# The roadblock effect

Yan Hu [*] [#]
*University of Edinburgh*

December 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 inland 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 inland 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, trade costs, modal choices
JEL classification: F14, R12, R40

[*] I am indebted to Diego Puga for his support and guidance. I am very 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, Stephan Maurer, Joan Monras, David Nagy, Andrei Potlogea, Stephen Redding, Daniel Sturm, Matthew Turner, Maisy Wong, Noam Yuchtman, Yanos Zylberberg 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, Bristol Applied Economics Meetings, IEB Workshop in Urban Economics, JIE Summer School 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

Trade costs affect relative city sizes by altering location's access to larger markets (Krugman, 1993). Trade costs change over time when new transport technologies become available (Pascali, 2017), and locations may benefit from these changes through increased market access. 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.

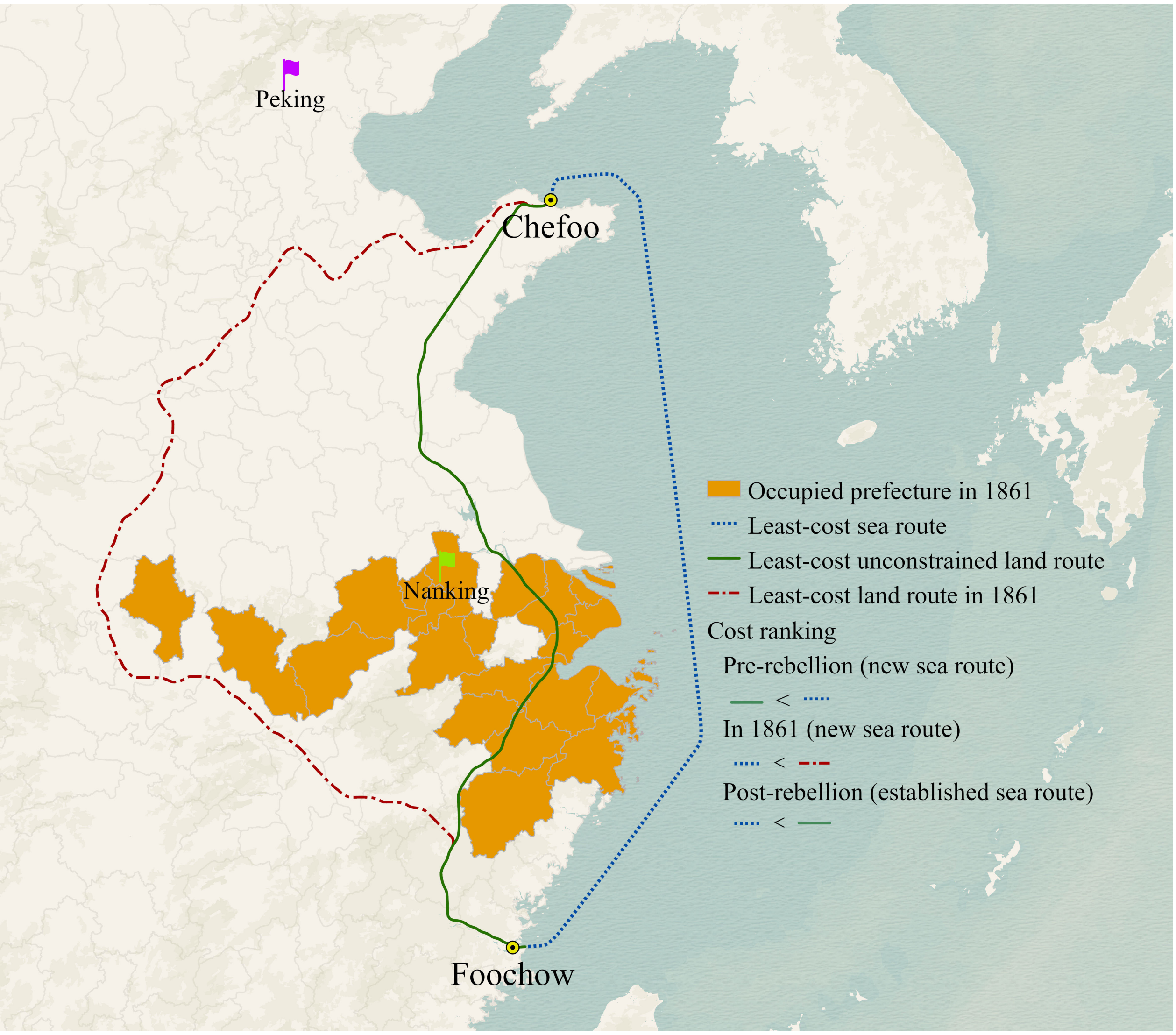
This study exploits the heterogeneous, and plausibly exogenous, incentives to adopt steamship transport through the route and modal changes induced by shocks of civil conflict elsewhere. The Taiping Rebellion was the largest peasant revolt in the history of China. The rebellion caused direct damage and casualties in the war-affected area, concentrated mainly in regions under rebel occupation. At the same time, the rebellion temporarily disrupted the transport network, leading to permanent changes in trade costs and market access through technological adoption and investment.

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. This in turn catalysed a shift of population towards port cities. An example of this can be seen in Figure 1. The figure shows the least-cost inland route (solid line) 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 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 by the two dashed line. According to my estimates presented below, the additional cost this detour entailed was large enough to instead incur the sunk cost of establishing a steamship route between these two cities. Although the unconstrained regular land route became once again usable after the rebellion, with the sea route already in operation, steamship trade remained a less costly option. Indeed the data show that there was active sea trade between the two ports immediately after the rebellion. In contrast, regular 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 the relative population size of locations occupied during the rebellion dropped permanently by 44.7%, while

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

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.

being located on the coast was associated with a 24.6% increase in population. This increase was even more pronounced in coastal cities that experienced 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, interregional trade 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. In general, the identification strategy also accounts for common shocks to sea transport over time, such as any changes in government attitudes towards maritime trade.

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.[2] Given only sea trade immediately after the rebellion is observed in the data, I add modal choice (by land or by sea) to an otherwise standard trade model.[3] 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 implication of the model is consistent with the reduced-form evidence that location pairs with higher increases in inland transport costs due to roadblocks had a higher post-rebellion probability of maritime trade.

The model yields a probability of sea trade after the rebellion for each location pair, which we observe in the data, given relative costs of land to sea transport. The relative cost of using different 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 methods,

[2] 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.

[3] 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.

variations in cost due to road types and the terrain, as well as 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 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 0.692 compared with 0.379 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, 0.358 instead of 0.692. I show that the "roadblock effect" can qualitatively and quantitatively account for the changes in relative population sizes following the economic geography framework of Redding and Sturm (2008).

The "roadblock effect" is not confounded by alternative factors contributing to the post-rebellion surge in maritime trade. Section 6 develops 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 trade prior to the rebellion were generally low and not systematically different for "roadblocked" and "non-roadblocked" pairs, the estimated "roadblock effect" remains similar. The prediction of the extended model for the level of sail trade before the rebellion aligns with historical figures, despite its cost advantage over inland transport, supporting the model's structure and underlying mechanisms in explaining modal choices.[4]

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. In fact, the framework allows flexibly considering other technological or institutional changes that potentially contributed to a higher tendency for sea trade before and after the rebellion. The quantification exercise shows that the combined contribution of other factors is relatively small, with approximately 78% of the increase in maritime trade after the rebellion attributable to the "roadblock effect." Then, in Section 7, I provide more details why 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 historical dependence of the spatial distribution of economic activities (e.g., Davis and Weinstein, 2002; Redding, Sturm and Wolf, 2011; Bleakley and Lin, 2012; Michaels and Rauch, 2017; Allen and Donaldson, 2020). A key contribution of my framework is that, in addition to the direct channel operating through temporary

[4]This holds in a context where the government neither supported nor banned domestic sea trade. More historical contexts are provided in Section 2.3.

population shocks, I highlight an indirect channel operating through temporary roadblocks and the incentives they provide to invest in new transport technologies.

In this regard, the paper connects to the literature on the importance of trade costs and market access in determining the distribution of economic activities across locations (e.g., Head and Mayer, 2004; Redding and Venables, 2004; Redding and Sturm, 2008; Faber, 2014; Baum-Snow, Brandt, Henderson, Turner and Zhang, 2017; Donaldson, 2018; Barjamovic, Chaney, Coşar and Hortaçsu, 2019; Heblich, Redding and Sturm, 2020; Bakker, Maurer, Pischke and Rauch, 2021; Ellingsen, 2021; Flückiger, Hornung, Larch, Ludwig and Mees, 2021; Maurer and Rauch, 2022; Tsivanidis, 2022). An important challenge when trying to evaluate the consequences of transport improvements is the endogeneity of infrastructure investments. Here, I model decision to adopt steamships using the exogenous variation in incentives to invest in sea transport driven by conflicts elsewhere.[5]

Hence, the paper contributes to a growing literature on endogenous transport costs and investment (e.g., Brancaccio, Kalouptsidi and Papageorgiou, 2020; Fajgelbaum and Schaal, 2020; Santamaria, 2020; Allen and Arkolakis, 2022; Wong, 2022; Ducruet, Juhász, Nagy and Steinwender, 2024), as well as to a growing body of work that considers transport mode and route choices (Hummels and Schaur, 2013; Allen and Arkolakis, 2014; Coşar and Demir, 2018; Glaeser and Poterba, 2020; Fan, Lu and Luo, 2023; Bonadio, 2024; Fuchs and Wong, 2024). Similar to the framework developed by Coşar and Demir (2018), I incorporate mode-specific variable costs and fixed costs. The sunk fixed costs identified in this paper are bilateral and intertemporal, highlighting the importance of relationship-specific investments or trade-promoting capital in facilitating trade (Combes, Lafourcade and Mayer, 2005; Head, Mayer and Ries, 2010; Marczinek, Maurer and Rauch, 2022).

This is closely related to a debate on the persistence of city locations that pertains to the failure to adopt new technologies, even when they are clearly superior to older ones. This may be due to the trade-off between the sunk costs of developing or 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 the opposite direction: The temporary blockage of inland trade routes during the Taiping Rebellion triggered substantial investments to facilitate the use of steamships. This had long-term consequences for the distribution of population across Chinese cities, in addition to the direct shocks of the rebellion.

[5] Juhász (2018) also looks at war-induced trade blockades 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 consisted mainly of agricultural products, dictated by local conditions. Consequently, locations responded to blockades by finding alternative trade routes rather than innovating in production.

## 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.[6] About one out of every six people in China (approximately 70 million) died during the rebellion (Ge, Hou and Zhang, 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 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).[7] 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.[8] 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 that had previously traversed the rebellion-occupied region. This disruption was substantiated by contemporary reports from local officials to the imperial court. The governor of Anhwei province wrote in 1854 (Liao, 2010):

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

[6] 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 localised.

[7] 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.

[8] A contemporary witness from the imperial capital reported shortages in food and a surge in prices due to the rebellion (Ni, 2005).

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 by light boundaries and shaded if occupied by the Taiping Rebellion. China's historical postal districts delimited by dashed lines, with district headquarters marked by pins.

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

> *Kuachow [in Yangchow] and Chinkiang are still under rebel occupation, trade routes are not accessible.*

In addition to reports of roadblocks from the Qing officials, the Chinese Maritime Customs Service took 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 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 established 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 Kiukiang's all 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 (which ruled from 1644 to 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, the recovery of population in coastal regions was sluggish.[12]

---

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

[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 despite some restrictions imposed by the Canton system (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.

