The roadblock effect

Yan Hu [*] [#]
University of Edinburgh

January 2024

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 transport 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 the sea and shifted population towards port cities.

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

[*]I am deeply indebted to Diego Puga for his support and guidance. I am especially grateful for Paula Bustos and Nezih Guner for their feedbacks. I thank Guillermo Caruana, Jorge De la Roca, Mauricio Drelichman, Keith Head, James Kung, Monica Martinez-Bravo, Joan Monras, Stephen Redding, Daniel Sturm, Matthew Turner, Maisy Wong, Noam Yuchtman and seminar and conference participants at CEMFI, the Meeting of the Urban Economics Association, LSE Economic History Graduate Seminar, Econometrics Society European Meeting, NEUDC, Virtual Economic History Seminar, CES North America Virtual Conference, RES and SES Annual Conference, GEP/CEPR Postgraduate Conference, University of Edinburgh, EEA-ESEM congress, Women in International Economics Conference for 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 Competividad (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.

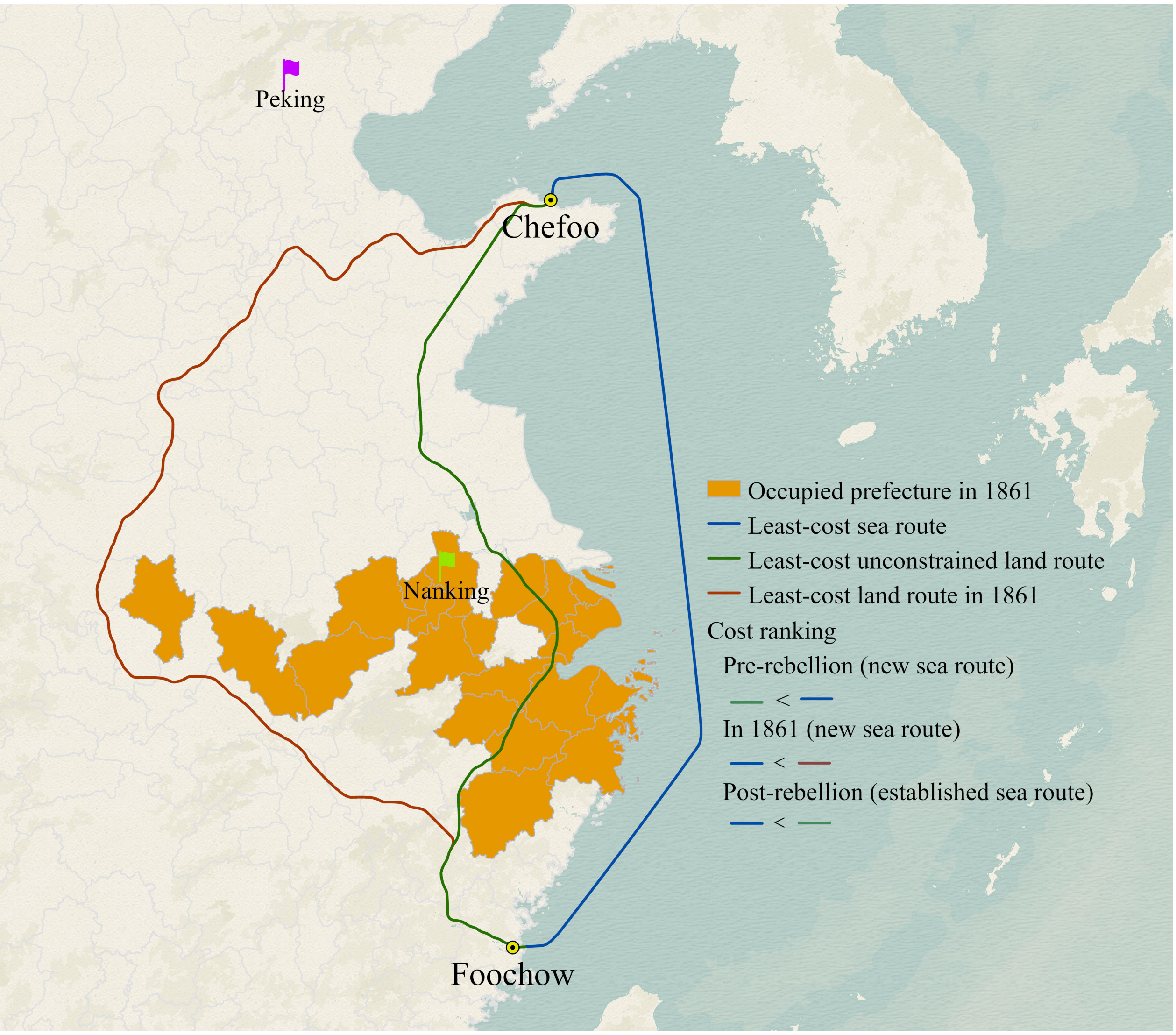
City locations and sizes reflect a combination of natural advantage and agglomeration economies. Natural endowments and technology differ across space, making some potential city locations more suitable than others. At the same time, localised increasing returns make locations that already have many 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 change over time as new transport 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). 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 inland 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 inland

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 I. 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 these, the cost of inter-modal transfer, and the cost of developing a new land or sea route, as estimated in *Panel B* of Table I.

route was not large enough to incur the sunk costs associated with establishing such a route.[1] The rebellion occupied many of the prefectures traversed by the usual inland 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,

[1] Section 2.3 discusses sunk investment costs for sea trade.

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 inland 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. I find that being occupied during the rebellion was associated with a permanent population loss of 59%, whereas being located on the coast was associated with a 22% population increase. This increase was even greater in coastal cities that saw a surge in domestic sea trade immediately after the rebellion.

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, inland 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 inland 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 land trade remained intact during the rebellion, when advances in transport technology became commonly available.[2] 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 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 Chinese Maritime Customs Service.[3] 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 is used to carry the cargoes. Instead, I add modal choice (by land or by

[2]In other words, the identification strategy "differences out" common shocks to sea transport over time, such as technological improvement, change of government attitudes towards maritime trade or 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 gained access to foreign trade and could have adopted the frontier technology of steamships for domestic trade through technological diffusion, I compare whether domestic trade between any two of these ports was more likely to use sea transport, depending on the heterogeneous increase in inland transport costs caused by the conflict.

[3]The Chinese 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.

sea) to an otherwise standard trade model.[4] Sea trade required fixed investments, and this likely held them back until the blockage of inland 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 transport.

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 inland 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 transport infrastructures 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 0.693 compared with 0.378 for "non-roadblocked" location pairs immediately after the rebellion in 1867. The respective probabilities in the data are 0.696 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, 0.364 instead of 0.693.

The "roadblock effect" is not confounded by alternative explanations for the post-rebellion surge in maritime trade. Section 6 builds and estimates an extended model, considering the potential reduction in investment costs for steamships on sea routes that could previously have sail trade. As the tendencies to use sailing ships prior to the rebellion were generally low and not systematically different for "roadblocked" and "non-roadblocked" pairs, the estimated "roadblock effect" remains

[4]The baseline model considers sea transport using steamships during and after the rebellion. Prior to the rebellion, there was no domestic steamship trade. The extended model considers sea transport by sail before the rebellion, the choice between sail and steam ships during and after the rebellion.

similar. The extended model also allows decomposing the relative contribution of the rebellion to the rise of maritime trade and the part resulting from the technological upgrade from sail to steam ships. The quantification exercise shows that the majority (74%) of the increase in maritime trade after the rebellion is attributable to the "roadblock effect." Section 7 further shows that the "roadblock effect" is not confounded by the impacts of treaty ports and foreign trade, potential changes in the inland transport network other than roadblocks, and that roadblocks are more likely to occur for distant location pairs.

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).[5] 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 and economic activities across cities (e.g., Redding and Venables, 2004; Faber, 2014; Donaldson, 2018; Bakker, Maurer, Pischke and Rauch, 2021) and within cities (e.g., Baum-Snow, Brandt, Henderson, Turner and Zhang, 2017; Tsivanidis, 2018; Heblich, Redding and Sturm, 2020). An important challenge when trying to evaluate the consequences of transport improvements is the endogeneity of infrastructure 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.,

[5] 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 the relocation of the air hub from Berlin to Frankfurt. After German reunification, air traffic did not return to Berlin due to the sunk costs of creating hubs. Donaldson and Allen (2020) develop a dynamic quantitative spatial model featuring path dependence as a result of reliance on past infrastructure investment.

Glaeser and Shapiro, 2002; Voigtländer and Voth, 2012; Dincecco and Onorato, 2016). While the previous studies have primarily 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.[6]

A debate related to the persistence of city locations also pertains to the failure of 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 inland 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.

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.[7] About one out of every six people in China (approximately 70 million) died during the rebellion (Ge et al., 1999).

The rebellion started as guerrilla warfare in the mountainous area of the Kuanghsi province in southwestern China (Figure 2). Peasants, discontented with the Qing dynasty's ever increasing taxes following a series of natural disasters and an economic crisis post the First Opium War, rose up against the ruling power. The uprising evolved into a full-fledged civil war with the Qing government when the rebels marched northeastward and occupied Nanking in 1853, establishing around it the "Taiping Heavenly Kingdom" in the mid and lower Yangtze region. The rebels' failed attempt to besiege the imperial capital, Peking, led to a protracted conflict centred on the rebellion-occupied area, resulting in massive destruction and civilian casualties. Some prefectures in the occupied

[6] Juhász (2018) also looks at war-induced trade blockade and technological adoption, where temporary protection from trade fostered domestic activity in mechanised cotton spinning in 19th-century France. The interregional trade in 19th-century China, however, consisted mainly of agricultural produces, mandated by local conditions. As a result, locations responded to blockade by finding alternative trade routes rather than adjusting production.

[7] Some scholars use the year 1850 as the starting point of the rebellion, while others use 1851. The military conflict at the outset remained small and local.

region lost up to 80% of their population (Cao and Li, 2000) and many places were razed to ground (Meyer-Fong, 2013).

Prior to the rebellion, the area around the Yangtze River, particularly its lower section, stood as 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).[8] 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. As a contemporary witness noted, the imperial capital was struck by food shortage and a surge in prices as a result of the rebellion (Ni, 2005). With its supplies through the regular routes combining land, the Yangtze River and the Grand Canal cut off, the Qing government was forced to ship tribute grain by sea to the imperial capital through Tientsin (Xia, 1995; Dai, 2012).

This paper aims to show that the modal change was not unique to the transport of tribute grain or limited to trade with Tientsin, but instead impacted on the broader exchange of goods throughout the entire country. The military confrontation between the rebels and the Qing government significantly disrupted inland transport (including inland waterways) that had previously traversed the rebellion-occupied region. This disruption is substantiated by contemporary reports from local officials to the imperial court. In 1852, the governor of Kiangxi province wrote to the emperor (Liao, 2010):

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

The governor of Anhwei province reported in 1854 (Liao, 2010):

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

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

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

In addition to accounts of roadblocks from Chinese officials, the Chinese Maritime Customs Service made note of an instance where roadblocks affected the adoption of steamship trade.[9] The opening of Kiukiang as a treaty port, located along the Yangtze River, enabled its direct steamship connections with seaports, as the Yangtze River meets the sea at Shanghai. Yet high expectations for the Kiukiang treaty port, "with four maritime ports close by from whence goods could be obtained so easily," were met with disappointment as trade continued along established inland routes. Due to

[8] Estimation in Section 4 shows that population growth in occupied region is slightly higher than the non-occupied region before the rebellion, but the difference is small and not statistically significant after controlling for observable characteristics. This generates, if any, an underestimation of the direct effect of the rebellion.

[9] *Reports on trade at the treaty ports in China for the years 1871-2 (Shanghai, 1874).*

Figure 2: The Taiping Rebellion and Imperial postal districts

Notes: Author's map based on the 1903 postal working map from the Chinese 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 "exorbitant rate of freight then charged," steamship transport offered little advantage over the existing inland routes, "unless it was that just about that time the trade of Ningpo was paralysed by the invasion of the Rebels into Chekiang."[10] The narrative is substantiated by data. Kiukiang had domestic steamship trade with Ningpo but not with any other maritime ports immediately after the rebellion.

