Communities of Commerce:

The Legacy of Chinese Immigration on Java, Indonesia

Quoc-Anh Do
Monash University

Sebastian Ellingsen
University of Bristol

Gedeon Lim
University of Hong Kong

October 27, 2024

PRELIMINARY AND INCOMPLETE

PLEASE DO NOT CITE WITHOUT PERMISSION

Abstract

Ethnic minorities have played an out-sized role in both historical and contemporary economic development. This paper studies one of the most prominent but understudied cases, in the world's most populous island: the economic legacy of ethnic Chinese immigration on Java, Indonesia in the early 21st century. Using an instrumental variables strategy, we exploit plausibly arbitrary 15th-century ethnic Chinese landing sites on the Northern coast of Java that determined initial Chinese presence but have not been accessible by sea since the mid 17th-century due to silting. We find a strong positive effect on local economic development. In districts with higher ethnic Chinese shares, villages are wealthier as measured by consumption, population density and nighttime light intensity. We trace this to the persistently important economic role of Chinese in trade and commerce. Districts with a historically higher ethnic Chinese share have, today, a higher (lower) employment share in services (agriculture), firms have higher sales and these positive effects can be traced to a larger financial industry. Furthermore, we find that subsectors more closely linked to traditional ethnic Chinese sectors continue to employ a larger share of individuals. Our findings suggest that the ethnic Chinese minority in Indonesia, despite their small size, have and continue to play an out-sized role in Indonesia's modern-day development with spillovers that benefit the entire local population.

*We thank Sam Bazzi and Alex Rothenberg for sharing their data. We thank Sheva Sonia Rahmani for excellent research assistance. Do: quoc.anh.do@monash.edu. Ellingsen: School of Economics, University of Bristol, 12A Priory Rd, Bristol, United Kingdom. sebastian.ellingsen@bristol.ac.uk. Lim: Faculty of Business and Economics. Pokfulam Rd., Hong Kong SAR. Email: gedeonl@hku.hk

1 Introduction

Ethnic minorities have played an out-sized role in historical and contemporary development. Examples abound of how small immigrant communities, both historically and in the present, have spurred on economic development by bringing innovation, and skills and complementing the skill sets of the native population. Yet, quantitative evidence is limited and the mechanisms remain poorly understood, especially in developing countries where the long-term impacts may differ greatly. In this paper, we seek to shed light on this question by exploring the role of Chinese minority communities in long-term economic development in Java, a business community that, globally, has been called one of the most powerful economic forces in the world (Redding, 2013).

Ethnic Chinese immigration into Indonesia began as early as the 13th century. These earlier waves were often a result of poorer individuals escaping from the lack of cultivable land, wars, and famines in southern China, primarily what is today the provinces of Guangdong and Fujian. In 1407, Zheng He's first expeditions mostly landed on the northern shore of Central Java around today's Semarang, before moving on to various settlements on the now-defunct Muria strait and then larger settlements in Eastern Java. While some areas in Eastern Java already had some Chinese communities before, the ones around Semarang and the Muria strait were mostly populated with Chinese after Zheng He's expeditions, including a fraction of the crew that decided to stay and build a Chinese community in Semarang. Over the subsequent six centuries, waves of Chinese immigrants continued to settle across Java.

While Chinese communities likely promoted long-term economic development, the observed association between Chinese communities and economic development can also be driven by sorting. Specifically, if settlers sought locations with particularly favorable fundamentals, or were forced to settle more marginal land, this could explain any correlation between the size of a Chinese community and economic development. To address this issue, we use an instrumental variable strategy motivated by features of Chinese settlement in Indonesia. The size of contemporary communities is shaped by historical settlements, which were in part driven by the choice of Chinese maritime portage sites on Java's northern coast. Many of these portage sites, chosen based on pre-industrial maritime connectivity with limited information about the hinterland's economic potential, have since become inaccessible from the sea due to shifts in the location of the coastline. As a result, their influence on contemporary economic development likely stems from their role

in shaping the location of Chinese communities. We challenge this assumption in several ways and provide evidence in support of this interpretation below. Furthermore, we adjust for different exposure to the portage sites driven by geography by implementing the procedure developed in Borusyak and Hull (2023).

We find a strong, precisely estimated impact of distance to the Muria strait on the size of the Chinese communities, as measured as the share of the population belonging to a Chinese community, across districts in 1930. Specifically, an increase in 1km distance to the Muria strait leads to a decrease in ethnic Chinese shares of about 9 percentage points in the preferred specification. Reassuringly, the association between distance to the strait is robust to controlling for differences in observable location fundamentals.

We next use the distance to the strait, adjusted for average exposure consistent with Borusyak and Hull (2023), as an instrument for the size of the Chinese community. We begin by exploring the impact of the size of the Chinese community by district on contemporary economic development. We find that a one percentage point increase in the size of the Chinese community leads to an around 7.6 percent increase in average household consumption today. Furthermore, we also find a 21.1 percent increase in the population density in 2000. We next explore the impact on nightlight luminosity following Henderson et al. (2012). Consistent with the previous results, we find a positive impact on nightlight luminosity. In particular, a one percentage point increase in 1930 Chinese shares leads to a 3.04 increase in night-time light intensity between 1992 and 2010. Last, we find strong evidence of structural transformation. Districts with 1 percentage point higher 1930 Chinese shares have 8.9 (6.1) percentage points lower (higher) employment share in agriculture (services). Taken together, the findings suggest a positive impact of the size of the Chinese community on contemporary economic development.

Why are districts with larger Chinese communities more economically developed? We consider several possibilities. Motivated by the literature we consider several mechanisms such as investment in human capital (see e.g. Rocha et al., 2017; Droller, 2018; González, 2020), the transmission of an entrepreneurial or market-oriented culture (Figlio et al., 2019), access to commercial networks and trade facilitated by Chinese communities (Johnson and Koyama, 2017). Exploring the effect of Chinese community size on sectoral employment, we find evidence of larger employment shares in small-scale commerce in the services sector. Taken together, these findings provide evidence that

Chinese communities have had a positive impact on economic development, possibly through the diffusion and establishment of cultural practices promoting market participation.