At the same time, domestic trade relied largely on a combination of land transport and inland waterways. 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, interregional 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 transport, whether by sail or steam, compared to inland routes, trade patterns remained remarkably unchanged for a decade before the Taiping Rebellion. 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 foreign 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. The launch of steamships was unique in the Chinese context because, unlike in some Western countries, it did not receive backing from government capital or public subsidisation (Liu, 1959).[14] During and immediately after the rebellion, coordinated and systematic investment in maritime trade was largely absent, with minimal state involvement. Consequently, the sunk costs required to initiate bilateral maritime trade had to be borne by merchants and shipping companies. According to the Chinese Maritime Customs Service, local "sporadic attempts were being made everywhere" (Banister, 1932).

On the other hand, the Qing government had long favoured and relied on inland trade, developing and maintaining an extensive inland transport network (Deng, 2009), including the costly upkeep of the Grand Canal (Dai, 2012). Having (almost) entirely shut down sea trade during the early years of the dynasty, many locations would face substantial sunk costs to restart maritime trade after the

[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, 69.6% "roadblocked" location pairs had sea trade, compared to 37.5% "non-roadblocked" location pairs. The extended model developed in Section 6 can match both the long-run pre-rebellion sail trade levels and the heterogeneous pattern of post-rebellion sea trade.

[14]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).

lifting of the sea ban.[15]

This paper focuses on the adoption of steamships within China's domestic trade. Accordingly, the associated trade costs pertain to establishing bilateral steamship links for domestic trade between treaty ports.[16] The bilateral fixed cost to establish a new trade route is related to the concept of "trading capital," introduced by Head, Mayer and Ries (2010). They show that independence leads not to an immediate, but to a steady decline in bilateral trade, suggesting a gradual depreciation of trade-enhancing capital. Beestermöller and Rauch (2018) classify trading capital into (1) physical capital, which in the case of roadblocks involves investigative trips to scout new routes and building logistics for sea trade, and (2) capital related to personal contacts and relationships.[17] With modal choices, locations may need to invest in establishing new contacts for sea trade while severing existing relationships with incumbent trade networks. If dismantling existing trade modes is costly, switching to alternatives may require substantial shocks.

The opening of treaty ports may have reduced the sunk costs needed to initiate sea trade, due to port-level investments made to facilitate foreign trade.[18] Therefore, the sunk cost that the paper identifies captures the costs of establishing a bilateral sea route for domestic trade after the treaty ports were established. In Section 5 and 6, I consider counterfactual scenarios in which treaty ports were established but roadblocks did not occur, to quantify the effect of roadblocks.

## 3. Data

To identify the "roadblock effect," whereby blocked inland routes during the Taiping Rebellion catalysed the reallocation of trade and populations to the coast, I assemble 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). Founded in 1854 by foreign consuls in Shanghai to collect unpaid tariffs during the Taiping Rebellion, the CMCS supervised tariff collection for

[15] This scenario contrasts with the global adoption of steamships in international trade. For many cross-country trading pairs, sea transport was the only feasible option, with sail trade already well-established. As steamships became cost-competitive, their adoption was relatively smooth. Pascali (2017) shows that the adoption of the steamship significantly impacted global trade patterns, contributing to approximately 50% of the trade boom during the first wave of globalisation.

[16] As of 1867, foreign goods flowing between any two Chinese mainland ports were practically zero, except for routes involving Shanghai, which are excluded from the analysis.

[17] In addition, Beestermöller and Rauch (2018) consider trading capital in the form of historical and cultural legacies, which explains the initial surplus trade between countries of the former Austro-Hungarian monarchy following the fall of the Iron Curtain and the end of their separation over four decades, as well as the subsequent decline of the surplus trade.

[18] At the port level, the values of foreign trade and the number of years since the treaty ports opened are similar between "roadblocked" and unaffected pairs. This holds true by looking at the origin, destination separately, and the sum or the product of the two.

foreign exports, imports, and domestic maritime trade at treaty ports. The statistics department of the CMCS published trade reports with strict standards, ensuring consistency across ports and over time. Cheng (1956) notes that the trade data from the CMCS is "the only reliable and comprehensive statistical information" to use in studying China's historical trade. 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. Maritime trade by Chinese sailing ships (e.g., junks) is not reported for the majority of years and ports. Inland trade is not reported.

To address the data limitation of only observing sea trade immediately after the rebellion, I embed modal and route choices into a standard trade framework to model the probability of observing sea trade. 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 foreign sailing ships 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 model whether locations had established active bilateral sea trade. The baseline model focuses on the 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 the intermediate inputs of transport costs by least-cost routes in 19th-century China. While the literature usually derives relative 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, established by the CMCS in 1896 to deliver mail services to the general public. The services expanded from postal district headquarters, typically a treaty port, to inland areas through an extensive delivery network (See Figure 2 for postal districts and their headquarters).

The Imperial postal office also introduced a transliteration system for Chinese place names known as postal romanisation. In this paper, historical sites are referred to using postal romanised place names as they appear in trade reports. The mapping from these postal names to their current locations is provided in Table A.1.

The inland transport cost is modelled to vary with road types (i.e., courier routes and navigable rivers) and the 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 roadblock measures for each location pair from 1853 to 1864, using occupied regions for each year and optimal transport routes.[20]

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 for 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 to the historical prefecture boundaries in 1820, obtained from the China Historical Geographic Information System (CHGIS V4).[21]

## 4. The evolution of population

As 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 fully recovered 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 an effect of the rebellion through the "roadblock effect," which can qualitatively and quantitatively explain the population relocation to coastal regions during this period.

[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, Kuanghsi 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.

[21]The China Historical Geographic Information System is a historical GIS project collaboratively developed by the Harvard Yenching Institute and Fudan Center for Historical Geography.

I estimate the following reduced-form relationship between population, occupation during the rebellion, and coastal location using an event-study specification:[22]

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

where $\text{pop}_{it}$ 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 at the outset of the Taiping Rebellion;[23] 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 $\gamma_t$ 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; $\psi_i$ is a prefecture fixed effect; $\kappa_t$ 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 $\epsilon_{it}$ is an error term. Each observation is one of the 179 prefectures from 13 provinces in the core region of 19th-century China in each data year.[24] Robust standard errors are clustered at the prefecture level.

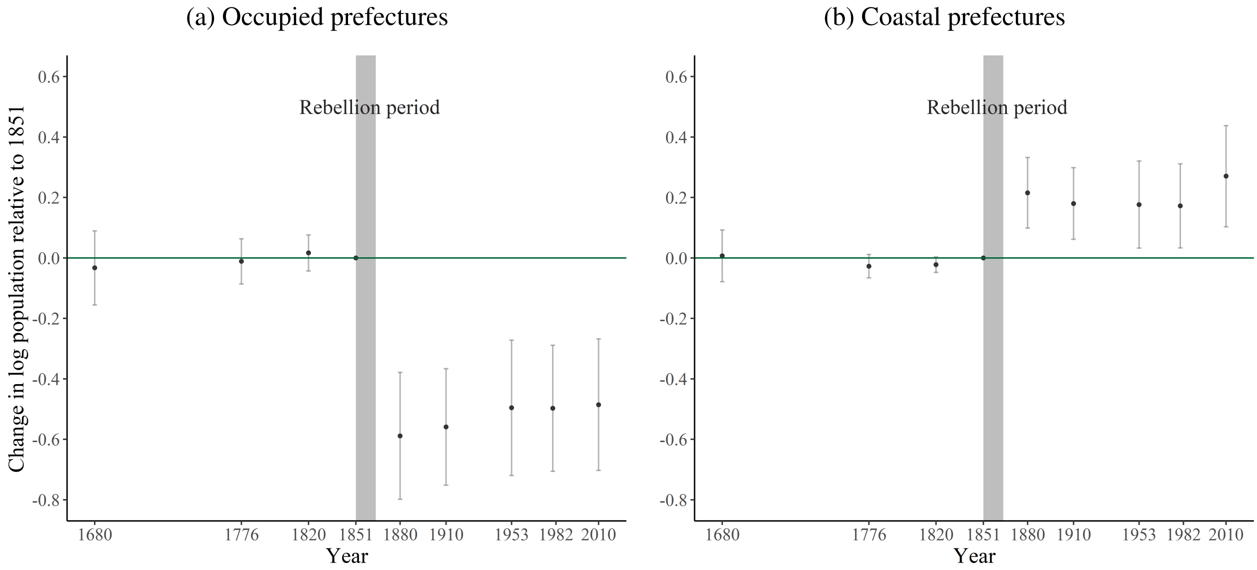
The coefficients of interest in Equation 1 are $\alpha_t$ and $\beta_t$, which are plotted in Figure 3 (Table C.1 shows the coefficient estimates). The vertical bars represent the 95 percent confidence intervals. The coefficient $\alpha_t$ has a "difference-in-difference" interpretation, where the first 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 excluded year 1851 for each preceding or succeeding data year. The left panel of Figure 3 plots the estimates of $\alpha_t$. The estimates 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

[22] The specification is similar to Heblich, Redding and Zylberberg (2024).

[23] In 1851, the rebellion activities remained localised and small-scale. Effectively, the 1851 data reflects the condition of prefecture populations just before the rebellion exerted its impact.

[24] 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.

Figure 3: Population evolution for occupied and coastal prefectures 1680-2010




*Notes*: Estimation based on Equation 1 (results in Table C.1). The left panel shows, for each data year, the estimated change in the log population for occupied prefectures compared to non-occupied prefectures, relative to their difference in 1851, corresponding to the coefficient estimate ($\alpha_t$) for the interaction between the occupied prefecture indicator and a year dummy. The right panel shows for, each data year, the estimated log population change for coastal prefectures compared to non-coastal prefectures, relative to their difference in 1851, estimated by the coefficient ($\beta_t$) of the 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).

after controlling for the time-varying effects of other locational characteristics.[25] 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 44.5% (58.9 log points) compared to non-occupied prefectures. Over time, the population in occupied prefectures caught up, as indicated by coefficients of smaller magnitudes in later years, but the gap has not closed. In 2010, the average population in occupied prefectures was still 38.5% (48.6 log points) lower than the non-occupied prefectures compared to their pre-rebellion level.

Beyond the direct impact on war-torn areas, the rebellion could have affected attractiveness of locations by shifting trade routes to the sea–a process that could otherwise have been slow even

[25] 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.

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 $\beta_t$. Similar to $\alpha_t$, the "difference-in-difference" coefficient $\beta_t$ 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 24.1% (21.6 log points). Although the effect attenuated over time, coastal population remained still 19.2% (17.6 log points) higher than non-coastal population a century later, compared to their difference prior to the rebellion. After 1949, 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.

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 $\text{post}_t$, 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. This allows controlling for differential population growth in occupied and coastal prefectures by including a linear time trend for each: $\lambda_o$ and $\lambda_c$. Consequently, the coefficients, $\alpha$ and $\beta$, can be interpreted as the average post-rebellion population change in occupied and coastal locations, after accounting for any pre-existing trends. I exclude observations from the data years 1980 and 2010, as government policies had a more pronounced impact on regional development after 1949, making the effect harder to interpret, but 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 relative population size of occupied prefectures after the rebellion. The estimated population decline is even slightly greater, about 44.7% (59.2 log points) on average, after accounting for the pre-existing population trends. The relative increase in coastal population was 24.6% (22 log points) on average 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. Note that this is based on a smaller sample of coastal locations with available trade data (i.e., treaty ports) and needs to be taken with a grain of salt. Table C.3 further incorporates the direct impact of opening treaty ports, and the results remain similar.

