

Communities of Commerce: The Legacy of Chinese Immigration on Java

Quoc-Anh Do

Monash University

Sebastian Ellingsen

University of Bristol

Gedeon Lim

University of Hong Kong

December 23, 2025

Abstract

This paper studies the economic legacy of ethnic Chinese immigration on Indonesia. Our instrumental variables strategy exploits 15th-century landing sites that shaped subsequent Chinese settlement but lost maritime access due to 17th-century silting. A one-percentage-point higher Chinese share in 1930 leads to 9.4% higher household consumption and 26% higher population today. Using manufacturing census data, we show strong effects on firm sales driven by improved financing through private networks. Effects attenuate sharply during the 1997-1998 crisis, then recover. Despite its small size, ethnic Chinese minorities have played an outsized role in Indonesia's modern development, with spillovers to the broader population.

Keywords: Ethnic minority; Immigrants; Ethno-occupational segregation; Conflict; Discrimination; Minority Resilience; Private Finance; Financing Networks.

JEL Codes: F63, N33, J15, Z13

*We are grateful for insights and comments from Ying Bai, Sam Bazzi, Sascha Becker, Adam Brzezinski, Zhiwu Chen, Ma Chicheng, Pierre van der Eng, Benjamin Faber, Maxime Gravouille, Masyhur Hilmy, Mark Hup, Saumitra Jha, Ruixue Jia, James Kung, Horacio Larreguy, Binging Li, Weijia Li, Joel Mokyr, Kieu-Trang Nguyen, Nathan Nunn, Jacopo Ponticelli, Ana Tur Prats, Jens Prüfer, Gianluca Russo, Uta Schönenberg, Sandra Sequeira, Michael Zheng Song, Kjetil Storesletten, Guo Xu, Zhihao Xu, Melanie Meng Xue, Felipe Valencia Caceido, David Yang. We thank seminar participants at Hong Kong U, Monash U, the ASREC Europe meeting, Australian Natural Experiment Workshop (Deakin U), Australian Labor Econometrics Workshop (U Melbourne), Australian Development Economics Workshop (UNSW), Australasian Cliometrics Workshop (UNSW), European Meeting of the Urban Economics Association, RIDGE Historical Development conference, Tsinghua/U Chicago Growth and Institutions workshop, Technology and Long-Term Development workshop (U Valencia), Social Conflict and Political Economy workshop (U Sydney), Workshop in Economic History at Uppsala U, Narratives and Culture in Historical Perspective workshop (LSE), the World Congress of the Econometric Society in Seoul, and the Annual SIOE Conference for helpful comments and suggestions. We thank Sam Bazzi, Pierre van der Eng, and Alex Rothenberg for sharing their data. We thank Sheva Sonia Rahmani for excellent research assistance. Do: Department of Economics, Monash University, 900 Dandenong Road, Caulfield East, VIC Australia. 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

Human history abounds with cases where hostility toward minority groups, especially those concentrated in “middleman” occupations such as credit and trade, has escalated from exclusion and expulsion to mass violence and genocide ([Grosfeld et al., 2020](#); [Anderton and Brauer, 2021](#); [Becker et al., 2022](#)). This hostility is frequently a result of narratives that portrays minorities in such occupations as being “unproductive” and “exploitative” of the majority’s “honest work” ([Bonacich, 1973](#); [Chua, 2004a](#)). This paper provides evidence against these claims by studying one of the largest ethnic minority groups in Southeast Asia: Chinese immigrants and their descendants in Java, Indonesia. We leverage the deep history of Chinese immigration to Java, dating back to the 15th century, to estimate strong, positive causal effects of Chinese communities on long-run, local economic development. Importantly, we trace the attenuation and recovery of these effects throughout recent decades of discrimination, socioeconomic crises, and subsequent recovery.

Chinese immigration to Southeast Asia dates back to more than half a millennium, motivated by both push and pull forces including policies and social events in mainland China as well as the commercial attraction in Southeast Asia ([Macaulay, 2021](#)). By the 1970s, the ethnic Chinese population in Southeast Asia numbered 15.8 million out of a total population of 308.9 million, with those in Indonesia numbering 3.2 million out of a total population of 129 million (2.5% share) ([Wu and Wu, 1980](#), p. 133). Chinese minorities are usually seen across the region as a key player in economic development, and in certain specific contexts, also in political competition ([Suryadinata, ed, 1997, 2004; Suryadinata, 2007](#)). We focus on the ethnic Chinese in Java, the world’s most populous island, where they represent a most tiny fraction and yet are considered crucial in the corporate world ([Wu, 1983](#)). Similar to most contexts in Southeast Asia, throughout modern Indonesian history, the ethnic Chinese have found themselves subject to various perpetual, deeply discriminative policies and conflicts, including massacres as recent as 1998.

We digitize granular historical data on the ethnic composition of Java from the 1930 Dutch East Indies Census (*Volkstelling, 1930*) at the end of the colonial period. This was a unique and reasonably truthful snapshot of the ethnic composition of Java: After independence, ethnicity quickly became a taboo topic and was removed from all censuses until the ethnic Chinese were given back most equal rights by 2000. We further digitize district maps from 1930 to match those data to all available modern Indonesian datasets.

We begin by documenting a strong link between a district’s contemporary economic development and its size of the Chinese community in 1930. A one-percentage-point increase in the size of the Chinese community is associated with a 3% increase in average household consumption and a 2.6 unit increase in night-light luminosity (relative to a district mean of 13.85). This association is robust to controlling flexibly for a battery of covariates based on geography, early development, and colonial extraction ([Dell and Olken, 2020](#); [De Zwart et al., 2022](#)), and is unlikely driven by omitted unobservables ([Cinelli and Hazlett, 2020](#)).

To address the common concern of endogeneity of the location of ethnic Chinese in 1930, notably due to their sorting into locations with higher growth potential, we propose an instrumental variable approach

motivated by the deep history of Chinese settlement in Java. Building on a milestone of Chinese immigration to Southeast Asia, the series of expeditions led by the early Ming dynasty's Admiral Zheng He in the early 1400s ([Suryadinata, 2005](#); [Dreyer, 2006](#)), we use a district's distance to the expeditions' landing sites in the Muria Strait on the Northern coastline of Java (*Pasisir*) as instrumental variable for its Chinese population share in 1930. This empirical strategy builds from the long tradition of the immigration networks instrument ([Card, 2001](#); [Munshi, 2014](#)). It draws further support from the expeditions' lack of navigational technology to precisely choose landing sites ([Pascali, 2017](#); [Rivers, 2012](#)) and lack of information regarding the Javanese hinterland, as well as the complete sedimentation of the Muria Strait since the 17th century that removed all of those locations' prior waterway advantages.¹ To further strengthen the causal interpretation of our IV strategy, we also implement two additional strategies to simulate counterfactual landing sites to only exploit the arbitrariness of the ones along the Muria Strait.

The IV exerts a strong, robust first-stage effect on the share of ethnic Chinese in 1930, which does not replicate for placebo cases of other groups of non-Chinese immigrants. The IV strategy estimates the long-term impacts of an increase of one percentage point of Chinese share in 1930 to be 9.4% more of modern household consumption, 26% more district population, and 14 percentage points more of urbanization rate. These long-run impacts have likely resulted from the impacts of Chinese presence on local growth rates, notably estimated at 3.6 points of luminosity from 1992 to 2010 out of a mean of 9.7 points growth during this period. The strong growth effects also come with strong effects of the Chinese community on the structural transformation of the economy away from agricultural to manufacturing and service jobs, at the range of 10 percentage points lower employment in agriculture for a one percentage point higher in Chinese share in 1930.²

In light of these large growth effects, we further test for heterogeneous effects against the backdrop of social conflicts that ethnic Chinese have faced, notably in 1998 when the Asian Financial Crisis of 1997 met the political vacuum around the fall of Suharto. This double crisis has led to a wave of assaults against ethnic Chinese throughout Indonesia (as [Grosfeld et al. \(2020\)](#)'s theory would have predicted). Quantitatively, we find a sudden drop of the positive effects of Chinese presence around the crisis. Following the recovery and the restoration of equal rights for the Chinese minority under Gus Dur's government in 1999-2001, we find a quantitatively strong recovery of the effects of Chinese presence.

To better understand these main results, we investigate the role of firm-level mechanisms. Using comprehensive administrative data of Indonesian manufacturing firms from 1980 to 2013, we find strong effects of historical Chinese presence on firms' sales, which works principally through the channel of private financial sources based on private networks instead of broad-based financial development. This channel was strongest before the crisis period of 1997-1998, was shut down during the crisis, and only partially recover thereafter, which is reminiscent of other qualitative studies of Chinese-founded businesses in the

¹This unique feature of the disruption of a historical intervention is reminiscent of the instrument in [Valencia Caicedo \(2019\)](#).

²The large magnitudes of the effects need to be considered with the perspective of a tiny average proportion of ethnic Chinese – at an average of close to 1% in 1930. The accumulation also takes place over a very long duration.

region (Suryadinata, ed, 2008; Suryadinata, 2015). This pattern is difficult to reconcile with explanations based solely on time-invariant locational fundamentals or persistent physical or institutional infrastructure. Instead, the findings are consistent with the view that Chinese communities function as economic intermediaries, leveraging dense social and commercial networks to facilitate investment and stimulate firm growth.³

We also check and find limited evidence for alternative mechanisms. The main results remain strong after carefully controlling for previously studied determinants of colonial extraction and early development (e.g., Dell and Olken, 2020). Importantly, we do not find a displacement (business stealing) effect on neighboring districts; instead, the estimated spillover towards neighbors is slightly positive. We also explore other potential mediators including market access, international trade, or international aid (Rauch, 2001; Rauch and Trindade, 2002), human capital (Becker et al., 2020; Arbatli and Gokmen, 2023), or the diffusion of cultural practices and norms (Valencia Caicedo, 2019; Dippel and Heblich, 2021). Our findings are most consistent with an important role played by occupational specialization among the Chinese in sectors with important complementarities with the broader economy, such as finance, transportation, and trading activity, as in Jha (2013, 2018). We hypothesize that a comparative advantage in middle-man occupations arose from the importance of kinship ties in organizing economic activity among Chinese migrants (Landa, 1998; Greif and Tabellini, 2017), giving Chinese communities an advantage in coordinating economic activity in contexts with limited enforcement of formal rules (Greif, 1989, 1993). This pattern, reinforced by both colonial and New Order policies, has persisted to the present day. In summary, our findings provide evidence that minorities with a comparative advantage in middle-man occupations can provide economic complementarities that benefit the majority population.

Our work relates to a sizable literature in social science on the long-term economic and political impacts of ethnic minorities in general (Acemoglu et al., 2011; Voigtländer and Voth, 2012; Hornung, 2014; Becker and Pascali, 2019) and occupationally segregated groups in particular (Bonacich, 1973; Anderson et al., 2017; Becker and Pascali, 2019; Grosfeld et al., 2020). A substantial body of research emphasizes economic competition between ethnic groups as a driver of ethnic tension (Bates, 1974; Caselli and Coleman, 2013; Mitra and Ray, 2014). However, less is known about the economic impacts of occupationally segregated minorities and existing research has primarily focused on Jewish communities in Europe (see e.g. Akbulut-Yuksel and Yuksel, 2015; Johnson and Koyama, 2017; D'Acunto et al., 2019; Huber et al., 2021; Do et al., 2024). Our paper is the first to quantitatively examine the role of Chinese communities in long-term economic development. Furthermore, our findings bridge both strands of literature by examining how the broader political context mediates the economic impact of such ethnic minorities. In this aspect, our paper also contributes to a recent literature that shows the long-run economic resilience and revival of economic elite groups through adversity, especially after a reversal of policies or rejection of discriminatory beliefs

³Using additional sources of data from before 1930, we find some evidence of firm development in Chinese communities even under Dutch rule. However, the magnitude and scope of historical evidence suggest that most of the development impacts of the presence of Chinese have materialized much later.

(Ager et al., 2021; Alesina et al., 2020; Do et al., 2024; Michelman et al., 2022).

More broadly, our paper also builds on a large literature in economics on the impact of ethnic diversity on economic development and social conflict (see e.g. Easterly and Levine, 1997; Alesina and La Ferrara, 2005; Bazzi et al., 2016, 2019; Montalvo and Reynal-Querol, 2021; Kok et al., 2025). Whether and under what conditions ethnic diversity promotes long-term economic development while sustaining inter-ethnic tolerance remains an open question. Jha (2013, 2014, 2018) highlight the importance of durable inter-ethnic complementarities combined with mechanisms to share the gains arising from such exchange. Grosfeld et al. (2020) show that large political shocks increase the vulnerability of middleman minorities by undermining their ability to sustain repeated interactions with other groups. Here, despite the organization of ethnic Chinese trading networks through community and kinship-based ties, both of which might have limited initial entry and ethnic majority gains (Jha, 2013), our results suggest that persistent state-led enforcement of surplus-sharing norms and institutions can, to a certain extent, lead to economic benefits for the majority and partially mediate the effects of minority complementarities on social conflict. Hence, we provide evidence of *how* sizable economic benefits from a minority presence can arise *even* when entry into minority networks is limited.

We also contribute to the growing literature on the effects of historical immigration on long-term development in host countries, especially the work by Sequeira et al. (2020) and Burchardi et al. (2021) on the long-run economic impacts of immigrants to the US. Previous studies have highlighted the importance of human capital (e.g., Rocha et al., 2017; Droller, 2018; González, 2020). Our findings show that low-skilled migration can promote long-term economic development in weakly institutionalized settings. Our findings suggest that the immigration of groups with strong kinship networks can provide a substitute for state institutions in such contexts. As such, our findings also build on the impact of family ties on long-term economic development (Alesina and Giuliano, 2014; Chen et al., 2022).