This paper contributes to several literatures. First, we contribute to a growing literature on the effects of historical immigration on long-run development. Previous studies have highlighted the importance of human capital in shaping immigrants' effects on economic development (Rocha et al., 2017; Droller, 2018; González, 2020; Burchardi et al., 2020; Sequeira et al., 2020). Notably, Burchardi et al. (2020) which finds that immigration into the US has had a net positive impact on development over a 40-year period, driven by highly educated innovators. Sequeira et al. (2020) show that low-skilled immigration in the US had positive effects in the early stages of industrialization. This paper focuses on a single minority group in a *developing country* and shows how *low-skilled* immigrants can drive local development through capital accumulation and local, social commercial networks and practices. In this sense, our findings also speak to evidence of the role of migrant communities in the diffusion of cultural practices and norms (Valencia, 2019; Dippel and Heblich, 2021).

Our work also relates to a rich literature on the long-term economic effects of ethnic and religious minorities (Jha, 2013; Hornung, 2014; Johnson and Koyama, 2017). There are notable parallels between the economic roles of Jewish communities in Europe and Chinese communities in Southeast Asia. In this regard, our paper builds on Johnson and Koyama (2017) who study the role of Jewish communities in promoting economic development in pre-industrial Europe. Our paper differs in emphasizing the long-term impact beyond the pre-industrial context. Methodologically, our paper also builds on Johnson and Koyama (2017) as Chinese settlement followed maritime trading routes, unlike the overland routes followed by Jewish communities in Europe. Finally, and more broadly, our paper contributes to the quantitative literature on determinants of long-term economic development in Indonesia (see e.g. van der Eng, 1992, 2010; Dell and Olken, 2020).[1]

The paper proceeds as follows. Section 2 discusses key aspects of various waves of ethnic Chinese out-migration to Indonesia starting from the 15th century till the early 20th century. We especially highlight those relevant to our identification strategy. Section 3 describes our data sources. Section 4 outlines our empirical strategy. Section 5.1 describes our main results and robustness. Section 6 explores the mechanisms. Section 7 concludes.

[1]There is a small related literature on economic history on Indonesia. In particular, Hup and de Zwart (2024) studies changes in wages of Chinese emigrants to Indonesia in the late 19th to early 20th century.

2 Historical Background: Chinese Immigration to Indonesia

This section describes the background of the historical immigration of ethnic Chinese into Indonesia over six centuries. In particular, we focus on three key facts that inform our main identification strategy: (1) Pre-15th century immigration; (2) The (lack of) choice over portage sites in the early 15th century that led to plausibly exogenous initial ethnic Chinese settlement patterns; (3) Post-15th century immigration.[2]

First, Chinese emigration to Java before the 15th century was highly limited.[3] While evidence of Chinese immigrants can be traced as far back as the 13th century, there is little written record of their presence throughout most of Java. Based on the only written account of Zheng He's expeditions in the 1400s by a member of the crew, Ma Huan (Ma Huan 1433, translated by Mills, 1970), the expeditions only encountered some Chinese on shore in Turban and Gresik in Easter Java.

Second, the appearance of large ethnic Chinese settlements on Java can be largely traced to Zheng He's voyage into the Dutch East Indies (as Indonesia was then called) in the early 15th century. Zheng He commanded one of the largest sea fleets in the world in his multiple voyages from China down to various parts of Southeast Asia. His travels are extensively documented in primary archival sources. In particular, contemporary ethnic Chinese settlement today can be traced to many of the portage sites along the north coast of Java where members of his crew and voyage landed. Figure 1 plots 1930 ethnic Chinese shares and, with a solid black line, Zheng He's sea passage along the historical north coast of Java, the *Muria Straits*. These portage sites are based on historical records that record the passage of Zheng He's fleet through the historical towns of Semarang, Demak, Koedoes, Pati, Rembang and Lasem, as portage sites. Importantly, the historiography suggests that these sites were unlikely to have been selected based on economic or locational fundamentals – seafarers then had little recourse to modern-day navigational equipment and Zheng He's fleet was visiting a hitherto unknown land. Furthermore and crucially, all of these portage sites are, today, in-accessible by water due to extensive silting from Mount Muria in 1657. This allays concerns that advances and changes in latter-day maritime trade and migration could be omitted factors driving any results we find on ethnic Chinese shares.

Third, the first main influx of post-15th century Chinese immigrants arose with the fall of the

[2]This section is largely based off Suryadinata (2005) unless stated otherwise.
[3]The discussion in this page draws extensively from Hoadley (1988) and Carey (1984)

Ming dynasty (1644) and the reopening of Chinese trade with Southeast Asia in 1683. This created ideal conditions for an increased flow of immigrants from the main southern maritime provinces like Hokkiens from Amoy in Fukien province and Kwong Fus (Cantonese) from Canton and Macau. Many Chinese assimilated very closely with the indigenous population, married local women, and adopted local customs and religions like Islam. An important reason was that virtually no Chinese women migrated overseas before the late 19th century. A possible second main influx of Chinese immigrants to Indonesia arose near the end of the fall of the Qing dynasty (1644 - 1911) where many ethnic Chinese fled to Southeast Asia due to famine and war.

Crucially, kinship ties and pre-existing clan networks in Southeast Asia were the key determinants of where latter-day Chinese immigrants chose to settle (Setijadi, 2023). Hence, we see in our data that initial portage sites in the early 15th-century, plausibly unrelated to modern-day locational fundamentals, have ethnic Chinese communities that are disproportionate in both size and share relative to 19th-century shipping and aviation routes.

The Historical Role of Ethnic Chinese in Indonesia Chinese migrants have played important economic and political roles. These included that of retailers, tax-collectors, tax-farmers as well as middlemen between Javanese princes and Dutch officials during the Dutch colonial period. At the extreme, a number of Chinese were even appointed as *tumenggung*, or retainers in the Islamic courts, and were assigned vassals and rice fields for support. Many of these Chinese took on local names and titles. The importance of the Chinese in pivotal economic roles can be seen in the following quotes. First, most tax collectors and farmers were Chinese. Along the north coast of Java, van Hoorn wrote that "there is not a river, harbor, bay, or creek navigable for shipping, which does not have a customs post (tollgate), of which the keeper is invariably a Chinese." There are also indications that whole provinces of the Mataram empire were farmed by the Chinese during this period. For example, the fertile, rice-producing areas of Pekalongan and Batang, amounting to 20 percent of the entire coastal population in 1704, were leased by the Sunan (provincial ruler) to a Chinese family who ruled under Javanese names.