The next section exploits the status of bilateral maritime trade immediately after the rebellion. By linking the heterogeneous incentives to adopt sea trade through roadblocks during the rebellion to the observed post-rebellion trade, I show 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. Motivated by reduced-form evidence (AppendixD) showing that location pairs experiencing greater increases in inland transport costs due to roadblocks were more likely to engage in sea trade following the rebellion, I develop a trade model incorporating modal and route choices to characterise the indirect effects of roadblocks on post-rebellion trade. 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. Section 5.4 uses a general equilibrium economic geography framework to quantify the effect of the rebellion on population.

The model's cost structure and estimation share common features with Coşar and Demir (2018). Both approach modal choices incorporating mode-specific variable and fixed costs. Estimation in both cases follows a two-step procedure: first, variable costs are parameterised and estimated, albeit using very different methods, and then these estimates, along with the model structure, are used to recover fixed costs by mode and additional model parameters. Finally, model-consistent predictions and counterfactuals are conducted.

### *5.1 Model set-up*

Consider a country with multiple cities (prefectures in historical China). 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 for tradeable goods and that firms face monopolistic competition, the profit function of a firm in origin $i$ exporting to destination $j$ at time $t$ can be written as:[26]

$$\pi_{ijt} = \frac{1}{\sigma}(\frac{\sigma}{\sigma-1})^{1-\sigma}(\frac{\tau_{ijt}^{m}c_{it}}{P_{jt}})^{1-\sigma}E_{jt}. \tag{3}$$

where the elasticity of substitution $\sigma > 1$. As standard in trade models, firm profit increases with the destination's total expenditure on tradeable goods ($E_{jt}$) and its Dixit-Stiglitz price index ($P_{jt}$), which proxies for the market's competitiveness. The profit decreases with the local cost of production ($c_{it}$).

Departing from the standard setting, I model the iceberg trade cost $\tau_{ijt}^{m}$ to vary with the transport mode $m$. Firms can choose between two transport modes: inland transport (including inland waterways), denoted as $l$, or sea transport, denoted as $s$.

In addition to making the trade cost a function of transport cost ($D_{ijt}^{m}$) that takes into account optimal route choices within each mode, I incorporate inter-temporal sunk investments for establishing new trade routes for sea transport ($I_t^s$) and for inland transport ($I_t^l$).[27] The mode-dependent iceberg trade costs are:

$$(\tau_{ijt}^{l})^{1-\sigma} = (D_{ijt}^{l}I_t^l)^{-\Gamma}e^{\eta_{l,ijt}}, \quad (\tau_{ijt}^{s})^{1-\sigma} = (D_{ijt}^{s}I_t^s)^{-\Gamma}e^{\eta_{s,ijt}} \tag{4}$$

for inland and sea transport respectively. For each mode, trade costs depend on the respective transport cost ($D$), any sunk investment cost ($I$) if needed, the elasticity of trade with respect to transport costs ($\Gamma$), and a shock term ($e$). 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.[28] 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:

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

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 $I^l$.

[26] Monopolistic competition is used to characterise the market structure here, consistent with the economic geography model in Section 5.4. Nevertheless, the "roadblock effect" operates more generally, such as under the Armington assumption. As will become evident later, the key assumption for modal choice is that the least-cost mode (taking into account sunk investment cost) is selected.

[27] $D_{ijt}^{m}$ is the least-cost route effective distance as in Donaldson (2018). Here I calculate the effective distance separately for sea and inland transport and therefore the effective distance is mode-specific.

[28] As $I^s$ is a common parameter multiplied by the transport cost for each location pair, it is implicitly assumed that the bilateral sunk investment cost is proportional to transport cost. This may include costs related to establishing a new trade route, such as exploratory trips, logistical preparations, and forming business contacts, which are likely to rise with the effective distance.

Trade costs are subject to shocks. The shocks to inland 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 from the model set-up, 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.[29] There was no domestic steamship trade on the mainland before the rebellion. 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 modelled to be based on the immediate factors of relative transport costs and sunk investment costs, without regard to potential cost savings for future sea trade. This may be justified by the uncertainty of the post-rebellion status of sea trade. In AppendixE, I develop a dynamic model where firms incorporate cost reductions for future sea trade. The implied sunk investment cost will be larger if firms are assumed to be forward-looking, but the estimated effect of the rebellion on 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 if a distinct transportation sector operates, offering transport services to production firms for a fee. 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 goods.[30]

Lastly, I do not make explicit assumptions about whether firms in the same location coordinate on the sunk investment. If firms do not coordinate and there is no complementarity in their individual investments, $I^s$ would represent the total sunk investment required to establish a new route. At the other extreme, if firms coordinate completely and contribute equally, the total sunk investment would be $I^s$ multiplied by the number of firms in each location. When there is a separate shipping company that 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

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

[30] 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 $\tau_{ijt}^l$, and it will only enter the market if $\tau_{ijt}^s < \tau_{ijt}^l$. 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. However, if firms transport their own manufactured goods, $I^s$ may represent the partial sunk cost undertaken by a single firm.

the baseline model entails. Therefore, the sunk investment cost estimated from the baseline model is likely a lower bound.

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 make the modal choices. The probability of sea trade immediately after the rebellion is:

$$\Pr(S_{ij,post} = 1) = \Phi\big(\gamma \ln(D^l/D^s)_{ij,post} - \mathbb{1}(S_{ij,reb} = 0)i^s\big), \tag{7}$$

where $\Phi$ denotes the cumulative distribution function of the standard normal distribution. The post-rebellion probability of sea trade depends first on the post-rebellion transport cost by land relative to sea, $(D^l/D^s)_{ij,post}$. The elasticity of trade with respect to the transport cost, $\Gamma$, normalised by the standard deviation of transport shocks, $\sqrt{\sigma_l^2 + \sigma_s^2}$, which I denote as $\gamma$, governs how responsive the usage of sea trade is to the relative cost of inland 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 $i^s = \gamma \ln I^s$, which measures the deterrence of sunk cost to sea trade. If sea trade was used during the rebellion, and therefore a sea route was established ($\mathbb{1}(S_{ij,reb} = 0) = 0$), saving the sunk investment cost in the post-rebellion period, the probability of using steamships thereafter increases.

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

$$\Pr(S_{ij,reb} = 1) = \Phi\big(\gamma \ln(D^l/D^s)_{ij,reb} + roadblock_{ij} \times i^l - i^s\big). \tag{8}$$

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 alternative inland routes entailed an investment cost $I^l$ (the transformed sunk investment cost $i^l = \gamma \ln I^l$ measures its contribution to sea trade probability). Second, as the new route deviated from the unconstrained optimal path, the relative cost of land transport compared to sea, $(D^l/D^s)_{ij,reb}$, increased. The two factors combined heightened the appeal of sea transport during the rebellion for "roadblocked" location pairs relative to "non-roadblocked" location pairs subject to similar technologies and policies, incentivising their investment for sea trade during the rebellion.[31]

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

[31] Sea transport costs are modelled as constant during roadblocks and immediately after the rebellion (approximately a 5-year window), and the estimated effect is nearly invariant to introducing cost changes over this short period. The overall reduction in sea transport costs over a longer timeframe, before and after the rebellion, is incorporated into the extended model in Section 6.

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:

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

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

Equation 10 highlights the "roadblock effect," whereby a temporary increase in land transport cost during the rebellion affects post-rebellion modal choices by facilitating investment in maritime transport during the rebellion.[32] It also shows that the effect of the rebellion on bilateral trade hinges crucially on whether its regular inland route became unusable when the occupied region hit some intermediate point on its connection, and how costly a detour was needed compared to that of sea alternatives. This, in turn, requires information about the availability of roads, rivers and terrain along different inland trade routes and the change during the rebellion. If we know these, 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. To quantify the inland transport costs and their change during the rebellion, it requires knowledge about the costs associated with different trade routes, which, in turn, depends on the cost of using various transport infrastructures underlying them. While the previous literature on transport costs usually infers them from freight rates or travel speed, lack of data on them in 19th-century China imposes a challenge.

To deal with this problem, I develop a novel method of estimating historical transport costs based on China's postal map in 1903. I allow the typical cost of different transport methods (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.

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

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 extended the mail delivery services from the postal headquarters, typically situated in a treaty port, to the 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 various transport infrastructures along different routes.[33]

I divide the surface of China into 0.1 degree by 0.1 degree cells (553 $\times$ 826 cells in total, a cell at 30 degree latitude covering approximately 90 $km^2$ area) and overlay it with rivers, courier routes, and the 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.[34]

I search for the relative transport costs that maximise the number of matches between actual and FMM allocations of prefectures to postal headquarters through iterations. In the best scenario, 173 out of 199 prefectures can be matched to the actual postal districts compared to 114 matches 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 derived from postal districts renders the transport cost by inland waterways. To obtain the transport cost by sea, I use the information about travel time for sea relative to river as the transport cost per distance by steamship, including factors such as coal consumption and crew wage, increases proportionally with travel time. According to the CMCS trade reports, a steamship trip between river ports Chinkiang and Hankow took half a week, while a trip between seaports Ningpo and Wenchow took 26 hours. With a distance ratio of 1.63 for the two trips, it implies that the per distance transport cost by sea is half that by river. Note that while steamships were widely used on inland waterways by 1903, this was not the case during and immediately after the rebellion in 1867. At that time, inland transport modes used instead sailing ships for navigation along waterways. Given the relative transport cost estimates for others, the transport cost of sailing is pinned down by matching the aggregate level of sail trade before the rebellion, detailed in Section 6.[35]

[33] 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).

[34] The FMM algorithm is based on Allen and Arkolakis (2014), where the "Accurate fast marching" Matlab toolbox by Dirk-Jan Kroon is used.

[35] It is possible that steamship technology continued to advance between 1867 and 1903. The estimated effect of roadblocks, however, is not sensitive to incorporating further advancements in steamship technology. Intuitively, this is because such advancements would impact both the "roadblocked" and unaffected pairs.

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 1 so that the unit transport costs by other methods is relative to that by flatland. The cost of using courier routes was 18% of that of non-courier land routes. The cost of using inland waterways by steamship and sailing was 8% and 14%, respectively, of the cost of non-courier land transport. Steamship transport by sea costed only 4% of that of non-courier land transport. An increase in terrain ruggedness by one standard deviation was associated with a 32% increase in transport cost. Regarding transition from land to water transport, I convert the cost into equivalence to distance travelled by waterways for each mode. The transition cost amounted to the cost of travelling for 47 kilometres by inland waterways or, equivalently, 165 kilometres by sea. For instance, for a steamship trip between Shanghai and Ningpo, about half the transport cost will be borne by transshipment.

To access the validity of these estimated costs, I estimate the elasticity of migration with respect to transport costs in historical China, with the estimated values of relative transport costs in Table I. The results are presented in AppendixF. 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., least-cost route effective distance) is also larger than the elasticity estimated by geographic distance.[36]

In a more general context, this method can be applied to estimate transport costs for any region and historical period, provided the availability of two inputs. The first involves the assignment of locations to larger areas, presumably optimised to minimise transport costs. The second input comprises the location of different transport infrastructures. 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, along with the fixed costs associated with transitions, as long as a location allocation based on cost minimisation is available.