Finally, our paper builds on the quantitative literature on determinants of long-term economic development in Indonesia (see e.g. van der Eng, 1992, 2010). Bestari et al. (2022) examine the impact of early European settlement in Indonesia on local economic development. Particularly relevant to our study, Hup and De Zwart (2024) examine the drivers of Chinese migration to the Dutch East Indies around the first half of the 20th century. Our paper differs in that we explore the impact of Chinese communities on long-term economic development. Relatedly, our findings also build on the literature on the legacy of colonialism on contemporary outcomes in general (Acemoglu et al., 2001, 2005) and in Indonesia in particular (Pepinsky, 2016; Dell and Olken, 2020; De Zwart et al., 2022; Kuipers, 2022; Lim, 2024). Our paper is most closely related to research on the legacies of colonial policies designed to strengthen ethnic occupational and social stratification (Kuipers, 2022). Our findings suggest that the occupational specialization of the Chinese had a persistent impact on economic development and social conflict that has lasted through centuries of de jure and de facto discrimination.

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. Section 3 describes our data sources. Section 4 describes our OLS empirical strategy and results. Section 5 outlines our instrumental variable approach, and presents the main results and mechanisms. Section 6 explores alternative mechanisms. Section 7 provides a discussion of the main findings. Section 8 concludes.

2 Chinese Communities on Java

In this section, we describe the relevant historical and institutional background of Chinese immigration to Indonesia, largely following [Suryadinata \(2005\)](#). We focus on three key facts that inform our empirical strategy: (i) Pre-19th century immigration; (ii) The determinants of early Chinese settlement patterns; (iii) Post-19th century immigration.

2.1 Chinese Immigration to Indonesia

Pre 19th Century Chinese Emigration. Chinese emigration to Java before the 15th century was limited.⁴ While some evidence of Chinese immigrants can be traced as far back as the 13th century following the failed Mongol invasion of Java, 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 ([Mills, 1971](#)), the expeditions only encountered very few Chinese on shore in Tuban and Gresik in Eastern Java.

The first major recorded involvement of Chinese in Java was surrounding Ming Admiral Zheng He's voyages into the archipelago (1405-1433). Zheng He's maritime expeditions were a series of seven state-sponsored voyages that projected the Ming dynasty's power across the Indian Ocean, all the way from Southeast Asia to East Africa. They were a historical milestone that established durable diplomatic, commercial, and migrant networks in Southeast Asia, leaving lasting Chinese settlements and trade links, while marking the high point of China's premodern naval capability before the Ming state turned inward and abandoned large-scale maritime expansion ([Dreyer, 2006; Suryadinata, 2005](#)).

Contemporary ethnic Chinese settlement today can be traced to many of the ports of call along the north coast of Java where members of his crew landed. Figure 1 plots 1930 ethnic Chinese shares and, with a solid black line representing Zheng He's sea passage along the historical *Muria Strait* on the north coast of Java separating the main island with Mount Muria, and small dots on the line representing the historical towns of Semarang, Demak, Koedoes, Pati, Rembang and Lasem, as ports of call.⁵ 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 ([Rivers, 2012](#)). It is further crucial that the entire

⁴The discussion in this page draws extensively from [Hoadley \(1988\)](#) and [Carey \(1984\)](#).

⁵Those landing locations are based on the local historian Liem Thian Joe's work ([Liem, 1933](#)) building from local sources. See [Suryadinata \(2005\)](#) in particular on the history of Semarang.

Muria Strait silted up in the 17th century due to heavy sedimentation from tributary rivers, to the point of becoming dry land. 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.

The influx of post-15th century Chinese immigrants arose with the fall of the Ming dynasty (1644) and the reopening of Chinese trade with Southeast Asia around 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, from where many Chinese in Indonesia originated ([Claver, 2014](#), p. 135). Further waves of influx of Chinese immigrants to Indonesia followed various policy pushes by the Qing dynasty (1644 - 1911) ([Macauley, 2021](#)), and increased sharply due to famine and war in the second half of the 19th century. Many Chinese assimilated closely with the indigenous population, married local women, and adopted local customs and religions, creating the Peranakan culture across Indonesia and Malaysia.

Crucially, kinship ties and pre-existing clan networks in Southeast Asia and on Java were the key determinants of where latter-day Chinese immigrants chose to settle ([Setijadi, 2023](#); [Hup and De Zwart, 2024](#)). Hence, we see in our data that initial port of calls 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 routes.

Late 19th - Early 20th Century Immigration. Chinese immigration accelerated between the mid-19th and 20th centuries when nearly 50 million ethnic Chinese and Indians emigrated to Southeast Asia ([McKeown, 2004](#)). During the colonial period, Chinese migrants played important economic and political roles, such as mediating economic exchange between the European and native populations ([Ongkokham, 1989](#)). By the 1970s, the ethnic Chinese population in Southeast Asia numbered 15.8 million out of a total population of 308.9 million, with those in Indonesia numbering 3.2 million out of a total population of 129 million (2.5% share) ([Wu and Wu, 1980](#), p. 133).

Dutch Colonial Reports (1897-1920) on annual immigration document several facts about the stock and flow of Chinese emigration. First, Chinese emigration to the Dutch East Indies, excluding contract labor and hence potentially under-counting the flow of emigrants, comprised about 20,000 – 40,000 per year and, to Java alone, about 4,000 – 10,000 per year during this period ([Booth, 1988](#); [Rush, 1991](#)). Second, Chinese emigration dominated all other country/ethnic groups. The annual flow of Chinese immigrants as a percentage of total resident Chinese population was around 4%. As a proportion of 1920 inhabitants, the sum of all Chinese immigrants that arrived from 1897-1920 comprised a small share of the total population at around 5 per 1,000 inhabitants ([Hup and De Zwart, 2024](#)). In the context of Java, [Hup and De Zwart \(2024\)](#) records that each region of West and East Java attracted more than 40,000 Chinese immigrants across this period relative to some Outer Island residencies, outside of East Sumatra and West Borneo, that attracted less than 1,000 Chinese emigrants. As a result of these migration flows, by 1930 about 1.2 million

inhabitants in the Dutch East Indies were classified as Chinese. This accounted for around 2% of the total population—with 64% of them having been born in Indonesia—and 120,000 as “Other Foreign Asians” which included a majority of Arabs and 30,000 Indians ([Volkstelling, 1930](#)).

Selection and Composition of Immigrants. While much of the literature, (especially on emigration to Southeast Asia in the late-19th century) has emphasized the unskilled nature of migrant labor from China (and India) to Southeast Asia driven by the booming cash crop estates and tin mine industry ([Huff and Caggiano, 2007; Kaur, 2004](#)), a substantial and persistent skill premium in Southeast Asia vis-a-vis other parts of the world in the early 20th century suggested that there was a strong demand for skilled workers ([Bassino and van der Eng, 2021](#)).

This was especially true for the large number of ethnic Chinese emigration flows to Java in the early 20th century that were likely to have been semi-skilled or had access to savings or credits through families and other extended networks ([Macauley, 2021; Hup and De Zwart, 2024](#)). Qualitatively, many of these migrants established themselves as skilled artisans and businessmen ([Booth, 1998; Rush, 1991; Die, 1943](#)). Quantitatively, [Van Leeuwen and Maas \(2011\)](#) finds that 63% of 19,000 Chinese emigrants (with recorded occupations) were “medium-skilled”, consisting of skilled craftsmen (3,200), shopkeepers and traders. Furthermore, [Hup and De Zwart \(2024\)](#) finds that the potential migrant surplus (the level of earnings subtracted by cost of living) of skilled Chinese emigrants to the Dutch East Indies was about 2-4 times higher than in the sending regions of southeast China. It is likely that this applied more broadly to Chinese emigration to Java rather than the Outer Islands given that regulations in the Outer Islands mainly restricted emigration flows there to indentured labor although it is not possible to estimate the proportion of unskilled migrants to Java during this period [Hup and De Zwart \(2024\)](#).

With regards to capital, the need for sufficient capital to make the migration journey suggests that many, if not all Chinese emigrants to Java would have hailed from extended families that had the ability to put together enough capital for their sons to afford the transportation fees to the Dutch East Indies. [Hup and De Zwart \(2024\)](#) estimates that the fee for an “entrance card” was initially minor but rose to fl. 25 after 1911, approximately (44 - 89) 14-40 days of (un-)skilled labor daily wages. In the 20th century, in terms of selection relative to the existing native population, while detailed data on the human capital of migrants does not exist, it would appear that the Chinese emigrant community (and “other foreign Asians/Easterns”) as a whole had established businesses and had higher average incomes than indigenous Indonesians from tax records ([De Zwart, 2022; Booth, 1988](#)).

2.2 The Economic and Political Position of Ethnic Chinese

Specialization During the Colonial Period Chinese involvement in trade and commerce dates back to before the arrival of Europeans ([Pan, 1999](#), p. 152). Under both VOC and Dutch colonial rule, Chinese communities played important roles as intermediaries between the local population and the Europeans

(Houben, 2002, p. 74). Throughout the period, the Chinese were active as tax farmers, as middlemen, in retail trade, in entrepreneurship, and in informal capital markets. As a result of their detailed local knowledge and access to kin-based organizations that facilitated commercial activity, both the VOC and the Dutch colonial state relied heavily on the Chinese to perform these economic functions. A key manner in which this collaboration was institutionalized was through the leasing out of tax farms and monopoly rights, most notably opium (Zanden and Marks, 2012). The revenues from this system were important, making up between 15 to 25 percent of the government revenue in the 19th century (Houben, 2002, p. 74). The economic activities of the Chinese communities thereby played an important role in financing the Dutch colonial state.

Beyond, the need of the Dutch to fill perform these economic role, the system greatly reduced the administrative burden and contributed to maintaining political stability. As a result, the Dutch had strong incentives to reinforce Chinese distinctiveness and the associated occupational segregation. Like other minority groups, the Chinese community was governed through a distinct administrative structure, overseen by local representatives (such as *kapitan* or *majoer*) endowed with administrative, executive, and judicial authority (Claver, 2014, p. 18). Various restrictions on the movement of Chinese in the countryside were also imposed through much of the colonial period to promote this segregation and prevent the Chinese from becoming too economically powerful. This system remained in place until the late 19th and changed after the abolition of the Dutch cultivation system and the gradual liberalization of the economy. While this represented an important loss of revenue among the Chinese, many were able to enter other sectors. As a result, we see that the commercial orientation of the Chinese communities persisted until the 1930s (Zanden and Marks, 2012, p. 78).

By 1935, the occupational structure of ethnic Chinese still differed substantially from the Javanese. Ethnic Chinese in Java were employed in three main occupations: (i) traders (57.7%); industry (20.8%); and farming, fishing, and mining (9.1%) (Pepinsky, 2016). Notably, only 0.5% were employed in the public sector. Within trade, the Chinese were largely concentrated in three main subgroups of (i) Miscellaneous small trading (46.8%); foodstuffs (22.3%); and textiles (16.0%).⁶ Importantly, only 5.1% of Chinese involved in trade were involved in banking and finance.⁷ It is clear by the 1930s the Chinese occupied distinct economic roles that set them apart from the majority population.

Contemporary Economic Status Ethnic Chinese continue to play an important economic role in Indonesia post independence. Today, Indonesia is the largest country in Southeast Asia and is the single country with the largest absolute number of ethnic Chinese (Suryadinata, ed, 2008). In terms of shares, however, the ethnic Chinese are one of the smallest ethnic groups in Indonesia, with estimates ranging from 1.5

⁶Small-scale trading is perhaps better understood in relation to the complete list of occupational subgroups: Foodstuffs; textiles; ceramics; Wood and bamboo products; vehicles; clothing; wholesale and distribution; other trade; banking and finance. Other foreign easterners, mainly Arabs, as will be explained below, were involved in the textile trade (48.8%); miscellaneous small trading (27.5%); and other trade (8.6%).

⁷Notably, many ethnic Chinese founders of the biggest Indonesian conglomerates today who also founded many banks, as will be discussed later, had their first businesses in small-scale trading (Borsuk, 2014).

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](#)). In particular, in contrast to the 5.1% share of Chinese involved in the financial sector in 1935, today about 72% of banks founded in Indonesia and whose capital is in Indonesia were founded by individuals of ethnic Chinese descent.⁸

3 Data

To explore the long-term impact of Chinese communities on economic development in Java, we construct a novel dataset of historical Chinese settlements that we match to a variety of contemporary and historical sources. We focus on Java for three key reasons. First, as the administrative and economic center of the Dutch East Indies, data on ethnic composition on Java are available at a more fine-grained level than in the Outer Islands. Second, the majority of ethnic Chinese in Indonesia, outside of East Sumatra and Western Borneo, emigrated to and reside in Java ([Hup and De Zwart, 2024](#); [Volkstelling, 1930](#)). Finally, the experience of Chinese communities in and outside of Java differed markedly in terms of the degree of social exclusion from native ethnic groups. Our main unit of analysis is therefore the Javanese districts of the Dutch East Indies used to report outcomes in the 1930 census. Below, we elaborate on the main data sources. Details are reported in Section A.1 of the Appendix.

3.1 Measuring Ethnic Composition in 1930

1930 Indonesian Population Census We use data on the ethnic composition from the 1930 Dutch East Indies Census ([Volkstelling, 1930](#)) to construct ethnic Chinese shares. This is our main measure of the historical size of the ethnic Chinese community. The 1930 Census was the first full-count census to record ethnicity and therefore provides the earliest and most disaggregated record of ethnic composition at the Dutch regency level (henceforth, the 1930 district level). After 1930, neither the Dutch nor the post-colonial Indonesian government collected data on ethnic composition until 2000, ostensibly as a means of nation-building and promotion of homogeneity among the Indonesian population ([Ananta et al., 2015](#); [Auwalin, 2020](#); [Van Klinken, 2003](#)). The 1930 Census primarily relied on “social criteria”, broadly defined as language spoken, customs, and habits to distinguish between different ethnic groups. The Chinese, who followed different religious beliefs (“Confucianism in the midst of a sea of Islam”), arranged marriage practices, and “clannish” associations and festivals were clearly set apart from the native population ([Van Klinken, 2003](#)).

We use ethnic Chinese shares in 1930 rather than 2000 as our main explanatory variable, as the identification of ethnic Chinese in 1930 was possibly more accurate than in 2000 for two main reasons. First, while the Dutch colonial authorities much preferred ethnic separation across occupations and locations,

⁸Authors' own calculations.

after independence both Sukarno and Suharto's regimes promoted widespread assimilation policies in discrimination against ethnic Chinese, leading to the fear to openly identify one's Chinese origin. Second, the 2000 census took place shortly after widespread anti-Chinese riots and killings across Indonesia during the Asian Financial Crisis and the fall of the Suharto regime in 1997-1998 ([Chua, 2004b](#); [Heryanto, 1998](#)), further blurring the veracity of the ethnicity question on the 2000 census.