The Chinese were also active as retailers in the profitable opium and textile traders, middlemen handling local commerce in raw cotton, salt, indigo, tobacco, ginger, silks, and other important commodities. Many eventually rose to become entrepreneurs and large traders who played important roles in developing Indonesia's nascent rice and sugar growing industries. Furthermore,

Chinese merchants came to play a crucial in informal capital markets already in the early 19th century (Zanden and Marks, 2012, p. 32).

The Contemporary Role of Ethnic Chinese in Indonesia Ethnic Chinese businesses have been called one of the most powerful economic forces in the world (Redding, 2013). In particular, Indonesia is the largest country in Southeast Asia and is the single country with the largest absolute number of ethnic Chinese (Suryadinata, 2008). In terms of shares, however, the ethnic Chinese are one of the smallest ethnic groups in Indonesia, with estimates ranging from 1.5 to 2 percent of the total Indonesian population of 230 million. Despite their small share, however, the ethnic Chinese in Indonesia have played a significant economic role. For example, the share of total private domestic capital owned by ethnic Chinese exceeds that of any other ethnic group in Indonesia (Efferin and Pontjoharyo, 2006).

3 Data

In this section, we describe our main data sources. We utilize various sources of rich individual and firm-level data at various points from the early 20th to early 21st century. First, we construct our main explanatory variable of district-level ethnic Chinese shares from the Indonesian Population Census (1930, 2000, 2010). We also construct individual-level economic sector participation rates from the 2010 Census. Second, we construct measures of firm-level profits, sales, and productivity from the Indonesian Manufacturing Census (*Statistik Industri*) (1980-2006).[4] Third, we construct measures of village-level household consumption, trust, and attitudes towards the state from the Indonesian Socio-Economic Surveys (*Susenas* 2000-2012). Fourth, nightlight intensity from NGHIS. Fifth, village census data (*PODES*).

Indonesian Population Census. First, we construct ethnic Chinese shares from 3 rounds of the Indonesian Population Census: 1930, 2000, and 2010. The 1930 census is uniquely suited for our purposes. It was the first to provide a complete enumeration of all Indonesians of Chinese descent at the municipal and sub-district level. We use this to construct our main explanatory variable, *Chinese Shares*, and for estimation purposes, aggregate this measure to the district level. Notably,

[4]We supplement this with two rounds of the Indonesian Firm Census in 2006 and 2016.

the 2000 census is the first census round, since 1930, in which individual-level ethnicity is recorded – all questions on ethnicity were banned under Suharto (1965-1999).[5]

Second, we construct measures of employment shares in various economic sectors using the 2010 Population Census. Crucially, these sectors include trade, transportation, and finance which the ethnic Chinese specialized in (Pepinsky, 2016). We estimate the effects on two levels. First, on aggregated shares in agricultural, manufacturing, and service. Second, on specific sectoral shares as informed by the historical and anthropological literature.

Figure 1 plots the spatial distribution of ethnic Chinese communities, in terms of ethnic shares, from the 1930 census. We document the following key facts. First, the majority of Chinese are concentrated around the Northern coast of Java. Visually, the impact of the Muria Straits, outlined with a solid black line, on Chinese settlement patterns is clear. In Central Java, we see a high ethnic Chinese concentration along the Muria Straits, which decreases as we move further away. Importantly, these settlements are situated along a coastline that has not existed since the early 18th century. Yet, ethnic Chinese concentration continues to persist in the 1930s, underscoring the importance of initial settlement patterns in determining latter-day ethnic settlements. Second, and more generally, the spatial distribution reflects traditional maritime shipping routes driven by wind and steam which resulted in fewer Chinese settling on the southern coast and in the interior of Java.

Socio-Economic Surveys We are interested in estimating the impact of Chinese presence, be it through business establishments, credit, loans, or other mechanisms highlighted in the [previous section], on the outcomes of local Indonesians of non-ethnic Chinese descent. Susenas covers a representative sample of households and individuals in each round. We pool all available rounds from 2000-2010 (following Dell & Olken 2020) to obtain household-level measures of consumption of local, non-Chinese individuals. Pooling all Susenas rounds, we construct our main measure of household-level income and welfare by taking the logarithmic of equivalent household consumption (Deaton, 1997) between 2000-2010.

Village Census (*PODES*) This gives us measures of population density in 2000, we supplement this with population counts from the 2010 Population Census.

[5]Having said that, we have found archival records of ethnic Chinese counts in the 1970s and are in the process of digitizing district-level measures of ethnic Chinese shares in 1970 as an additional robustness check.

Night-lights We use average night-time light intensity from NGHIS, aggregated to either the village or district-level polygons in 2000.

Manufacturing Census (*Statistik Indonesia*) We use all rounds from 1980-2006. The Manufacturing Census covers all firms that employ minimally 20 to 30 employees and is limited to firms involved in any form of manufacturing activity. Hence, we supplement this with the Economic Census in 2016 (described below). THe Manufacturing Census gives us measures of firm-level sales, profits, and productivity which we aggregate to the district level.

Economic Census 2016 We use the latest round of the Indonesian Economic Census. This dataset comprises the universe of small and medium firms in rural areas (>= [20-30] employees) and a representative sample of similarly sized firms in urban areas. The key advantage of this dataset is that it surveys both manufacturing and non-manufacturing firms, and classifies these firms into 19 disaggregated economic sector categories. Firm-size measures here, however are largely limited to number of employees. Hence, we supplement this with various rounds of the Manufacturing Census (*Statistik Indonesia*).