The paper proceeds to provide quantitative evidence of the "roadblock effect." My estimates of transport costs during the rebellion also confirm that inland transport between Kiukiang and Ningpo was the one most affected by roadblocks, amongst all its trade pairs.

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.

Early in its reign, the Qing dynasty (1644-1912) implemented a sea ban on the mainland with the aim of severing supplies to the Zheng regime in Taiwan, established by loyalists of the previous Ming dynasty (Gu, 1983). As a part of the policy, coastal residents were evacuated to areas situated 15 to 25 kilometres inland (Gu, 1983). The ban was lifted when the Zheng family surrendered in 1683, and former coastal residents were allowed to return (Huang, 1986). Thereafter, aside from a brief sea ban imposed between 1717 and 1727, maritime trade with foreign countries was acquiesced.[11] Nevertheless, sea trade with foreign destinations did not regain its former glory from the previous dynasties, and the recovery of population in coastal regions was sluggish.[12]

At the same time, domestic trade relied on a combination of land transport and inland waterways (including canals). Sea trade, on the other hand, was limited to specific commodities (e.g., the shipment of soy beans from the northeast to Shanghai), and its usage fluctuated substantially from year to year (Liao, 2010). Despite sea transport being more cost-effective, trade was predominantly conducted through inland routes.[13]

In 1842, following its defeat in the First Opium War, China was forced to open treaty ports, in addition to Canton, for trade with Western countries. The establishment of these treaty ports also brought in the new sea transport technology of steamships. Despite the evident efficiency of sea

[10] The imposition of a high freight rate would be consistent with the pricing strategy of a monopolistic shipping company that bears sunk investment costs for creating new trade routes.

[11] 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 (Huang, 1986). Trade with non-Western countries was not subject to the Canton rule.

[12] Figure 3 shows the estimation result that the increase in coastal population relative to non-coastal population was limited between 1680 and 1851.

[13] The Qing statesman Wei (1826) estimated that sea transport using sail was about one-third the cost of canal shipping and much faster. However, only 25.6% of grain trade was carried out by sail in the early Qing period (Deng, 2009). Immediately after the rebellion in 1867, 69.6% "roadblocked" location pairs had sea trade, compared to 37.5% "non-roadblocked" location pairs. The model developed in Section 6 can match both pre- and post-rebellion trade patterns.

transport, whether by sail or steam, compared to inland routes, trade patterns remained remarkably unchanged. Although designated as one of the four ports to open in 1842, Foochow persisted in sending its celebrated Bohea tea from the nearby Wuyi Mountains on a lengthy overland route to Canton for export. It was not until over ten years later, with the outbreak of the Taiping Rebellion blocking the regular inland route to Canton, that Bohea tea began to be shipped directly from Foochow in 1853 (Liu, 2016).

2.3 Sea trade and sunk investment

The inertia in using established inland routes may be due to the large sunk costs usually associated with establishing new trade routes, especially for sea trade. Robert Hart, the long-time Inspector General of the Chinese Maritime Custom Service (the CMCS), outlined the requirements for ships to travel along the coast of China, which included "on the voyage, warning to be given of dangers," "in nearing a port, assistance to be given by experts acquainted with local peculiarities, tides, currents," "in the waters of the port, spots to be avoided to be marked," and "in the anchorage, order and regularity to be ensured."[14] Before systematic and coordinated efforts of infrastructural investment were made later by the CMCS to facilitate coastwise trade, the sunk investment required to initiate bilateral maritime trade would need to be borne by merchants or shipping companies alike, who needed to undertake investigative trips to establish new sea routes and lay down primitive navigational aids.[15] In addition to infrastructure investment, including the training of personnel capable of operating and managing shipping, the establishment of new trade routes typically requires significant logistical preparation. This involves setting up for the storage and movement of goods between different points along the route.

In a similar fashion that sunk costs can trap economic activities in unfavourable locations (e.g., Bleakley and Lin, 2012), the presence of sunk investments can lock trade in inefficient transport modes or routes, affecting economic geography indirectly. The launch of steamships was unique in the Chinese context because, unlike in some Western countries, it was not supported by government capital or public subsidisation (Liu, 1959).[16] Instead, the Qing government had long favoured and relied on inland trade, by building and maintaining an extensive inland transport network (Deng, 2009), including the costly keeping of the Grand Canal (Dai, 2012). Having (almost) entirely shut down sea trade in the early years of the dynasty means that for many locations to restart maritime trade after the lifting of the sea ban, the sunk costs they needed to incur would be substantial. The

[14] Circular No. 10 of 1868, in *Inspector general's circulars: first series: 1861-1875 (Shanghai, 1879).*

[15] According to the CMCS, local "sporadic attempts were being made everywhere" for shipping before the creation of Marine Department of the Customs in 1868 to "control, systematise, and put on a secure financial basis these uncoordinated efforts" (Banister, 1932).

[16] Steamships were introduced by foreign firms in treaty ports. These private firms did not have access to aids from neither foreign nor Chinese government, and instead sought financing in the Chinese money market, sometimes by promising rebates and preferential treatments to shippers in return for their investment (Liu, 1959).

dearth of government initiative discouraged private enterprises from embracing sea transport in domestic trade, notwithstanding its lower transport cost.

When steamship technology was introduced in China through treaty ports, further increasing the superiority of sea transport over land, the adoption staggered until the rebellion made it a necessity. The paper proceeds to show that locations that had a head start in adopting steamships were those whose inland trade routes were severely curtailed during the rebellion and as a result had strong incentives to invest in sea trade. The initial take-up of sea trade during the conflict acted as the coordination cue for transitioning and diverting state resources from land to sea trade.

3. Data

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, it requires data on trade, historical transport costs, population and measures of roadblocks due to the rebellion. I describe each one of them as follows.

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 the Chinese Maritime Customs Service (henceforth referred to as the CMCS). The CMCS was a bureaucratic agency founded in 1853 by foreign consuls in Shanghai to collect tariffs that went unpaid when the Qing government became overwhelmed by the Taiping Rebellion.The 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 the CMCS are believed to be "the only reliable and comprehensive statistical information" to use in studying China's historical trade (Cheng, 1956). Keller, Li and Shiue (2012) made an excellent introduction of the trade data from the 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. The CMCS trade data differentiate exporting and re-exporting and thus I can use direct export to measure trade between treaty ports. However, one limitation of the CMCS data is that domestic maritime trade covers only treaty ports, which were ports opened for trade with Western countries, and only the part of maritime trade carried by steamship (foreign and domestic) and foreign sailing ships (e.g., lorchas). 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 address the data limitation that we only observe sea trade immediately after the rebellion, I construct a trade model that incorporates modal and route choices, which helps to infer missing trade based on relative transport costs. As less than 1% of coastwise traffic was borne by Chinese ships (Hsiao, 1974), it was highly unlikely that two treaty ports with active sea trade would have used only Chinese junks and not steamers or lorchas in 1867. Therefore, I use the information from the CMCS trade report to construct the use of sea trade between any two treaty ports in 1867 and to model whether two locations had established active trade by sea. The simple baseline model focuses solely on adoption of steamships as there was no domestic steamship trade before the rebellion and the vast majority of post-rebellion maritime trade was carried by steamship. The extended model also takes into account sail trade before the rebellion and the choice between sail and steamships during and after the rebellion. The estimated effects of roadblocks from the two models are similar.

Estimation of the model requires 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 listed 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 headquarters).[17] [18]

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 refer 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 routes (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).[19]

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 annual data of roadblocks between location pairs from 1853 to 1864.[20]

---

[17]The headquarters of each postal district was usually a treaty port with the customs commissioner acting as the postmaster, except for two inland postal districts: Peking and Chinan.

[18]There were no major changes in land transport technology in 19th-century China, and the country's railways were still in their infancy. By 1903, only regional railways operated around the capital, Peking, in the North. These rail lines connected Chengting to Peking, Peking to Tientsin and Chinchow, and Chinchow to Chinan. For the estimation, postal districts potentially affected by the railway network are excluded (i.e., Peking, Tientsin, Newchwang, and Chinan postal districts).

[19]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). They consisted of inland waterways and overland roads. When coded rivers overlap with courier routes, they are modelled to have river transport costs.

[20]Although the earliest uprising took place in Kweiping, Kuangshi in 1851, the impact was local and limited. The rebels grew in power when they advanced to the middle and lower Yangtze region, where they occupied the first prefecture, Chinkiang, in 1853. The last occupied prefectures fell to the Qing government in 1864.

To look at population evolution before and after the rebellion, I use the 2018 update of historical prefecture population data of Cao (2000). The data were constructed based on government-administered census records, and cross-validated 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. Population data for historical prefectures are available for the pre-rebellion period in 1680, 1776, 1820, and 1851, as well as the post-rebellion period in 1880, 1910, and 1953. I supplement the historical population data with the 1982 county census and the 2010 township census, and map modern population data to the historical prefecture boundaries in 1820 from the China Historical Geographic Information System (CHGIS V4), constructed by the Harvard Yenching Institute and Fudan Center for Historical Geography.

4. The evolution of population

As a motivation, I begin my analysis in this section by examining the long-term evolution of the population of Chinese prefectures (cities and their 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, particularly those where domestic sea trade thrived. 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 can explain population relocation to coastal regions during this period.

I estimate the following reduced-form relationship between population, occupation during the rebellion, and coastal location using an event-study specification:

$$\ln \text{pop}_{it} = \sum_{t=1680}^{2010} \alpha_t (\mathbb{I}_t \times \mathbb{O}_i) + \sum_{t=1680}^{2010} \beta_t (\mathbb{I}_t \times \mathbb{C}_i) + \sum_{t=1680}^{2010} (X_i \times \gamma_t) + \psi_i + \kappa_t + \epsilon_{it} \,, \tag{1}$$

where pop_{it} denotes the population of prefecture i in year t; $\mathbb{I}_t$ is an indicator for the data year t (taking value 1 if the data year is t and 0 otherwise); the excluded category is the data year 1851, just before the Taiping Rebellion started; the data years 1680, 1776, 1820 belong to the pre-rebellion period and the data years 1880, 1910, 1953, 1982 and 2010 belong to the post-rebellion period; $\mathbb{O}_i$ is an indicator variable that takes value 1 if the prefecture was occupied during the rebellion; $\mathbb{C}_i$ is an indicator variable that takes value 1 if the prefecture is coastal; X_i is a vector of prefecture characteristics and γ_t are their time-varying coefficients; while location fixed effects will capture the effect of locational characteristics on population levels, the idea is to further control for location traits that can have a time-varying effect on population, especially if these characteristics correlate with rebellion occupation and coastal location; these are the level of human capital and a set of geographic controls including terrain ruggedness, agricultural suitability for major crops in historical China (i.e., rice, wheat and foxtail millet) and new world crops (i.e., potato, maize and sweet potato), and whether the Grand Canal went through a prefecture; ψ_i is a prefecture fixed effect; κ_t is a data year dummy, controlling for secular changes in population across all prefectures over time and differential growth for different gaps between data years; and ϵ_{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.[21] Robust standard errors are clustered at the prefecture level to adjust for heteroskedasticity and within-prefecture correlation over time.

The coefficients of interest in Equation 1 are α_t and β_t, which are plotted in Figure 3 (Table C.1 shows the coefficient estimates). The vertical bars represent the 95 percent confidence intervals. The coefficient α_t has a "difference-in-difference" interpretation, where the first difference compares the natural logarithm of population in occupied and non-occupied prefectures and the second difference compares the evolution of the first difference relative to the year 1851 (i.e., just before the rebellion started) for each preceding or succeeding data year. The estimates are plotted on the left panel of Figure 3. The estimated coefficients for the pre-rebellion data years are all small and statistically insignificant. This indicates that evolution of population in occupied and non-occupied prefectures was similar before the rebellion after controlling for the time-varying effects of other locational characteristics.[22] However, the coefficients for the post-rebellion data years decline sharply, capturing the additional population drop in occupied prefectures relative to non-occupied prefectures after the rebellion. For instance, the estimated coefficient for the year 1880 indicates that immediately following the rebellion, population in occupied prefectures dropped by an additional 59% compared to non-occupied prefectures. Over time population in occupied prefectures caught up as indicated by smaller coefficients (in absolute terms) in later years but the gap has not closed. In 2010, the average population in occupied prefectures was still 49% lower than the non-occupied prefectures compared to the pre-rebellion level.