We also digitize the 1930 share of “other foreign Easterners”, 88% of which were Arabs ([Pepinsky, 2016](#)). Java has long maintained trade and tributary relations with China, India, and the Arabian Peninsula. In Section 4 and 5, we use the share of “other foreign Easterners” in 1930 as a control variable and placebo test respectively.⁹

3.2 Additional Data Sources

Registry of Businesses 1890-1940. We obtain the number and type of Chinese businesses across various economic sectors from a Dutch colonial business registry, the “Directory of Agricultural and Trading Companies in Colonial Indonesia” (*Handboek voor Cultuur- en Handelsondernemingen in Nederlandsch-Indië*).¹⁰ This directory is *not* a complete listing of all ethnic Chinese businesses—it lists all businesses that were *voluntarily* incorporated by owners as limited liability companies under company law, which by 1855 applied to firms of ethnic Chinese entrepreneurs in Indonesia ([Liem, 1952](#)). The key advantage of registration was *limited liability*—company owners would be liable for the company’s actions only up to the value of their investment at registration, protecting owners’ personal assets in cases of bankruptcy.¹¹ To that end, this dataset is best interpreted as capturing a relatively complete picture of Chinese businesses on the right-tail—that operated in Indonesia until 1940—given that it excludes unregistered ventures ([Van der Eng, 2022](#)).¹²

We restrict our analyses to three main variables for which data are relatively more complete: (i) number of ethnic Chinese firms; (ii) (shareholder) equity; and (iii) economic sector. Businesses in the Handbook are classified as ethnic Chinese on the basis of their headquarters, the name of their company, and, most importantly, whether the names of owners and directors were ethnic Chinese. Shareholder equity refers to funds that shareholders invested in the company at incorporation or at agreed moments after incorporation until full subscription and are based on the public announcement of incorporation of the venture ([Van der Eng, 2022](#)). Economic sectors are based on names of companies and brief descriptions in their articles of

⁹An important distinction between Arabs and Chinese is in terms of religion and social exclusion ([Pepinsky, 2016](#)): most Arabs practice Islam and having Arab ancestry is widely viewed as a mark of prestige for native Muslims in Java. Hence, Arab communities in Java have experienced relatively little discrimination and have integrated much more easily into native society.

¹⁰This dataset was significantly improved by and described thoroughly in [Van der Eng \(2022\)](#).

¹¹Other advantages were the ability to raise external funding through formal bank lending, issuing bonds, and selling shares on regulated stock exchanges in Jakarta, Semarang, and Surabaya ([Van der Eng, 2022](#)).

¹²[Van der Eng \(2022\)](#) raises a few key examples: (i) An incomplete listing of industrial companies found 3,482 being operated by Chinese entrepreneurs—double the number of registered companies in the Handbook—1,619 of which employed more than 5 people ([Vleming, 1926](#)); (ii) An inventory of manufacturing enterprises identified 3,520 firms whereas the 1940 Handbook records 2,156 firms across all economic sectors; (iii) the database identifies 140 Chinese firms registered in Jakarta in 1925 but [Vleming \(1926\)](#) estimated total number of “not too small” Chinese ventures in Jakarta to be 2,200.

association and should be interpreted with caution, given that companies often try to incorporate as much diversification as possible, given the prohibitive costs of amending articles of association and hence, many Chinese firms were multipurpose and could actually have straddled multiple sectors. For example, the growth in agricultural businesses from 22% to 46% from 1910 to 1940 was mainly a result of sugar factories, which could reasonably be categorized as manufacturing instead ([Van der Eng, 2022](#)).

Contemporary Economic Outcomes. To capture broad measures of economic development at the district level, we primarily rely on three sources. The *Village Census* (PODES) gives us measures of population density in 2000. We supplement this with population counts from the 2010 Population Census. In a similar vein, we use average night-time light intensity from *National Historical Geographic Information System* (NGHIS), aggregated to either the village or district-level polygons in 2000. Together, this also captures the level of economic activity and population density at a sufficiently granular level. We construct measures of sectoral employment shares in various economic sectors using the 2010 Population Census. Crucially, these sectors include trade, transportation, and finance, which were sectors in which the ethnic Chinese specialized ([Pepinsky, 2016](#)). We estimate the effects on two levels. First, on aggregated shares in agriculture, manufacturing, and services. Second, on specific sectoral shares as informed by the historical and anthropological literature.

We also construct and use household consumption data from *Susenas* (Indonesian Socio-Economic Surveys) to estimate the overall impact of Chinese presence (be it through business establishments, credit, loans, trust, or other mechanisms) on the economic prosperity of local Indonesians of non-ethnic Chinese descent. In each round, *Susenas* covers a representative sample of households and individuals. We pool all available rounds from 2000-2010 ([Dell and Olken, 2020](#)) to obtain household-level measures of consumption of local, non-Chinese individuals, and construct our main measure of household-level income and welfare by taking the logarithm of equivalent household consumption ([Deaton, 1997](#); [Dell, 2010](#)).

Contemporary Firm-level Data. To explore firm-level outcomes, we use the *Manufacturing Census* (*Statistik Indonesia*). We used all rounds between 1980 and 2006. It covers all firms that employ minimally 20 to 30 employees and is limited to firms involved in any form of manufacturing activity. The Manufacturing Census gives us measures of firm-level sales, profits, productivity, and sources of finance, which we aggregate to the district level. In particular, we use this data in Section 6 to understand how these outcomes evolve over time, in response to changes in anti-Chinese sentiment and social conflict.

We also use the latest round of the Indonesian Economic Census in 2016. This dataset comprises the universe of small and medium firms in rural areas ($\geq 20\text{-}30$ 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 a number of employees. All firm-level results on

productivity, wages, and related outcomes are derived from the Manufacturing Census (Statistik Indonesia), while results on cross-sectoral industry shares and composition (i.e., structural transformation) are based on the Economic Census.

Other data. Finally, we also include a variety of variables capturing differences in the intensity of colonial extraction across districts. To this end, we include geocoded information on the location of sugar mills from [Dell and Olken \(2020\)](#). We use data on the location of opium concessions granted to the Chinese by district from [Kuijpers \(2022\)](#). In both these cases, we calculate the distance of each district to these locations. Furthermore, we include information on the number of Chinese officers by district using information from [Cribb \(1997\)](#). To account for historical market access and infrastructure, we use information on the historical railroad as well as the Great Postal Road. Lastly, we include a range of information on location fundamentals such as elevation, agricultural potential, and distance to the coastline. For data on agricultural productivity we use data from [Bazzi et al. \(2020a\)](#). As a robustness check, we also include data on caloric potential from [Galor and Özak \(2015\)](#) and [Galor and Özak \(2016\)](#). Data on Chinese development finance is from [Dreher et al. \(2021\)](#). Data on communist vote shares in the 1950s legislative elections are from [Bazzi et al. \(2020b\)](#).

4 Chinese Communities and Long-Term Development

In this section, we examine the impact of Chinese communities on long-term local economic development by comparing outcomes across districts with different sizes of the Chinese community in 1930. We begin by splitting the district depending on whether the Chinese population in 1930 was above or below the median. We then compare the mean, median, and interquartile range of various measures of economic development across these two samples. These results are depicted in Figure 2. From the figure, we see that the population density in 1930, the population density in 2000, average income, and nightlight luminosity are all higher for districts with larger Chinese populations. The differences in means are statistically significant at conventional levels for all the outcomes. As such, we see that districts with larger Chinese communities are more economically developed according to these metrics.

4.1 Estimation Strategy

We continue by examining whether the outcomes differ significantly across districts with varying Chinese population shares and whether these differences are robust to alternative explanations. Our starting point is the following regression

$$Y_i = \gamma_p + \beta C_i + X'_i \delta + Z'_i \gamma + u_i, \quad (1)$$

where Y_i denotes the outcome in district i . We measure most outcomes between 1990 and 2010, except

population density from 1930. C_i denotes the share of ethnic Chinese of district i 's total population in 1930, measured in percentage from 0 to 100. γ_p denotes provincial fixed effects, and u_i denotes the error term, estimated with adjustment for clustering at the regency level.¹³

A range of district-level factors may have influenced both the size of the Chinese community and long-term economic development through multiple channels. We consider two broad explanations. First, differences in agricultural potential, the disease environment, or the ease of participating in trade might all have shaped the location choices of Chinese migrants. We therefore include a vector of location fundamentals X_i , which includes observable location characteristics such as agricultural potential, distance to the coast, elevation, and proximity to historical urban centers. Second, in light of the role Chinese communities played in the colonial economy, the historical presence of Chinese might be correlated with the intensity of colonial extraction, which subsequently shaped the spatial distribution of economic activity (see e.g. [Dell and Olken, 2020](#); [Kuipers, 2022](#)). To address this, Z_i contains information on the intensity of colonial extraction, such as the distance to the opium concessions, sugar mills, and the location of railroads in the early twentieth century.

The parameter of interest in Equation 1, β , captures the difference in the outcomes between districts that differ by one percentage point in the Chinese population share, holding observable characteristics fixed. This interpretation is based on several assumptions. First, we interpret the variables in X'_i as pre-determined and not affected by human intervention and Z'_i to be pre-determined from the perspective of the Chinese population distribution in 1930. Second, we assume that the salience of ethnic identity is unrelated to factors that subsequently shape long-term economic development. Violations of this assumption could arise if the reporting of ethnic identity were correlated with underlying conditions, such as ethnic animosity or prevailing levels of economic development, that themselves affect long-run outcomes. (see e.g. [Jia and Persson, 2021](#)). However, we think this is less likely to be the case for 1930 than other years, in which data is available at a granular level for reasons elaborated upon in section 3. Classical measurement error in C_i could lead to attenuation in the parameter of interest. Lastly, we assume that the contemporary measurements of location fundamentals are highly correlated with the historical counterpart (e.g. elevation), and that we observe the most important factors that shaped the Chinese settlement pattern historically. We explore this last issue further in Section 4.3.

4.2 Chinese Shares in 1930 and Contemporary Development

We first consider the association between the Chinese population size and various broad measures of economic development. The results are presented in Table 2. We regress five key outcomes of long-run economic development at the district level: (i) Population density in 1930; (ii) Population density in 2000; (iii) Percentage of villages classified as urban in 2010; (iv) Nighttime light intensity in 2010; (v) Average

¹³There are 54 regencies in the main sample. Clustering by the coarser level of residency (with very few clusters), as shown in Table A.1, does not affect the qualitative messages.

household consumption (averaged across 2000-2011). Panel A presents the estimates without the inclusion of any controls. Panel B adds provincial fixed effects and Panel C includes geographic controls of altitude, longitude, logarithm of distance to the Northern coast, logarithm of total area (1930), logarithm of total dry agricultural land, total wet agricultural land (both in 1963) and foreigner shares (1930).¹⁴ Finally, Panel D includes the quadratic polynomial of logarithm of distance to each major provincial capital in Java (Batavia, Semarang, and Surabaya) and to the historical Majapahit urban centers of Tuban, Gresik, and the Majapahit capital at Trowulan, which may have affected both Chinese settlement choices and long-term development via governance and state capacity ([Campante and Do, 2014](#); [Bai and Jia, 2023](#)).¹⁵

For most specifications, we find a large and precisely estimated effect of the size of the Chinese community and the various measures of economic development. Column (1) reports the change in population density by district in 1930 with respect to the share of the Chinese population. From the table, we see that the coefficient ranges from approximately 0.07 to around 0.19. This means that districts with a one percentage point difference in the Chinese population share (a 100 percent increase from the mean of around 1 percent) differ in population density by between 6 and 19 percent. However, this result should be interpreted with caution since the population density and the Chinese share in 1930 are potentially jointly determined. In Column (2), we consider the population density in 2000. We see that the magnitudes are similar but slightly larger. As such, we see that districts with different rates of Chinese population differ substantially in population density in both 1930 and 2000.

We continue by examining the impact on the urban population share in Column (3). We find that a one percentage point increase in the Chinese population share increases the share of villages designated urban between 4 and 8 percentage points. The estimates are again precisely estimated across the different specifications. In Column (4), we also consider the impact on nightlight luminosity. For this outcome, we also find positive effects, although the precision is more sensitive to the inclusion of controls. Lastly, we consider the impact on the average level of consumption in the district. Again, we find a positive effect on the level of consumption. The impact is more modest and ranges between a 1 and 6 percent change for a one percentage point difference in the Chinese share. Taken together, the findings in Table 2 suggest that districts with larger Chinese populations in 1930 are more economically developed today. In the next section, we explore the robustness of these findings in more detail.

4.3 Sensitivity Analysis

In this section, we discuss the main robustness checks and challenges facing the above analysis. An important concern arises from the economic role Chinese communities played during the colonial era, acting at times as mediators between the European and indigenous populations. As such, agglomerations of Chinese communities may be correlated with the colonial presence, which could have affected economic develop-

¹⁴We exclude latitude in the controls, since the distance to the Northern coast is conceptually much similar and statistically highly correlated with latitude.

¹⁵Those cities, together with Surabaya, were highlighted in Ma Huan's chronicles of Zheng He's expeditions to Java ([Mills, 1971](#)).

ment in a myriad of ways (see, e.g., [Dell and Olken, 2020](#)). In Table A.2 , we therefore test for robustness to the addition of key controls such as whether a district contains (i) a sugar factory; (ii) the communist vote share in the 1950s; (iii) distance to colonial railroads; (iv) distance to the Great Postal Road; (v) distance to opium concessions. All in all, we find little evidence that these factors explain the association between economic development and the location of Chinese communities.

Taken together, most of the estimated effects are robust to including the various controls. This suggests that the estimated impacts are unlikely to be primarily driven by omitted variable bias. [Cinelli and Hazlett \(2020\)](#) develop a more systematic approach to this form of sensitivity analysis. Following this approach, we calculate the robustness value, quantifying the explanatory power unobserved confounders would need to account for the effect sizes we estimate, or render these statistically insignificant against a null of the effect being zero. We implement this approach with the baseline specification, including the main geographic controls. In summary, the approach suggests that the effects are unlikely to be driven by unobserved confounders. With the exception of consumption, we find robust values ranging from 30 to 40 percent. As a result, unobserved confounders that explain between 30 to 40 percent of the variation in the Chinese population shares and the various outcomes can account for a true effect size of zero. We think the omission of such confounders from Equation 1, as this is comparable to the combined effect of all observable geographic covariates. For the case of consumption, we cannot rule out that the findings are driven by an unobserved confounder.