4 Empirical Strategy

We begin our exploration of the long-run association between the size of Chinese communities on contemporary economic development using the following regression model,

$$\ln(Y_{pd}) = \gamma_p + \delta \ln(C_{pd}) + X'_{pd}\gamma + \varepsilon_{pd}, \tag{1}$$

where Y_{pd} is the level of economic development in province p and district d, and C_{pd} denotes the size of the Chinese community in district d as measured by the share of Chinese of the total population in 1930. X'_{pd} includes observable location fundamentals such as agricultural productivity and proximity to large urban centers, that might have affected the formation of Chinese communities and economic development. δ_p is a province-fixed effect that captures unobserved heterogeneity at the province level. Finally, ε_{pd} is an unobserved error term, potentially correlated within provinces.

In Equation 1, δ approximates the average difference in economic development across districts with where the Chinese community differs in size, keeping unobserved heterogeneity at the

province level and observed location fundamentals fixed. Two broad factors influence the size of δ. First, differences in contemporary development can be *caused* by the communities through human capital, cultural persistence, or improved commercial linkages to China. Second, the observed association might be influenced by the sorting of settlers into locations that differ in unobserved characteristics. This could happen if Chinese immigrants could only settle more marginal land, with lower agricultural productivity. This might have stunted subsequent industrialization, and urbanization in districts with larger Chinese communities. Furthermore, classical measurement error could attenuate the estimated observed association between Chinese community size and contemporary development.

To isolate the causal impact of the size of Chinese communities, we therefore implement an instrumental variable strategy that leverages two key features of Chinese settlement in Indonesia. Since the size of Chinese communities is highly persistent, contemporary settlement is driven by historical settlement which in turn was influenced by the choice of Chinese maritime portage sites on the Northern coast of Java. We hypothesize that the locations chosen by the initial pioneers in turn influenced subsequent migrants' location choices. The locations of these portage sites were primarily driven by pre-industrial maritime connectivity and were chosen with limited information about the economic potential of the hinterland. We therefore assume that locations with varying distances to these portage sites are similar in unobserved characteristics. Importantly, due to soil erosion and resulting shifts in the location of the coastline, the initial portage sites have not been accessible by sea since the early 18th-century. As a result, the portage sites are unlikely to be important locations for trade or destinations for Chinese migration. Taken together, we assume that any long-term impact of proximity to these historical portage sites is mediated through its effect on the formation of Chinese communities.

We consider the following model:

$$\ln\left(Y_{pd}\right) = \alpha_p + \beta \ln\left(C_{pd}\right) + X'_{pd}\omega + u_{pd}, \tag{2}$$

where the variables are defined as above. β now denotes the causal impact of changes in the Chinese community size on contemporary economic development in a district. To estimate β, we estimate the first stage relationship between the distance to the portage site and the size of the Chinese

community

$$\ln(C_{pd}) = \mu_p + \rho \ln(\mathrm{D}_{pd}) + X'_{pd}\lambda + \psi_{pd}, \tag{3}$$

where D_d denotes the distance to the portage site for district d. Following the above discussion, we assume that the distance to the portage sites is unrelated to unobserved determinants of economic development and only affects contemporary development through its impact on the location of Chinese communities. We discuss the plausibility of these assumptions below.

In Equation 2, β captures the difference in economic development across locations with different distances to the historical Chinese portage locations, keeping unobserved determinants of contemporary economic development constant. However, given the localized nature of the shock, some locations would be more likely to be treated regardless of which segment of the coastline actually served as portage locations. Regardless of which segment of the coastline the portage site emerged, due to its central location, it would have had a higher average exposure to the shock under any treatment assignment.

To address this concern, we follow Borusyak and Hull (2023) and calculate the *expected* treatment of each district that we denote $\mathbb{E}[\ln(\mathrm{D}_{pd})]$ where the expectation is taken over alternative treatment assignment procedures over all the potential portage sites. Specifically, for each coastal segment, we simulate the assignment process of Chinese portage sites and average the distance over all potential assignments. For each district, we then subtract the expected distance to a portage location from the realized distance. Figure 2 provides a visual depiction of this process and Section A.1 discusses this process in further detail. The resulting first stage specification is therefore given by

$$\ln(C_{pd}) = \theta_p + \eta \ln\left(\mathrm{D}'_{pd}\right) + X'_d\phi + \upsilon_{dp}, \tag{4}$$

where $\mathrm{D}'_{pd} = \mathrm{D}_{pd} - \mathbb{E}[\mathrm{D}_{pd}]$. Finally, we cluster standard errors at the 1930 Dutch Regency level.[6]

5 Results

This section presents our main results. Throughout, we report Kleibergen-Paap F-statistics and report all other relevant weak-instrument robust test statistics. We begin by estimating our first-stage

[6] In the 1930s, Java was largely divided into the following administrative tiers. From the top: Residency, Regency, District, and Sub-District. Modern-day administrative tiers, from the top, are Province, District, and Sub-District. 1930 Regencies are roughly equivalent in size and importance to large modern-day districts.

results of distance to 15th-century portage sites on 1930 ethnic Chinese shares. We then investigate the effects of 1930 ethnic Chinese shares on (i) wealth and economic development, as measured by village-level household consumption, nightlight intensity, and district-level population density; (ii) structural transformation. In the subsequent section, we present results on key mechanisms including increases in district-level economic specialization in the financial and services sector using firm-level counts from the 2016 Economic Census and individual-level economic participation data from the full-count 2010 population census.

5.1 Main Results

First-stage. Table 1 presents our first-stage results from estimating equation (2). Column (1) shows the raw correlation between ethnic Chinese shares in 1930 and our recentered IV. An increase in 1km distance leads to a decrease in ethnic Chinese shares of about 19.2pp. Column (2) shows our preferred specification: the correlation between ethnic Chinese shares in 1930 and our recentered IV, conditional on a battery of predetermined covariates.[7] The coefficient nearly halves in size but, reassuringly, remains statistically significant at the 1% significance level. Having said that, we note that the Kleibergen-Papp F-stat for our instrument in our later results falls slightly below 10 and hence we report additional weak-instrument robust test statistics throughout. In general, p-values for under-identification remain below 0.05.

