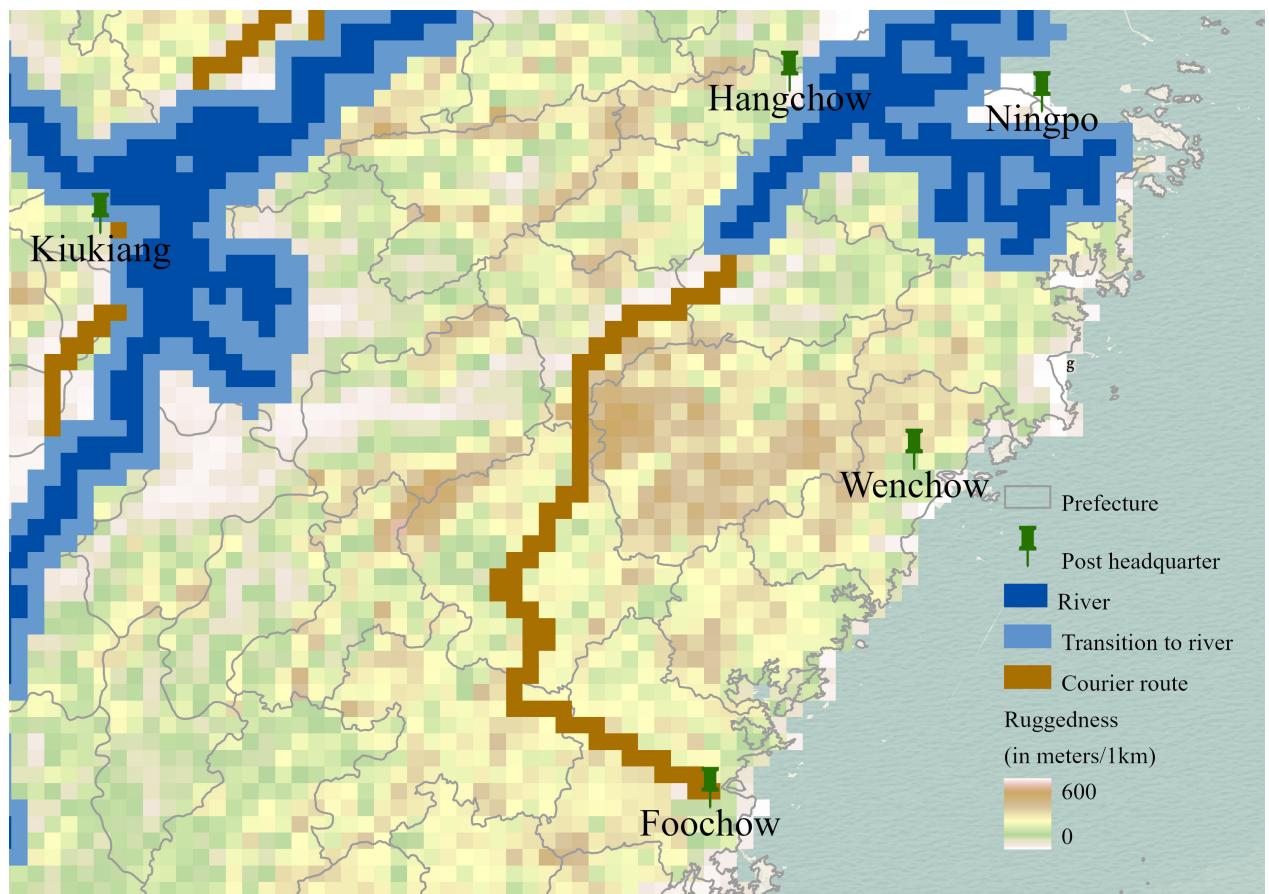


Figure B.1: Transport cost in 19th-century China

Notes: 0.1 degree by 0.1 degree cells covering the land area of China overlaid with rivers, courier routes, terrain ruggedness, transition between land and inland waterways.

## AppendixC. Additional results on population

In this section, I provide additional results on population evolution. First, I show that population increase in coastal prefectures after the rebellion was robust to exclusion of treaty ports. Second, I show that the sharp increase of coastal population occurred immediately after the rebellion, not one period before or after.