4.4 Discussion

So far, we have documented that districts with larger Chinese populations in 1930 are more developed today. For most of the outcomes, the magnitudes are sizable. However, there are two important considerations. First, we are considering outcomes in levels after at least 70 to 80 years. As such, the annual average growth rates implied by the difference in levels are fairly small. Second, a one percentage point change in the Chinese population share is a relatively large difference, accounting for an approximately 100 percent increase from the sample mean. Nevertheless, our findings point to sizable differences in economic development across these districts.

Two primary factors can account for the magnitude of our estimates. First, differences in contemporary development can be *caused* by the communities through a range of possible channels, e.g. human capital spillovers, the diffusion of culture or norms, or occupational specialization in sectors of the economy with growth-enhancing complementarities. Second, the observed association between the size of the Chinese community and contemporary economic development might be determined by the sorting of Chinese migrants into locations that differ in unobserved characteristics. It is *a priori* unclear if this would attenuate our estimates or bias them upwards. For example, if Chinese immigrants were only able to settle more marginal land with lower agricultural productivity upon arrival, it might have affected subsequent industrialization and urbanization in these locations. Alternatively, if Chinese migrants settled in locations with

higher commercial potential, we might have observed higher rates of growth in those districts also without Chinese immigrants. While the robustness checks in the previous section suggest that sorting is unlikely to account for the entire effect, we will attempt to examine this mechanism in more detail in the next section.

5 The Deep Roots of Chinese Settlements

To investigate the channels underlying the patterns documented in Section 4, our empirical strategy builds on a selection of relatively arbitrary landing sites of Zheng He’s expeditions in the 15th century. Specifically, we use distance to landing sites in the Muria Strait as instrumental variable for Chinese population share in 1930.

This strategy leverages two key features of Chinese settlement across Indonesia discussed in Section 2. First, the persistence of Chinese communities over time hints at the typical mechanism behind the often-used immigration networks instrument (Card, 2001; Munshi, 2014, 2020), according to which subsequent waves of immigrants tend to first flock to earlier settlements. Macauley (2021) highlights the importance of this pull factor in the consolidation of Chinese communities across Southeast Asia.

Second, we pick the locations along the Muria Strait as they are relatively arbitrary, underdeveloped arrival sites, likely unknown to Zheng He’s crews before their voyages.¹⁶ Their arbitrariness comes in part from the lack of sufficient technology to determine longitudes in maritime navigation until the 19th century (Pascali, 2017; Rivers, 2012). This was reinforced by the underdevelopment of Semarang and other Muria Strait landing sites at that time, compared with the major ports and urban centers of the Kingdom of Majapahit such as Turban and Gresik. Last, even if the locations of initial landing sites may still be chosen for their access to maritime trade, the potential of the hinterland was certainly not understood. To further exploit this arbitrariness, we will compare Semarang and landing sites on the Muria Strait with hypothetical landing sites near Cirebon, the best deep-water port in Java.

In addition, the Muria Strait became silted up and turned into land in the 17th century. The initial landing sites therefore are unlikely to have locational fundamentals particularly amenable to trade in subsequent centuries, strengthening the case of the exclusion restriction.¹⁷ Taken together, we assume that any long-term impact of proximity to these initial settlement locations is mediated through its effect on the formation of Chinese communities once we condition on observable locational fundamentals.

5.1 Estimation Strategy

Based on the above discussion and the discussion in Section 2, we seek to estimate the long-term impact of the historical Chinese population using an instrumental variable approach. Our starting point is the

¹⁶We do not pick other well-known cities of the Majapahit Kingdom visited by Zheng He as recorded by the crew’s chronicler Ma Huan (Mills, 1971), including Turban, Gresik, Surabaya, and the Majapahit capital at Trowulan. Instead, we include a polynomial control of the distance to each of those cities.

¹⁷This can also be seen in the case of Semarang. While an important port developed there in the 18th century, it was not suited for steamships, which had deeper drafts. Its importance therefore diminished in the second half of the 19th century (Onghokham, 1989).

following homogeneous effect model,

$$Y_i = \alpha_p + \rho C_i + X'_i \omega + \varepsilon_i, \quad (2)$$

where Y_i again captures the outcome in district i , C_i denotes the share of ethnic Chinese of district i 's total population in 1930, and α_p denotes provincial fixed effects. We control for an extensive set of covariates X_i that may have affected Chinese settlement locations as well as long-term development. They include geographic controls of altitude, longitude, logarithm of distance to the Northern coast, logarithm of total area (1930), logarithm of total dry agricultural land, total wet agricultural land (both in 1963) and foreigner shares (1930), plus the quadratic polynomial of logarithm of distance to each major provincial capital (Batavia, Semarang and Surabaya) and to the historical Majapahit urban centers of Tuban, Gresik, and the Majapahit capital at Trowulan. Those variables may have largely influenced early Chinese settlement location all the way from the 15th century until modern day, as well as local economic conditions via governance and state capacity (Campante and Do, 2014; Bai and Jia, 2023). We correct standard error estimates for potential clustering at the regency level to avoid issues of spatially correlated error terms. We interpret ρ as the *causal* impact of the Chinese community size on contemporary economic development in a district, conditional on both observed and unobserved district characteristics,

To deal with the potential issues of Chinese immigrants' sorting and other omitted variables, we use the natural logarithm of each district's geodesic distance to landing sites on the Muria Strait $\ln(D_i)$ as instrumental variable for C_i , its share of Chinese in 1930. The method's causal interpretation relies on the strength of this instrument as well as its excludability, understood in this case as $\mathbb{E}(\varepsilon_i | D_i, X_i) = 0$.

To operationalize this empirical strategy, we need to avoid the extraordinary influence of Batavia (Jakarta nowadays) as the major attraction of a large part of the population and economic activities, leading to a large number of Chinese living relatively near Batavia in West Java. Hence we choose to be cautious by excluding the entire province of West Java.¹⁸

5.1.1 Assessing the Empirical Strategy

We begin by considering the instrument's strength. Table 3's Panel B presents our first-stage result from estimating equation (2) on the benchmark sample of 242 districts. An increase in the distance to the Muria Strait significantly reduces the Chinese population share, with a corresponding t-statistic of 3.67 and the Kleibergen-Paap weak-instrument-robust F-statistic of 13.75 ($= 3.67^2$). In other specifications, depending on the matched samples, the Kleibergen-Paap F-statistic may get smaller than the commonly used threshold of 10. For robustness, we thus report the Anderson-Rubin test of the significance of the main coefficient, as this test is robust to weak instrument concerns (Keane and Neal, 2023).

We further examine the exclusion restriction indirectly by estimating regressions similar to the first

¹⁸Table A.3 shows the OLS results on the larger sample that includes West Java, and Tables A.4 and A.5 show IV results on that extended sample. Overall, while the IV is much weaker on the extended sample due to the concern about Batavia mentioned above, the estimates still show broadly similar qualitative messages.

stage using alternative observable outcomes. We focus on other ethnic groups, represented in the 1930 census as Native, European, and other foreigners. Table 4 shows that the shares of all other minority groups were not affected by the instrument, which indicates that any potential violation of the exclusion restriction unlikely passes through other ethnic groups. (The strong positive effect on the share of natives is naturally expected as a near mirror image of the effect on the share of Chinese.)

In subsections 5.2.1 and 5.2.2, we present two additional empirical strategies that deepens our IV approach to deal with particularly concerns about the potential confoundedness of distance measures.

5.2 Effects on Contemporary Economic Development

Table 3 presents results on our main measures of contemporary economic development. Column (2) estimates a large and positive effect of 1930 ethnic Chinese shares on population density in 2000. Specifically, a 1 p.p. increase in 1930 ethnic Chinese shares leads to 26% higher population density in 2000. The corresponding effect on urbanization rate shown in column (3) is 13.28 percentage points, and that on night-time luminosity is 5.57 points out of a mean of 13.87. Last, such effect of the 1930 share of ethnic Chinese on contemporary household consumption is 9.4%.

Two points of caution are in order regarding those large magnitudes. First, as demonstrated in Table 1, the average ethnic Chinese share in 1930 is approximately 0.9%, and a 1 standard deviation increase in the ethnic Chinese share is around 1.2%. As such, a one percentage point increase in the Chinese share is substantial. Second, the magnitude corresponds to a low difference in average annual growth rate, as these differences likely have accrued over many decades.

Column (1) shows that the effects on the 1930 population density are relatively smaller and statistically insignificant. We interpret this as important supporting evidence that initial ethnic Chinese settlements were not more likely to have been established in places that were more positively selected on (un)observables for development and growth and that positive effects of ethnic Chinese settlements did not yet accrue in the early pre-industrialization period of early 20th-century Indonesia.

We further test whether the results in Table 3 are sensitive to controlling for important historical determinants of development in Indonesia, by including the presence of sugar factory (Table A.9), the Communist votes in the 1950s (Table A.10), distance to the colonial railroad system (Table A.11), distance to the great postal road (Table A.12), and distance to the opium concessions' boundaries (Table A.13). The results remain largely similar, thus strengthening the causality of the estimated effects.

5.2.1 Robustness Strategy Using Counterfactual Landing Sites near Cirebon

We further deepen the IV approach in two ways to deal with potential confounders of the IV strategy. The essential concern is that, while the historical landing sites can be arbitrary due to lack of longitudinal knowledge in navigation technology (Pascali, 2017; Rivers, 2012), the distance to those sites may incorporate

information that correlates with potential determinants of development. It is thus important to try to isolate the arbitrary source of variation as much as possible.

First, we propose to consider a particular set of counterfactual landing sites. This approach takes advantage of the arbitrary landing sites of Zheng He's expeditions by considering a set of counterfactual landing sites that could have been picked by the expeditions, and then compare the hypothetical consequences of those landing sites on the construction of the IV with the actual IV constructed from the distance to historical landing sites.

For this purpose, we choose the natural counterfactual landing site of Cirebon, right at the limit between Central Java and West Java. Cirebon is the key deep-water port in Central and East Java, and has also attracted a lot of Chinese immigrants by 1930 (Figure 1). It was unlikely a landing site by Zheng He's expeditions (see [Salmon and Lombard \(1981\)](#) on the "Chinese chronicles of Cirebon and Samarang"). By letting Cirebon play the role of Semarang, we recreate four hypothetical landing sites on the Northern coastline to the East of Cirebon that resemble the actual landing sites of Demak, Koedoes, Pati, and Waroe (Rembang) to the city of Semarang. We calculate the counterfactual geodesic distance from each district in our sample to those counterfactual landing sites.

Next, we run similar regressions to the specifications in [1](#) and [2](#) with added fixed effects to compare only locations with certain distances to the historical sites with other locations with similar distances to the counterfactual sites. The methodological details are explained in section [5.2.1](#).

Using this method, the OLS results in Table [A.6](#) and the IV results in Table [A.7](#) closely follow the main OLS ([Table 2](#)) and IV results ([Table 3](#)). They thus underscore the credibility of the long-run effects of the Chinese communities.

5.2.2 Robustness Strategy Using Recentered IV

Another concern of our main IV strategy is that, even if the location of the landing sites is randomly assigned, those sites are still constrained to be on a specific segment of the Northern coastline of Java, and that restriction would impose certain additional information on the calculation of distances from all districts in our sample. Consequently, districts that are closer to the Northern coast tend to have a shorter distance to the Muria Strait, and using this variation in our empirical analysis may entail confounding factors that could have produced the main results without a real impact of the Chinese communities.

This concern is akin to the more general setting in [Borusyak and Hull \(2023\)](#)'s examination of instruments as nonlinear functions of an arbitrary treatment. We follow [Borusyak and Hull \(2023\)](#)'s method by calculating the *expected* treatment of each district, denoted $\mathbb{E} [\ln(D_{pd})]$ where the expectation is taken over alternative treatment assignment procedures over all the potential initial settlement sites, and then extracting it from the instrument to compute the recentered IV as $D_{pd}^r = D_{pd} - \mathbb{E} [D_{pd}]$. This IV is then applied in the same specification as in equation [2](#).

To compute $\mathbb{E} [\ln(D_{pd})]$, we construct alternative sets of landing sites in a similar way to our construc-

tion of the counterfactual landing sites around Cirebon in section 5.2.1, starting at any arbitrary point on the segment of the Northern coastline that spans from Cirebon to Surabaya,¹⁹ then averaging the distance over the entire segment. More details of this calculation is elaborated in section A.3 .

The results are shown in Table A.8 . They are slightly smaller but still quantitatively close to the results in Table 3. The first stage's strength is still reasonably good, with a Kleibergen-Paap weak-instrument-robust F-statistic of 12.59. This robustness exercise thus lends additional credibility to our main interpretations of the long-run effects of the Chinese communities.

5.2.3 Nighttime Light Intensity and Growth

Next, we focus on the outcome variable of nighttime light intensity between 1992 and 2010 as our main proxy for development, for two key advantages. First, it gives us a balanced panel of districts over time from 1992 to 2010. Second, and conditional on being a relatively robust measure for sub-national economic development (Henderson et al., 2012), it allows us to study geographically granular changes in growth and development over time.

Table 5 presents results where our dependent variable is the level of nightlights at the beginning of 3 decades (1992, 2000, and 2010). Across all 3 columns, districts with higher Chinese shares have a positive and statistically significantly higher nighttime light intensity and the effects appear to be larger in magnitude in 2010 vis-a-vis 1992 and 2000. In terms of magnitudes, the positive coefficient of 1.91 in Column (2) suggests that a 1 p.p. higher ethnic Chinese share in 1930 is associated with a 6.07% higher GDP per capita, using Henderson et al. (2012)'s estimated elasticity of GDP per capita to lights of 0.277 (as $\frac{1.91}{8.71} \times 0.277 = 0.0607$).