Contemporary Development. Table 2 estimates our results on three measures of contemporary economic development. Column (1) estimates that a 1p.p. increase in 1930 ethnic Chinese shares leads to a 7.6% increase in average household consumption today. Similarly, Column (2) estimates that a 1p.p. increase in 1930 ethnic Chinese shares leads to a 21p.p. higher population density in 2000. Finally, Column (3) estimates the effects of 1930 Chinese shares on changes in night-time light intensity (Henderson et al., 2012) between 1992, the first year for which night-time light intensity data is available, and 2010. A 1p.p. increase in 1930 Chinese shares leads to a 3.04 increase in night-time light intensity from 1992-2010. Compared to the mean, our estimates suggest that increases in nighttime light intensity, in districts with 1p.p. higher ethnic Chinese shares, was nearly a third higher.

[7] These covariates are altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan, Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude.

Structural Transformation. We continue by examining if the higher economic development levels documented above may have accelerated structural transformation (see e.g. Herrendorf et al., 2014). Table 3 estimates results on employment shares in agricultural, manufacturing, and services in 2010. Columns (1), (2), and (3) show that a district with 1p.p. higher ethnic Chinese shares in 1930, has on average, 8.9p.p. fewer individuals working in the agricultural sector, 2.8p.p. more individuals working in the manufacturing sector, and 6.1p.p. more individuals working in the services sector. Taken together, our results suggest that districts with higher ethnic Chinese shares are, today, more economically developed with a higher (lower) share of individuals working in services (agricultural). Are there specific sectors within the service industry that experienced positive effects from increases in Chinese presence? We turn to this when exploring mechanisms in Section 6.

5.2 Robustness

Our results are robust to alternative specifications. In particular, Table A.1 shows that results on contemporary development are robust to the inclusion of West Java. Similarly, Table A.2 checks for the robustness of our results on structural transformation. Column (1) shows that estimates on agricultural employment shares remain negative and statistically significant. Columns (2) and (3), show that estimates on manufacturing and service employment shares remain positive but become statistically insignificant. This is possibly due to the weaker strength of our instrument and is expected given that growth in West Java is strongly driven by economic and locational fundamentals that have less to do with plausibly exogenous, historical, ethnic Chinese settlement. For example, Jakarta and the largest concentration of special economic zones in Indonesia, with large local and foreign investment, are located in both the interior and along the northern coast of West Java.

6 Key Mechanisms

Historical ethnic Chinese settlement has positive and persistent effects on development today that can be observed in consumption, population density, and higher nightlight growth. What explains these persistent effects? The economic and historical literature suggests 3 main mechanisms through which Chinese communities could have led to higher economic development: (1) Human capital; (2) Cultural transmission; (3) Market access and trade; and (4) Economic specialization and

complementarities. We show estimates for our results on (4), and are in the process of collecting data to test mechanisms (1), (2), and (3).

Human capital. Squicciarini and Voigtländer (2015) and Gennaioli et al. (2013) argue that elite human capital may be particularly important for transmitting knowledge and ideas. Anecdotal sources suggest that contemporary human capital in Chinese communities is relatively higher but the situation regarding historical human capital is unclear. Relatedly, Rocha et al. (2017) documents the persistence of human capital from immigrants in Brazil with higher levels of schooling and shows that there is a positive effect on long-term development.

Cultural transmission. Figlio et al. (2019) look at the impact of the cultural transmission of a "Long Term Orientation" mindset of American immigrants to their children on educational outcomes. Here, the question would be whether there was a transmission of a culture of entrepreneurship and hard work and the importance of these channels beyond that of human capital, to the natives.

Market access and trade. Johnson and Koyama (2017) show that cities with Jewish communities in pre-industrial Europe experienced greater benefits from an increase in market access across Europe between 1400 and 1850. Cities with Jewish communities were better able to take advantage of the intensive margin and translate greater market density into growth. Similarly, Chinese communities have played an out-sized role as tax-farmers, tax collectors, and entrepreneurs since the early 18th century (Carey, 1984; Hoadley, 1988) which is suggestive of a similar explanation for the link between Chinese communities and economic growth in Indonesia.

Economic specialization and complementarities. A rich literature documents that the ethnic Chinese were often, from the mid-20th century onwards, excluded from both the military and politics (Setijadi, 2023). This resulted in many ethnic Chinese remaining in their traditional roles of trade, finance, and commerce. Given this, we hypothesize that districts with higher ethnic Chinese shares might have experienced larger firm growth due to ethnic Chinese specialization in these sectors.

In the Indonesian context, firm panel data does not exist. As a proxy, Column (1) of Table 4 regresses average firm sales, calculated by pooling all rounds of the Indonesian Manufacturing Census (1986-2006), on ethnic Chinese shares. We find that a 1p.p. increase in 1930 ethnic Chinese shares leads to a 29.1p.p. increase in firm sales. Column (2) regresses the number of financial firms

from the 2016 Manufacturing Census on ethnic Chinese shares and finds that districts with 1p.p. higher ethnic Chinese shares in 1930 have nearly 65p.p. more finance firms.

Next, we turn to individual-level employment in various service sub-sectors that either have had a historical Chinese presence or are likely to be upstream and/or downstream to sectors where the ethnic Chinese have had a dominant presence. Table 5, Columns (1)-(6), regresses the number of individuals employed in Trade, Food and Accommodation (*FoodAccomm*), Transport, Information Communication and Technology, Finance, and Educational, Health and Social Services (*EduHealthSoc*) at the village-level. We find that higher ethnic Chinese shares are positively associated with higher employment in all sub-sectors except for ICT and Finance.[8]

7 Conclusion

This paper presents novel evidence of the effects of historical ethnic Chinese immigration and settlement on the trajectory of Indonesia's long-run economic development. We show that the ethnic Chinese have played a central economic role, disproportionate to their numbers, and disparities in economic development across much of Java can be traced to differences in initial ethnic Chinese settlement patterns. These long-run differences show up, crucially, in non-ethnic Chinese villages and non-ethnic Chinese individual-level measures of economic welfare and prosperity throughout Java and we provide evidence that the Chinese, in line with much of the qualitative historical literature, played an important role in financing and middle-men trade of businesses that continues to be apparent today. The role of the ethnic Chinese in Indonesian development has been under-studied in much of the existing literature and our paper corrects this.