Column (1) of Table C.1 replicates the results of column (2) of Table I, which shows an increase of population in coastal prefectures by 28.6% on average after the rebellion. In column (2) of Table C.1, I exclude coastal prefectures that had treaty ports. The result shows an average 26% increase in population in coastal prefectures that did not have any treaty port. In column (3), I further include an interaction term between the post-rebellion dummy and an indicator for whether the prefecture was next to a treaty port but was not coastal. The coefficient estimates show that population increase in inland prefectures adjacent to treaty ports was more modest, 14%, than the coastal prefectures that were not treaty ports, 28% in comparison. This evidence was consistent with the story that coastal locations gained prominence after the rebellion because of improvement in market access when adopting sea trade instead of mere spillovers from the treaty ports. Column (4) replicates the results of column (3) of Table I by including additional controls interacted with the post-rebellion dummy. The estimation results with controls from columns (5) and (6) paint a similar picture: the increase in population of coastal prefectures that were not treaty ports was also substantial after the rebellion and larger than inland prefectures that were next to treaty ports.

Next I conduct an alternative estimation specification for the population evolution between 1680 and 2010 to show that the relocation of population in coastal locations was immediately after the rebellion, and not one period before or after. In particular, instead of having a post-rebellion dummy and its interactions, I include indicators for each year and their interactions with rebellion occupation and coastal locations. Similarly as before, prefecture fixed effects and provincial trends are included. The robust standard errors are clustered at the prefecture level to adjust for heteroskedasticity and within-prefecture correlation over time. The indicator for the initial year 1680 and its interactions were dropped with inclusion of prefecture fixed effects. Therefore, the interpretation of the coefficients of year dummies is the increase in population relative to the 1680 level and the coefficients of interactions is the change relative to the 1680 distribution.

The estimation results are shown in Table C.2. Row (1) presents coefficient estimates for the year dummies, which show population increase in respective years relative to the 1680 level. Row (2) presents estimates for the interactions between year dummies and indicators for whether the prefecture was occupied during the rebellion. Before the rebellion, the positive coefficients indicate that the prefectures that were later occupied by the rebellion had had higher population growth than the non-occupied prefectures relative to their initial distribution in 1680. The occupied prefectures by the rebels situated in the economic heartland of China then, the affluent Yangtze region, which had an upward population growth trend before the rebellion. However, as column (4) of row (2)

shows, occupied prefectures suffered a much larger population loss during the rebellion than the other prefectures: a 43% loss compared to the 1680 distribution and a 64% loss compared to the pre-rebellion 1851 distribution.

Regarding population in coastal locations, the results in row (3) show that its growth was sluggish before the rebellion despite that in 1683 the Qing government lifted sea ban and allowed coastal population previously evacuated to inland areas to return. Coastal population increased by 2.6% in 1776, and by an additional 1.9% in 1820 and another 2.9% in 1851. In 1820, the coastal prefectures had 7.4% more population than the non-coastal prefectures relative to their 1680 distribution although the difference is not statistically significant at the 10% level. However, immediately after the rebellion, population in coastal prefectures saw a sizeable increase of 33% compared to the 1620 distribution and 25% compared to 1851. The coefficients for later years show that although the population loss in occupied prefectures and gain in coastal prefectures got attenuated over time, the effect was persistent for more than a hundred years.<sup>50</sup>

Table C.2: Population evolution relative to the year 1680

|                     | Pre-rebellion period |                     |                     |             | Post-rebellion period |                      |                     |                     |                     |
|---------------------|----------------------|---------------------|---------------------|-------------|-----------------------|----------------------|---------------------|---------------------|---------------------|
|                     | 1776<br>(1)          | 1820<br>(2)         | 1851<br>(3)         | 1880<br>(4) | 1910<br>(5)           | 1953<br>(6)          | 1982<br>(7)         | 2010<br>(8)         |                     |
| (1) Year            | 0.610***<br>(0.047)  | 0.855***<br>(0.070) | 1.005***<br>(0.088) |             | 0.905***<br>(0.103)   | 1.129***<br>(0.116)  | 1.404***<br>(0.138) | 2.018***<br>(0.160) | 2.263***<br>(0.169) |
| (2) Year × occupied | 0.099**<br>(0.045)   | 0.163**<br>(0.063)  | 0.210***<br>(0.076) |             | -0.434***<br>(0.116)  | -0.373***<br>(0.113) | -0.297**<br>(0.130) | -0.289**<br>(0.134) | -0.225<br>(0.144)   |
| (3) Year × coastal  | 0.026<br>(0.041)     | 0.045<br>(0.053)    | 0.074<br>(0.062)    |             | 0.327***<br>(0.073)   | 0.260***<br>(0.076)  | 0.263***<br>(0.089) | 0.238***<br>(0.087) | 0.398***<br>(0.098) |