The permanent change in the spatial distribution of population after the Taiping Rebellion stands in contrast to findings of Davis and Weinstein (2002), where Japanese cities quickly reverted to their relative sizes after the Allied bombings. I hypothesise that this is because the Taiping Rebellion coincided with an era of changing comparative advantage of locations when the potential of maritime trade arose, particularly with the introduction of steamships. In addition, the rebellion itself might have played a role in incentivising the adoption of sea transport (evidence to follow in Section 5). In other words, beyond the direct impact on war-torn areas, the rebellion could have influenced the

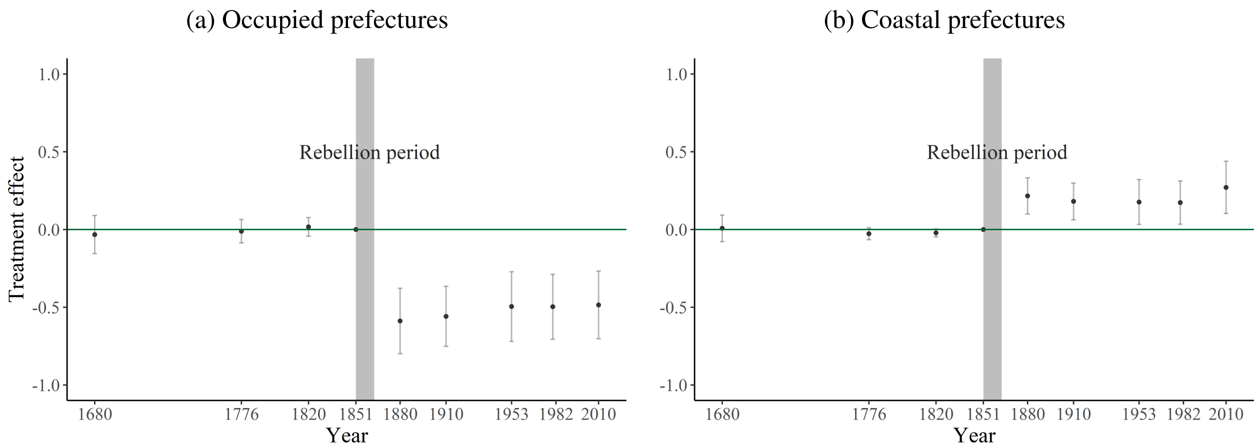
[21]As Qing troops were redeployed from different parts of the empire to quell the Taiping Rebellion, a military vacuum emerged, sparking uprisings in the periphery regions in Yunnan, Shaanxi, Gansu, Sichuan, and Guizhou. These periphery provinces were also inhabited by non-Han minority groups and are therefore excluded from the analysis. Nomadic areas of the Qing dynasty including 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, where the population data for 1982 are missing. Despite sharing the same name, this prefecture is not related to the rebellion.

[22]A negative coefficient for a pre-rebellion data year indicates a higher population growth in occupied prefectures before the rebellion. This leads to an underestimation, if any, of the effect of the rebellion on population loss in occupied prefectures using the post-rebellion coefficients. Table C.2 shows estimation results with an alternative simple difference-in-difference specification of population change before and after the rebellion, which allows controlling explicitly for the pre-rebellion trends.

relative attractiveness of locations by shifting trade routes to the sea-a process that could otherwise have been slow even when the superior technology of steamships became available.

Locations that stood to benefit from sea trade would be coastal prefectures, and the evolution of coastal population is captured by the coefficient β_t. Similar to α_t, the "difference-in-difference" coefficient β_t measures the additional change in population of coastal prefectures compared to non-coastal ones relative to their difference in 1851. As the right panel of Figure 3 shows, the coefficients are close to zero before 1851, meaning that evolution of population between non-coastal and coastal locations was similar before the rebellion. However, the positive and significant coefficient in 1880 indicates that immediately after the rebellion, population in coastal prefectures experienced a relative increase of about 22%. Although the effect attenuated over time, coastal population remained still 18% higher than non-coastal population a century later, compared to their difference prior to the rebellion. After 1949, as government policies increasingly influenced regional development in China, with early periods prioritising inland provinces and later reforms favouring coastal provinces, the population in coastal prefectures shrank slightly in 1980 but expanded further in 2010.

Figure 3: Estimated log population change in occupied and coastal prefectures



Notes: Estimation based on Equation 1. The left panel shows estimated coefficient (α_t) of an interaction between the occupied prefecture indicator and a year dummy. The right panel shows estimated coefficient (β_t) of an interaction between the coastal indicator and a year dummy. Vertical lines indicate 95% confidence intervals with standard errors clustered by prefecture. The specification includes prefecture and year fixed effects, and controls for the time-varying effects of other determinants of population (i.e., agricultural suitability, terrain ruggedness, canal access and human capital).

We can also estimate the average population change after the rebellion with the following specification:

$$\ln \text{pop}_{it} = \alpha(\text{post}_t \times \mathbb{O}_i) + \beta(\text{post}_t \times \mathbb{C}_i) + \sum_{t=1680}^{1953} (X_i \times \gamma_t) + \psi_i + \kappa_t + \lambda_o(\mathbb{O}_i t) + \lambda_c(\mathbb{C}_i t) + \epsilon_{it} \ , \quad (2)$$

where variables are defined in the same way as in Equation 1, except for post_t, which is an indicator variable denoting whether a data year t is after the rebellion. Instead of interacting rebellion occupation or coastal indicator with each of the data year, Equation 2 interacts them with a post-rebellion indicator. Therefore, we can control for differential population growth in occupied and coastal prefectures by including a linear time trend for each: λ_o and λ_c. Consequently, the coefficients α and β can be interpreted as the average population change in occupied and coastal locations after the rebellion after accounting for pre-existing trends of the population. I exclude observations from the data years 1980 and 2010, as government policies had a more pronounced impact on regional development after 1949, although the results including these years would be very similar.

The estimation results are displayed in Table C.2 of AppendixC. Consistent with the findings in Figure 3, the result shows a substantial drop in the population of occupied prefectures after the rebellion. The estimated population loss is even slightly greater, about 59% on average, after accounting for the pre-existing population trends. The average increase in coastal population was 22% after the rebellion, after controlling for the direct effect of the rebellion on the occupied region and the pre-existing trends. Furthermore, port cities that witnessed a surge in domestic maritime trade following the rebellion, rather than foreign maritime trade, experienced the highest population growth after the rebellion, although this part of the result is based on small sample of ports where trade data are available and needs to be taken with a grain of salt.

Fortunately, the trade data include more details, such as the status of bilateral maritime trade immediately after the rebellion. By exploiting the heterogeneous incentives to adopt sea trade through roadblocks during the rebellion and linking them to the observed post-rebellion bilateral trade, the next section shows that population relocation to coastal regions can be a result of the "roadblock effect."

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 sea trade immediately 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. I use counterfactual predictions to quantify the effect of the rebellion on maritime trade.

5.1 Model set-up

Consider a country with J cities (prefectures in historical China), indexed by $j = 1, 2....J$. 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. If we assume that consumers have constant elasticity of substitution utility and that firms face monopolistic competition, the profit function of a firm in location j exporting its product to i at time t can be written as:[23]

$$\pi_{ijt} = (1-\alpha)(\frac{\tau^m_{ijt} c_{jt}}{\alpha P_{it}})^{1-\epsilon} Y_{it}. \tag{3}$$

The elasticity of substitution is denoted as α, which falls between 0 and 1. To keep the notation concise, I introduce $\epsilon = 1/(1-\alpha)$. As standard in trade models, 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 τ^m_{ijt} to vary with transport mode m. Firms can either choose inland transport (including inland waterways) or sea transport, which I denote l and s respectively.

In addition to making the trade cost a function of the transport cost D^m_{ijt} that takes into account optimal route choice within each mode, I incorporate intertemporal investment for establishing new trade routes for sea transport, I^s_t, and inland transport, I^l_t.[24] The mode-dependent iceberg trade costs are:

$$(\tau^l_{ijt})^{1-\epsilon} = (D^l_{ijt} I^l_t)^{-\gamma} e^{\eta_{l,ijt}}, \quad (\tau^s_{ijt})^{1-\epsilon} = (D^s_{ijt} I^s_t)^{-\gamma} e^{\eta_{s,ijt}} \tag{4}$$

for inland 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.[25] 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 an indicator of whether sea trade occurs in period t. We have:

$$I^s_t = \begin{cases} I^s, & \text{if } S_{ij,t-1} = 0; \\ 1, & \text{if } S_{ij,t-1} = 1. \end{cases} \tag{5}$$

[23]While monopolistic competition is used to characterise the market structure here, as it will become evident later, the results do not rely on this assumption. The key assumption is that the least-cost transport mode (taking into account sunk investment cost) will be chosen.

[24]D^m_{ijt} is the lowest-cost route effective distance using the terminology of Donaldson (2018). Here I calculate the effective distance separately for sea and inland transport and therefore the effective distance is mode-specific.

[25]The implicit assumption is that sunk investment is proportional to the transport cost. Sunk costs associated with establishing a new sea route such as investigative trips, and setting up navigational aids and staging posts increase with distance.

If a 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 inland route, it is assumed that firms need to incur an investment cost of I^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:

$$\eta_{l,ijt} \sim \mathcal{N}(0, \sigma_l^2), \quad \eta_{s,ijt} \sim \mathcal{N}(0, \sigma_s^2). \tag{6}$$

There are several simplifying assumptions I make about the baseline model. Before deriving the estimation equations, I discuss how these assumptions can affect the results.

First, the baseline model assumes that sea transport uses the new technology of steamships between treaty ports during and after the rebellion.[26] There was no domestic steamship trade before the rebellion.[27] 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. Because the probability of sail trade was low before the rebellion, the results change little.

Second, in the model, firms are assumed to be myopic. In other words, their decision to initiate sea transport during the rebellion is based on the immediate factors of relative transport costs and sunk investment costs, but without regard to potential cost savings from the investment for future sea trade. This may be justified by the uncertainty of the post-rebellion status of sea trade. In AppendixD, I develop a dynamic model where firms do incorporate cost reductions for future sea trade. The implied sunk investment cost will be larger if we assume firms are forward-looking, but the estimated effect of the rebellion on the adoption of sea trade is very similar compared to the baseline model.

Third, in the model, firms transport their own manufactured goods, aiming to maximise profit by minimising transport costs, incorporating also the sunk costs. The results remain unchanged when a distinct transportation sector provides transport services and charges production firms freight for transportation. Intuitively, in the absence of arbitrage opportunities, the problem of transport companies is equivalent to that faced by firms responsible for transporting their own manufactured

[26] Steamship technology was available and became the dominant mode of sea transport during and after the rebellion when treaty ports were established.

[27] Steamship trade existed between Hong Kong, Macau and Canton, and between Hong Kong and Shanghai before the rebellion. As (parts of) Hong Kong and Macau had become colonies, bilateral trade with Hong Kong and Macau was not treated as domestic trade by the CMCS and is not included in the sample.

goods.[28]

Lastly, I do not make explicit assumptions about whether firms in the same location coordinate on the sunk investment. If firms do not coordinate at all, I^s would represent the total sunk investment needed to establish a new route. At the other extreme, if firms coordinate completely, the total sunk investment would be I^s multiplied by the number of firms. When there is a separate shipping company and it incurs the sunk investment, I^s would still be the total sunk investment. If firms or the shipping company are forward-looking, the implied sunk investment cost will be larger than what the baseline model entails. In summary, the sunk investment cost estimated from the baseline model can be considered a lower bound of the total sunk investment cost.