[8]This is possibly because firms in ICT and Finance employ a smaller number of individuals on average.

References

Borusyak, Kirill and Peter Hull, "Nonrandom Exposure to Exogenous Shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

Burchardi, Konrad B, Thomas Chaney, Tarek Alexander Hassan, Lisa Tarquinio, and Stephen J Terry, "Immigration, innovation, and growth," Working Paper 27075, National Bureau of Economic Research 2020.

Carey, Peter, "Changing Javanese perceptions of the Chinese communities in Central Java, 1755-1825," *Indonesia*, 1984, (37).

Deaton, A, "The Analysis of Household Surveys: A Microeconometric Approach to Development Policy," 1997.

Dell, Melissa and Benjamin A Olken, "The development effects of the extractive colonial economy: The dutch cultivation system in java," *The Review of Economic Studies*, 2020, *87* (1), 164–203.

Dippel, Christian and Stephan Heblich, "Leadership in Social Movements: Evidence from the "Forty-Eighters" in the Civil War," *American Economic Review*, 2021, *111* (2), 472–505.

Droller, Federico, "Migration, Population Composition and Long Run Economic Development: Evidence from Settlements in the Pampas," *The Economic Journal*, 2018, *128* (614), 2321–2352.

Efferin, Sujoko and Wiyono Pontjoharyo, "Chinese Indonesian business in the era of globalization: Ethnicity, culture and the rise of China," *Southeast Asia's Chinese Businesses in an Era of Globalization: Coping with the Rise of China, Institute of Southeast Asian Studies, Singapore*, 2006, pp. 102–160.

Figlio, David, Paola Giuliano, Umut Özek, and Paola Sapienza, "Long-term orientation and educational performance," *American Economic Journal: Economic Policy*, 2019, *11* (4).

Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer, "Human capital and regional development," *The Quarterly journal of economics*, 2013, *128* (1).

González, Felipe, "Immigration and Human Capital: Consequences of a Nineteenth Century Settlement Policy," *Cliometrica*, 2020, *14* (3), 443–477.

Henderson, J Vernon, Adam Storeygard, and David N Weil, "Measuring economic growth from outer space," *American economic review*, 2012, *102* (2), 994–1028.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi, "Growth and Structural Transformation," in "Handbook of Economic Growth," Vol. 2, Elsevier, 2014, pp. 855–941.

Hoadley, Mason C, "Javanese, Peranakan, and Chinese elites in Cirebon: Changing ethnic boundaries," *The Journal of Asian Studies*, 1988, *47* (3).

Hornung, Erik, "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia," *American Economic Review*, 2014, *104* (1), 84–122.

Hup, Mark and Pim de Zwart, "Brain Gain in Late Colonial Indonesia? New Evidence on Chinese Migration and Wages," Technical Report, CEPR Discussion Papers 2024.

Jha, Saumitra, "Trade, institutions, and ethnic tolerance: Evidence from South Asia," *American political Science review*, 2013, *107* (4), 806–832.

Johnson, Noel D and Mark Koyama, "Jewish communities and city growth in preindustrial Europe," *Journal of Development Economics*, 2017, *127*.

Pepinsky, Thomas B, "Colonial migration and the origins of governance: Theory and evidence from Java," *Comparative Political Studies*, 2016, *49* (9), 1201–1237.

Redding, Gordon, *The spirit of Chinese capitalism*, Vol. 22, Walter de Gruyter, 2013.

Rocha, Rudi, Claudio Ferraz, and Rodrigo R Soares, "Human Capital Persistence and Development," *American Economic Journal: Applied Economics*, 2017, *9* (4), 105–136.

Sequeira, Sandra, Nathan Nunn, and Nancy Qian, "Immigrants and the Making of America," *The Review of Economic Studies*, 2020, *87* (1), 382–419.

Setijadi, Charlotte, *Memories of Unbelonging: Ethnic Chinese Identity Politics in Post-Suharto Indonesia*, University of Hawaii Press, 2023.

Squicciarini, Mara P and Nico Voigtländer, "Human capital and industrialization: Evidence from the age of enlightenment," *The Quarterly Journal of Economics*, 2015, *130* (4), 1825–1883.

Suryadinata, Leo, *Admiral Zheng He and Southeast Asia*, Vol. 286, Institute of Southeast Asian Studies, 2005.

_____ , *Ethnic Chinese in Contemporary Indonesia*, Institute of Southeast Asian Studies, 2008.

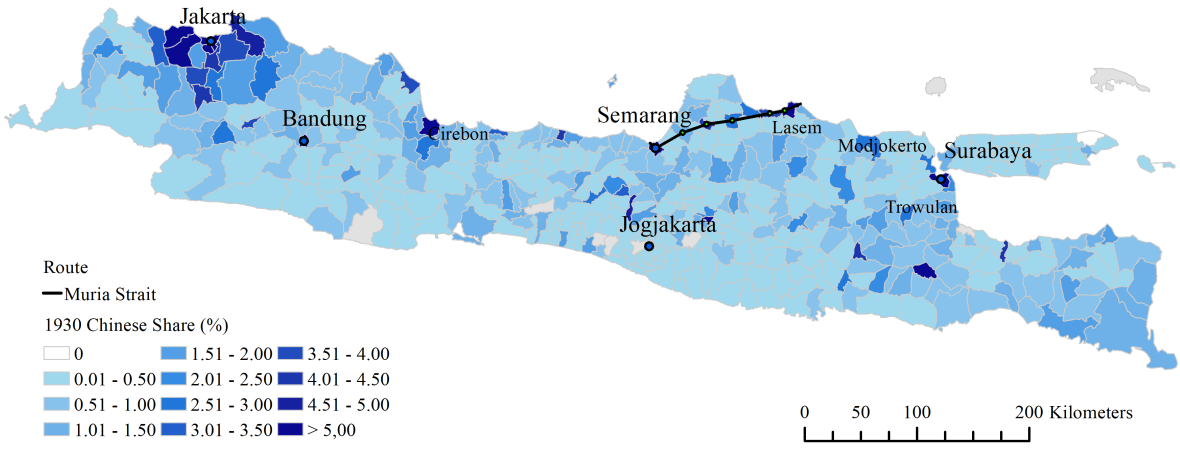
Valencia, Felipe, "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, 2019, *134* (1), 507–556.

van der Eng, Pierre, "The Real Domestic Product of Indonesia, 1880–1989," *Explorations in Economic History*, 1992, *29* (3), 343–373.