*Notes:* The table presents the coefficient estimates from one panel data regression. The dependent variable is the natural logarithm of population of one prefecture in one data year (1680, 1776, 1820, 1880, 1910, 1953, 1982 or 2010). Prefecture fixed effects and provincial linear time trends are included but not reported. Robust standard errors clustered by prefecture are reported in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels.

<sup>50</sup>Coastal population saw an enlarging 39.8% gap over the non-coastal population in 2010. However, this was almost entirely driven by treaty ports (see interactions between the year 2010 and coastal dummies for treaty ports in columns (4) and (5) of Table I and for non-treaty ports in columns (2) and (3) of Table C.1). Jia (2014) discusses potential channels for the population increase in treaty ports after 1982.

Table C.1: Prefecture population evolution 1680–2010: exclude treaty ports

|                                   | Dep. var.: Log (population) |                      |                      |                      |                      |                      |
|-----------------------------------|-----------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                                   | (1)                         | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
| Post-rebellion                    | -0.323***<br>(0.037)        | -0.327***<br>(0.038) | -0.345***<br>(0.043) | -0.104<br>(0.174)    | -0.150<br>(0.180)    | -0.206<br>(0.179)    |
| Post-reb. × occupied              | -0.575***<br>(0.091)        | -0.529***<br>(0.106) | -0.521***<br>(0.106) | -0.546***<br>(0.102) | -0.501***<br>(0.117) | -0.489***<br>(0.116) |
| Post-reb. × coastal               | 0.286***<br>(0.052)         | 0.257***<br>(0.068)  | 0.280***<br>(0.069)  | 0.315***<br>(0.057)  | 0.262***<br>(0.074)  | 0.292***<br>(0.075)  |
| Post-reb. × next to port          |                             |                      | 0.138*<br>(0.082)    |                      |                      | 0.185**<br>(0.091)   |
| Post-reb. × initial human capital |                             |                      |                      | -0.044**<br>(0.022)  | -0.046*<br>(0.025)   | -0.044*<br>(0.025)   |
| 1910                              | 0.038*<br>(0.019)           | 0.020**<br>(0.010)   | 0.027**<br>(0.011)   | 0.020**<br>(0.009)   | 0.020**<br>(0.010)   | 0.028***<br>(0.011)  |
| 1953                              | 0.046*<br>(0.027)           | 0.035<br>(0.023)     | 0.049*<br>(0.025)    | 0.028<br>(0.022)     | 0.035<br>(0.023)     | 0.051**<br>(0.025)   |
| 1982                              | 0.481***<br>(0.035)         | 0.451***<br>(0.029)  | 0.444***<br>(0.032)  | 0.445***<br>(0.028)  | 0.448***<br>(0.029)  | 0.444***<br>(0.032)  |
| 2010                              | 0.551***<br>(0.039)         | 0.527***<br>(0.034)  | 0.531***<br>(0.038)  | 0.523***<br>(0.033)  | 0.528***<br>(0.035)  | 0.535***<br>(0.037)  |
| 1910 × occupied                   | 0.055<br>(0.042)            | 0.076<br>(0.049)     | 0.071<br>(0.048)     | 0.068*<br>(0.040)    | 0.075<br>(0.049)     | 0.069<br>(0.048)     |
| 1953 × occupied                   | 0.121*<br>(0.066)           | 0.098<br>(0.072)     | 0.087<br>(0.072)     | 0.131**<br>(0.065)   | 0.096<br>(0.072)     | 0.084<br>(0.072)     |
| 1982 × occupied                   | 0.122*<br>(0.070)           | 0.143*<br>(0.080)    | 0.144*<br>(0.080)    | 0.146**<br>(0.066)   | 0.143*<br>(0.079)    | 0.141*<br>(0.079)    |
| 2010 × occupied                   | 0.179**<br>(0.087)          | 0.137<br>(0.095)     | 0.131<br>(0.095)     | 0.194**<br>(0.083)   | 0.132<br>(0.093)     | 0.123<br>(0.093)     |
| 1910 × coastal                    | -0.068***<br>(0.022)        | -0.055**<br>(0.022)  | -0.060***<br>(0.023) | -0.052***<br>(0.018) | -0.052**<br>(0.023)  | -0.057**<br>(0.023)  |
| 1953 × coastal                    | -0.068<br>(0.042)           | -0.082*<br>(0.043)   | -0.089**<br>(0.045)  | -0.051<br>(0.043)    | -0.074<br>(0.044)    | -0.083*<br>(0.046)   |
| 1982 × coastal                    | -0.095*<br>(0.049)          | -0.102*<br>(0.059)   | -0.090<br>(0.061)    | -0.061<br>(0.044)    | -0.089<br>(0.057)    | -0.079<br>(0.058)    |
| 2010 × coastal                    | 0.065<br>(0.073)            | 0.007<br>(0.088)     | 0.012<br>(0.091)     | 0.093<br>(0.070)     | 0.020<br>(0.083)     | 0.022<br>(0.085)     |
| 1910 × near port                  |                             |                      | -0.059***<br>(0.017) |                      |                      | -0.070***<br>(0.017) |
| 1953 × near port                  |                             |                      | -0.107***<br>(0.039) |                      |                      | -0.132***<br>(0.040) |
| 1982 × near port                  |                             |                      | 0.046<br>(0.064)     |                      |                      | 0.031<br>(0.063)     |
| 2010 × near port                  |                             |                      | -0.034<br>(0.082)    |                      |                      | -0.063<br>(0.083)    |
| Post-reb. × geographic controls   |                             |                      |                      | Yes                  | Yes                  | Yes                  |
| Prefecture fixed effects          | Yes                         | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| Provincial time trends            | Yes                         | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| Number of observations            | 1,610                       | 1,368                | 1,368                | 1,565                | 1,350                | 1,350                |
| Number of prefectures             | 179                         | 152                  | 152                  | 174                  | 150                  | 150                  |
| R-squared                         | 0.839                       | 0.855                | 0.856                | 0.867                | 0.859                | 0.860                |