I proceed to derive the estimation equations from the baseline model set-up. We can first obtain the post-rebellion sea trade probability. At the beginning of each period, firms observe origin-destination specific shocks to inland and sea transport and choose either inland or sea transport. The probability of sea trade immediately after the rebellion is therefore:

$$\Pr(S_{ij,post} = 1) = \Phi\Big(\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,post}}{D^s_{ij,post}} - \mathbb{1}(S_{ij,reb} = 0) i^s\Big), \tag{7}$$

which depends first on the post-rebellion transport cost by land relative to sea ($D^l_{ij,post}/D^s_{ij,post}$). The elasticity of trade with respect to the transport cost γ, 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 $i^s = (\gamma/\sigma) \ln I^s$, 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 a sea route was 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 steamships afterwards.

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

$$\Pr(S_{ij,reb} = 1) = \Phi\Big(\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,reb}}{D^s_{ij,reb}} + roadblock_{ij} \times i^l - i^s\Big). \tag{8}$$

Equation 8 shows how the rebellion affects 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 $roadblock_{ij} = 1$), finding

[28] For example, if a monopolistic shipping company operates on each bilateral route and competes with inland transport companies, the optimal price the shipping company charges if it enters the market will be the cost of inland transport τ^l_{ijt}, and it will only enter the market if $\tau^s_{ijt} < \tau^l_{ijt}$. Note, however, that the interpretation of sunk cost may change. If a single shipping company is in charge of the transportation from j to i, I^s is the total sunk cost for establishing a new sea route, whereas if firms transport their own manufactured goods, I^s may represent the sunk cost for an average firm.

alternative inland routes entailed an investment cost I^l (the transformed sunk investment cost $i^l = (\gamma/\sigma)\ln I^l$ 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 ($D^l_{ij,reb}/D^s_{ij,reb}$) increased. Both raised the relative desirability of using sea transport during the rebellion, which induced "roadblocked" location pairs to incur the sunk investment cost for sea trade during the rebellion.

The post-rebellion probability of sea trade can be linked 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:

$$\begin{aligned} \Pr(S_{ij,post}=1) =& \Pr(S_{ij,reb}=1)\Pr(S_{ij,post}=1|S_{ij,reb}=1) \\ &+ \Pr(S_{ij,reb}=0)\Pr(S_{ij,post}=1|S_{ij,reb}=0). \end{aligned} \tag{9}$$

Combining Equations 7, 8 and 9, the post-rebellion probability of sea trade is:

$$\begin{aligned} \Pr(S_{ij,post}=1) =& \Phi\Big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,reb}}{D^s_{ij,reb}} + \text{roadblock}_{ij}\times i^l - i^s\Big)\Phi\Big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}}\Big) \\ &+ \Big(1-\Phi\Big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,reb}}{D^s_{ij,reb}} + \text{roadblock}_{ij}\times i^l - i^s\Big)\Big)\Phi\Big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}} - i^s\Big), \end{aligned} \tag{10}$$

which is a function of (1) whether the locality pairs had roadblocks, (2) the relative cost of inland transport compared to sea during and after the rebellion, and (3) the parameters of the model (γ/σ, i^s, and i^l).

Equation 10 highlights the "roadblock effect," whereby a temporary increase in land transport cost during the rebellion affect post-rebellion modal choices by facilitating investment in maritime transport.[29] It also shows that the effect of the rebellion on trade hinges crucially on whether the regular inland route became unusable when the occupied region hit some intermediate point on its connection, and how lengthy a detour was needed compared to the cost of sea alternatives. This requires in turn knowledge about the availability of roads, rivers and terrain along different inland trade routes and their costs relative to 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 part presents the intermediate step of obtaining estimates of relative transport costs in 19th-century China.

5.2 Estimation of historical transport costs

The increase in inland transport costs due to roadblocks compared to sea alternatives is a key determinant of the differential impacts of the rebellion on steamship take-up in different locations.

[29] The post-rebellion probability of maritime trade is a weighted average of probabilities $\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}}\big)$ and $\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}} - i^s\big)$. The weight on the higher probability $\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}}\big)$ becomes larger when roadblocks occur.

To quantify the inland transport costs and how the rebellion affected them, one needs knowledge about the costs of using different trade routes, which depends, in turn, on the cost of using various transport infrastructure underlying these routes. While the previous literature on transport costs usually infers relative costs 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., non-courier route, courier route and waterway), variations in cost due to terrain, as well as a fixed cost 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.

I divide the surface of China into 0.1 degree by 0.1 degree cells (553×826 cells in total, a cell at 30 degree latitude covering approximately 90 km^2 area) and overlay it with rivers, courier routes, 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 route cell, a non-courier 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.

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 renders the transport cost by inland waterways.[30] To obtain the transport cost by sea, I use the information about travel time for sea transport relative to river as the 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, a 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.

[30]A trip from prefectures to their postal headquarters would not involve maritime transport and therefore the transport cost by sea cannot be estimated based on the assignment.

The estimation results of transport costs are summarised in *Panel A* of Table I. The cost of travelling by land without any road infrastructure and with 0 ruggedness is normalised to one so that the unit transport costs by other modes is relative to that by flatland. 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 that by non-courier land transport and the cost by sea was only 4% of that by non-courier land transport. 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. For instance, for a sea trip between Shanghai and Ningpo, about half the transport cost will be borne by transshipment.

Using the estimated historical transport costs, in AppendixE, 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 magnitude of the elasticity using transport costs (i.e., lowest-cost route effective distance) is also larger than the elasticity estimated by distance.[31]

In a broader context, this method can be applied to estimate transport costs for any region and any historical period, provided the availability of two inputs. The first requirement involves the assignment of locations to larger areas, typically optimised to minimise transport costs. The second input is the location of different transport infrastructure. While roads may change or disappear over the course of a lengthy history, terrain and rivers generally remain stable. Consequently, it becomes feasible to estimate relative transport costs for different terrains, rivers, and sea, as well as the fixed costs associated with transitions, as long as the assignment of locations is accessible.

5.3 Model estimation and prediction

With the intermediate input of transport costs in 19th-century China, we can now proceed to estimate the model parameters and use them to predict the probability of sea trade with and without roadblocks.

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 were not occupied. For each trade pair in each year 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. If roadblocks happened in more than one year and the required detours were different as occupied regions changed over time, I use the maximum increase in the inland transport cost due to roadblocks to calculate the relative inland transport cost during the rebellion.[32]

[31] Distance is a proxy for transport cost and may therefore be subject to the classical measurement error. As a result, the coefficient of distance may suffer from attenuation bias.

[32] The estimated effects are similar if the average detour instead of the maximum is used (not shown).

The estimation sample comprises all mainland Chinese treaty ports that were opened before or during the rebellion, forming the same set of mainland treaty ports that were active in 1867. Trade between two river ports is excluded from the analysis as it does not involve sea transport.[33] 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.[34]

With the estimated relative transport costs of land compared to sea alternatives, accounting for their increase during the rebellion due to roadblocks, we can obtain estimates of the parameters of the model using the post-rebellion sea trade probability derived in Equation 10, by maximising the likelihood of observing maritime trade immediately after the rebellion in the data.

The estimation results of the baseline model are presented in *Panel B* of Table I. As expected, the elasticity of using sea transport with respect to the ratio of land to sea transport cost is positive, indicating that location pairs are more likely to resort to maritime trade when the cost of inland transport is comparatively high. The cost associated with initiating sea trade is substantial, with the estimate implying that the cost of using a new sea route is approximately 20 times that of an existing one ($I^s = \exp(i^s \times \sigma/\gamma)$). Compared to sea transport, the investment cost for inland transport is modest, with a cost ratio of using a new inland route over an existing one being about 4.6 ($I^l = \exp(i^l \times \sigma/\gamma)$).

Using the parameter estimates from *Panel B*, we can make predictions regarding the probability of maritime trade after the rebellion, and these predictions are presented in *Panel C*. The observations are categorised into two groups: locality pairs that encountered roadblocks during the rebellion and those 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) displays the prediction from the model, which aligns well with the observed modal choices. The average predicted probability of maritime trade is 0.693 for "roadblocked" pairs and 0.378 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) remains the same as in column (3). For "roadblocked" pairs, however, their counterfactual probability of

[33] In other words, the analysis includes trade between seaports and trade between a seaport and a Yangtze River port that does not involve any road transit. The CMCS classifies exports from a seaport transferred to a Yangtze port by an inland route as transit trade instead. Steamship trade between Yangtze ports and seaports was permitted after 1861 (Fan, 1985), following the Second Opium War, and the estimation accounts for this timing.

[34] Throughout the rebellion, Shanghai found itself surrounded by occupied prefectures, rendering inland transport to the city impossible according to my estimate (i.e., $D^l_{ij,reb}/D^s_{ij,reb}$ would be infinite). Simultaneously, Shanghai held a unique position as the only treaty port engaged in maritime trade with all other mainland treaty ports in 1867. To show that the estimated "roadblock effect" is not solely driven by this exceptional example, it is dropped from the estimation. Despite Chinkiang not being occupied in the latter part of the rebellion, being situated next to the rebellion capital, it was still caught by warfare and suffered significant damage. This could explain the 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." Between 1871-1872, it started to export rice to the southern provinces by sea, which "was quite a new feature."

Table I: Estimation of roadblock effect: Trade cost and model prediction

Panel A: Estimation results on transport costs

Unit cost relative to flat land		Punishment of ruggedness by one std.
Flat land (without courier route)	1.00	32% increase in transport cost
Courier route	0.18	Transition cost from land to waterways
Inland waterways	0.08	Equivalent to 82.5 km by inland waterways
Sea	0.04	Equivalent to 165 km by sea

Panel B: Estimation results on trade model parameters

	Parameter	Value
Elasticity of probability of sea trade w.r.t. land/sea cost ratio	γ/σ	0.733
		(0.185)
Sunk cost of new sea route (transformed)	i^s	2.195
		(0.454)
Sunk cost of new land route (transformed)	i^l	1.120
		(0.393)
Observations		88

Panel C: 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	0.364
			(0.021)	(0.020)
Unaffected pairs	32	0.375	0.378	0.378
			(0.045)	(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 4.6 ($I^l = \exp(i^l \times \sigma/\gamma)$). 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 inland 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.

maritime trade would have been only 0.364 without roadblocks, a figure close to that of unaffected pairs. This suggests that, had it not been for the rebellion, the adoption of sea trade for both groups would have been similar.[35] Roadblocks significantly enhance the use of sea trade after the rebellion, increasing the probability by approximately 90% from 0.364 to 0.693.

6. Model extension

In this section, I extend the model developed in Section 5.1 by incorporating the possibility of pre-rebellion sail trade and the post-rebellion choice between sail and steamships. Although the pre-rebellion bilateral trade by sail is unobserved, the model structure enables the estimation of the probability of sail trade between location pairs before the rebellion. The extended model considers the likelihood that certain location pairs engaged in sail trade before the rebellion, thus potentially avoiding the sunk investment associated with establishing a new sea route when steamships were introduced. Additionally, it allows to break down the contribution to the increase in sea trade into components attributed to roadblocks and those stemming from the technological upgrade from sail to steamships, as further discussed in Section 7.

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

$$(\tau_{ijt}^{sail})^{1-\epsilon} = (D_{ij}^{sail} e^{\iota_{kijt}} I_t^s)^{-\gamma} e^{\eta_{s,ijt}}, \tag{11}$$

where D_{ijt}^{sail} represents the transport cost of sailing ships and $e^{\iota_{kijt}}$ denotes a shock or a preference parameter for sail trade for firm k following a normal distribution with a mean of zero and a standard deviation of σ_{sail}. The trade costs by land and steamship are the same as before:

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

where the superscript for sea transport s in the baseline model is replaced by a new one specifically for steamships. It is further assumed that firm owners with a log utility choose either sea or inland transport mode after observing location specific shocks to transport costs ($e^{\eta_{l,ijt}}$ and $e^{\eta_{s,ijt}}$) and make sunk investment if needed. Afterwards, firm-specific shocks to sailing ships ($e^{\iota_{kijt}}$) are realised. When both technologies of sea transport become available, a portion of firms exporting from location i to j will opt for sail over steamships. After incorporating sailing ships, the new transport cost by sea can be summarised as follows:

$$D_{ijt}^{s} = \begin{cases} D_{ijt}^{sail}, & \text{before rebellion;} \\ (D_{ijt}^{steam})^{\theta} (D_{ijt}^{sail})^{1-\theta}, & \text{during and after rebellion;} \end{cases} \tag{12}$$

where $\theta = \Phi(\frac{\ln(D_{ijt}^{sail} / D_{ijt}^{steam})}{\sigma_{sail}})$ is the probability that a firm prefers steam to sail ships conditional on sea transport.