_____ , "The sources of long-term economic growth in Indonesia, 1880–2008," *Explorations in Economic History*, 2010, *47* (3), 294–309.

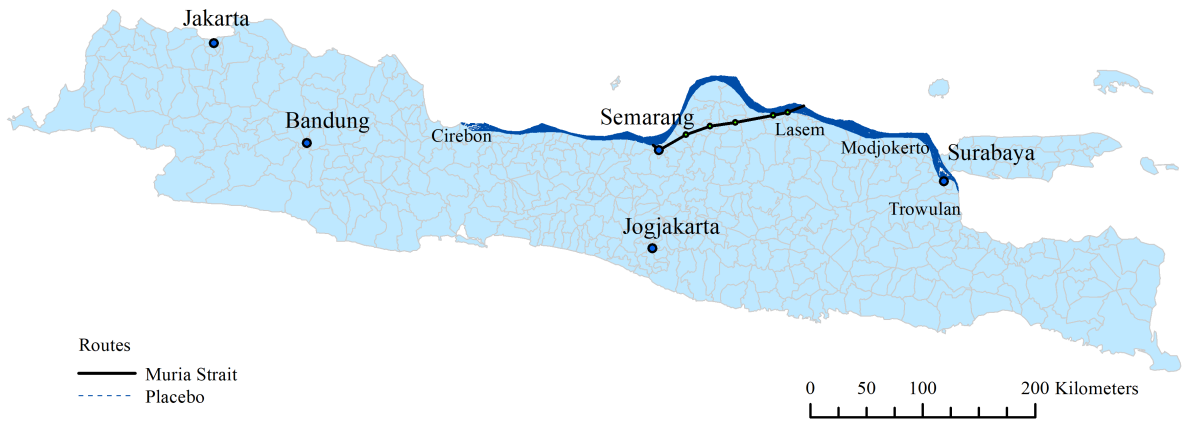
van Zanden, J. L. and Daan Marks, *An Economic History of Indonesia: 1800-2010* number 109. In 'Routledge Studies in the Growth Economies of Asia (2005).', Hoboken: Taylor and Francis, 2012.

Figure 1: The Distribution of Ethnic Chinese Shares at the 1930 District-Level on Java



This figure depicts the underlying variation in our treatment variable: the share of Chinese by district-level across 1930 districts. The solid black line traces the Muria Straits across historical ethnic Chinese portage sites in the early 15th-century as informed by historical sources. Source: Indonesian Population Census of 1930.

Figure 2: IV Strategy: Recentered IV based on Distance to Muria Straits



This figure illustrates our instrumental variable strategy. The black solid line traces out the actual Muria Straits, along which actual 15th-century ethnic Chinese portage sites were established. We obtain our main treatment variable, distance to Muria Straits, by calculating the distance between the centroid of each 1930 district and the Muria Straits. The dotted blue lines trace out 493 placebo routes that we obtain by keeping route length constant and shifting the route along the northern coast of Java. We calculate the distance between each 1930 district and each placebo route. Finally, for each 1930 district, we compute our recentered IV (Borusyak and Hull, 2023) by subtracting the average distance across all placebo routes from the distance to the actual Muria Straits.

Table 1: First Stage

	(1) Chinese Shares 1930	(2) Chinese Shares 1930
Dist. to Muria Straits, Recentered	-0.192**	-0.090***
	(0.084)	(0.029)
FE	Province	Province
Cluster	Regency	Regency
West Java	No	No
Controls	No	Yes
R2	0	1
Observations	248.00	242.00
Mean	0.94	0.91

Notes: Regression of ethnic chinese shares in 1930 (%) on our recentered IV of difference between actual log distance to muria strait and the average of log distances to placebo routes. Column (1) estimates the raw correlation and Column (2) estimates the conditional correlation upon inclusion of a battery of predetermined covariates of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the district-level. Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Historical Chinese Shares and Contemporary Development

	(1) consumption ('00-10)	(2) pop. density ('00)	(3) Δ light intensity ('92-10)
Chinese Shares 1930	0.076***	0.211***	3.044**
	(0.027)	(0.067)	(1.488)
FE	Province	Province	Province
Cluster	Regency	Regency	Regency
West Java	No	No	No
R2	0.940	0.677	0.499
Observations	31467	242	242
Mean	11.59	-7.16	9.71
F-stat	10.22	9.39	9.39
Underidentification Test, AR p-value	0.012	0.037	0.173
Underidentification Test, SW p-value	0.011	0.007	0.025

Notes: Regressions of, respectively, the natural logarithm of equivalent household income (pooling all years from the Indonesian Socio-Economic Survey *(Susenas)* 2000-2010), natural log of population density (Indonesian Village Census, 2000), and night-time light intensity (2000) (NHGHIS), on 1930 ethnic Chinese shares. All regressions include controls of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the village-level for column (1) and district-level for columns (2) and (3). Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Historical Chinese Shares and Structural Transformation

	(1) Agric. Shares (%, 2010)	(2) Manuf. Shares (%, 2010)	(3) Service Shares (%, 2010)
Chinese Shares 1930	-8.919***	2.811	6.108*
	(3.090)	(4.162)	(3.299)
FE	Province	Province	Province
Cluster	Regency	Regency	Regency
West Java	No	No	No
R2	0.246	0.202	0.177
Observations	12770	12770	12770
Mean	43.66	21.95	34.39
F-stat	2.21	2.21	2.21
Underidentification Test, AR p-value	0.204	0.621	0.009
Underidentification Test, SW p-value	0.011	0.397	0.097