Notes: Columns (1) and (4) replicate the results of columns (2) and (3) of Table I, which include all coastal prefectures. Columns (2), (3), (5) and (6) only include coastal prefectures without treaty ports. In columns (3) and (5), initial human capital is measured by the natural logarithm number of graduates from the Imperial Exam in the 1840s. The geographic controls are an indicator for whether the Grand Canal passed through, terrain ruggedness, and categorical variables for agricultural suitability for major crops in historical China. Robust standard errors clustered by prefecture are reported in parentheses. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels.

## AppendixD. Estimation of historical migration gravity

The following equation is used for the estimation of migration gravity:

$$\log L_{ijt} = \eta_{it} + \psi_{jt} + \kappa \log D_{ijt} + \nu_{ijt}.$$

The dependent variable is the natural logarithm of the migration flow from location  $i$  to  $j$  at time  $t$ . The origin and destination fixed effects,  $\eta_{it}$  and  $\psi_{jt}$  respectively, can vary over time.  $\kappa$  measures the elasticity of migration flows to migration costs  $D_{ijt}$ .  $\nu_{ijt}$  is an error term.

Historical migration flows are rarely available. Here I construct a new data set of migration flows in historical China based on studies that trace the origin of modern villages. Historically, many villages in China were founded by clans from outside areas when the local living conditions worsened. By virtue of the tradition of keeping a rich family history, Cao (1997) was able to collect information about migration villages regarding the place of origin of their founders and the formation time frame. We can thus use the number of villages at location  $i$  founded by people from location  $j$  at time  $t$  to proxy for the migration flow from  $i$  to  $j$  at  $t$ .<sup>51</sup>

The following tables present the estimation results on the migration gravity using the number of migration villages to proxy for migration flows. Table D.1 shows results by using transport costs estimated in Section 5.2 to measure migration cost. Table D.2 shows results using distance.<sup>52</sup> Column (1) of Table D.1 shows that migration flows respond negatively to transport cost, as expected, and the elasticity is about -1.5. This is also similar to migration elasticities estimated by Allen and Arkolakis (2018) using international migration between 1960 and 2000 and intranational migration within the United States between 1850 and 2000.