### *5.3 Model estimation and prediction: Trade*

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 after the rebellion and for each year during the rebellion when the two locations were not occupied.[37] The post-rebellion inland transport costs,

[36] Geographic distance is a proxy for transport cost and may therefore be subject to the classical measurement error. Consequently, the coefficient of distance may suffer from attenuation bias.

[37] For most analyses, the same inland transport cost applies before and after the rebellion. For robustness, I also calculate the inland transport cost after the rebellion, assuming a permanent change in the Yellow River due to its shift during the rebellion (see Section 7). The baseline model only needs transport costs during and after the rebellion.

$D^l_{ij,post}$ and sea costs $D^s_{ij}$, are calculated by identifying the least-cost routes using FMM and the transport cost parameters estimated in Section 5.2.

For each location pair in each year during the rebellion, I identify whether their inland trade needed a detour as the occupied region hit some intermediate point on their unconstrained optimal inland route. If so, I find the new optimal inland route to stay away from the occupied region. In cases where roadblocks occurred in multiple years 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 inland transport cost, $D^l_{ij,reb}$, during the rebellion.[38] [39]

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 did not involve sea transport.[40] Two extreme cases, Shanghai, which had active maritime trade with all other treaty ports, and Chinkiang, which had no trade with any other sea ports in 1867, are excluded from the model estimation, although the results including them would be similar.[41]

With the estimated relative transport costs of land compared to sea alternatives, accounting for their change 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 were more likely to resort to maritime trade when the cost of inland transport was relatively high. The cost associated with initiating sea trade was substantial, estimated at approximately 20 times the cost of an existing sea route ($I^s = \exp(i^s/\gamma)$). In contrast, the investment cost for inland transport was comparatively modest, with a cost ratio of using a new inland route over an existing one being about 5 ($I^l = \exp(i^l/\gamma)$).

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

[39] Because the maximum roadblock occurred for all location pairs after both locations had become treaty ports, the baseline model uses sea transport cost corresponding to steamship cost.

[40] 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.

[41] Throughout the rebellion, Shanghai found itself surrounded by occupied prefectures, rendering inland transport to the city impossible according to my estimate (i.e., $(D^l/D^s)_{ij,reb}$ 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" and started to export rice to the southern provinces by sea, which "was quite a new feature." Estimation and prediction of the model including the two ports render very similar results with an arbitrarily large $(D^l/D^s)_{ij,reb}$ for Shanghai's roadblock.

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 | |
| Inland waterways-steam | 0.08 | Transition cost from land to waterways |
| Inland waterways-sail | 0.14 | Equivalent to 47km by inland waterways-sail |
| Sea-steam | 0.04 | Equivalent to 165km by sea-steam |

*Panel B*: Estimation results on trade model parameters

| | Parameter | Value |
|---|---|---|
| Elasticity of probability of sea trade w.r.t. land/sea cost ratio | $\gamma$ | 0.709 |
| (normalised) | | (0.180) |
| Sunk cost of new sea route | $i^s$ | 2.124 |
| (transformed) | | (0.443) |
| Sunk cost of new land route | $i^l$ | 1.164 |
| (transformed) | | (0.403) |
| 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 | 0.692 | 0.358 |
| | | | (0.021) | (0.020) |
| Unaffected pairs | 32 | 0.375 | 0.379 | 0.379 |
| | | | (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 ($I^s = \exp(i^s/\gamma)$). The implied cost ratio of new inland route over existing one is 5 ($I^l = \exp(i^l/\gamma)$). Standard errors in parentheses. In *Panel C*, the model prediction is based on parameter estimates of *Panel B*. The prediction with roadblocks in column (3) provides the mean estimated probability for city prefecture pairs for which the least-cost 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. McFadden's pseudo R-squared value is 0.195. Bootstrap standard errors using 500 replicates in parentheses.

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 0.692 for "roadblocked" pairs and 0.379 for "non-roadblocked" pairs. In column (4), I calculate probabilities of maritime trade under a counterfactual scenario in which the roadblocks did not occur. 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 maritime trade would have been only 0.358 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.[42] Roadblocks significantly enhanced the use of sea trade after the rebellion, increasing the probability by approximately 93%, from 0.358 to 0.692. Note that the counterfactual removes only roadblocks, but not other technological or institutional changes that occur simultaneously for the "roadblocked" and unaffected pairs.

### *5.4 Model calibration and simulation: Population*

Motivated by the reduced-form results in Section 4 that population shifts following the Taiping Rebellion, I show in the previous parts that roadblocks incentivised the adoption of sea transport, using steamship in particular, in domestic trade. To complete the analysis, this part examines whether the "roadblock effect" can not only qualitatively, but also quantitatively, explain the change in relative city sizes after the Taiping Rebellion, in conjunction with the direct impact of the war. With population data available at the prefecture level before and after the rebellion, I follow the multi-region economic geography framework of Redding and Sturm (2008) to examine the direct effect (through war destruction) and indirect effect (through trade costs) of the rebellion on relative city sizes.

The production technology and consumer preferences are the same as defined in Section 5.1. Further, it is assumed that each location $i$ has an inelastic supply of an exogenous stock of a non-tradeable amenity, $H_i$. Consumers spend a constant expenditure share, $\mu$, on tradeable goods and the remaining share on non-tradeable amentities. In addition, each consumer supplies one unit of labour inelastically, and there is a fixed cost in terms of labour of producing tradeable goods. Equilibrium city sizes are determined by real wage equalisation across cities.

The model features agglomeration and dispersion forces. Both firms and consumers desire to locate near large markets, but better market access is counterbalanced by goods competition and congestion in non-tradeable amenities. The relative strength of agglomeration forces in shaping city sizes and the extent to which changes in trade costs affect relative city sizes depend on the

[42] "Roadblocked" pairs and "non-roadblocked" pairs are similar in terms of their pre-rebellion populations, years since the port was signed to open as a treaty port, the value of foreign trade in 1867. This holds true looking at the origin/destination separately, as well as using the sum or the logarithm sum of origin and destination for each pair.

parameters of the model. These include the elasticity of substitution between tradeable varieties ($\sigma$), the share of tradeables in expenditure ($\mu$), and the (effective) distance elasticity of trade ($\Gamma$ in Section 5.1). In addition to the change in market access present in Redding and Sturm (2008), the Taiping Rebellion caused war damages, which are modelled as a reduction in local non-tradeable amenities in the occupied region, for simplicity.[43]

Table II: Parameter configurations with the smallest deviations between simulation and estimation

| Elasticity of substitution $\sigma$ (1) | Tradeables share $\mu$ (2) | Distance elasticity $\Gamma$ (3) | Fraction of reduced non-tradeables $\delta$ (4) | Simulated treatment coastal locations (5) | Simulated treatment occupied locations (6) | Square root of sum of squared deviations (7) |
|---|---|---|---|---|---|---|
| 3.0 | 0.65 | 0.6 | 0.40 | 0.243 | -0.435 | 0.010 |
| 5.5 | 0.80 | 1.0 | 0.20 | 0.246 | -0.436 | 0.010 |
| 3.5 | 0.70 | 0.6 | 0.40 | 0.249 | -0.439 | 0.010 |
| 5.5 | 0.80 | 0.6 | 0.40 | 0.242 | -0.434 | 0.011 |
| 6.0 | 0.80 | 0.8 | 0.35 | 0.233 | -0.457 | 0.014 |
| 2.5 | 0.55 | 0.8 | 0.35 | 0.226 | -0.448 | 0.015 |
| 6.0 | 0.80 | 0.7 | 0.40 | 0.234 | -0.460 | 0.016 |
| 3.5 | 0.70 | 0.5 | 0.45 | 0.224 | -0.448 | 0.018 |
| 2.5 | 0.55 | 0.7 | 0.40 | 0.224 | -0.451 | 0.018 |
| 3.0 | 0.65 | 1.1 | 0.15 | 0.259 | -0.454 | 0.020 |

*Notes*: The table shows the simulated rebellion treatment on population using the multi-region economic geography model of Redding and Sturm (2008), compared to the reduced-form estimated treatment in Figure 3. Each row represents a parameter configuration. Columns (5) and (6) report the simulated relative changes in the coastal and occupied location populations between 1851 and 1880, given the parameter configuration from columns (1) to (4). Column (7) calculates the square root of the sum of squared deviation between the simulated and estimated treatments. The estimated changes for coastal and occupied locations are 0.241 and -0.445, respectively, which correspond to 21.6 and 58.9 log points in Table C.1.

I examine whether there are plausible parameter values for which the model can quantitatively account for the change in relative city sizes after the rebellion. To do this, I search over possible values of parameters $\sigma$, $\mu$, and $\Gamma$ from the existing literature, and a fraction of reduced non-tradeable amenities, $\delta$.[44] For each parameter configuration, I calibrate the model to the pre-rebellion city sizes in 1851 and then simulate the new relative sizes in 1880, incorporating changes in trade costs

[43] As the model only requires location population data, it can accommodate one exogenous feature of the location (e.g., non-tradeable amenities). Nevertheless, the loss in local non-tradeable amenities due to war shocks can be conceptualised using the theory of Allen and Donaldson (2020), as war shocks reduced past population and infrastructure, or Takeda and Yamagishi (2024), as changes in trade routes may have shifted expectations away from the old equilibrium, hindering the restoration of war damages.

[44] Specifically, I search over an elasticity of substitution in the range of 2 to 6.5 (with a step size of 0.5), a share of expenditure on tradeables from 0.5 to 0.8 (with a step size of 0.05), a distance elasticity of trade from 0.5 to 1.5 (with a step size of 0.1), and a share of non-tradeable amenities destroyed from 0 to 0.9 (with a step size of 0.05). The search is restricted to the case where a unique equilibrium exists with $\sigma(1-\mu) > 1$.

due to the new steamship links observed in 1867. The population sample used for calibration and simulation is the same as that used in Section 4. The reduced-form estimates indicate that, in relative terms, coastal population increased by 24.1%, while population in the occupied region declined by 44.5% after the rebellion. I search for the model parameters that best match these two key moments of the data.[45]

Table II reports the ten parameter combinations with the smallest deviations between the simulated and estimated rebellion treatments for coastal and occupied locations. Under these plausible parameters, the model simulations closely match the estimated population changes for both coastal and occupied locations.[46] The magnitude of the change in coastal population is in line with the changes in trade costs and market access brought about by the adoption of steamship transport in domestic trade. This provides support that the "roadblock effect" can quantitatively account for the relative increase of coastal population after the rebellion.

## 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 steam ships. 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 could have engaged in sail trade before the rebellion and thus could potentially spare the sunk investment associated with establishing a new sea route when steamships were introduced. 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, or even from institutional changes.

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 sailing vessels with an iceberg trade cost:

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

where $D_{ijt}^{sail}$ represents the transport cost of sailing 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 $\sigma_{sail}$. As a benchmark, it is assumed that the sunk cost for sea trade, represented by the ratio of the new sea routes over existing ones $I^s$, does not vary across the sail and steamship eras. This

[45] I use the population changes between 1851 and 1880, but the long-run population changes between 1851 and 1953 estimated in Table C.2 are also similar (24.6% and 44.7%, respectively).