[35] The pre-rebellion population for “roadblocked” origin or destination, as well as the product or sum of origin and destination population, is slightly lower and not statistically different from “non-roadblocked” pairs (not shown).

Pre-rebellion sea trade can only use sail. Therefore, the probability of sea transport in time t before the rebellion is:

$$\begin{aligned}\Pr(Sail_{ijt} = 1) =& \Pr(Sail_{ij,t-1} = 1)\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,pre}}{D^{sail}_{ij}}\big) \\ &+ \big(1 - \Pr(Sail_{ij,t-1} = 1)\big)\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,pre}}{D^{sail}_{ij}} - i^{s}\big).\end{aligned} \tag{13}$$

Suppose that the probability of bilateral trade by sail has reached a steady state before the rebellion. This probability can be derived by setting $\Pr(Sail_{ijt} = 1)$ equal to $\Pr(Sail_{ij,t-1} = 1)$ for any location pair i and j. Denote the steady state probability of sail trade as ρ_{ij}.[36] We have:

$$\rho_{ij} = \frac{\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,pre}}{D^{sail}_{ij}} - i^{s}\big)}{1 + \Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,pre}}{D^{sail}_{ij}} - i^{s}\big) - \Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,pre}}{D^{sail}_{ij}}\big)}. \tag{14}$$

Next, the post-rebellion probability of sea trade can be obtained after considering the pre-rebellion probability to use sailing ships:

$$\begin{aligned}\Pr(S_{ij,post} = 1|Sail) =& \rho_{ij}\Big(\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,reb}}{D^{s}_{ij,reb}} + \text{roadblock}_{ij} \times i^{l}\big)\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,post}}{D^{s}_{ij,post}}\big) \\ &+ \Big(1 - \Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,reb}}{D^{s}_{ij,reb}} + \text{roadblock}_{ij} \times i^{l}\big)\Big)\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,post}}{D^{s}_{ij,post}} - i^{s}\big)\Big) \\ &+ (1 - \rho_{ij})\Big(\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,reb}}{D^{s}_{ij,reb}} + \text{roadblock}_{ij} \times i^{l} - i^{s}\big)\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,post}}{D^{s}_{ij,post}}\big) \\ &+ \Big(1 - \Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,reb}}{D^{s}_{ij,reb}} + \text{roadblock}_{ij} \times i^{l} - i^{s}\big)\Big)\Phi\big(\frac{\gamma}{\sigma}\ln\frac{D^{l}_{ij,post}}{D^{s}_{ij,post}} - i^{s}\big)\Big).\end{aligned} \tag{15}$$

While Equation 15 may appear convoluted, ultimately, the post-rebellion probability of sea trade depends on the relative land transport costs, their change during the rebellion due to roadblocks, and parameters of the model, as in the baseline model. Consequently, we can similarly estimate the model parameters by maximising the likelihood of observing post-rebellion sea trade.[37] The

[36] Note that a single realisation of trade does not reach a steady state because subject to shocks, at any given time, either sea or inland transport can be chosen. However, in expected or probabilistic terms, the use of sail trade between two locations will be stable.

[37] The underlying assumption in the extended model is that if sailing ships are in use before the rebellion, the transition to steamship trade during the rebellion does not require additional investment. In contrast, the baseline model operates on the premise that adopting the novel technology of steamships requires new investment separate from sailing ships. In reality, it's possible that the two methods of sea transport involve both shared and separate investment. The estimation of an intermediate case, assuming only a fraction of the investment costs are deductible, yields similar outcomes (not shown).

additional input in the extended model is the relative transport cost of sail to steam ship, used to calculate the sea transport cost D^s_{ijt}. My preferred sail to steam cost ratio is 2, which justifies the share of grain trade by sea (25.6%) before the rebellion according to Deng (2009).[38] I also estimate the model under different cost ratios of sail to steam ships. The lower bound of the cost ratio is one, indicating that sail ships have the same variable transport cost as steam ships. The upper bound is 2.25, indicating that sail ships have the same variable cost as courier routes. The results are present in Table A.2.

The estimation results of the extended model are presented in Table II. The parameter estimates in *Panel A* are comparable to those in Table I, falling within the 90% confidence interval of the estimates under the baseline model. Compared to the baseline model, the estimated elasticity of trade is smaller in the extended model, but the implied sunk costs for new sea routes and inland routes are larger. Intuitively, this is because the extended model factors in the possibility that some location pairs may incur smaller or no sunk cost in establishing a new trade route by sea if they previously had sail trade, while the baseline model would calculate an average sunk investment cost, including these pairs with potentially reduced or zero costs.

In *Panel B* of Table II, I show the predicted mean probability of sea trade using the corresponding parameters in *Panel A*. Column (5) indicates that "roadblocked" pairs are predicted to have a slightly lower probability of sea trade by sail before the rebellion. Hence, 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 trade by sail before the rebellion. The implied "roadblock effect" is also consistent with the baseline model. Under the counterfactual scenario in which the rebellion did not happen, the probability of maritime trade would be only 0.365 (column (4)) for "roadblocked" pairs, similar to unaffected location pairs. This counterfactual indicates that roadblocks increase the probability of sea trade by 87%, rising from 0.365 to 0.684.

The extended model, incorporating the pre-rebellion probability of sail trade, allows the decomposition of the relative contributions of technological change and roadblocks to the post-rebellion surge in sea trade. This decomposition naturally hinges on the cost ratio between sail and steam transport. When the cost of sail to steam ratio is low, indicating that steamships are not significantly superior to the incumbent technology, the increase in sea trade due to technological improvement would be small, whereas the contribution of roadblocks would be relatively large. Conversely, if steamships are much more superior to sail technology, the relative contribution of technological

[38] Deng (2009) estimated that the combined grain trade by inland navigation and sea was approximately 80 million piculs (1 picul $\approx$ 60 kilogrammes), with sea trade accounting for 20.5 million piculs. Using Wang and Huang (1989)'s classification of grain export and import regions during the Qing dynasty, I calculated the probability of sea trade using sail for each grain trade pair based on their transport costs via sea and inland routes. Monte Carlo simulation is used to calculate the average share of sea trade conducted by sail. The volume of trade was assumed to be proportional to the origin and destination population and inversely proportional to the transport cost, following the prediction of a naive gravity equation. The share of sea trade was found to be 25.6% with a cost ratio of 2. For comparison, the share would be 36% with a cost ratio of 1 and 23.6% with a cost ratio of 2.25.

Table II: Estimation of roadblock effect: Extended model prediction

Panel A: Estimation results on trade model parameters		
	Parameter	Value
Elasticity of probability of sea trade w.r.t. land/sea cost ratio	γ/σ	0.542
		(0.116)
Sunk cost of new sea route (transformed)	i^s	2.062
		(0.386)
Sunk cost of new land route (transformed)	i^l	1.537
		(0.485)
Observations		88

Panel B: Mean probability of sea trade before and after the rebellion					
	Observations	Data	Model prediction		
			Post-rebellion sea		Pre-rebellion sail
			With roadblock	W/o roadblock	
	(1)	(2)	(3)	(4)	(5)
Roadblocked pairs	56	0.696	0.684	0.365	0.255
			(0.018)	(0.021)	(0.018)
Unaffected pairs	32	0.375	0.387	0.387	0.290
			(0.047)	(0.047)	(0.042)

Notes: In *Panel A*, the estimates imply that the cost ratio of new sea route over existing route is 45 ($I^s = \exp(i^s \times \sigma/\gamma)$). The implied cost ratio of new land route over existing one is 17 ($I^l = \exp(i^l \times \sigma/\gamma)$). Standard errors in parentheses. In *Panel B*, the model prediction is based on parameters estimates of *Panel A*. The prediction with roadblocks in column (3) provides the mean estimated probability for city prefecture pairs for which the least costly inland route was roadblocked by the rebellion, taking into account the pre-rebellion sail trade, whether the required detour was costly enough to incur the sunk cost of starting to trade by sea between those locations if pre-rebellion sea trade was not established, 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. Column (5) provides the mean estimated pre-rebellion probability of sail trade. Bootstrap standard errors using 500 replicates in parentheses.

improvement would be more substantial, although this holds true for both "roadblocked" pairs and "non-roadblocked" pairs.

Table II reveals that even with a cost ratio of 2 between sail and steam ships, which is towards the upper limit of the advantage of steamships, the implied contribution of roadblocks to the rise of sea trade during this period remains substantial. The predicted sea trade probability by sail before the rebellion for "roadblocked" pairs is, on average, 0.255. If roadblocks did not happen, their post-rebellion sea trade probability would have been 0.365. The increase in the probability from 0.255 to 0.365 represents the contribution of technological change, which is only 11 percentage points. The majority of the surge in maritime trade (74%) comes from roadblocks, raising the sea trade probability from 0.365 to 0.684. Similarly, for "non-roadblocked" pairs, the model predicts a probability of sea trade increasing by 0.097, from 0.29 to 0.387, with the introduction of steamship. Section 7 provides a detailed discussion of the role of technological change and whether it confounds the "roadblock effect."

7. Alternative explanations

This section discusses some alternative explanations for the rise of maritime trade during this period and provides evidence that the "roadblock effect" is not confounded by them.

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: 15% immediately after the rebellion.[39][40] 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 steamship for domestic trade, the identification is through comparing whether domestic trade between any two of them is more likely to use sea transport depending on the heterogeneous increase in their inland transport costs due to the conflict.

Another potential channel through which the introduction of treaty ports could shape domestic trade patterns is transit trade for re-exporting domestic goods or importing foreign goods. Before the rebellion, exporting to foreign destinations was mainly through Canton (today's Guangzhou) and Shanghai. If the regular routes from one treaty port location to major entrepôts, Canton for example, became blocked due to the rebellion, the location could opt to directly export its goods to foreign destinations. A notable instance is the direct export of Bohea tea from Foochow during the Taiping Rebellion, which used to take a lengthy detour to Canton. Even if a location intended to re-export its goods through Canton, the effects of roadblocks on modal choices for domestic and transit trade would work in an analogous and parallel manner. While there may still be scale economies for adopting steamship trade when transit trade was present, I show that the magnitudes of the "roadblocked effect" are similar after excluding Canton and Shanghai in Tables A.3 and A.4.

Technological improvement One might question whether the post-rebellion adoption of sea trade was due to the availability of the new steamship technology rather than the rebellion. The emphasis of my findings is that even when a new and superior technology is commonly available, its adoption

[39]The population growth in coastal prefectures without treaty ports is expected to be smaller for several reasons. One potential reason is that these prefectures were, on average, smaller in size. Also, they were unable to use steamships in the early days. Faced with roadblocks, these prefectures had the incentive to shift to sea transport using sail, but this incentive would not have been as strong as it would have been if steamship technology were accessible.

[40]Additionally, the population change in treaty ports correlates positively with domestic sea trade value, but not with foreign trade value, although the evidence is more suggestive than conclusive due to the limited sample size for which the trade value data are available.

is not guaranteed. Holding (new) technology constant, I show that location pairs subject to shocks to their inland route were much more likely to adopt steamship compared to location pairs unaffected by the rebellion. The counterfactual probability of using steamship without rebellion for "roadblocked" pairs is similar to "non-roadblocked" pairs. These findings demonstrate that the rebellion played a catalysing role in 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 heterogeneous war exposure, it is true that the *level* of sea trade observed 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 aggregate level of pre-rebellion grain trade by sail), Table II shows that the counterfactual probability of sea trade for "roadblocked" pairs without the rebellion would be 0.365, an increase of 0.11 from the pre-rebellion probability of sea trade with sail. The majority of increase in sea trade (0.319/0.429=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.