The upward nature of the estimates in columns (2) to (4) points us to the question of the effect of Chinese shares on long-run growth. Indeed, we find in column (5) that a 1p.p. increase in Chinese shares is associated with a 3.657 increase in the growth of nighttime light intensity over the period. This substantial growth effect has likely taken place throughout the 20th century, as shown by the long-run effect on population density in column (1).

5.2.4 Chinese Communities through Social Conflicts

The estimated positive effects of Chinese shares over the long-run (1992-2010) point in the opposite direction of the many discriminatory policies targeting ethnic Chinese throughout modern Indonesian history, and likely mask important, intermediate fluctuations in macroeconomic conditions due to attendant episodes of social conflict (see Section 2). As targets of discriminatory policies under both Sukarno's and Suharto's various governments, ethnic Chinese became scapegoats at the tumultuous period following the 1997 Asian Financial Crisis and the unpredictable end of Suharto's regime. This turning episode of coin-

¹⁹This range covers the historical passage of Zheng He's expeditions to Java, throughout Central Java all the way to Surabaya, where the Brantas river that runs through the Majapahit capital in Trowulan pours into to the sea at Surabaya.

ciding economic and political crises, as predicted by [Grosfeld et al. \(2020\)](#), culminated in large-scale riots and targeted assaults against ethnic Chinese across Indonesia in May 1998, which have resulted in the massacre of at least a thousand deaths of ethnic Chinese ([Purdey, 2006](#)). The riots subsided in 1999, and the second president after the fall of Suharto, Abdurrahman Wahid (also known as Gus Dur), reversed most discriminatory policies against the Chinese minority in 1999-2001.²⁰

The frequency of social conflicts over these different periods are highlighted in Table [A.17](#). Based on reports of conflicts by UNSFIR, there was a surge of conflicts in districts with more Chinese presence in a single year of 1998 (statistically significant at 5% by the AR p-value). After 1998 the effect subsides, yet there remains significantly more conflicts in those areas (statistically significant at 10% by the AR p-value), as both recorded by UNSFIR and by ACLED. We do not see any effect of the Chinese share on officially classified anti-Chinese conflicts, because naming those conflicts as such was a taboo before 2000.

To examine the variation across those periods, Table [6](#) replicates results in Table [5](#) by separately regressing the annualized growth of nighttime light intensity (i.e., the change in light intensity divided by the number of ellapsed years) between (i) 1992-1997 (under the Suharto regime until the onset of the Asian Financial Crisis and targeted conflict episodes against the ethnic Chinese); (ii) 1997-1998 (the financial crisis and the assaults against ethnic Chinese); and (iii) 1999-2013 (the recovery period with the abolishment of most anti-Chinese discriminatory policies) on 1930 ethnic Chinese shares.

The table shows that a higher ethnic Chinese share in 1930 leads to (i) a strong, statistically significant positive effect on the growth of nighttime light intensity from 1992 to 1996 (ii) a substantial negative growth effect from 1996 to 1997 and from 1997 to 1998; and (iii) a *positive* growth rate of nighttime light intensity from 1998 to 2010. The magnitudes of the corresponding estimates on annual growth are larger than the estimate in column (5) of Table [5](#), showing the importance of the ups and downs during this long duration from 1992 to 2010. When we lower the unit of analysis to the village level, the results remain largely consistent, as shown in Table [A.14](#).

These results also show evidence of how discriminatory periods where a minority is socially prosecuted as scapegoats can lead to strongly negative economic outcomes, but these effects can be reversed and became substantially positive upon the abolishment of these policies, as previously shown in case of Jewish firms in France ([Do et al., 2024](#)) and rich families in China ([Alesina et al., 2020](#)). This result also contributes to the recent literature on the strong long-run resilience of an economic elite even in adverse contexts ([Ager et al., 2021](#); [Michelman et al., 2022](#)).

5.2.5 Chinese Communities and Structural Transformation

So far, we have presented evidence that the Chinese communities exert strong positive effects on local economic development over the long term. Given their magnitude, these findings hint at a potential effect of

²⁰Abdurrahman Wahid has a mixed native and Chinese ancestry, and is rumored to have ancestors that can be traced back to Zheng He's crews as well as the apostles that brought Islam to Indonesia.

the Chinese communities on the structural transformation of the Indonesian economy. We explore this possibility in Table 7, where we apply the specification in equation 2 on the employment shares of agriculture, manufacturing, and services across districts in 2010.

Column (1) shows that a 1 p.p. increase in ethnic Chinese shares leads to a 10.315 percentage point reduction in the agricultural employment share. This is from a sample mean of approximately 42 percent. In Columns (2) and (3), we consider the effects on employment shares in manufacturing and services. We find that a district with 1 p.p. higher ethnic Chinese shares in 1930, has on average, 5.8 p.p. more individuals working in the manufacturing sector, and 4.53 p.p. more individuals working in the services sector. As such, the findings suggest that both the manufacturing and services sectors are larger at the expense of the agricultural sector. While the two effects are not separately statistically significant, possibly due to an issue of weak instruments, the weak-instrument-robust Anderson-Rubin test clearly indicates a strong effect of the Chinese share on employment in the services (significant at 5%).

Taken together, our results suggest that districts with higher ethnic Chinese shares have become more economically developed with a higher (lower) share of individuals working in services (agricultural). Furthermore, these effects are unlikely to be mechanically driven by the occupational specialization of the Chinese communities themselves, as they only make up a small fraction of the total employment in most districts. Lastly, the results remain robust when we focus on the village as the level of analysis, as reported in Table A.15 .

5.3 Mechanism: Effects on Firms in Chinese Communities

Given the strong effects of Chinese communities on long-run economic development and structural transformation, we expect to find that those effects build upon the development of firms in Indonesia. This firm-level channel is particularly relevant in light of the narrative that Suharto leveraged wealthy Chinese financiers (*cukongs*) to develop the Indonesian economy since independence but especially from the 1990s (Borsuk, 2014).²¹ In particular, financial sector reforms of 1988 opened the banking sector to new, privately owned banks which were affiliated to ethnic Chinese conglomerates (Schwarz, 2018).

In this section, our analysis exploits administrative data covering the universe of Indonesian firms from 1980 to 2013, allowing us to trace how historical settlement patterns operate through distinct mechanisms across Indonesia's development trajectory. We first present general results throughout the available data from 1980 to 2013, before again dissecting it into three major periods, namely the first period of development under Suharto from 1980 to 1996, the second period around the economic and political crises from 1997 to 1999, and the third period from 2000 until the end of the available data in 2013.

²¹Case in point, Liem Sioe Liong, the founder of the Salim Group, Indonesia's biggest conglomerate, was well-acquainted with Suharto, even before his ascension to the presidency. First, during Indonesia's struggle for independence, when Liem was a small-time trader in the hills of Central Java and Suharto was part of the Indonesian Republican Army. Second, when Suharto was posted to Semarang in 1956 (Borsuk, 2014). More broadly, in response to rising anti-conglomerate sentiments, Suharto invited 31 leaders of Indonesia's largest conglomerates (29 of which were ethnic Chinese) to his ranch in West Java in 1990, where in spite of his having played a central role in creating these conglomerates, he proceeded to admonish and urge them to contribute to Indonesia's "growth, national stability and equitable distribution of wealth" in front of national television. (Schwarz, 2018).

5.3.1 Chinese Population Shares and Firm Performance

Table 8 reports the benchmark IV specifications that estimate the long-run effects of the share of Chinese in each district in 1930 on their firms' performance today. While the effect of the share of Chinese in 1930 on more firms accounted for in the Manufacturing Census (firms of a minimum of 20 employees in the manufacturing sector) is not statistically significant (column 1), the corresponding effect on average firm sale shown in column 2 is both statistically and economically strong. One p.p. increase in the share of Chinese is estimated to imply an increase of 53% ($= \exp(0.426) - 1$) in firm-average sales.

Naturally, the larger firms in places with more Chinese tend to dominate along many performance indicators. The more relevant question is whether they stand out in any dimension after controlling for their size. In columns (3) to (8), we examine those other dimensions, namely profits, capital (both market-price estimations and book values), total assets, and exports status and shares, all in the IV specification in equation 2 while controlling for the logarithm of firm sales.

The results show that Chinese communities do not make firms stand out on any other dimension but firm sales. If anything, those firms have lower book values of capital, although slightly higher values of market-estimated capital and total assets. This points to the possibility that they have made better investments in capital, and on average may get slightly higher profits out of them, although the effect on profit is not statistically significant. It is also noteworthy that those firms do not enjoy a stronger status of exporters, and in fact have a smaller share of exported goods.

Growth Dynamics. We move on to examine firm performances over the three periods surrounding the wave of assaults against ethnic Chinese in 1998. Similar to the results from Table 6, firms in Chinese communities are larger (Table A.16) and grow faster (Table A.18) during the pre-crisis period. During the crisis period from 1997 to 1999, they suffer more than firms elsewhere, and experience negative growth. After the crisis, firms in Chinese communities grow even faster and regain their cross-sectional advantage.

5.3.2 Chinese Population Shares and Firms' Investment Sources

Seeking to detect firms' growth advantages, we continue to examine the effect of the Chinese communities on various sources of firm's investment in Table 9. There remains no evidence of the narrative of favorable financing for firms in more Chinese-populated communities. While those firms tend to have a stronger investment profile with a 14 p.p. higher ratio of investment per sales (significant at 5% according to the AR p-value), they have a lower share of official bank loans, from both domestic and foreign sources, and receive less funding from the government and FDI sources (although all of those estimates are not statistically significant.) Moreover, stock market financing is significantly less of a financial source for those firms, a result that is consistent with potential discrimination on the stock market (similar to Do et al. (2024)'s evidence of antisemitism on the stock market).

Strikingly, the only category where firms in Chinese communities have a substantial advantage is financing from *private* sources. The effect size is substantially large: private-source investment is 21 p.p. higher (as a share of sales). These sources of finance refer to that from private networks of financiers – a unique feature of the Chinese communities that has been highlighted in all accounts of the Chinese in Southeast Asia, especially in comparison with the Jews in Europe (Suryadinata, ed, 2008; Suryadinata, 2015). Hence this result is consistent with ethnographic accounts of Chinese businesses in Southeast Asia and, more broadly, provides one potential explanation for the strong resilience of Chinese communities in the face of discrimination.

To highlight the financial privileges from those private networks, we can turn to general access to financing. Table A.24 shows that households in places with more Chinese presence have no advantage in financial access across all categories considered in the household survey *Susenas*. If anything, households from those districts are at a small disadvantage of 2.5% less likely to have access to business loans, and are 2.6% more likely to sell their belongings in financial distress (both numbers significant at 5% by the AR p-value). There is thus a clear distinction between the private financing networks that firms enjoy in those areas and the general level of financial development.

Investment Dynamics. Table A.19 further focuses on the variation of firms' private sources of investment over the three major periods around the wave of assaults against ethnic Chinese in 1998. The advantages in terms of private investment created in Chinese communities were particularly strong in the first period before the Asian crisis. These advantages are consistent with the possible effects of the financial sector liberalization in 1988, especially in terms of opening entries to ethnic Chinese banks (Schwarz, 2018). However, they collapsed completely during the crisis period. After the crisis, we do not see much evidence of a strong recovery as we have seen in sales – one key caveat is that after 2001 we only observe one more year of data in 2006 on investment sources.

This pattern is inconsistent with alternative explanations based on permanent regional characteristics unrelated to Chinese settlement patterns. If advantages stem from immutable factors like geography or resource endowments, we should observe persistent effects even during political turmoil. Instead, the collapse and recovery patterns strongly suggest that the advantages operate through social and economic networks concentrated in areas with historical Chinese presence – networks that can be disrupted by targeted ethnic violence and institutional shocks.

After the crisis period, networks of ethnic Chinese capital likely never fully recovered. This was possibly due to the flight of ethnic Chinese capital and entrepreneurs to other countries in the region (e.g. Singapore and Hong Kong) in the aftermath of the crisis (Borsuk, 2014).

5.3.3 Firms' Other Advantages

Stronger investment among firms in Chinese communities should also produce advantages on other dimensions. We focus on labor advantages in Tables A.20 , A.21 , and A.22 . Chinese share in 1930 generates positive effects on employment, the total wage bill, and also firm's average wage during the period before the Asian crisis in 1997. This effect is further consistent with the effect of Chinese share in 1930 on labor productivity shown in Table A.23 . The combination of rising wages and labor productivity suggests firms in areas with greater Chinese presence pursue capital-intensive, skill-biased production strategies rather than just labor-intensive growth.

We again find that the double crisis period of 1997-1999 wipes away a large part of the advantages accrued to Chinese communities. Unlike the results on firm sales, we do not find much evidence of a recovery, in part because after 2001 data on firm's investment, employment, and wages were only available for the year 2006.

5.3.4 Interpretation of Results on Firms

These patterns collectively indicate that historical Chinese settlement affects firm outcomes through three distinct phases: (1) a broad-based advantage phase (1980-1996) where firms in districts with more Chinese presence can take advantage of strong access to private sources of financing; (2) a collapse phase (1997-1999) when political violence targeting ethnic Chinese disrupts region-specific advantages across all channels; and (3) a selective recovery phase (2000-2013) suggesting limited recovery that is less reliant on advantages from ethnic Chinese networks.

The symmetric rise-fall-partial recovery pattern, combined with the comprehensive nature of the collapse during the double crisis, provides evidence that the mechanisms operate primarily through social and economic networks associated with Chinese settlement rather than persistent physical or institutional infrastructure. The natural experiment of the double crisis demonstrates that regional advantages due to Chinese presence can be eliminated when ethnically-targeted violence disrupts local networks, then partially reconstituted when institutional conditions stabilize—but with altered composition favoring internal over external financing channels. While we cannot directly identify firm ownership or ethnicity in these data, the strong interaction between Chinese population shares and ethnic violence suggests networks linking co-ethnic entrepreneurs, financiers, and business communities play a central role in transmitting historical settlement patterns into contemporary firm outcomes.

5.4 Historical Development of Chinese Communities

How far back do the positive effects of Chinese communities on firm performance date? To explore this question in more depth, we consider data on historical firm characteristics among Chinese firms dating back to the late 19th century. This was a period of economic liberalization, during which Chinese communities

across Java diversified their economic activities to more sectors and entrepreneurial activities ([Zanden and Marks, 2012](#)).