[46] As noted by Redding and Sturm (2008), the effect of changes in trade costs and market access on city sizes depends on the strength of agglomeration and dispersion forces, which is captured by $\sigma(1-\mu)$, and the distance elasticity $\Gamma$, rather than on the individual values of $\sigma$ or $\mu$. Consequently, the ten parameter configurations with the smallest deviations feature distinct individual values for $\sigma$ and $\mu$. However, all ten combinations result in similar magnitudes for $\sigma(1-\mu)$, ranging from 1.05 to 1.2.

assumption is relaxed later. The trade costs by land and steamship are modelled in the same way as before:

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

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 incur any sunk investments if needed. Afterwards, firm-specific shocks to sailing ships ($e^{\iota_{kijt}}$) are realised. When both technologies of sea transport are available, a fraction of firms exporting from location $i$ to $j$ will choose sail over steam ships. After incorporating the option of sailing ships, the new sea transport cost can be summarised as follows:

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

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

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

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

where the pre-rebellion relative transport cost, $D^l/D^s$, is modelled as the ratio between inland and sail transport costs. Suppose that the probability of bilateral trade by sail has reached a steady state prior to the rebellion. This probability can be derived by setting $\Pr(Sail_{ijt} = 1)$ equal to the probability in the previous time period $\Pr(Sail_{ij,t-1} = 1)$ for any location pair $i$ and $j$. Denoting the steady state probability of sail trade as $\rho_{ij}$, we have:[47]

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

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

[47] 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.

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

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.[48]

In addition, the extended model explicitly takes into account the change in sea transport costs from sail to steam ship, which is the additional input in the extended model compared to the baseline model. I use the sail-to-steam cost ratio, 1.75, that matches the aggregate share of grain trade by sea (25.6%) before the rebellion.[49] [50] Note that the lower bound of the cost ratio is 1, indicating that sailing vessels have the same variable transport cost as steamships. The upper bound is 2.25, indicating that sailing ships have the same variable cost as courier routes. The cost ratio that matches the pre-rebellion sea trade falling within this range further supports that the model's mechanism and structure can account for the low level of sea trade before the rebellion, in a scenario where the government neither prohibited nor supported domestic sea trade and maintained inland transportation network.

The estimation results of the extended model are presented in Table III. The parameter estimates in *Panel A* are comparable to those estimated from the baseline model in Table I. 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 certain location pairs may face reduced or no sunk costs when establishing a new sea route by steamship, if they had previously engaged in sail trade. In such instance, the baseline model

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

[49]Deng (2009) estimated that the combined grain trade by inland navigation and sea amounted to 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 simulate the average share of sea trade given a transport cost by sail. The volume of trade is assumed to be proportional to the origin and destination population and inversely proportional to the transport cost, following the prediction of a naïve gravity equation. The resulting share of sea trade is found to be 25.2% with a sail to steam cost ratio of 1.75. For comparison, the share would be 34.3% with a cost ratio of 1 and 21.7% with a cost ratio of 2.25.

[50]Estimate results of the model under different cost ratios of sail to steam ships are also present in Table A.2.

Table III: 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 | $\gamma$ | 0.524 |
| (normalised) | | (0.111) |
| Sunk cost of new sea route | $i^s$ | 2.044 |
| (transformed) | | (0.388) |
| Sunk cost of new land route | $i^l$ | 1.604 |
| (transformed) | | (0.504) |
| 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 | 0.682 | 0.361 | 0.273 |
| | | | (0.018) | (0.021) | (0.018) |
| Unaffected pairs | 32 | 0.375 | 0.388 | 0.388 | 0.312 |
| | | | (0.047) | (0.047) | (0.043) |

*Notes*: In *Panel A*, the estimates imply that the cost ratio of new sea route over existing route is 49 ($I^s = \exp(i^s/\gamma)$). The implied cost ratio of new inland route over existing one is 21 ($I^l = \exp(i^l/\gamma)$). 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- cost 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 had not been 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. McFadden's pseudo R-squared value is 0.181. Bootstrap standard errors using 500 replicates in parentheses.

calculates an average sunk investment cost for sea trade across all localities engaging in post-rebellion sea trade, including even those with potentially reduced or zero costs owing to prior involvement in sail trade.

In *Panel B* of Table III, 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 where roadblocks did not occur, the probability of maritime trade would have been only 0.361 for "roadblocked" pairs, similar to that of the unaffected location pairs (column (4)). This counterfactual indicates that roadblocks increased the probability of sea trade by approximately 89%, rising from 0.361 to 0.682.

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 maritime 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 were 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 were much more superior to the incumbent, the relative contribution of technological improvement would be more substantial, although this holds true for both "roadblocked" pairs and "non-roadblocked" pairs.

Table III reveals that with a cost ratio of 1.75 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 is still substantial. The predicted sea trade probability by sail before the rebellion for "roadblocked" pairs is, on average, 0.273. If roadblocks did not occur, their post-rebellion sea trade probability would have been 0.361. The probability increase from 0.273 to 0.361 represents the contribution of technological change, which is only 8.8 percentage points. The majority of the surge in maritime trade (78%) comes from roadblocks, raising the sea trade probability from 0.361 to 0.682. Similarly, for "non-roadblocked" pairs, the model predicts a probability of sea trade increase of 7.6 percentage points, from 0.312 to 0.388, with the introduction of steamships.

It has been maintained thus far that the bilateral sunk cost parameter for sea trade ($I^s$ ) remains unchanged over time. If the sunk cost drops before and after the rebellion, the implied transport cost by sail would be smaller. A reduction in sunk investment cost maps to a sail to steam transport cost ratio to match the level of pre-rebellion sail trade. For instance, if the sunk cost reduces by 50% before and after the rebellion, the implied cost ratio of sail to steam ships would be 1.5, instead of 1.75. The combination of reduction in transport cost and sunk investment cost, due to technological or even institutional changes brought by treaty ports, can be flexibly incorporated in this framework, fuelling a general increase in maritime trade over time. While it is not possible to pinpoint the reduction in sunk cost and transport cost separately, the reduction in sunk cost during this period can be bounded. If steamship had the same transport cost as sail, the implied sunk investment cost before the treaty port era would have been 10 times higher, despite the unlikelihood of both scenarios.

Importantly, introducing a change in sunk costs has little bearing on the "roadblock effect." It alters the interpretation of the sunk investment cost estimated in Section 5, which may have been larger had treaty ports or the technological change not taken place. However, this has minimal impact on the estimated magnitude of the "roadblock effect" or its relative contribution to the rise of maritime trade during this time. Section 7 delves deeper into the role of technological change and its implication for 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*** First, I examine whether the "roadblock effect" is confounded by opening up treaty ports to trade with Western countries. While steamship was a foreign technology, introduced via the establishment of treaty ports, and exclusively operated by foreign companies as of 1867, the paper isolates the effect of the roadblocks on maritime trade from the broader advancements in sea transport technology, institutional changes, or shifts in attitudes towards sea trade that made possible the adoption of steamships. When steamship technology became accessible between treaty port pairs, the demand for steamships varied substantially, driven by the heterogeneous increase in trade costs along their regular inland trade networks compared to sea costs during the rebellion. For simplicity, in the models, production firms make their transport mode decision with the associated sunk costs. As explained in Section 5, this is isomorphic to a separate (foreign) shipping company offering its services when incurring investment costs, and production firms determining their usage of the shipping services based on the price. Despite foreign companies' overall interest in supplying and profiting from steamship shipping, this enthusiasm was not consistently met with demand. This discrepancy is exemplified by the case of Kiukiang, as discussed in Section 2, the comparison between "roadblocked" and "non-roadblocked" location pairs, and counterfactual analysis for the "roadblocked" pairs. While it is possible that foreign traders or couriers might have a vested interest in pursuing steamship trade for foreign trade between specific locations and domestic trade could benefit consequently, I show that in practice, this did not happen nor affect the "roadblock effect."

Treaty ports can influence economic development through other channels. When conducting reduced-form analysis on population, I directly control for any population changes in a prefecture upon the opening of a treaty port. The post-rebellion population changes in coastal locations remain robust, with an average increase of 22.9% (equivalent to 20.6 log points) (Table C.3). Coastal prefectures without treaty ports also experienced substantial population growth immediately after the rebellion, averaging 16.5%. Furthermore, post-rebellion population changes in treaty ports strongly and positively correlate with domestic maritime trade, but not with foreign trade. However, this latter finding should be interpreted as suggestive due to the limited sample size.

Regarding the results on trade, I only compare bilateral trade between treaty ports due to data availability. Although all treaty ports had access to foreign trade, steamship technology, and similar institutional arrangements, identification is achieved by comparing whether domestic trade between any two of them was more likely to use sea transport, depending on the heterogeneous increase in their relative inland transport costs due to the conflict. Treaty ports, together with the technological or institutional changes they introduced, would affect the interpretation of the entailed trade costs estimated from the model, but not the estimated "roadblock effect."

Focusing on bilateral trade, a 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. Foreign traders or couriers may be particularly interested in pursuing steamship trade between specific locations, which may consequently benefit domestic trade. In theory, if these considerations were orthogonal to roadblocks, they do not contaminate the estimated "roadblock effect." In practice, the use of transit trade *by sea* was very limited, except for Shanghai (which is excluded for most analysis).[51] I proceed to discuss the status of transit trade before and after the rebellion and show robustness results that transit trade is unlikely to shape the "roadblock effect."

A notable example of what happened to foreign trade during the rebellion was the re-routing of Bohea tea, a significant Chinese export to the world in the 19th century. Before the rebellion, foreign export primarily occurred through Shanghai and Canton (today's Guangzhou). Despite the nearby treaty port of Foochow, which had been signed to opened about ten years prior to the rebellion, Bohea tea persisted in being transported overland to Canton for foreign export. During the rebellion in 1853, the export of Bohea tea finally shifted to direct shipping from Foochow to foreign destinations when the regular inland route to Canton was blocked. If Canton had functioned as an entrepôt for re-export by sea, roadblocks might have similarly incentivised the use of sea transport between Foochow and Canton, for both foreign transit trade and domestic trade. Yet, since all treaty ports could export directly to foreign destinations and Canton did not serve as such entrepôt for sea transit, roadblocks would have had a separate effect on domestic trade from foreign trade.

In practice, the only ports on the mainland that actively engaged in transit trade by sea were Shanghai and, to a much less extent, Amoy in 1867.[52] The goods re-exported from other ports "are not sufficiently important in amount to be noticed."[53] The magnitudes of the "roadblocked effect" are also similar after excluding Amoy, in addition to Shanghai, shown 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 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. The gap in their adoption of sea trade was much larger than the overall increase in maritime trade during that period. These findings demonstrate that the rebellion played a catalysing role in the adoption of the new transport technology.

[51] The analysis also excludes non-mainland ports, such as treaty ports in Hong Kong and Macau.

[52] The ratio of foreign transit trade (re-export of foreign imports and re-export of Chinese goods for foreign export) over net total trade (net imports of foreign and Chinese goods plus exports of Chinese good) was only 7% for Amoy. Furthermore, its re-export activity was primarily directed towards ports in Taiwan.