Notes: Regression of the industry shares of working productive-aged males in 2010 on ethnic Chinese shares in 1930. Observations at the village-level. All regressions include controls of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the village-level for column (1) and district-level for columns (2) and (3). Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Firm Sales and Economic Specialization I: Financial Firms

	(1) firm sales ('86-'06)	(2) # of Finance Firms ('16)
Chinese Shares 1930	0.291*	0.658***
	(0.169)	(0.120)
FE	Province	Province
Cluster	Regency	Regency
West Java	No	No
R2	0.475	0.465
Observations	242	242
Mean	15.93	3.48
F-stat	9.39	9.39
Underidentification Test, AR p-value	0.069	0.002
Underidentification Test, SW p-value	0.003	0.019

Notes: Regression of the natural log of firm sales, averaged at the district-level, and the natural log of number of firms in the financial sector, at the district-level, on ethnic Chinese shares (1930). All regressions include controls of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the district-level. Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Economic Specialization II: Employment in Service Sub-Sectors

	(1) Trade	(2) FoodAccomm	(3) Transport	(4) ICT	(5) Finance	(6) EduHealthSoc
Chinese Shares 1930	57.975***	8.330*	27.005***	0.985	2.993	44.122*
	(19.333)	(4.609)	(9.733)	(2.018)	(2.785)	(23.604)
FE	Province	Province	Province	Province	Province	Province
Cluster	Regency	Regency	Regency	Regency	Regency	Regency
West Java	No	No	No	No	No	No
R2	0.299	0.170	0.268	0.235	0.202	0.268
Observations	12770	12770	12770	12770	12770	12770
Mean	143.94	16.37	66.43	6.20	11.21	151.37
F-stat	2.49	2.49	2.49	2.49	2.49	2.49
Underidentification Test, AR p-value	0.133	0.286	0.050	0.692	0.479	0.271
Underidentification Test, SW p-value	0.027	0.179	0.029	0.523	0.327	0.111

Notes: Regression of total number of productive-aged male workers in various sub-sectors of the service industry, on ethnic Chinese shares (1930). All regressions include controls of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the village-level. Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix Tables and Figures

Table A.1: Contemporary Development: Robustness to Inclusion of West Java

	(1) consumption ('00-10)	(2) pop. density ('00)	(3) Δ light intensity ('92-10)
Chinese Shares 1930	0.058**	0.116	3.062*
	(0.028)	(0.097)	(1.651)
FE	Province	Province	Province
Cluster	Regency	Regency	Regency
West Java	Yes	Yes	Yes
R2	0.930	0.651	0.338
Observations	47872	344	344
Mean	11.66	-7.09	9.76
F-stat	7.46	7.01	8.41
Underidentification Test, AR p-value	0.058	0.307	0.198
Underidentification Test, SW p-value	0.113	0.351	0.043

Notes: Regressions of, respectively, the natural logarithm of equivalent household income (pooling all years from the Indonesian Socio-Economic Survey *(Susenas)* 2000-2010), natural log of population density (Indonesian Village Census, 2000), and night-time light intensity (2000) (NHGHIS), on 1930 ethnic Chinese shares. All regressions include controls of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the village-level for column (1) and district-level for columns (2) and (3). Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Structural Transformation: Robustness to Inclusion of West Java

	(1) Agric. Shares (%, 2010)	(2) Manuf. Shares (%, 2010)	(3) Service Shares (%, 2010)
Chinese Shares 1930	-7.074**	4.129	2.945
	(2.972)	(4.847)	(3.754)
FE	Province	Province	Province
Cluster	Regency	Regency	Regency
West Java	Yes	Yes	Yes
R2	0.234	0.110	0.230
Observations	19706	19706	19706
Mean	39.65	22.35	38.01
F-stat	1.50	1.50	1.50
Underidentification Test, AR p-value	0.354	0.596	0.301
Underidentification Test, SW p-value	0.207	0.359	0.449

Notes: Regression of the industry shares of working productive-aged males in 2010 on ethnic Chinese shares in 1930. Observations at the village-level. All regressions include controls of altitude, linear and squared polynomials of the natural logarithm of distance to coastlines, to Batavia (present-day Jakarta), to Semarang, Tuban, Gresik, Trowulan (the historical capital of the Majapahit empire, one of the largest pre-colonial empires in Southeast Asia), Surabaya, the natural logarithm of total area (1930), total dry agricultural land (1963), and total wet agricultural land (1963) and longitude. Observations are at the village-level for column (1) and district-level for columns (2) and (3). Standard errors clustered at the Dutch regency level and displayed in parantheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.1 Data Appendix on IV Construction. Distance to Muria Straits, Re-centered IV

We start by digitizing 1930 district areas based on map scans from *Atlas van tropisch Nederland*, a Dutch archival map on Indonesia's topography and 1930s administration. We trace the borders of each district to create GIS polygons that can contain spatial data.

Following (Borusyak and Hull, 2023), we construct placebo routes that are similar in length to the total length of the Muria Straits. The Muria Strait route comprises stops in the districts of Semarang, Demak, Koedoes, Pati, Waroe (Rembang), and Binangoen (Lasem). These are the main set of historical portage 15th-century portage sites – we triangulate these from various historical sources. We calculate the length of the Muria Strait by pinning down the centroids of the Muria Strait districts, based on modern-day GIS boundaries and historical sources. We then calculate the geodisic distance passing through each and every district centroid, starting from Semarang to Binangoen to obtain the total length of the Muria Straits.

From this exercise, we compute two objects. First, we compute *Distance to Muria Straits* as the geodisic distance between the Muria Strait route and 1930 district centroids. Our recentered IV, at the 1930 district-level, is obtained by subtracting the average of all placebo distances from this.

Second, we create placebo routes that are similar in the distance length of the Muria Strait, along the northern coast of Java. The routes are placed within a 1000-meter interval, starting from the border between West and Central Java and ending at Surabaya. We obtain a total of 493 placebo routes. We then calculate the average distance of each district centroid to all the placebo routes and take its logarithmic form. Lastly, we compute our recentered IV by taking the difference between the logarithmic of *Distance to Muria Straits*, and the logarithmic of the average distance of each district's centroid across all 493 placebo routes.