Table [A.25](#) shows reduced-form results that link our IV to firm-level data. As those data precede the 1930 census, we consider only the intention to treat in this section. We find that distance to historical landing sites, controlling for the full set of covariates in equation [2](#), reduces both the number of firms, the number of Chinese firms, and their equity sizes in a district. As such, the positive effects on economic development appear to precede the period of accelerated growth in the 1960s. The positive effects on contemporary economic development can therefore be driven by both the Chinese communities themselves since 1930, and the increased economic development they had induced even before independence.

To further examine quantitatively the historical development of Chinese communities under Dutch colonial rules since the 19th century, we consider similar reduced-form style regressions with other historical outcome variables. Table [A.26](#) confirms that Dutch authorities did strengthen the class of Chinese officers in districts with more Chinese presence all the way back to 1867 (especially at the lower ranks), an effect that is not found among non-Chinese non-European officers.

Interestingly, the better business environment in Chinese communities has not conducted to a higher share of prominent Chinese in business and finance areas, as defined in [Suryadinata \(2015\)](#)'s collection of prominent Chinese in recent Indonesian history from 1941 to 1984 (Table [A.28](#)). Those communities do produce more prominent Chinese (Table [A.27](#)), but only in cultural domains (Table [A.29](#)). Taken together, this set of evidence hints that the big contribution of Chinese communities to development is more likely materialized after independence, although it may have taken some historical root since the Dutch colonial period.

6 Alternative Mechanisms

So far, we have documented that a larger historical Chinese population share is associated with higher levels of long-term economic development and firm performance. What explains these effects? To account for the observed patterns, a mechanism should satisfy three conditions. First, they need to be localized in that they affect the development of the district, not the entire country or regions as a whole. Second, considering the small Chinese population share, there needs to be positive spillovers to the wider economy to account for the magnitudes of the effects. Finally, it needs to account for the time-varying impacts of the Chinese presence. The economic and historical literature suggests various mechanisms through which Chinese communities could have led to higher economic development: (1) spillovers; (2) human capital; (3) market access and international trade; (4) Economic specialization and complementarities; and (5) cultural transmission. These channels satisfy the three conditions to varying degrees, and multiple mechanisms are likely to operate simultaneously. We therefore examine each potential channel in turn.

6.1 Negative Spillovers to Neighboring Districts

In Table A.30, we first consider whether better development outcomes at districts with more Chinese presence occur at the expense of neighboring districts. We do so by estimating equation 2 in dyadic form on a sample of dyads (i, j) that includes a pair of each original district i with all of its neighbors j (the sample size thus expands from 242 to 1,131). In addition to the IV for district i as used in specification 2, we also include the corresponding IV for the neighboring district j .

Overall, we do not find evidence of negative spillovers to neighboring districts. As such, we interpret the effect mainly as promoting growth in the district, rather than a change in the location in economic activity.

6.2 Human Capital and Education

A sizable literature emphasizes the role of human capital in mediating the impact of migration on host countries (Hornung, 2014; Rocha et al., 2017). Anecdotal sources suggest that contemporary human capital in Chinese communities is relatively higher. This is a pattern dating back at least to the 1920s (see e.g. Zanden and Marks, 2012, p. 123).²² Table A.31 presents estimates on educational attainment and shows that individuals living in districts with a higher ethnic Chinese share are not more likely to have bachelor's or master's degrees, and if anything, are slightly less likely to finish high school. However, none of these results are statistically significant. Turning to the supply side, we find that these areas have a slightly larger number of high schools and universities. However, this is to some extent to be expected in light of the higher density in these districts. Taken together, in light of the mixed evidence, we conclude that the results presented in sections 4 and 5 are unlikely to have been driven by human capital spillovers from ethnic Chinese communities.²³

6.3 International Trade, International Aid, and Market Access

We further explore whether the effects can be driven by the role of the Chinese communities in promoting trade and the development of local market access (see e.g. Johnson and Koyama, 2017). As early as 1822, trade with China was significant, making up ten percent of imports from outside Indonesia (see Zanden and Marks, 2012, p. 37). We explore the question in three ways. First, the results on exporting activities among firms shown previously in Table 8 do not show that firms in districts with more Chinese presence are more involved in exporting.

Second, we consider the role of market access. The economic benefits brought by Chinese communities

²²As pointed out in Zanden and Marks (2012), in the 1920 census, only 6.52 per cent of male Indonesians were able to read and write, while the same number for "Foreign Asiatics" was 58.09 (see Zanden and Marks, 2012, p. 123).

²³These findings are seemingly at odds with the findings in Hup and De Zwart (2024), who document that Chinese migrants to Indonesia in the early 20th century were positively selected. However, we consider the impact of the entire stock of Chinese in 1930, only some of whom arrived in the early 20th century. Furthermore, our outcome captures the education spillovers to the district as a whole, grouping both the Chinese and non-Chinese populations.

historically might have given rise to sunk investments in these districts. As a result, it is possible that part of the contemporary effects of Chinese communities documented above are driven by subsequent investments in transportation infrastructure. If this is the case, we would expect to see higher market access today in locations with larger Chinese communities. To explore this issue, we construct a measure of market access by combining data on economic activity and the modern data transportation network and then use this measure as an outcome (see section A.4). The results in Table A.32 do not show evidence of this channel.

Third, we ask whether Chinese communities receive more favorable investments and aids, especially from China. Generally, during the period prior to the double crisis of 1997-1998 where the effects of Chinese presence was by far the strongest (see section 5.3), international investments and aids from China were unlikely substantial, since China was still quite underdeveloped. To explore subsequent periods, we consider China's foreign aids to Indonesia from 2003 to 2009, separately in terms of infrastructure projects and non-infrastructure causes (mostly humanitarian) in Table A.33 .

Chinese aids indeed favor areas with more Chinese presence, even after controlling for development in 1998. This can be a form of ethnic favoritism in development aid (Dreher et al., 2019). However, this favoritism is only present among infrastructure projects, while the effect is opposite for non-infrastructure projects. Furthermore, using Chinese aid as mediator does not change substantially the estimated effects of Chinese share, confirming that Chinese aid plays no significant role in explaining our main findings.

6.4 Other Political and Cultural Mechanisms

As the Chinese primarily originated from more commercially oriented coastal regions of Southern China (Macauley, 2021; Wu and Wu, 1980), we consider the role of cultural spillovers of market-oriented values and beliefs or ethnic tolerance in the majority population. Using information on social attitudes from the *Susenas*, Table A.34 shows no significant evidence of this mechanism.

Lastly, one may think Chinese immigrants may have helped build stronger state capacity where they live. This mechanism is unlikely, as Chinese were barred from all public sector jobs since independence until the end of the 20th century. We verify state's public investments in those areas in Table A.35 . Chinese communities receive on average significantly less government grants per head. Generally, there is no evidence that they get better support from the central government.

7 Discussion

Our findings are consistent with the view that the positive effects of the Chinese population size are driven by the degree of occupational specialization in middle-man occupations, facilitating firm growth and access to finance. In 1875, 48% of the Chinese population was engaged in some form of commercial activity, in 1930, this had increased to 58% (Claver, 2014, p.139). 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, which are likely to be undersupplied in contexts with limited legal enforcement. The question then arises, what drove the underlying comparative advantage that induced the Chinese to sort into such occupations at a higher rate? There are two particularly plausible channels: a productivity advantage in these occupations or policies that limited occupational mobility among the Chinese. Below, we consider the plausibility of each explanation in turn.

First, it is plausible that ethnic Chinese had a productivity advantage in middleman occupations during the colonial period. Important contributions in economics highlight the importance of social organization in facilitating economic exchange in contexts with weak or limited enforcement of formal laws (Greif, 1989, 1993). Kinship ties, which are particularly important for organizing economic activity (Zhongguo et al., 1992, p. 91) and deep-rooted in the context of China (Greif and Tabellini, 2017), can help to overcome commitment problems and foster trust in such contexts (Landa, 1998).²⁴ While Javanese kinship ties were far less extensive, Chinese kinship networks also facilitated access to trading networks and supported the growth of larger firms through capital pooling and risk sharing (Claver, 2014). Kinship ties conferred an advantage that did not erode over time as they are non-expropriable and not easily replicated by the native population (Jha, 2013, 2018).

A second possibility is that a comparative advantage in middleman occupations emerged in response to colonial policy. Several policies might have limited occupational mobility among the Chinese during the colonial period. One example was the limits put on labor mobility during the 19th century. This system was implemented after the massacre of Chinese in Batavia in the 1740s, was rigorously enforced since the 1830s, and lasted until the early 1900s (Ongkokham, 1989). The system did not apply to certain occupations, such as the tax farmers, who could move freely to collect taxes in their tax farms. Policies that might have contributed to limiting the occupational mobility were also present during the early years of independence, and the New Order regime (for examples of such policies, see Wu and Wu, 1980, p. 173). Such policies could induce the Chinese to sort into certain occupations.

We remain agnostic about the relative importance of the two mechanisms; however, the persistent occupational segregation of the Chinese across time and space suggests a central role for productivity advantages. First, the presence of Chinese in major ports in Southeast Asia and Java predates the arrival of the Europeans (Pan, 1999, p. 152). Second, occupational segregation was widespread across Southeast Asia, found in areas colonized by different powers (Malaysia and Vietnam) and countries that remained independent alike (Thailand) (Wu and Wu, 1980). Third, it remained fairly stable within Java over time, although policies towards the Chinese underwent substantial changes over time. Lastly, it has largely persisted to this day, despite more than eight decades of independence. However, Dutch colonial policy likely reinforced the ethnically segregated occupational pattern. In addition, the initial pattern of specialization was

²⁴Enke (2019) provides quantitative evidence that kinship tightness is particularly high in the context of China.

likely further reinforced by endogenous responses to the initial pattern in the form of social organization ([Wu and Wu, 1980](#)), skills, and capital that enabled this advantage to persist.

Kinship networks enabled Chinese communities to organize more complex forms of economic activity and exchange. Our findings suggest that this promoted economic activity in the entire locations where they were present. Yet, this complementarity proved insufficient to promote peaceful coexistence. We can understand this in light of the findings and framework in ([Jha, 2013, 2018](#)). In contrast to trading networks based on a shared religious affiliation, integration into a different kinship or clanship network is very costly ([Landa, 1998](#)). This suggests that middleman occupations in the context of Java were characterized by high barriers of entry, resulting in high markups and rents in these occupations that were shared by the Chinese and colonial elites ([Zanden and Marks, 2012](#), p. 66). This is exemplified by the existence of policies that sought to limit the exploitation of the Javanese by the Chinese in the hinterland. While this arrangement benefited the Chinese during stable periods, it also fostered resentment, making the Chinese vulnerable to scapegoating during episodes of political and economic instability, such as the 1998 riots.

8 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.

An important limitation of this study is that we cannot fully explain why a comparative advantage in middleman occupations emerged during the colonial period; such patterns may reflect both discriminatory barriers and inherent productivity differences across sectors. Future work examining Chinese communities in other settings where colonial institutions varied could help assess external validity and shed further light on the mechanisms behind these long-run patterns.

References

- Acemoglu, Daron, Simon Johnson, and James A. Robinson**, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001, 91 (5), 229.
- _____, _____, and **James A. Robinson**, "Institutions as a Fundamental Cause of Long-Run Growth," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Elsevier, 2005, pp. 385–472.
- _____, **Tarek A. Hassan, and James A. Robinson**, "Social Structure and Development: A Legacy of the Holocaust in Russia," *The Quarterly Journal of Economics*, May 2011, 126 (2), 895–946.
- Ager, Philipp, Leah Platt Boustan, and Katherine Eriksson**, "The Intergenerational Effects of a Large Wealth Shock: White Southerners after the Civil War," *American Economic Review*, 2021, 111 (11), 3767–3794.
- Akulut-Yuksel, Mevlude and Mutlu Yuksel**, "The Long-Term Direct and External Effects of Jewish Expulsions in Nazi Germany," *American Economic Journal: Economic Policy*, 2015, 7 (3), 58–85.
- Alesina, Alberto and Eliana La Ferrara**, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, 43 (3), 762–800.
- _____, and **Paola Giuliano**, "Family Ties," in "Handbook of Economic Growth," Vol. 2, Elsevier, 2014, pp. 177–215.
- Alesina, Alberto F., Marlon Seror, David Y. Yang, Yang You, and Weihong Zeng**, "Persistence Despite Revolutions," Working Paper 27053, National Bureau of Economic Research 2020. Revised March 2021.
- Allen, Treb and Costas Arkolakis**, "Economic Activity across Space: A Supply and Demand Approach," *Journal of Economic Perspectives*, May 2023, 37 (2), 3–28.
- Ananta, Aris, Evi Nurvidya Arifin, M Sairi Hasbullah, Nur Budi Handayani, and Agus Pramono**, *Demography of Indonesia's Ethnicity*, Institute of Southeast Asian Studies, 2015.
- Anderson, Robert Warren, Noel D. Johnson, and Mark Koyama**, "Jewish Persecutions and Weather Shocks: 1100–1800," *The Economic Journal*, 2017, 127 (602), 924–958.
- Anderton, Charles H. and Jürgen Brauer**, "Mass Atrocities and Their Prevention," *Journal of Economic Literature*, 2021, 59 (4), 1240–1292.
- Arbatlı, Cemal Eren and Gunes Gokmen**, "Human Capital Transfers and Sub-national Development: Armenian and Greek Legacy in Post-expulsion Turkey," *Journal of Economic Growth*, March 2023, 28 (1), 1–43.
- Auwalin, Ilmiawan**, "Ethnic Identity and Internal Migration Decision in Indonesia," *Journal of Ethnic and Migration Studies*, 2020, 46 (13), 2841–2861.
- Bai, Ying and Ruixue Jia**, "The Economic Consequences of Political Hierarchy: Evidence from Regime Changes in China, 1000–2000 CE," *Review of Economics and Statistics*, 2023, 105 (3), 626–645.
- Bassino, Jean-Pascal and Pierre van der Eng**, *The First East Asian Economic Miracle: Wages, Living Standards and Foundations of Modern Economic Growth in Southeast Asia, 1880-1938*, Australian National University, 2021.
- Bates, Robert H.**, "Ethnic Competition and Modernization in Contemporary Africa," *Comparative Political Studies*, 1974, 6 (4), 457–484.
- Baum-Snow, Nathaniel, J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt**, "Does Investment in National Highways Help or Hurt Hinterland City Growth?," *Journal of Urban Economics*, January 2020, 115, 103124.

- Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong**, "Skill transferability, migration, and development: Evidence from population resettlement in Indonesia," *American Economic Review*, 2016, 106 (9), 2658–2698.
- _____, _____, _____, and _____, "Unity in diversity? How intergroup contact can foster nation building," *American Economic Review*, 2019, 109 (11), 3978–4025.
- _____, **Gabriel Koehler-Derrick, and Benjamin Marx**, "The Institutional Foundations of Religious Politics: Evidence from Indonesia," *The Quarterly Journal of Economics*, 2020, 135 (2), 845–911.
- _____, **Masyhur Hilmy, and Benjamin Marx**, "Religion, Education, and the State," *NBER Working Paper* 27073, 2020.
- Becker, Sascha O. and Luigi Pascali**, "Religion, Division of Labor, and Conflict: Anti-semitism in Germany over 600 Years," *American Economic Review*, May 2019, 109 (5), 1764–1804.
- _____, **Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya**, "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review*, May 2020, 110 (5), 1430–1463.
- _____, **Sharun Mukand, and Ivan Yotzov**, "Persecution, Pogroms and Genocide: A Conceptual Framework and New Evidence," *Explorations in Economic History*, 2022, 86, 101471.
- Bestari, Rossi Rizki, Yusuf Reza Kurniawan, and Muhammad Halley Yudhistira**, "The Long-term Effects of Early European Settlement on Local Development: Evidence from Indonesia," *World Development*, 2022, 158, 105977.
- Bonacich, Edna**, "A Theory of Middleman Minorities," *American Sociological Review*, 1973, 38 (5), 583–594.
- Booth, Anne**, "Living Standards and the Distribution of Income in Colonial Indonesia: A Review of the Evidence," *Journal of Southeast Asian Studies*, 1988, 19 (2), 310–334.
- _____, *The Indonesian Economy in the Nineteenth and Twentieth Centuries: A History of Missed Opportunities* A Modern Economic History of Southeast Asia (MEHSA), London, UK: Macmillian Press, 1998.
- Borsuk, Richard**, *Liem Sioe Liang's Salim Group*, Flipside Digital Content Company Inc., 2014.
- 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 2021.
- Caicedo, Felipe Valencia**, "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, 2019, 134 (1), 507–556.
- Campante, Filipe R. and Quoc-Anh Do**, "Isolated Capital Cities, Accountability, and Corruption: Evidence from US States," *American Economic Review*, 2014, 104 (8), 2456–2481.
- Card, David**, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, 19 (1), 22–64.
- Carey, Peter**, "Changing Javanese Perceptions of the Chinese Communities in Central Java, 1755–1825," *Indonesia*, 1984, (37).
- Caselli, Francesco and Wilbur John Coleman**, "On the Theory of Ethnic Conflict," *Journal of the European Economic Association*, 2013, 11, 161–192.

- Chen, Zhiwu, Chicheng Ma, and Andrew J Sinclair**, "Banking on the Confucian Clan: Why China Developed Financial Markets so Late," *The Economic Journal*, May 2022, 132 (644), 1378–1413.
- Chua, Amy**, *World on Fire: How Exporting Free Market Democracy Breeds Ethnic Hatred and Global Instability*, New York: Anchor Books, 2004.
- Chua, Christian**, "Defining Indonesian Chineseness under the New Order," *Journal of Contemporary Asia*, 2004, 34 (4), 465–479.
- Cinelli, Carlos and Chad Hazlett**, "Making Sense of Sensitivity: Extending Omitted Variable Bias," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2020, 82 (1), 39–67.
- Claver, Alexander**, *Dutch Commerce and Chinese Merchants in Java: Colonial Relationships in Trade and Finance, 1800-1942*, Vol. 291 of *Verhandelingen van het Koninklijk Instituut voor Taal-, Land en Volkenkunde*, Leiden and Boston: Brill, 2014.
- Cribb, Robert**, *Historical Atlas of Indonesia*, London: Curzon Press, 1997.
- De Zwart, Pim**, "Inequality in Late Colonial Indonesia: New Evidence on Regional Differences," *Clio-metrica*, 2022, 16 (1), 175–211.
- , **Daniel Gallardo-Albarrán, and Auke Rijpma**, "The Demographic Effects of Colonialism: Forced Labor and Mortality in Java, 1834–1879," *The Journal of Economic History*, 2022, 82 (1), 211–249.
- Deaton, Angus**, *The analysis of household surveys: a microeconometric approach to development policy*, World Bank Publications, 1997.
- Dell, Melissa**, "The persistent effects of Peru's mining mita," *Econometrica*, 2010, 78 (6), 1863–1903.
- 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.
- der Eng, Pierre Van**, "Chinese Entrepreneurship in Indonesia: A Business Demography Approach," *Business History*, 2022, 64 (4), 682–703.
- Die, Ong Eng**, *Chineezen in Nederlandsch-Indie: Sociografie van een Indonesische Bevolkingsgroep*, Van Gorcum, 1943.
- Dippel, Christian and Stephan Hebllich**, "Leadership in Social Movements: Evidence from the "Forty-Eighters" in the Civil War," *American Economic Review*, 2021, 111 (2), 472–505.
- Do, Quoc-Anh, Roberto Galbiati, Benjamin Marx, and Miguel A. Ortiz Serrano**, "J'Accuse! Anti-semitism and financial markets in the time of the Dreyfus Affair," *Journal of Financial Economics*, 2024, 154, 103809.
- Dreher, Axel, Andreas Fuchs, Bradley Parks, Austin Strange, and Michael J. Tierney**, "Aid, China, and Growth: Evidence from a New Global Development Finance Dataset," *American Economic Journal: Economic Policy*, 2021, 13 (2), 135–174.
- , ——, **Roland Hodler, Bradley C. Parks, Paul A. Raschky, and Michael J. Tierney**, "African Leaders and the Geography of China's Foreign Assistance," *Journal of Development Economics*, 2019, 140, 44–71.
- Dreyer, Edward L.**, *Zheng He: China and the Oceans in the Early Ming Dynasty, 1405–1433*, New York: Pearson Longman, 2006.
- Droller, Federico**, "Migration, Population Composition and Long Run Economic Development: Evidence from Settlements in the Pampas," *The Economic Journal*, 2018, 128 (614), 2321–2352.

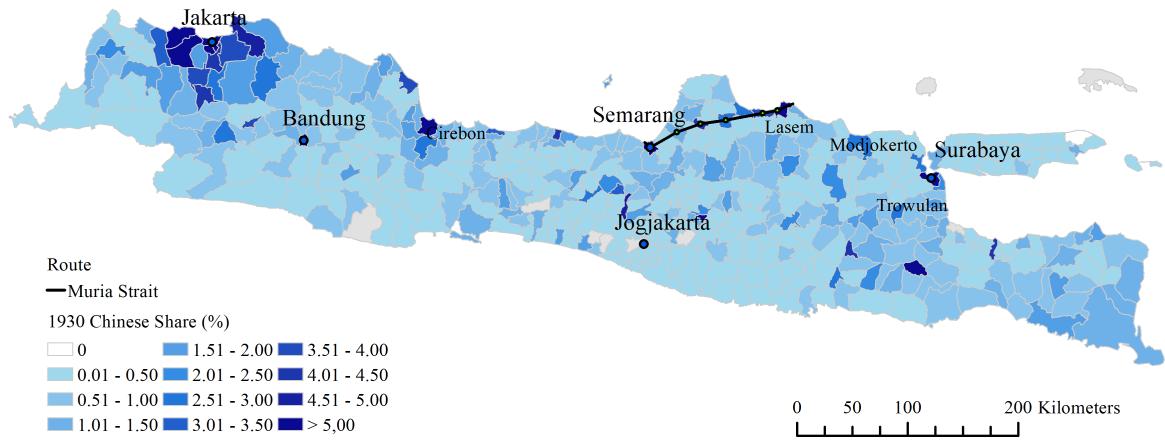
- D'Acunto, Francesco, Marcel Prokopcuk, and Michael Weber**, "Historical Antisemitism, Ethnic Specialization, and Financial Development," *The Review of Economic Studies*, May 2019, 86 (3), 1170–1206.
- Easterly, William and Ross Levine**, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 1997, 112 (4), 1203–1250.
- 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.
- Enke, Benjamin**, "Kinship, Cooperation, and the Evolution of Moral Systems," *The Quarterly Journal of Economics*, May 2019, 134 (2), 953–1019.
- Galor, Oded and Ömer Özak**, "Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability," *Manuscript*, 2015.
- _____ and _____, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, 106 (10), 3064–3103.
- González, Felipe**, "Immigration and Human Capital: Consequences of a Nineteenth Century Settlement Policy," *Cliometrica*, 2020, 14 (3), 443–477.
- Greif, Avner**, "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders," *The Journal of Economic History*, December 1989, 49 (4), 857–882.
- _____ , "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *The American Economic Review*, 1993, 83 (3), 525–548. Publisher: American Economic Association.
- _____ and Guido Tabellini, "The Clan and the Corporation: Sustaining Cooperation in China and Europe," *Journal of Comparative Economics*, February 2017, 45 (1), 1–35.
- Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya**, "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, January 2020, 87 (1), 289–342.
- Henderson, J Vernon, Adam Storeygard, and David N Weil**, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, 102 (2), 994–1028.
- Heryanto, Ariel**, "Ethnic Identities and Erasure," *Southeast Asian identities: Culture and the politics of representation in Indonesia, Malaysia, Singapore, and Thailand*, 1998, pp. 95–114.
- 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.
- Houben, Vincent J.H.**, "Java in the 19th Century: Consolidation of a Territorial State," in V.J.H. Houben and J.Th. Lindblad, eds., *The Emergence of a National Economy*, Leiden, Netherlands: KITLV Press, 2002, pp. 56–81.
- Huber, Kilian, Volker Lindenthal, and Fabian Waldinger**, "Discrimination, Managers, and Firm Performance: Evidence from "Aryanizations" in Nazi Germany," *Journal of Political Economy*, 2021, 129 (9), 2455–2503.
- Huff, Gregg and Giovanni Caggiano**, "Globalization, Immigration, and Lewisian Elastic Labor in Pre-World War II Southeast Asia," *The Journal of Economic History*, 2007, 67 (1), 33–68.

- 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.
- _____, ""Unfinished business": Historic Complementarities, Political Competition and Ethnic Violence in Gujarat," *Journal of Economic Behavior & Organization*, 2014, 104, 18–36.
- _____, "Trading for Peace," *Economic Policy*, 2018, 33 (95), 485–526.
- Jia, Ruixue and Torsten Persson**, "Choosing Ethnicity: The Interplay Between Individual and Social Motives," *Journal of the European Economic Association*, April 2021, 19 (2), 1203–1248.
- Johnson, Noel D and Mark Koyama**, "Jewish Communities and City Growth in Preindustrial Europe," *Journal of Development Economics*, 2017, 127.
- Kaur, Amarjit**, "Wage Labour in Southeast Asia since 1840," *Neg Yobk: Palgbafe Macmillan*, 2004.
- Keane, Michael P. and Timothy Neal**, "Instrument Strength in Structural Econometric Models," *Journal of Econometrics*, 2023, 235 (2), 344–368.
- Klinken, Gerry Van**, "Ethnicity in Indonesia," in "Ethnicity in Asia," New York, NY, USA: Routledge, 2003, pp. 88–111.
- Kok, Chun Chee, Gedeon J. Lim, Danial Shariat, Abu Siddique, and Shunsuke Tsuda**, "Interethnic Proximity and Political Development," 2025. Available as a working paper on SSRN and IZA.
- Kuipers, Nicholas**, "The Long-Run Consequences of The Opium Concessions for Out-Group Animosity on Java," *World Politics*, 2022, 74 (3), 405–442.
- Landa, Janet Tai**, *Trust, Ethnicity, and Identity: Beyond the New Institutional Economics of Ethnic Trading Networks, Contract Law, and Gift-exchange*, Ann Arbor, Michigan, USA: University of Michigan Press, 1998.
- Leeuwen, Marco HD Van and Ineke Maas**, *HISCLASS: A Historical International Social Class Scheme*, Universitaire Pers Leuven, 2011.
- Liem, Thian Joe**, *Riwajat Semarang*, Semarang: Ho Kim Yoe, 1933.
- Liem, Twan Djie**, "De Distribuerende Tusschenhandel der Chinezen op Java [The Distributive Intermediate Trade of the Chinese in Java]," *Nijhoff*, 1952.
- Lim, Gedeon J.**, "Political Elites and Land Rents: Evidence from Indonesia," Technical Report, The University of Hong Kong 2024. Pre-submission manuscript, available at https://github.com/gedeon89.github.io/gedeonlim.com/Lim_Bengkok_edited_pre_sub.pdf.
- Macaulay, Melissa**, *Distant Shores: Colonial Encounters on China's Maritime Frontier*, Princeton, New Jersey, USA: Princeton University Press, 2021.
- McKeown, Adam**, "Global Migration, 1846–1940," *Journal of World History*, 2004, 15 (2), 155–189.
- Michelman, Valerie, Joseph Price, and Seth D. Zimmerman**, "Old Boys' Clubs and Upward Mobility Among the Educational Elite," *The Quarterly Journal of Economics*, 2022, 137 (2), 845–909.
- Mills, J.V.G.**, *Ying-yai sheng-lan: 'The Overall Survey of the Ocean's Shores'* [1453], Cambridge, UK: Cambridge University Press, 1971.
- Mitra, Anirban and Debraj Ray**, "Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India," *Journal of Political Economy*, 2014, 22 (4).