[53] *Returns of trade at the ports in China open by treaty to foreign trade for the year 1867 (Shanghai, 1868).*

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. Here, technological progress is understood in a broad sense, encompassing all factors that increased the likelihood of bilateral sea trade with the introduction of treaty ports and steamships, including institutional or policy changes that reduced the cost of adopting sea trade. 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. Table III shows that the counterfactual probability of sea trade for "roadblocked" pairs without the rebellion would be 0.361, an increase of 0.088 from the pre-rebellion probability of sea trade with sail. Note that this counterfactual only removes roadblocks, but not other technological or institutional contributor to trade over time. The relative contribution of roadblocks to the rise of sea trade is calculated as $\frac{0.682-0.361}{0.682-0.273}$, where the numerator calculates the difference between post-rebellion probability of sea trade of "roadblocked" pairs and its post-rebellion counterfactual without roadblocks (but with technological change) and the denominator represents the overall increase in its probability of sea trade before and after the rebellion. Therefore, the majority of increase in sea trade ($0.321/0.409 \approx 78\%$) can be attributed to the "roadblock effect." Similarly, for "non-roadblocked" pairs, their estimated increase in the probability of sea trade due to technological advance is only 0.076.

***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.[54] 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 model predictions incorporating the increase in inland transport costs due to potential changes in the Yellow River yield very similar results to the baseline predictions reproduced in column (1).

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

Second, I estimate the effect by dropping all observations susceptible to potential changes from the Yellow River shifting course, including those affected by both roadblocks and the river. This yields a smaller sample of 62 location pairs. This sub-sample consists only of "roadblocked" pairs affected by rebellion occupation, but not by potential changes due to the Yellow River. Their post-rebellion probability of sea transport stands at 0.656 in the data, significantly and statistically higher than 0.383 of the "non-roadblocked" pairs. While the model prediction based on parameters estimated from the baseline sample slightly overestimates the post-rebellion probability for this sub-sample of "roadblocked" pairs (23 out of 32 pairs instead of 21 out of 32), the sub-sample of "non-roadblocked" pairs exhibits a comparable probability of sea trade after excluding observations susceptible to potential changes in the river and canal.

In summary, the "roadblock effect" remains robust when incorporating potential increases in transport costs due to changes from the Yellow River. 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 connecting more distant location pairs are more likely to traverse any given region, including those affected by the 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, rendering it an edge 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 accounted for the relative transport costs and the cost of transition, and that they are primary determinants of modal choices, this should not bias the estimated effect of roadblocks. Hence, instead of relying on geographic 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 time 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" captures 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, as well as absolute differences in longitude and latitude.[55] Table A.5 shows that, following this matching procedure, the two groups exhibit comparability in both geodesic distance and differences in longitude and latitude. Additionally, the transport costs by sea and land, which

[55]Because "roadblocked" observations include both longer distance and shorter distance trade pairs while "non-roadblocked" pairs consist mainly of shorter distance trade pairs, I match each "non-roadblocked" pair with a "roadblocked" pair, rather than vice versa.

are inputs of the model but not directly used for propensity score matching, are also balanced after the matching procedure.

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 the corresponding "non-roadblocked' pairs. This suggests that the previously estimated "roadblock effect" was not driven by those "roadblocked" observations that are farther apart than "non-roadblocked" pairs. Moreover, the model prediction, based on parameters estimated using the baseline sample, aligns well with the mean probability of sea trade for this sub-sample. This further indicates 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. The counterfactual probability for this sub-sample of "roadblocked" pairs without the rebellion would be 0.397, which indicates an approximately 84% increase in the probability of sea trade due to roadblocks. Both the data and the model indicate a slightly higher probability of sea trade for this sub-sample of "roadblocked" pairs that are less distant. This is because long-distance inland trade can more often use courier routes.

## 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 and thus increasing their market access. However, new technologies typically require large sunk investments, which can hold them back. The 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. I find that the relative population size in the war-ravaged locations shrank by 44.7% (59.2 log points) after the rebellion. On the other hand, the population in coastal locations experienced a 24.6% (22 log points) relative increase. I provide evidence that this is because, in addition to the direct impact of war destruction, the rebellion affected populations indirectly through trade routes and transport technologies. Prior to the rebellion, the use of sea transport was limited despite its potential. Even when superior maritime technology became available, its adoption was not guaranteed. I show that the spatial variations in maritime trade immediately after the rebellion was driven by blocked inland routes and the feasibility of sea alternatives. The rebellion disrupted regular inland trade routes, forcing some location pairs to search for alternatives, and this 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 and quantify the "roadblock effect." The model estimates find a substantial cost associated with initiating maritime trade. The counterfactual analysis using the model suggests that the likelihood of post-rebellion maritime trade would have been much lower in the absence of roadblocks. The "roadblock effect" can quantitatively account for the population changes through changes in trade costs and market access.

The findings of the paper highlight persistence in population and trade routes, technological inertia due to sunk costs, the critical role of incentives in driving technology adoption, and population changes through trade costs and technology.

## References

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

**_ and _** , "Modern spatial economics: a primer," in "World Trade Evolution," Routledge, 2018, pp. 435–472.

**_ and _** , "The Welfare Effects of Transportation Infrastructure Improvements," *The Review of Economic Studies*, 02 2022, *89* (6), 2911–2957.

**_ and Dave Donaldson**, "Persistence and Path Dependence in the Spatial Economy," Working Paper 28059, National Bureau of Economic Research November 2020.

**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.

**Barjamovic, Gojko, Thomas Chaney, Kerem Coşar, and Ali Hortaçsu**, "Trade, merchants, and the lost cities of the Bronze age," *The Quarterly Journal of Economics*, 2019, *134* (3), 1455–1503.

**Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A Turner, and Qinghua Zhang**, "Roads, railroads, and decentralization of Chinese cities," *The Review of Economics and Statistics*, 2017, *99* (3), 435–448.

**Beestermöller, Matthias and Ferdinand Rauch**, "A dissection of trading capital: Trade in the aftermath of the fall of the iron curtain," *Journal of Economic History*, 2018, *78* (2), 358–393.

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

**Bonadio, Barthélémy**, "Ports vs. roads: infrastructure, market access and regional outcomes," *CESifo Working Paper*, 2024.

**Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou**, "Geography, Transportation, and Endogenous Trade Costs," *Econometrica*, 2020, *88* (2), 657–691.

**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.

**Combes, Pierre-Philippe, Miren Lafourcade, and Thierry Mayer**, "The trade-creating effects of business and social networks: evidence from France," *Journal of International Economics*, 2005, *66* (1), 1–29.

**Coşar, A Kerem and Banu Demir**, "Shipping inside the box: Containerization and trade," *Journal of International Economics*, 2018, *114*, 331–345.

**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.

**Donaldson, Dave**, "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 2018, *108* (4-5), 899–934.

**Ducruet, César, Réka Juhász, Dávid Krisztián Nagy, and Claudia Steinwender**, "All aboard: The effects of port development," *Journal of International Economics*, 2024, *151*, 103963.

**Ellingsen, Sebastian**, "Long-distance trade and long-term persistence," *Working paper*, 2021.

**Faber, Benjamin**, "Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System," *The Review of Economic Studies*, 2014, *81* (3), 1046–1070.

**Fajgelbaum, Pablo D and Edouard Schaal**, "Optimal transport networks in spatial equilibrium," *Econometrica*, 2020, *88* (4), 1411–1452.

**Fan, Baichuan**, *Zhonguo lunchuan hangyunye de xingqi [The rise of China's steamship industry]*, Sichuan People's Press, 1985.

**Fan, Jingting, Yi Lu, and Wenlan Luo**, "Valuing domestic transport infrastructure: A view from the route choice of exporters," *The Review of Economics and Statistics*, 2023, *105* (6), 1562–1579.

**Flückiger, Matthias, Erik Hornung, Mario Larch, Markus Ludwig, and Allard Mees**, "Roman Transport Network Connectivity and Economic Integration," *The Review of Economic Studies*, 07 2021, *89* (2), 774–810.

**Fuchs, Simon and Woan Foong Wong**, "Multimodal transport networks," *National Bureau of Economic Research*, 2024.

**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 James M Poterba**, "Economic analysis and infrastructure investment," *National Bureau of Economic Research*, 2020.

**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.

**Head, Keith and Thierry Mayer**, "Market potential and the location of Japanese investment in the European Union," *The Review of Economics and Statistics*, 2004, *86* (4), 959–972.

**_, _, and John Ries**, "The erosion of colonial trade linkages after independence," *Journal of International Economics*, 2010, *81* (1), 1–14.

**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.

**_, _, and Yanos Zylberberg**, "The distributional consequences of trade: Evidence from the grain invasion," *National Bureau of Economic Research*, 2024.

**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.

**Hummels, David L and Georg Schaur**, "Time as a trade barrier," *American Economic Review*, 2013, *103* (7), 2935–2959.

**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.

**Marczinek, Maximilian, Stephan Maurer, and Ferdinand Rauch**, "Trade persistence and trader identity-evidence from the demise of the Hanseatic League," *CEPR Discussion Paper No. DP16957*, 2022.

**Maurer, Stephan and Ferdinand Rauch**, "Economic geography aspects of the Panama Canal," *Oxford Economic Papers*, 03 2022, *75* (1), 142–162.

**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.

**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," *The 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.

**Pascali, Luigi**, "The wind of change: Maritime technology, trade, and economic development," *American Economic Review*, 2017, *107* (9), 2821–54.

**Redding, Stephen J and Anthony J Venables**, "Economic geography and international inequality," *Journal of International Economics*, 2004, *62* (1), 53–82.

_ **and Daniel M Sturm**, "The costs of remoteness: Evidence from German division and reunification," *American Economic Review*, 2008, *98* (5), 1766–97.

__ , __ , **and Nikolaus Wolf**, “History and industry location: evidence from German airports,” *The Review of Economics and Statistics*, 2011, *93* (3), 814–831.

**Santamaria, Marta**, “Reshaping Infrastructure: Evidence from the division of Germany,” *Working paper*, 2020.

**Takeda, Kohei and Atsushi Yamagishi**, *The economic dynamics of city structure: Evidence from Hiroshima’s recovery*, Centre for Economic Performance, London School of Economics and Political, 2024.

**Tsivanidis, Nick**, “Evaluating the impact of urban transit infrastructure: Evidence from bogota’s transmilenio,” *Working paper*, 2022, *18.*

**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.

**Wong, Woan Foong**, “The round trip effect: Endogenous transport costs and international trade,” *American Economic Journal: Applied Economics*, 2022, *14* (4), 127–166.