The survey by Cao (1997) focuses primarily on migration villages in Kiangsi province, but also includes nearby prefectures in other provinces. To the extent that migration to prefectures outside Kiangsi may have different patterns, in column (2), I restrict the sample to only migration flows to Kiangxi province. The resulting elasticity is about -1.3. For all specifications, the origin and destination fixed effects are allowed to vary by time, but the distance elasticity itself can also vary over time. Therefore, I divide the sample into two periods in column (3) and column (4). In column (3), I restrict the sample to migration villages formed before 1722 and in column (4) to villages

---

<sup>51</sup>The implicit assumption here is that the initial population when the villages were founded is similar across villages. For a subset of villages where their recent population is available, I calculate the implied annual population growth rate to confirm that this is a reasonable assumption.

<sup>52</sup>Instead of the exact year of formation, we only know if the migration time falls between 1643 and 1722, 1643 and 1796, 1723 and 1796 or 1796 and 1912. When pooling observations for different time frames, I divide each observation by the number of years for the time frame it belongs to, to account for that the number of migrants/migration villages will be larger for longer time periods. The destinations are counties. For the majority of villages, only the origin province is known and for others the prefectures or counties are known. For observations that only the origin province is available, I calculate transport cost/distance to the nearest prefecture in the province. The estimate of elasticity using the average cost/distance to prefectures in the province generates statistically significant but much larger elasticity (in absolute value) than the rest of the sample where information about the prefecture/county is available.

formed after 1723. The results show that over time migration flows respond more to transport costs as the elasticity becomes more negative.

Table D.2 shows the results on migration elasticities using distance instead of transport costs. In the extreme case that distance does not correlate with transport costs and if we believe that migration flows correspond solely to transport costs, then the relationship between migration and distance will be null. In the intermediate scenario that distance is a noisy predictor of transport costs, the negative association between migration flows and transport costs using distance as a proxy will be attenuated. Results in Table D.2 suggests this is the case. Distance approximates transport costs reasonably well and migration flows still decrease with distance, but the estimated migration elasticity using distance is reduced by 50% compared to using transport cost.

Table D.1: Migration gravity: transport cost

|                           | Dep. var.: Log (number of villages) |                      |                     |                      |
|---------------------------|-------------------------------------|----------------------|---------------------|----------------------|
|                           | All<br>(1)                          | Kiangsi only<br>(2)  | 1643-1722<br>(3)    | 1723-1912<br>(4)     |
| Log (transport cost)      | -1.464***<br>(0.385)                | -1.306***<br>(0.333) | -0.975**<br>(0.451) | -1.679***<br>(0.499) |
| Origin fixed effects      | Yes                                 | Yes                  | Yes                 | Yes                  |
| Destination fixed effects | Yes                                 | Yes                  | Yes                 | Yes                  |
| Observations              | 113                                 | 93                   | 45                  | 56                   |
| R-squared                 | 0.884                               | 0.829                | 0.816               | 0.914                |

*Notes:* Columns (1) and (2) include time variant origin and destination fixed effects. The time frame for 12 observations is between 1643 and 1796 and therefore they are not included in columns (3) and (4). Robust standard errors are reported in parenthesis. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels.

Table D.2: Migration gravity: distance

|                           | Dep. var.: Log (number of villages) |                      |                    |                     |
|---------------------------|-------------------------------------|----------------------|--------------------|---------------------|
|                           | All<br>(1)                          | Kiangsi only<br>(2)  | 1643-1722<br>(3)   | 1723-1912<br>(4)    |
| Log (distance)            | -1.016***<br>(0.385)                | -0.873***<br>(0.333) | -0.773*<br>(0.451) | -0.965**<br>(0.397) |
| Origin fixed effects      | Yes                                 | Yes                  | Yes                | Yes                 |
| Destination fixed effects | Yes                                 | Yes                  | Yes                | Yes                 |
| Observations              | 113                                 | 93                   | 45                 | 48                  |
| R-squared                 | 0.893                               | 0.837                | 0.830              | 0.914               |

Notes: Columns (1) and (2) include time variant origin and destination fixed effects. The time frame for 12 observations is between 1643 and 1796 and therefore they are not included in the columns (3) and (4). Robust standard errors are reported in parenthesis. \*\*\*, \*\*, and \* indicate significance at the 1, 5, and 10 percent levels.