Alternative changes in inland transport Another potential mechanism that can confound the "roadblock effect" is the shifting course of the Yellow River in 1855, which resulted in flooding and paralysis of the northern part of the Grand Canal.[41] While the flooding had been a recurring issue throughout the history, the 1855 diversion of the river was seen as one of the repercussions of the rebellion. Faced with escalating military expenses, the Qing government struggled to allocate sufficient funds for the river maintenance (Jia, 2009; Dai, 2012). Nevertheless, I show that the "roadblock effect" is not primarily driven by this alternative impact of the rebellion.

First, I incorporate the potential change of the Yellow River in computing the transport costs. In particular, I assume that, starting from 1856, the abandoned section of the Yellow River and the part of the Grand Canal susceptible to flooding by the new river (i.e., the part in Shantung province) became unusable. To make estimation results comparable, I maintain the same classification of "roadblocked" pairs as in the previous analysis. Hence, location pairs that could be affected by both roadblocks and the Yellow River are classified as "roadblocked," whereas those affected solely by the Yellow River are deemed "non-roadblocked." Column (4) of Table A.3 shows that the model that incorporates the increase in inland transport costs due to potential changes in the Yellow River and the Grand Canal yields predictions very similar to the baseline predictions in Column (1).

Second, I estimate the effect by dropping all observations susceptible to potential changes in the river and canal, including those that could be affected by both roadblocks and the river and canal.

[41]The Yellow River, second only to the Yangtze River in length in China, derives its name from the yellow sediments carried into its middle stream from loess deposits.

This results in a smaller sample of 62 location pairs. This sub-sample of "roadblocked" pairs was affected only by rebellion occupation and unaffected by potential changes in the Yellow River and the Grand Canal. Their post-rebellion probability of sea transport stands at 0.639 in the data, significantly and statistically higher than the "non-roadblocked" pairs. While the model prediction based on parameters estimated from the full sample slightly overestimates the post-rebellion probability for this sub-sample of "roadblocked" pairs (25 out of 36 pairs instead of 23 out of 36), the sub-sample of "non-roadblocked" pairs exhibits a comparable probability of sea trade after excluding observations susceptible to potential changes in the canal and river.

In summary, the "roadblock effect" remains robust when considering potential increases in transport costs due to changes in the river and canal. It also holds true when restricting observations to those solely affected by rebellion occupation. This provides evidence that the "roadblock effect" is not primarily driven by the shifting course of the Yellow River during the rebellion.

Geography Trade routes between more distant location pairs carry a higher probability of traversing any given region, including those affected by rebellion. At the same time, regions that are farther apart may be more inclined to adopt sea trade. Naturally, sea transport offers a lower variable cost compared to inland transport, providing an advantage for long-distance trade. Additionally, a fixed cost associated with transition between land and sea transport further enhances the advantage of sea transport in longer distance trade. To the extent that the models in Sections 5 and 6 have correctly account for the effect of relative transport costs and the cost of transition and that they are primary determinants of transport modes, this should not impact the estimated effect of roadblocks. That is why, rather than relying on geodesic distance or borrowing transport cost parameters from other contexts, this paper develops a method in Section 5.2 to precisely estimate transport costs specific to China during this period.

However, a concern may arise that other factors could influence why location pairs farther apart engage in sea trade, and the estimated "roadblock effect" is capturing these factors instead. To address this, I narrow my analysis to a smaller sample comprising only "roadblocked" pairs with similar geodesic distance, as well as differences in latitude and longitude, to "non-roadblocked" pairs. Using propensity score matching, I pair each "non-roadblocked" pair with a "roadblocked" pair based on log distance, absolute differences in longitude, and latitude.[42] Table A.5 shows that, following this matching process, the two groups exhibit comparability in both geodesic distance and differences in longitude and latitude. Additionally, the transport costs by sea and land, which are inputs of the model but not directly used for propensity score matching, are also balanced after the matching procedure.

[42]Because "roadblocked" observations include both longer distance and shorter distance trade pairs while "non-roadblocked" pairs consist mainly of shorter distance trade pairs, I employ a matching approach where each "non-roadblocked" pair is matched with a "roadblocked" pair, rather than vice versa.

Column (6) of Tables A.3 and A.4 reveals that for the matched "roadblocked" pairs, their probability of sea trade after the rebellion is 0.75 in the data, significantly higher than the 0.375 probability for "non-roadblocked' pairs. The difference in sea trade probability suggests that the "roadblock effect" is not driven by "roadblocked" observations that are farther apart than "non-roadblocked" pairs. Moreover, the model prediction, based on parameters estimated using the full sample, aligns well with the mean probability of sea trade for this sub-sample. 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 an 82% increase in the probability of sea trade.[43]

8. Concluding remarks

The concentration of population in a location can persist even when the location is no longer advantageous due to technological changes. The introduction of new transport technologies, such as steamships, has the potential to enhance the relative attractiveness of coastal locations by reducing trade costs with other locations, thus increasing their market access. However, new technologies usually require large sunk investments, which can hold them back. Failure to adopt superior technologies can, in turn, trap populations in sub-optimal locations.

This paper uses the Taiping Rebellion in 19th-century China to study population changes and technological adoption following a large but temporary shock to city sizes. While in many cases relative city sizes remain persistent after war shocks, the Taiping Rebellion resulted in a permanent loss of about 59% in population for war-ravaged locations and a 22% increase in coastal locations. 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 use of sea transport was limited despite its potential. This was due to the persistence of population away from the coastal locations and the lack of state investment in sea trade. Even when superior maritime technology became available, adoption was not guaranteed. I show that the spatial variations in maritime trade immediately after the rebellion was driven by blocked land routes and the feasibility of sea alternatives. The rebellion disrupted regular inland trade routes, forcing some location pairs to search for alternatives, and it triggered substantial investment to facilitate sea trade, notably using steamships. After the rebellion, with sunk investment already incurred, many trade routes permanently shifted to the 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 the "roadblock effect." The model estimates find a

[43] Both the data and the model indicate a slightly higher probability of sea trade for this subsample of "roadblocked" pairs that are less distant. This is because long-distance inland trade can more often use courier routes.

substantial cost associated with initiating maritime trade. The counterfactual analysis using the model suggests that the likelihood of maritime trade post-rebellion would be much lower in the absence of roadblocks.

The findings of the paper highlight 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, "Modern spatial economics: a primer," in "World Trade Evolution," Routledge, 2018, pp. 435–472.

Bakker, Jan David, Stephan Maurer, Jörn-Steffen Pischke, and Ferdinand Rauch, "Of mice and merchants: Connectedness and the location of economic activity in the Iron Age," *The Review of Economics and Statistics*, 09 2021, *103* (4), 652–665.

Banister, T. Roger., *The coastwise lights of China an illustrated account of the Chinese Maritime Customs Lights Service*, Shanghai: Inspectorate General of Customs, 1932.

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.

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

Brezis, 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.

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.

_ **and Treb Allen**, "Persistence and path dependence in the spatial economy," *NBER Working Paper*, 2020, *28059.*

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, *Zhonguo 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.

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.

Hsiao, Liang-Lin, *China's foreign trade statistics, 1864-1949*, East Asian Research Center, Harvard University distributed by Harvard University Press, 1974.

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 Tongwaxiang," *Journal of Historical Science*, 2009, (12), 56–63.

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, Kwang-Ching, "Steamship Enterprise in Nineteenth-Century China," *The Journal of Asian Studies*, 1959, *18* (4), 435–455.

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.

_ **and Nancy Qian**, "The potato's contribution to population and urbanization: evidence from a historical experiment," *The Quarterly Journal of Economics*, 2011, *126* (2), 593–650.

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.

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.

Wang, Yehchien and Guoshu Huang, "Shiba Shiji Zhongguo Liangshi Gongxu de Kaocha [An Investigation into the China's Suppy-Demand of Grain in the Eighteenth Century]," *Collection on Rural Economic History of Modern China*, 1989.

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.

Online appendices

AppendixA. Table appendix

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

Postal name (Romanisation)	Current location (Pinyin)	Postal name (Romanisation)	Current location (Pinyin)	Postal name (Romanisation)	Current location (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	Kuanghsi	Guangxi	Tientsin	Tianjin
Fukien	Fujian	Kueichow	Guizhou	Wenchow	Wenzhou
Hankow	Hankou	Kweiping	Guiping	Yangchow	Yangzhou

Table A.2: Estimation of roadblock effect with different sail to steam cost ratios

	(1)	(2)	(3)	(4)	(5)	(6)
Sail to steam cost ratio	1	1.25	1.5	1.75	2	2.25
Panel A: Estimation on model parameters						
γ/σ (elasticity of trade)	0.446	0.475	0.502	0.524	0.542	0.561
	(0.096)	(0.101)	(0.106)	(0.111)	(0.116)	(0.121)
i^s (transformed sunk cost of new sea route)	1.875	1.977	2.017	2.044	2.062	2.084
	(0.421)	(0.399)	(0.392)	(0.388)	(0.386)	(0.386)
i^l (transformed sunk cost of new land route)	1.641	1.782	1.685	1.604	1.537	1.477
	(0.605)	(0.570)	(0.532)	(0.504)	(0.485)	(0.470)
Panel B: Prediction of mean probability of sea trade						
Roadblocked pairs:						
Pre-rebellion	0.363	0.318	0.294	0.273	0.255	0.237
	(0.020)	(0.020)	(0.019)	(0.018)	(0.018)	(0.017)
Post-rebellion with rebellion	0.651	0.677	0.680	0.682	0.684	0.686
	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.018)
Post-rebellion w/o rebellion	0.363	0.350	0.355	0.361	0.365	0.369
	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)
Observations	58	56	56	56	56	56
Non-roadblocked pairs:						
Pre-rebellion	0.420	0.363	0.336	0.312	0.290	0.271
	(0.048)	(0.045)	(0.044)	(0.043)	(0.042)	(0.041)
Post-rebellion	0.420	0.392	0.391	0.388	0.387	0.385
	(0.048)	(0.046)	(0.047)	(0.047)	(0.047)	(0.047)
Observations	30	32	32	32	32	32

Notes: *Panel A* shows estimates of parameters from the extended model developed in Section 6 under 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 from baseline model

	(1) Baseline sample	(2) Excl. Tientsin.	(3) Excl. Canton	(4) Change river	(5) Excl. river	(6) Matched sample
Roadblocked pairs:						
Data	0.696	0.705	0.674	0.696	0.639	0.750
Prediction with roadblock	0.693	0.709	0.697	0.697	0.701	0.729
	(0.021)	(0.023)	(0.025)	(0.021)	(0.027)	(0.026)
Prediction w/o roadblock	0.364	0.382	0.365	0.369	0.381	0.402
	(0.199)	(0.024)	(0.023)	(0.020)	(0.030)	(0.027)
Observations	56	44	46	56	36	32
Non-roadblocked pairs:						
Data	0.375	0.346	0.417	0.375	0.385	0.375
Prediction	0.378	0.395	0.395	0.380	0.391	0.378
	(0.045)	(0.048)	(0.051)	(0.045)	(0.043)	(0.046)
Observations	32	26	24	32	26	32

Notes: Predictions of the baseline model in Section 5 with different robustness specifications. Column (1) replicates the prediction of Table I using the baseline sample as a benchmark. Column (2) excludes trade with Tientsin, whose trade may be directly affected by the civil war. Column (3) excludes major ports for foreign trade (i.e., Canton in addition to Shanghai). Column (4) accounts for potential changes in the Yellow River and the Grand Canal during the war when calculating transport costs. Column (5) excludes all trade pairs that could potentially be affected by changes in the Yellow River and the canal. Column (6) uses a subsample matched in geographical distance, as well as differences in latitude and longitude. Bootstrap standard errors using 500 replicates in parentheses.