**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 | | | | | | |
| $\gamma$ | 0.446 | 0.475 | 0.502 | 0.524 | 0.542 | 0.561 |
| normalised elasticity of trade | (0.096) | (0.101) | (0.106) | (0.111) | (0.116) | (0.121) |
| $i^s$ | 1.875 | 1.977 | 2.017 | 2.044 | 2.062 | 2.084 |
| transformed sunk cost of new sea route | (0.421) | (0.399) | (0.392) | (0.388) | (0.386) | (0.386) |
| $i^l$ | 1.641 | 1.782 | 1.685 | 1.604 | 1.537 | 1.477 |
| transformed sunk cost of new land route | (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 from the baseline model: Alternative sample

| | (1) Baseline sample | (2) Excl. Tientsin. | (3) Excl. Entrepôt | (4) River change | (5) Excl. river | (6) Matched sample |
|---|---|---|---|---|---|---|
| *Roadblocked pairs*: | | | | | | |
| Data | 0.696 | 0.705 | 0.646 | 0.696 | 0.656 | 0.750 |
| Prediction with roadblock | 0.692 | 0.709 | 0.673 | 0.697 | 0.716 | 0.730 |
| | (0.021) | (0.024) | (0.024) | (0.021) | (0.027) | (0.026) |
| Prediction w/o roadblock | 0.358 | 0.377 | 0.336 | 0.366 | 0.383 | 0.397 |
| | (0.020) | (0.024) | (0.022) | (0.020) | (0.029) | (0.027) |
| Observations | 56 | 44 | 48 | 52 | 32 | 32 |
| *Non-roadblocked pairs*: | | | | | | |
| Data | 0.375 | 0.346 | 0.364 | 0.375 | 0.385 | 0.375 |
| Prediction | 0.379 | 0.375 | 0.353 | 0.381 | 0.396 | 0.379 |
| | (0.045) | (0.048) | (0.062) | (0.045) | (0.045) | (0.045) |
| Observations | 32 | 26 | 22 | 36 | 26 | 32 |

*Notes*: Predictions of the baseline model in Section 5 with different robustness specifications. Column (1) reports the prediction presented in 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 ports that can be affected by transit trade (i.e., Amoy, 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 sub-sample 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 from the extended model: Alternative sample

| | (1) Baseline sample | (2) Excl. Tientsin. | (3) Excl. Entrepôt | (4) River change | (5) Excl. river | (6) Matched sample |
|---|---|---|---|---|---|---|
| *Roadblocked pairs*: | | | | | | |
| Data | 0.696 | 0.705 | 0.646 | 0.696 | 0.656 | 0.750 |
| Prediction with roadblock | 0.682 | 0.698 | 0.665 | 0.687 | 0.703 | 0.715 |
| | (0.018) | (0.019) | (0.020) | (0.017) | (0.023) | (0.022) |
| Prediction w/o roadblock | 0.361 | 0.381 | 0.337 | 0.366 | 0.389 | 0.403 |
| | (0.021) | (0.025) | (0.023) | (0.021) | (0.031) | (0.028) |
| Observations | 56 | 44 | 48 | 56 | 32 | 32 |
| *Non-roadblocked pairs*: | | | | | | |
| Data | 0.375 | 0.346 | 0.364 | 0.375 | 0.385 | 0.375 |
| Prediction | 0.388 | 0.383 | 0.363 | 0.390 | 0.405 | 0.388 |
| | (0.047) | (0.050) | (0.065) | (0.047) | (0.048) | (0.048) |
| Observations | 32 | 26 | 22 | 32 | 26 | 32 |

*Notes*: Predictions of the extended model in Section 6 with different robustness specifications. Column (1) reports the prediction of Table III 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 ports that can be affected by transit trade (i.e., Amoy, 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 sub-sample 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 | 0.362 | 0.458 | 0.392 | 0.152 | -0.075 | 0.022 |
| | (0.462) | (0.871) | (0.110) | (0.182) | (0.091) | (0.432) | (0.856) | (0.115) | (0.201) | (0.101) |
| Constant | 2.925 | 4.419 | 6.273 | 1.065 | 3.049 | 2.925 | 4.419 | 6.273 | 1.065 | 3.049 |
| | (0.323) | (0.537) | (0.095) | (0.174) | (0.071) | (0.325) | (0.540) | (0.095) | (0.175) | (0.072) |
| Observations | 88 | 88 | 88 | 88 | 88 | 64 | 64 | 64 | 64 | 64 |
| R-squared | 0.079 | 0.257 | 0.381 | 0.093 | 0.154 | 0.018 | 0.003 | 0.027 | 0.002 | 0.001 |

*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

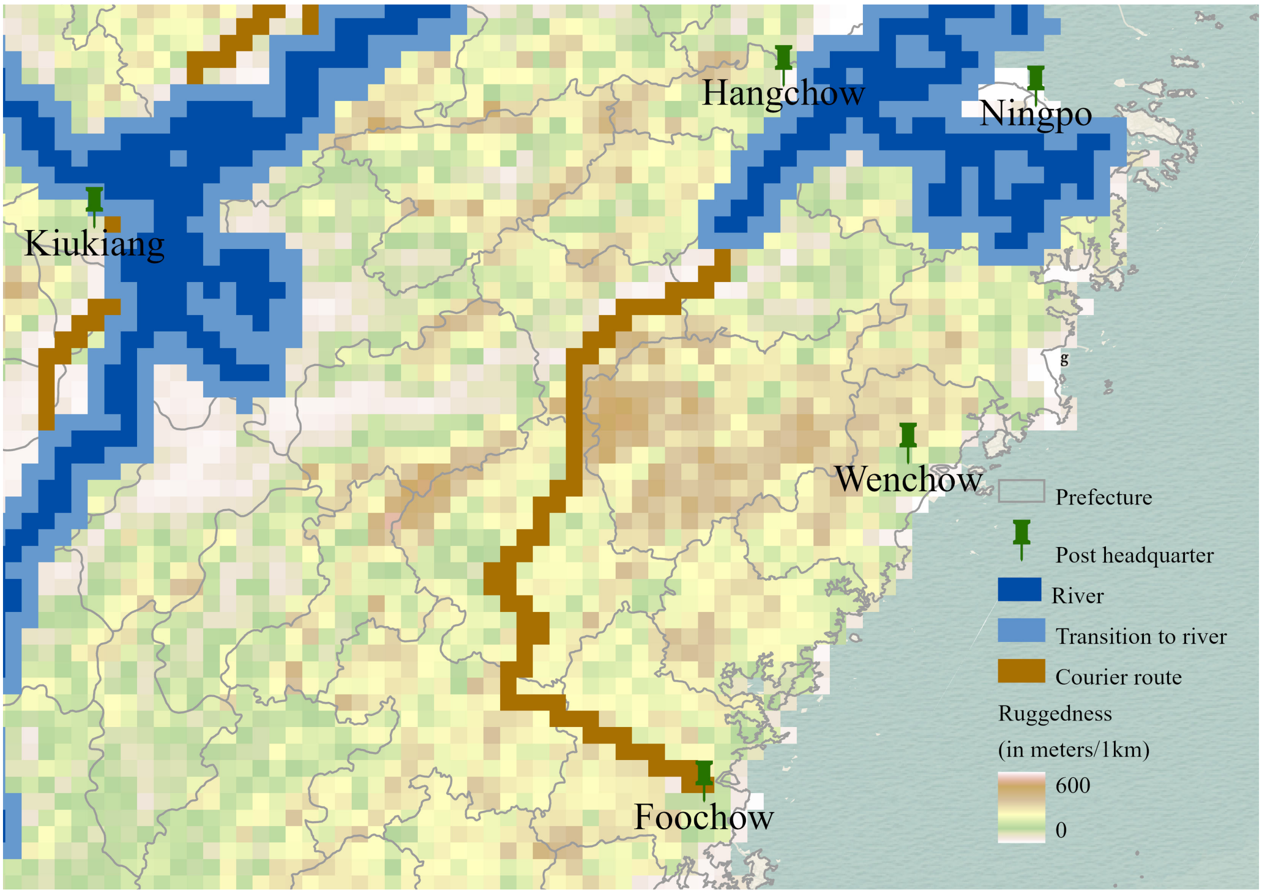



Figure B.1: Transport cost in 19th-century China

*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 capture the population change of a treated group relative to the control group compared to their difference in 1851. 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.[56] 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 a relative 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. Additionally, by replacing interactions of year indicators with a post-rebellion indicator, it is possible to explicitly control for differential population growth trends in occupied and coastal prefectures, respectively.

The estimation result shown 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 48.7% (66.8 log points) after the rebellion. Column (2) further includes a post-rebellion and coastal indicator interaction to estimate the change in population in coastal locations after the rebellion. The estimated coefficient shows a relative 24.2% (21.7 log points) increase in coastal population after the rebellion.

[56] These include the level of human capital and a set of geographic controls (i.e., 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 zero or 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 can be driven by some characteristics of coastal areas, other than being coastal. While prefecture fixed effects account for population differences across prefectures in levels, these characteristics might exert a differential impact on population following a large shock to population sizes. For instance, coastal areas may boast higher agricultural productivity on average and thus attract migration after the rebellion. More generally, 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 | -0.670 | -0.620 |
| | (0.081) | (0.079) | (0.101) | (0.090) | (0.119) |
| Post-reb. × coastal | | 0.217 | 0.220 | 0.294 | 0.360 |
| | | (0.055) | (0.061) | (0.065) | (0.090) |
| Post-reb. × coastal × domestic trade | | | | 0.563 | 0.296 |
| | | | | (0.195) | (0.143) |
| Post-reb. × coastal × foreign trade | | | | -0.170 | -0.114 |
| | | | | (0.121) | (0.073) |
| 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 | 0.629 | 0.634 | 0.697 | 0.605 | 0.682 |

*Notes*: 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. In columns (4) and (5), trade val. is the natural logarithm of domestic/foreign 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). Robust standard errors clustered by prefecture are reported in parentheses.

Therefore, in column (3), I include further a variety of locational traits (i.e., the level of human capital and a set of geographical controls) interacted with all year indicators.[57] The positive coefficient estimate of the interaction between the post-rebellion and coastal indicators remains robust after accounting for differential impacts of these additional locational characteristics on

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

population across different data years, showing a similarly 24.6% (22 log points) increase in the relative population level in coastal prefectures after the rebellion.

To provide suggestive evidence that population relocations are linked to a shift in trade activities towards the sea during the rebellion's disruption of inland trade, 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 maritime trade, respectively, relative to their averages in 1867.[58] These trade values, primarily derived from steamship trade, are used to examine whether the rise in population in coastal locations was associated with a surge in steamship trade.

Since trade data are exclusively available at treaty ports, who were authorised to trade with Western countries and typically larger in scale, the number of observations drops. For this subset of large ports, the post-rebellion $\times$ coastal interaction indicates an average population increase of about 34.2% (29.4 log points) after the rebellion. Furthermore, the triple interaction term (post-rebellion$\times$coastal$\times$domestic trade) indicates that if the domestic maritime trade of a treaty port was 10% higher than the average treaty port, its post-rebellion population was on average 5.6% higher. On the other hand, the association between coastal population and foreign maritime trade was negative and statistically insignificant.[59] This offers suggestive evidence that the rise in coastal population during this period was driven by the pattern of domestic maritime 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.

In Table C.3, I show that the post-rebellion population increase in coastal locations is robust to controlling for the opening of treaty ports, by incorporating a level change in prefecture population when a treaty port opens in its location. The estimated relative population increase in coastal location remains similarly at 22.9% (20.6 log points) after the rebellion.

[58]The CMCS denominated the value of trade in Customs Taels for most ports in 1867, with the exception of southern ports, which reported their trade value in Mexican dollars. I convert these values to Customs Taels using the contemporary conversion rate.

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

Table C.3: Prefecture population evolution 1680–1953: Control for treaty ports

| | Dep. var.: Ln (population) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Post-reb. × occupied | -0.668 | -0.662 | -0.592 | -0.670 | -0.619 |
| | (0.081) | (0.080) | (0.102) | (0.092) | (0.119) |
| Post-reb. × coastal | | 0.211 | 0.206 | 0.293 | 0.371 |
| | | (0.058) | (0.066) | (0.068) | (0.097) |
| Post-reb. × coastal × domestic trade | | | | 0.563 | 0.296 |
| | | | | (0.195) | (0.144) |
| Post-reb. × coastal × foreign trade | | | | -0.170 | -0.122 |
| | | | | (0.124) | (0.079) |
| Occupied linear time trend | 0.0005 | 0.0004 | 0.0004 | 0.0006 | 0.0004 |
| | (0.0002) | (0.0002) | (0.0003) | (0.0003) | (0.0004) |
| Coastal linear time trend | 0.0008 | -0.0002 | -0.0002 | -0.0003 | -0.0004 |
| | (0.0003) | (0.0002) | (0.0002) | (0.0006) | (0.0006) |
| Year indicators × controls | | | Yes | | Yes |
| Treaty ports × post-open | | Yes | Yes | 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 | 0.629 | 0.634 | 0.697 | 0.605 | 0.682 |

*Notes*: 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. In columns (4) and (5), trade val. is the natural logarithm of domestic/foreign 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). Robust standard errors clustered by prefecture are reported in parentheses.

## AppendixD. Reduced-form evidence on the effect of roadblocks

This section provides reduced-form evidence on the effect of roadblocks on post-rebellion maritime trade. The analysis predicts the post-rebellion maritime trade between location pairs with relative transport costs and their change due to roadblocks. Table D.1 reports the results using the linear probability and the probit model, with explanatory variables in natural logarithms. Table D.2 shows the corresponding results using explanatory variables in levels.

The dependent variable is an indicator for domestic maritime trade between two treaty ports, taking the value 1 if the two ports had bilateral maritime trade in 1867 and 0 otherwise. Roadblocks increased inland transport costs and, therefore, may have incentivised the use of maritime trade during the rebellion. Due to the inter-temporal linkage of trade, temporary roadblocks can even have a persistent effect on post-rebellion trade. Furthermore, the effect is expected to be larger if the increase in inland transport costs due to roadblocks was higher.

To investigate this possibility, I regress the indicator for bilateral maritime trade on the roadblock-induced change in inland transport costs, measured by the changes in the natural logarithm of inland transport costs due to roadblocks, based on travel costs in *Panel A* of Table I. For each location pair, I calculate the transport cost associated with the unconstrained optimal route and compare that with the cost of the constrained optimal route that stays way from the rebellion-occupied region. Columns (1) to (4) of Table D.1 present the results using the linear probability model. The coefficient estimate in column (1) indicates that a 10 percent increase in inland transport costs due to roadblocks was associated with approximately a 4.21 percentage point increase in the probability of post-rebellion maritime trade. In column (2), I control for the relative transport cost, defined as the natural logarithm of the transport cost ratio between unconstrained inland and sea optimal routes. The coefficient is positive as expected, indicating that locations were more likely to use maritime trade when inland transport costs were relatively high. The coefficient for the effect of roadblocks remains similar after controlling for relative transport costs. Given their relative transport cost, locations that experienced a 10 percent increase in inland transport costs during the rebellion had a 4.66 percentage point higher probability of using sea trade after the rebellion. In column (3), I consider alternative changes in inland transport during the rebellion. As explained in Section 7, the Yellow River shifted in 1855, causing the northern part of the Grand Canal to become paralysed. I calculate the increase in inland transport costs due to potential changes in the Yellow River.[60] The effect of roadblocks remains robust to this alternative change in the inland transport network. In column (4), I further include geographical controls, specifically the differences in latitude, longitude, and the geographic distance between two locations. The coefficient for the effect of roadblocks remains similar with geographical controls.

[60]The coefficient on this alternative change is, however, negative and insignificant. One explanation may be that the cost increase due to this alternative channel was of a much smaller magnitude than that caused by roadblocks. Affected pairs were often able to find nearby inland transport networks available.

Columns (5) to (8) show the results with probit specifications, which paint a similar picture. The increase in inland transport costs due to roadblocks was associated with a higher probability of post-rebellion sea trade, taking into account the unconstrained relative inland transport costs, the increase in inland costs due to the Yellow River shifting course, and geographic differences.

Table D.2 reports the results with explanatory variables in levels, showing robustness of the effect of roadblocks with alternative specifications for transport costs.

Table D.1: Roadblocks and post-rebellion trade

| | Dep. var.: Prob. of post-rebellion maritime trade | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | LPM | | | | Probit | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Δ ln (land cost) | 0.421 | 0.466 | 0.473 | 0.444 | 1.314 | 1.377 | 1.519 | 1.809 |
| | (0.125) | (0.123) | (0.097) | (0.149) | (0.587) | (0.507) | (0.377) | (0.537) |
| ln (land/sea cost) | | 0.275 | 0.272 | 0.291 | | 0.760 | 0.733 | 0.978 |
| | | (0.076) | (0.079) | (0.084) | | (0.257) | (0.264) | (0.297) |
| Riverchange | | | Yes | Yes | | | Yes | Yes |
| Geographic controls | | | | Yes | | | | Yes |
| Observations | 88 | 88 | 88 | 88 | 88 | 88 | 88 | 88 |
| R-squared | 0.110 | 0.247 | 0.248 | 0.326 | 0.092 | 0.193 | 0.197 | 0.301 |

*Notes*: The linear probability models in columns (1) to (4) and the probit models in columns (5) to (8). *Ln (land/sea cost)* calculates the natural logarithm of the relative inland transport costs based on travel costs in *Panel A* of Table I. Δ *ln (land cost)* calculates the roadblock-induced increase in the natural logarithm of inland transport costs. *Riverchange* is the increase in the natural logarithm of the inland transport cost resulting from the shifting of the Yellow River. Geographic controls include log distance and differences in latitude and longitude. Columns (5) to (8) report the pseudo R-squared values. Robust standard errors in parentheses.

Table D.2: Roadblocks and post-rebellion trade, in levels

| | Dep. var.: Prob. of post-rebellion maritime trade | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | LPM | | | | Probit | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Δ Land cost | 0.013 | 0.013 | 0.012 | 0.010 | 0.039 | 0.042 | 0.039 | 0.051 |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.015) | (0.015) | (0.013) | (0.018) |
| Land/sea cost | | 0.030 | 0.032 | 0.038 | | 0.086 | 0.088 | 0.131 |
| | | (0.012) | (0.012) | (0.012) | | (0.037) | (0.037) | (0.051) |
| Riverchange | | | Yes | Yes | | | Yes | Yes |
| Geographic controls | | | | Yes | | | | Yes |
| Observations | 88 | 88 | 88 | 88 | 88 | 88 | 88 | 88 |
| R-squared | 0.143 | 0.228 | 0.237 | 0.379 | 0.118 | 0.185 | 0.189 | 0.358 |

*Notes*: The linear probability models in columns (1) to (4) and the probit models in columns (5) to (8). *Land/sea cost* calculates the relative inland transport costs based on travel costs in *Panel A* of Table I. Δ *Land cost* calculates the roadblock-induced increase in the inland transport costs. *Riverchange* is the increase in the inland transport cost resulting from the shifting of the Yellow River. Geographic controls include geographic distance and differences in latitude and longitude. Columns (5) to (8) report the pseudo R-squared values. Robust standard errors in parentheses.

## AppendixE. 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 $p = \Phi(\gamma \ln(D^l/D^s)_{ij,post})$ 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 $q = \Phi(\gamma \ln(D^l/D^s)_{ij,post} - i^s)$ 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:

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

Note that compared to Equation 8, the dynamic model has two additional terms $\beta q i^s$ and $\beta(p - q)\gamma \ln(D^l/D^s)_{ij,post}$, 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 $\beta$, the discount factor, is larger than 0. Formally, we can write down the post-rebellion probability of sea trade using Equation 9, and it becomes:

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

Table E.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 | $\gamma$ | 0.477 |
| (normalised) | | (0.094) |
| Sunk cost of new sea route | $i^s$ | 2.068 |
| (transformed) | | (0.404) |
| Sunk cost of new land route | $i^l$ | 1.275 |
| (transformed) | | (0.470) |
| 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 | 0.680 | 0.360 |
| | | | (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 76 ($I^s = \exp(i^s/\gamma)$). The implied cost ratio of new inland route over existing one is 14 ($I^l = \exp(i^l/\gamma)$). 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-cost 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 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. McFadden's pseudo R-squared value is 0.183. Bootstrap standard errors using 500 replicates in parentheses.

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

Table E.1 shows the estimation result of the dynamic model when we set the discount factor $\beta$ 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 76 times that of an existing one, while the cost ratio for a new inland route versus an existing one is around 14. However, the impact of roadblocks remains consistent with the baseline model, regardless of whether firms are forward-looking or not. The model predicts a 0.68 probability of sea trade for "roadblocked" pairs after the rebellion, compared to a 0.36 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.

## AppendixF. 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, $\eta_{it}$ and $\psi_{jt}$ respectively, can vary over time. $\kappa$ measures the elasticity of migration flows to migration costs $D_{ijt}$. $\nu_{ijt}$ is an error term.

Historical migration flows are rarely available. I construct a new dataset of migration flows in historical China by leveraging a study that traces the origins of present-day villages. Throughout history, many villages in China were established by clans from external regions during periods of deteriorating local living conditions. Due to the tradition of maintaining detailed family histories, Cao (1997) managed to collect and assemble 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 $j$ founded by people from location $i$ at time $t$, I proxy for the migration flow from $i$ to $j$ at $t$.[61]

The following tables present the estimation results on migration gravity, using the number of migration villages as a proxy for migration flows. Table F.1 shows results by using transport cost parameters estimated in Section 5.2 to measure migration costs. Table F.2 shows results using geographic distance.[62] [63]

Column (1) of Table F.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

[61] The implicit assumption here is that the initial population when the (survived) villages were founded was similar. 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.

[62] Instead of the exact year of formation, we only know whether the migration time falls between 1643 and 1722, 1643 and 1796, 1723 and 1796, or between 1796 and 1912. To account for that the number of migrants/migration villages will be larger for longer time periods, each migration pair is divided by the total number of years of the corresponding time frame.

[63] 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 and the estimated elasticity is similar to that estimated from the sample where information about the prefecture/county is available. The elasticity estimate using the average cost/distance to different prefectures in the province is also negative, however, the estimated magnitude and variance are much larger.

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

Table F.2 presents the results on migration elasticities using geographic distance instead of transport costs. In the extreme scenario where distance does not correlate with transport costs, and if we assume that migration flows are solely influenced by transport costs, the relationship between migration and distance would be null. In the intermediate scenario, where distance is considered an imperfect 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 F.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 31.4% compared to using transport costs.

Table F.1: Migration gravity: Transport cost

| | Dep. var.: Ln (number of villages) | | | |
|---|---|---|---|---|
| | All<br>(1) | Kiangsi only<br>(2) | 1643-1722<br>(3) | 1723-1912<br>(4) |
| Ln (transport cost) | -1.481 | -1.267 | -1.002 | -1.545 |
| | (0.356) | (0.307) | (0.440) | (0.431) |
| Origin fixed effects | Yes | Yes | Yes | Yes |
| Destination fixed effects | Yes | Yes | Yes | Yes |
| Observations | 113 | 93 | 45 | 56 |
| R-squared | 0.893 | 0.836 | 0.823 | 0.917 |

*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 columns (3) and (4). Robust standard errors in parenthesis.

Table F.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 columns (3) and (4). Robust standard errors in parenthesis.