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

	(1) Baseline sample	(2) Excl. Tientsin.	(3) Excl. Canton	(4) Change river	(5) Excl. river	(6) Matched sample
Roadblocked pairs:						
Data	0.696	0.705	0.674	0.696	0.639	0.750
Prediction with roadblock	0.684	0.699	0.687	0.688	0.694	0.716
	(0.018)	(0.020)	(0.021)	(0.018)	(0.024)	(0.023)
Prediction w/o roadblock	0.365	0.385	0.367	0.369	0.385	0.407
	(0.021)	(0.025)	(0.024)	(0.021)	(0.031)	(0.028)
Observations	56	44	46	56	36	32
Non-roadblocked pairs:						
Data	0.375	0.346	0.402	0.375	0.385	0.375
Prediction	0.387	0.381	0.403	0.388	0.399	0.386
	(0.047)	(0.050)	(0.054)	(0.047)	(0.045)	(0.048)
Observations	32	26	24	32	26	32

Notes: Predictions of the extended model in Section 6 with different robustness specifications. Column (1) replicates the prediction of Table II using the baseline sample as a benchmark. Column (2) excludes trade with Tientsin, whose trade may be directly affected by the civil war. Column (3) excludes major ports for foreign trade (i.e., Canton in addition to Shanghai). Column (4) accounts for potential changes in the Yellow River and the Grand Canal during the war when calculating transport costs. Column (5) excludes all trade pairs that could potentially be affected by changes in the Yellow River and the canal. Column (6) uses a subsample matched in geodesic distance, as well as differences in latitude and longitude. Bootstrap standard errors using 500 replicates in parentheses.

Table A.5: Geography and transport cost

	Full sample					Matched sample				
	(1) Long.	(2) Lat.	(3) Dist.	(4) Sea	(5) Land	(6) Long.	(7) Lat.	(8) Dist.	(9) Sea	(10) Land
Roadblock	1.358	5.419	0.756	0.442	0.386	0.458	0.392	0.152	-0.075	0.043
	(0.462)	(0.871)	(0.110)	(0.182)	(0.089)	(0.432)	(0.856)	(0.115)	(0.201)	(0.099)
Constant	2.925	4.419	6.273	1.065	3.071	2.925	4.419	6.273	1.065	3.071
	(0.323)	(0.537)	(0.095)	(0.174)	(0.071)	(0.325)	(0.540)	(0.095)	(0.175)	(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 indicate the mean gap in longitude and latitude, as well as distance, sea transport, and land transport costs for “non-roadblocked” location pairs. The roadblock indicator shows differences in the aforementioned dimensions for “roadblocked” pairs relative to the “non-roadblocked” pairs. Robust standard errors in parentheses.

AppendixB. Figure appendix

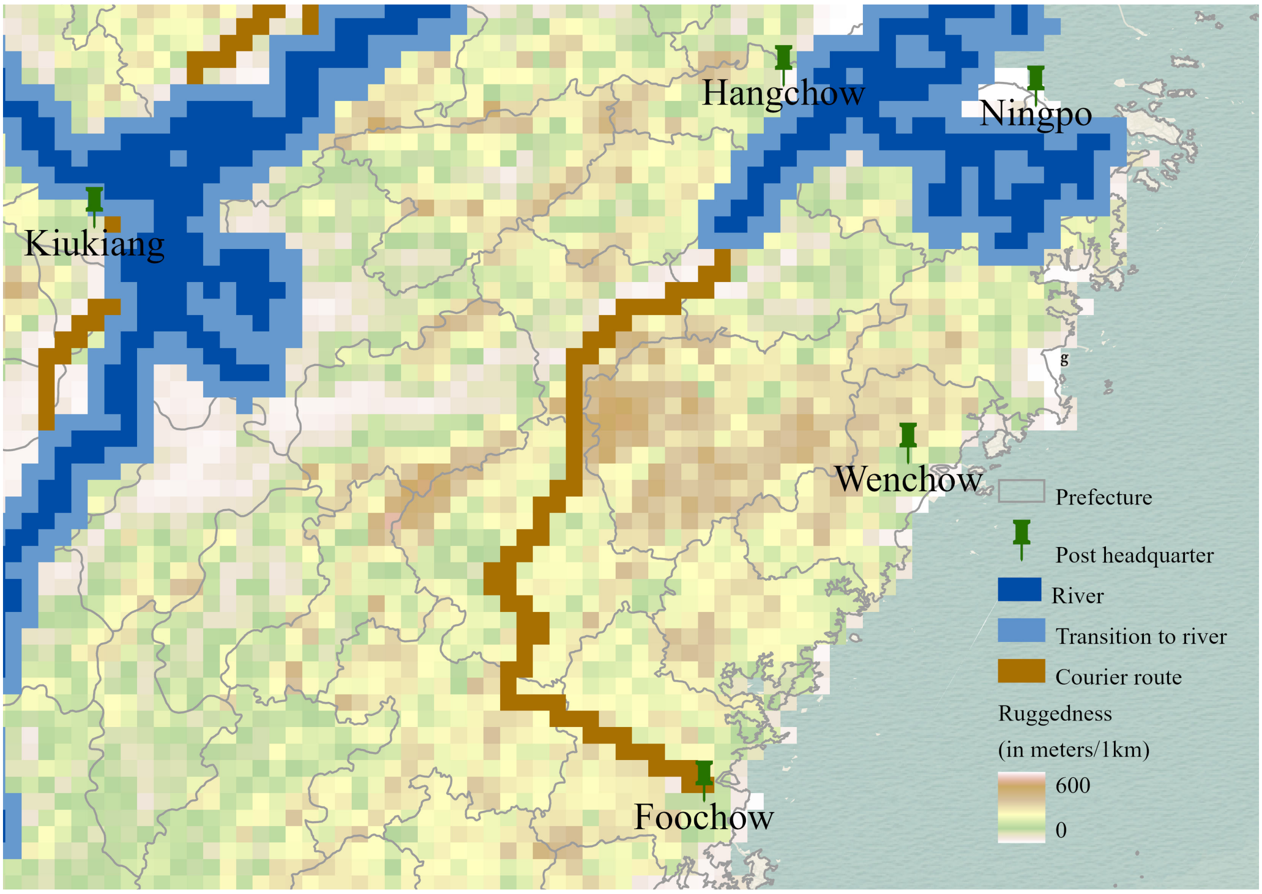


Figure B.1: Transport cost in 19th-century China

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. Estimation of population evolution

This section presents the estimation results for the evolution of prefecture populations. Table C.1 shows the reduced-form relationship between population, occupation during the rebellion and coastal location using event-study specifications. The excluded category is the data year 1851. Therefore, the coefficients represent the population change of a treated group relative to the control group compared to their difference in 1851, just before the Taiping Rebellion started. Prefecture fixed effects are included to control for time-invariant prefecture heterogeneity in population, and time fixed effects are included to account for secular trends.

Column (1) shows the results when only the interactions between year indicators and the indicator for occupied prefectures during the rebellion are included in the regression. Column (2) additionally includes the interactions between year indicators and the indicator for coastal locations. Column (3) further includes the interactions between year indicators and other locational characteristics, capturing any time-varying effects of these characteristics on population.[44] Column (3) corresponds to the estimation results of Equation 1 and its coefficients estimates are plotted in Figure 3. The results show that the population evolution of occupied and non-occupied prefectures was similar before the rebellion. However, occupied prefectures suffered a substantial population loss, which persisted until 2010. Meanwhile, coastal locations witnessed an increase in population immediately after the rebellion.

I proceed to estimate the average change in population after the rebellion using Equation 2. Similarly, this specification includes prefecture fixed effects to account for time-invariant prefecture heterogeneity in population and year fixed effects to capture secular trends over time. Additionally, by replacing interactions with year indicators with a post-rebellion indicator, we can explicitly control for differential population growth in occupied and coastal prefectures, respectively.

The estimation in column (1) includes the interaction between the post-rebellion indicator and the occupied indicator. The coefficient estimate of this interaction indicates that the population in occupied prefectures experienced an additional drop of approximately 67% after the rebellion. Column (2) further includes a post-rebellion and coastal interaction to estimate the change in coastal population after the rebellion. The estimated coefficient shows a relative 22% increase in coastal population after the rebellion.

[44] The included locational characteristics consist of the level of human capital and a set of geographic controls, encompassing terrain ruggedness, agricultural suitability of major crops and new world crops, and whether the prefecture had access to the Grand Canal.

Table C.1: Prefecture population evolution 1680–2010: event study

	Dep. var.: Ln (population)		
	(1)	(2)	(3)
1680 × occupied	-0.059	-0.058	-0.033
	(0.040)	(0.040)	(0.062)
1776 × occupied	-0.044	-0.044	-0.011
	(0.022)	(0.022)	(0.038)
1820 × occupied	-0.019	-0.020	0.017
	(0.016)	(0.016)	(0.030)
1880 × occupied	-0.680	-0.670	-0.589
	(0.091)	(0.089)	(0.106)
1910 × occupied	-0.642	-0.635	-0.559
	(0.074)	(0.074)	(0.098)
1953 × occupied	-0.603	-0.597	-0.496
	(0.084)	(0.084)	(0.113)
1982 × occupied	-0.618	-0.614	-0.497
	(0.078)	(0.078)	(0.106)
2010 × occupied	-0.587	-0.576	-0.486
	(0.087)	(0.084)	(0.110)
1680 × coastal		0.005	0.007
		(0.044)	(0.043)
1776 × coastal		-0.013	-0.027
		(0.018)	(0.020)
1820 × coastal		-0.015	-0.022
		(0.012)	(0.013)
1880 × coastal		0.240	0.216
		(0.053)	(0.059)
1910 × coastal		0.159	0.180
		(0.054)	(0.060)
1953 × coastal		0.142	0.176
		(0.073)	(0.073)
1982 × coastal		0.098	0.172
		(0.071)	(0.070)
2010 × coastal		0.251	0.270
		(0.089)	(0.085)
Year indicators × controls			Yes
Year fixed effects	Yes	Yes	Yes
Prefecture fixed effects	Yes	Yes	Yes
Number of observations	1,610	1,610	1,565
Number of prefectures	179	179	174
Number of coastal prefectures	35	35	35
R-squared	0.803	0.808	0.851

Notes: Controls in column (3) include the level of human capital, proxied by the natural logarithm number of graduates from the Imperial Exam in the 1840s, and a set of geographic controls including terrain ruggedness, agricultural suitability from FAO GAEZ v4 for major crops in historical China (i.e., rice, wheat and foxtail millet) and new world crops (i.e., potato, maize and sweet potato), and whether the Grand Canal went through a prefecture. Observations drop from column (2) to column (3) due to unavailability of the human capital measure. Robust standard errors clustered by prefecture are reported in parentheses.

Population growth in coastal regions immediately after the rebellion could be driven by other characteristics of coastal areas. While prefecture fixed effects account for population differences across prefectures in levels, these characteristics might exert a differential impact on population growth following a large shock to population sizes. For instance, coastal areas may have higher agricultural productivity, attracting migration after the rebellion. In a broader sense, as locational fundamentals evolve over time, prefectures with specific characteristics may undergo varying levels of population growth at different points in time.

Table C.2: Prefecture population evolution 1680–1953

	Dep. var.: Ln (population)				
	(1)	(2)	(3)	(4)	(5)
Post-reb. × occupied	-0.668	-0.659	-0.592	-0.681	-0.633
	(0.081)	(0.079)	(0.101)	(0.090)	(0.124)
Post-reb. × coastal		0.217	0.220	0.365	0.361
		(0.055)	(0.061)	(0.066)	(0.102)
Post-reb. × coastal × domestic trade				0.379	0.173
				(0.067)	(0.095)
Post-reb. × coastal × foreign trade				-0.089	-0.078
				(0.047)	(0.050)
Occupied linear time trend	0.0005	0.0004	0.0004	0.0006	0.0004
	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0004)
Coastal linear time trend	0.0008	-0.0002	-0.0001	-0.0003	-0.0004
	(0.0003)	(0.0002)	(0.0002)	(0.0006)	(0.0006)
Year indicators × controls			Yes		Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Prefecture fixed effects	Yes	Yes	Yes	Yes	Yes
Number of observations	1,253	1,253	1,218	1,064	1,029
Number of prefectures	179	179	174	152	147
Number of coastal prefectures	35	35	35	8	8
R-squared	0.623	0.633	0.696	0.607	0.682

Notes: In columns (4) and (5), trade val. is the natural logarithm of domestic maritime trade value relative to its average in 1867, which is only available for 8 treaty ports out of 35 coastal prefectures (and the interaction is 0 for all 144 non-coastal prefectures). In columns (3) and (5), controls include a proxy for human capital, an indicator for whether the Grand Canal passed through, terrain ruggedness, and categorical variables for agricultural suitability from FAO GAEZ v4 for major crops in historical China and new world crops. Robust standard errors clustered by prefecture are reported in parentheses.

Therefore, in column (3), I include a variety of locational traits (i.e., the level of human capital and a set of geographical controls) interacted with all year indicators.[45] The positive coefficient estimate of the interaction between the post-rebellion indicator and coastal dummy remains robust after accounting for any differential impacts of those locational characteristics on population across

[45] This specification is similar to Nunn and Qian (2011).

different data years, showing a similarly 22% increase in the relative population level in coastal prefectures after the rebellion.

To provide suggestive evidence that population relocations might be linked to a shift in trade activities toward the coast during the rebellion's disruption of regular inland trade routes, column (4) introduces an additional interaction. The post-rebellion $\times$ coastal interaction is further interacted with the natural logarithm of the value of domestic and foreign trade by sea relative to its average in 1867, as reported by the CMCS.[46] This trade value, mainly consisting of steamship trade, is used to examine whether port cities experiencing a surge in steamship trade witnessed an increase in population after the rebellion.

Since trade data are only available at treaty ports, which opened to trade with Western countries and were usually large ports, the number of observations drops. In this subset of large ports, the post-rebellion $\times$ coastal interaction indicates an average population increase of about 37% after the rebellion. Furthermore, the post-rebellion$\times$coastal$\times$domestic trade triple interaction indicates that a 1% increase in domestic sea trade, as reported by the CMCS, was associated with a 0.38% rise in coastal population after the rebellion. However, the association between coastal population and foreign maritime trade was much weaker, as indicated by the post-rebellion$\times$coastal$\times$foreign trade triple interaction.[47] This suggests that the increase in coastal population during this period was primarily influenced by patterns of domestic sea trade.

On the basis of column (4), column (5) further controls for locational characteristics interacted with year dummies. The overall picture remains consistent, indicating that the positive association between population and domestic maritime trade is not confounded by other locational characteristics that might simultaneously correlate with maritime trade.

[46]The value of trade recorded by the CMCS was denominated in Customs Taels, a silver-based monetary unit uniform across all ports in 1867, except for the southern ports that reported their trade value in Mexican dollars. I convert these values to Customs Taels using the contemporary conversion rate.

[47]While both the volume of trade and population growth could be greater in larger cities, using the value of trade per capita yields very similar results.

AppendixD. Dynamic sunk investment decision

This section develops a two-period dynamic trade model with modal and route choice, which allows firms to take into account that their sunk investment for sea transport during the rebellion can also affect their transport cost after the rebellion to various degrees. The model developed in Section 5.1 is a specialised case where the discount factor $\beta = 0$. In other words, in the baseline model, firms are myopic about their investment decisions, which may be justified by the uncertainty of the status of sea trade after the rebellion.

In this more generalised setting, the iceberg trade costs for inland and sea transport have the same structure as in Equation 4, 5 and 6. Additionally, forward-looking firm owners decide whether or not to start steamship trade during the rebellion, taking into account that the sunk investment made during the rebellion can permanently reduce the sea transport cost afterwards. Assume that firm owners have a log utility with a discount factor $\beta \in [0, 1]$, and therefore they will minimise the effective transport cost $\gamma \ln(D^m_{ij,t} I^m) - \eta^m_{ijt}$ for the whole period. If firms make sunk investment for sea transport during the rebellion, the expected transport cost they face for the whole period effectively is:

$$\gamma \ln D^s_{ij,reb} + \sigma i^s - \eta^s_{ij,reb} + \beta\big(p(\gamma \ln D^s_{ij,post}) + (1-p)(\gamma \ln D^l_{ij,post})\big), \tag{D1}$$

where $p = \Phi(\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,post}}{D^s_{ij,post}})$ is the probability of using sea transport when the sunk investment has already been made. The expected life-time transport cost of sticking to the inland route during the rebellion is:

$$\gamma \ln D^l_{ij,reb} + roadblock_{ij} \times \sigma i^l - \eta^l_{ij,reb} + \beta\big(q(\gamma \ln D^s_{ij,post} + \sigma i^s) + (1-q)(\gamma \ln D^l_{ij,post})\big), \tag{D2}$$

where $q = \Phi(\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,post}}{D^s_{ij,post}} - i^s)$ is the probability of using sea transport when the sunk investment has not yet been made. During the rebellion, after observing transport cost shocks $\eta^l_{ij,reb}$ and $\eta^s_{ij,reb}$, forward-looking firm owners will invest and start sea transport if $D1 < D2$. Therefore, the probability of sea trade during the rebellion is:

$$\Pr(S_{ij,reb} = 1) = \Phi\Big(\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,reb}}{D^s_{ij,reb}} + roadblock_{ij} \times i^l - (1-\beta q) i^s + \beta(p-q)\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,post}}{D^s_{ij,post}}\Big).$$

Note that compared to Equation 8, the dynamic model has two additional terms $\beta q i^s$ and $\beta(p - q)\frac{\gamma}{\sigma} \ln \frac{D^l_{ij,post}}{D^s_{ij,post}}$, which take into account the cost savings of (1) future sunk investment for sea transport and (2) a higher probability of using sea transport in the post-rebellion era, respectively, by investing in sea trade during the rebellion. Given the same level of sunk investment cost, firms are more likely to start sea trade during the rebellion if they are forward-looking. Therefore, the implied sunk cost from the observed trade will be larger if β, the discount factor, is larger than 0. Formally, we can

Table D.1: Estimation of roadblock effect: dynamic model prediction

Panel A: Estimation results on trade model parameters		
	Parameter	Value
Elasticity of probability of sea trade w.r.t. land/sea cost ratio	γ/σ	0.483
		(0.094)
Sunk cost of new sea route (transformed)	i^s	2.106
		(0.408)
Sunk cost of new land route (transformed)	i^l	1.218
		(0.460)
Observations		88

Panel B: 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	0.682	0.368
			(0.019)	(0.021)
Unaffected pairs	32	0.375	0.385	0.385
			(0.047)	(0.047)

Notes: In *Panel A*, the estimates imply that the cost ratio of new sea route over existing route is 78 ($I^s = \exp(i^s \times \sigma/\gamma)$). The implied cost ratio of new land route over existing one is 12 ($I^l = \exp(i^l \times \sigma/\gamma)$). Standard errors in parentheses. In *Panel B*, the model prediction is based on parameters estimates of *Panel A* for the dynamic model with the discount factor equal to 0.9. 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 factoring in the potential cost savings for future sea trade, 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.

write down the post-rebellion probability of sea trade using Equation 9, and it becomes:

$$\begin{aligned}\Pr(S_{ij,post}=1) =& \Phi\Big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,reb}}{D^s_{ij,reb}} + \text{roadblock}_{ij}\times i^l - (1-\beta q)i^s + \beta(p-q)\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}}\Big)p \\ &+\Big(1-\Phi\Big(\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,reb}}{D^s_{ij,reb}} + \text{roadblock}_{ij}\times i^l - (1-\beta q)i^s + \beta(p-q)\frac{\gamma}{\sigma}\ln\frac{D^l_{ij,post}}{D^s_{ij,post}}\Big)\Big)q,\end{aligned}$$

which is the same as Equation 10 if $\beta = 0$.

Table D.1 shows the estimation result of the dynamic model when we set the discount factor β to be 0.9. As expected, the implied cost for initiating sea trade is higher when firms are forward-looking, taking into account potential future cost savings for sea trade for making the sunk investment during the rebellion. The estimate indicates that the cost of a new sea route is approximately 78 times that of an existing one ($I^s = \exp(i^s \times \sigma/\gamma)$), while the cost ratio for a new inland route versus an existing one is around 12. However, the impact of roadblocks remains consistent with the baseline model, regardless of whether firms are forward-looking or not. The model predicts a 0.682 probability of sea trade for "roadblocked" pairs after the rebellion, compared to a 0.368 probability in a counterfactual scenario where roadblocks did not occur. This counterfactual probability of sea trade for "roadblocked" pairs closely mirrors that of non-affected pairs, standing at 0.385.

AppendixE. Estimation of historical migration gravity

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

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

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, η_{it} and ψ_{jt} respectively, can vary over time. κ measures the elasticity of migration flows to migration costs D_{ijt}. ν_{ijt} is an error term.

Historical migration flows are rarely available. Here I construct a new dataset of migration flows in historical China by leveraging research that traces the origins of modern villages. Throughout history, many villages in China were established by clans from external regions during periods of deteriorating local living conditions. By virtue of the tradition of keeping a rich family history, Cao (1997) was able to collect information on migration villages, including the place of origin of their founders and the timeframe of their formation. Using the number of villages at location i founded by people from location j at time t, I proxy for the migration flow from i to j at t.[48]

The following tables present the estimation results on migration gravity, using the number of migration villages as a proxy for migration flows. Table E.1 shows results by using transport costs estimated in Section 5.2 to measure migration cost. Table E.2 shows results using distance.[49] [50]

Column (1) of Table E.1 shows that migration flows responded negatively to transport costs, as expected, with an elasticity of about -1.5. This elasticity is 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. While the research of Cao (1997) predominantly focuses on migration villages in Kiangsi province, it also encompasses nearby prefectures in other provinces. To account for potential differences in migration patterns to prefectures outside Kiangsi, Column (2) restricts the sample to migration flows to solely Kiangxi province, yielding an elasticity of approximately -1.3. In 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: column (3) focuses on migration villages formed before 1722, and column (4) on villages formed after 1723. The results show that over

[48]The implicit assumption here is that the initial population when the villages were founded was similar across villages. For a subset of villages where their recent population is available, I calculated the implied annual population growth rate and confirmed that this is a reasonable assumption.

[49]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, each observation is divided by the number of years within its corresponding time frame. This adjustment accounts for that the number of migrants/migration villages will be larger for longer time periods.

[50]The destinations are identified at the county level. For the origins, the province is known for the majority of villages, while for others, the prefecture or the county is known. In cases where only the origin province is available, I calculate transport cost/distance to the nearest prefecture in the province. Notably, the elasticity estimate using the average cost/distance to different prefectures in the province is statistically significant but substantially larger (in absolute value) than the rest of the sample, where information about the prefecture/county is available.

time, migration flows became more responsive to transport costs with the elasticities becoming increasingly negative.

Table E.2 presents the results on migration elasticities using distance instead of transport costs. In the extreme scenario where distance does not correlate with transport costs, and if we assume that migration flows correspond are solely influenced by transport costs, the relationship between migration and distance would be null. In the intermediate scenario, where distance is considered a noisy predictor of transport costs, the negative association between migration flows and transport costs using distance as a proxy would be attenuated. The results in Table E.2 suggest this may be the case. Distance reasonably approximates transport costs, as migration flows still exhibit a decrease with distance. However, the estimated migration elasticity using distance is reduced by 50% compared to using transport cost.

Table E.1: Migration gravity: transport cost

	Dep. var.: Ln (number of villages)			
	All (1)	Kiangsi only (2)	1643-1722 (3)	1723-1912 (4)
Ln (transport cost)	-1.464 (0.385)	-1.306 (0.333)	-0.975 (0.451)	-1.679 (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. For 12 observations, the only information available is that the foundation year falls within the period between 1643 and 1796, and therefore they are not included in the columns (3) and (4). Robust standard errors in parenthesis.

Table E.2: Migration gravity: distance

	Dep. var.: Ln (number of villages)			
	All	Kiangsi only	1643-1722	1723-1912
	(1)	(2)	(3)	(4)
Ln (distance)	-1.016	-0.873	-0.773	-0.965
	(0.385)	(0.333)	(0.451)	(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. For 12 observations, the only information available is that the foundation year falls within the period between 1643 and 1796, and therefore they are not included in the columns (3) and (4). Robust standard errors in parenthesis.