- Montalvo, José G. and Marta Reynal-Querol**, "Ethnic Diversity and Growth: Revisiting the Evidence," *The Review of Economics and Statistics*, 2021, 103 (3), 521–532.
- Munshi, Kaivan**, "Community Networks and the Process of Development," *Journal of Economic Perspectives*, 2014, 28 (4), 49–76.
- _____, "Social Networks and Migration," *Annual Review of Economics*, 2020, 12, 503–24.
- Onghokham**, "Chinese Capitalism in Dutch Java," *Southeast Asian Studies*, 1989, 27 (2), 156–176.
- Pan, Lynn**, *The Encyclopedia of the Chinese Overseas*, Cambridge, Massachusetts, USA: Harvard University Press, 1999.
- Pascali, Luigi**, "The wind of change: Maritime technology, trade, and economic development," *American Economic Review*, 2017, 107 (9).
- Pepinsky, Thomas B**, "Colonial Migration and the Origins of Governance: Theory and Evidence from Java," *Comparative Political Studies*, 2016, 49 (9), 1201–1237.
- Purdey, Jemma**, *Anti-Chinese Violence in Indonesia, 1996–1999*, Honolulu: University of Hawai'i Press, 2006.
- Rauch, James E**, "Business and Social Networks in International Trade," *Journal of Economic Literature*, December 2001, 39 (4), 1177–1203.
- Rauch, James E. and Vitor Trindade**, "Ethnic Chinese Networks in International Trade," *The Review of Economics and Statistics*, February 2002, 84 (1), 116–130.
- Rivers, Philip J.**, "New Lamps for Old: Modern Nautical Terms for Ancient Marine Practices and the Navigation of the Zheng He Voyages," *Journal of the Malaysian Branch of the Royal Asiatic Society*, June 2012, 85 (1), 85–98.
- Rocha, Rudi, Claudio Ferraz, and Rodrigo R Soares**, "Human Capital Persistence and Development," *American Economic Journal: Applied Economics*, 2017, 9 (4), 105–136.
- Rush, James**, "Placing the Chinese in Java on the Eve of the Twentieth Century," *Indonesia*, 1991, pp. 13–24.
- Salmon, Claudine and Denys Lombard**, *Les chroniques chinoises de Cirebon et de Semarang*, Paris: École Française d'Extrême-Orient, 1981.
- Schwarz, Adam**, *A nation in waiting: Indonesia's search for stability*, Routledge, 2018.
- 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.
- Suryadinata, Leo**, *Admiral Zheng He and Southeast Asia*, Vol. 286, Institute of Southeast Asian Studies, 2005.
- _____, *Understanding the Ethnic Chinese in Southeast Asia*, Singapore: Institute of Southeast Asian Studies, 2007.
- _____, *Prominent Indonesian Chinese: Biographical Sketches*, Institute of Southeast Asian Studies, 2015.
- _____, ed., *Ethnic Chinese as Southeast Asians*, Singapore: Institute of Southeast Asian Studies, 1997.
- _____, ed., *Ethnic Relations and Nation-Building in Southeast Asia: The Case of the Ethnic Chinese*, Singapore: Institute of Southeast Asian Studies, 2004.

- _____, ed., *Ethnic Chinese in Contemporary Indonesia*, Singapore: ISEAS–Yusof Ishak Institute, 2008.
- 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* Routledge Studies in the Growth Economies of Asia (2005), Hoboken, NY, USA: Taylor and Francis, 2012.
- Varshney, Ashutosh, Mohammad Zulfan Tadjoeddin, and Rizal Panggabean**, "Creating Datasets in Information-Poor Environments: Patterns of Collective Violence in Indonesia, 1990–2003," *Journal of East Asian Studies*, December 2008, 8 (3), 361–394.
- Vleming, JL**, *Het Chineesche Zakenleven in Nederlandsch-Indië* number 730, Landsdrukkerij, 1926.
- Voigtländer, Nico and Hans-Joachim Voth**, "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," *The Quarterly Journal of Economics*, 2012, 127 (3), 1339–1392.
- Volkstelling**, *Census of 1930 in the Netherlands Indies. Vol. VII Chinese and Other Non-indigenous Orientals in the Netherlands Indies*, Batavia: Landsdrukkerij, 1930.
- Wu, Yuan-Li**, "Chinese Entrepreneurs in Southeast Asia," *American Economic Review: Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association*, 1983, 73 (2), 112–117.
- Wu, Yuan-li and Chung-Hsi Wu**, *Economic Development in Southeast Asia; the Chinese Dimension*, Stanford University, Stanford, California, USA: Stanford University Press, 1980.
- Zhongguo, Fei Xiaotong's Xiangtu, Gary G. Hamilton, and Wang Zheng**, *From the Soil: The Foundations of Chinese Society*, Berkeley, California, USA: University of California Press, 1992.

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



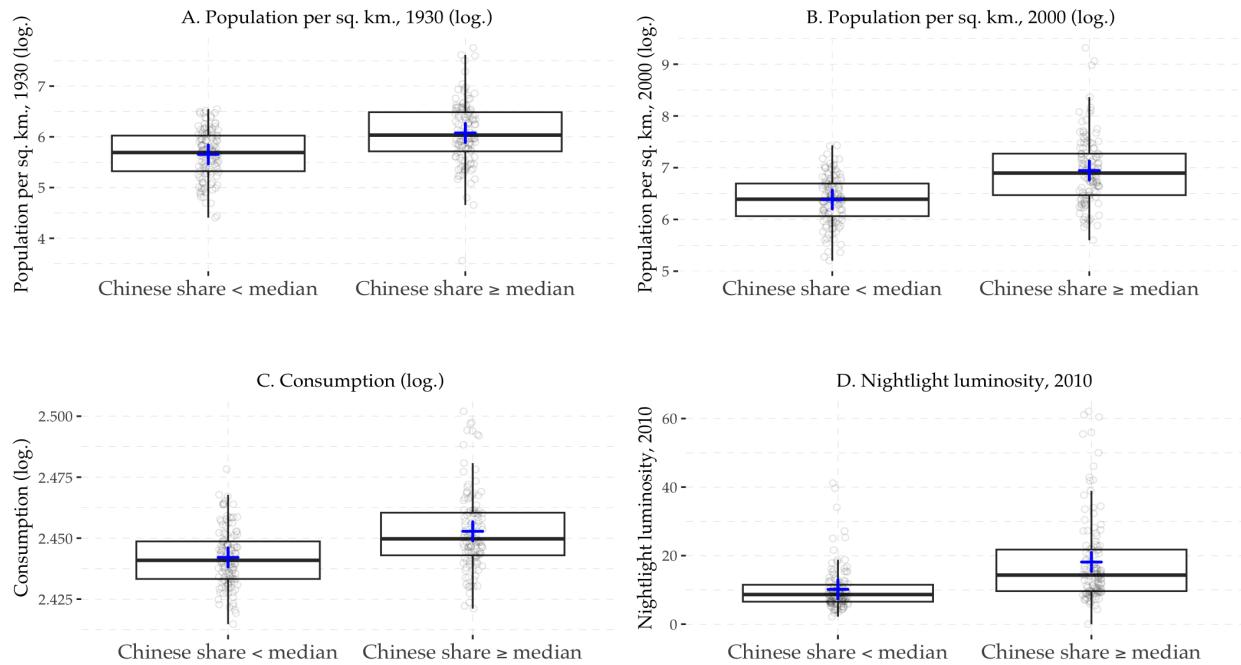
Notes: This figure depicts the underlying variation in our treatment variable: the share of Chinese by district-level across 1930 districts. The solid black line traces the Muria Strait across historical landing sites of Zheng He's expeditions in the early 15th-century. Source: Indonesian Population Census of 1930.

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



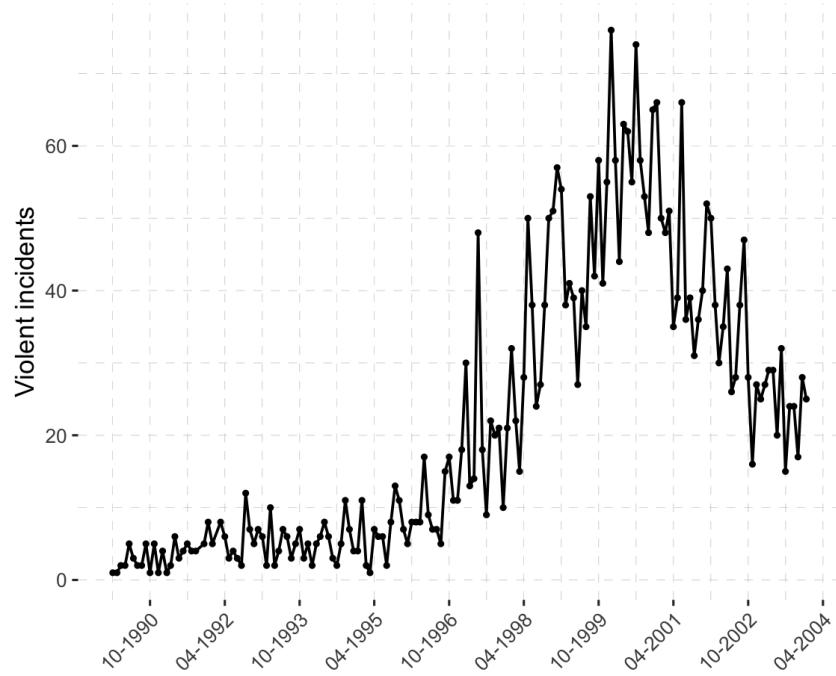
Notes: This figure illustrates our instrumental variable strategy. The black solid line traces out the actual Muria Strait, along which actual 15th-century ethnic Chinese landing sites were established. We obtain our main treatment variable, distance to the Muria Strait, by calculating the distance between the centroid of each 1930 district and the Muria Strait. 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 Strait.

Figure 3: Outcomes for districts with above/below median Chinese population shares in 1930.



Notes: The figure depicts the distribution of the main district-level outcomes. The figures compare the main outcomes for districts with low Chinese shares (below the median) to districts with a high share of the Chinese population (above the median). The horizontal lines depict the median and the first and third quartiles. The blue cross depicts the sample mean. Panel A. depicts the population density in 1930. Panel B. depicts the population density in 2000. Panel C. depicts the average income. Panel D. depicts the night light luminosity in 2010. The scatter plots are jittered.

Figure 4: Number of violent events in Indonesia, 1993-2003.



Notes: The figure depicts the number of violent events by month across Indonesia between 1993 and 2003. See [Varshney et al. \(2008\)](#) for details. Source: UNSFIR.

Table 1: SUMMARY STATISTICS - DISTRICT LEVEL

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Observations	Mean	Standard Deviation	Min	Max	25th Percentile	Median	75th Percentile
Chinese Share	242	0.91	1.21	0.0085	11.9	0.23	0.60	1.09
Native Share	242	98.7	1.93	79.3	100.0	98.7	99.22	99.7
European Share	242	0.25	0.69	0	7.06	0.020	0.09	0.23
Other Asian Share	242	0.092	0.26	0	1.93	0.0027	0.02	0.061
European & Other Asian Share	242	0.34	0.84	0	8.86	0.027	0.12	0.32
Ln(Distance to Muria Strait)	242	11.2	2.85	-10.1	12.8	11.3	11.81	12.1
Sugar Factory Dummy	242	0.18	0.38	0	1	0	0.00	0
Communist Votes (1955–1957)	241	0.26	0.13	0.044	0.50	0.15	0.23	0.37
Ln(Distance to Colonial Railroads)	242	8.31	1.26	2.06	10.8	7.59	8.49	9.26
Ln(Distance to Postal Roads)	242	10.1	1.43	4.11	12.0	9.35	10.51	11.1
Ln(Distance to Opium Boundary)	242	10.8	1.22	6.96	12.2	10.0	11.13	11.8
Ln(Population Density) (1930)	242	5.86	0.59	3.56	7.75	5.46	5.87	6.25
Ln(Population Density) (2000)	242	6.66	0.62	5.20	9.32	6.28	6.62	7.00
Urbanization Rate (2011)	242	26.4	22.7	0	100	8.86	21.85	37.5
Household Consumption (2001–2011)	242	11.6	0.17	11.2	12.2	11.4	11.54	11.7
Ln(Density Growth) (1930–2000)	242	0.79	0.29	-0.32	2.29	0.60	0.78	0.92
Light Intensity (1992)	242	4.15	6.18	0	57.6	0.71	2.58	5.56
Light Intensity (2000)	242	8.50	6.83	0	58.6	5.19	6.59	9.15
Light Intensity (2010)	242	13.9	10.2	0	62.1	7.69	10.38	15.8
Nightlights Growth (1992–2010)	242	9.71	5.75	0	37.6	6.28	8.06	11.4
Altitude	242	577.1	600.7	1	5337	117	500.00	820
Ln(Distance to Coast)	242	6.82	4.66	0	11.2	0	9.48	10.3
Ln(Total District Area) (1930)	242	5.53	0.57	3.56	7.84	5.13	5.52	5.91
Ln(Dry Land Area) (1963)	242	10.5	0.49	8.68	11.7	10.3	10.56	10.7
Ln(Wet Rice Area) (1963)	242	10.4	0.43	9.25	11.2	10.1	10.35	10.6

Table 2: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT

Panel A: OLS without Control

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.191*** (0.026)	0.255*** (0.029)	8.666*** (1.328)	4.409*** (0.553)	0.063*** (0.008)
R-squared	0.205	0.323	0.275	0.337	0.251
Mean (Dep. Var.)	5.865	6.665	26.868	14.104	11.560
Standard Deviation (Dep. Var.)	0.595	0.635	23.371	10.739	0.177

Panel B: Province Fixed Effects

Chinese Share (percent)	0.189*** (0.026)	0.254*** (0.030)	8.719*** (1.299)	4.454*** (0.520)	0.063*** (0.008)
R-squared	0.225	0.334	0.286	0.372	0.254
Mean (Dep. Var.)	5.865	6.665	26.868	14.104	11.560
Standard Deviation (Dep. Var.)	0.595	0.635	23.371	10.739	0.177

Panel C: Geographical Controls & Foreigner Shares

Chinese Share (percent)	0.066* (0.033)	0.100** (0.048)	4.221* (2.426)	1.214 (0.906)	0.010 (0.012)
R-squared	0.567	0.601	0.439	0.557	0.427
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172

Panel D: Urban Center Controls

Chinese Share (percent)	0.085** (0.032)	0.170*** (0.050)	7.234** (3.032)	2.648** (1.118)	0.032** (0.014)
R-squared	0.600	0.684	0.580	0.749	0.595
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Observations	242	242	242	242	242

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.130 (0.093)	0.261*** (0.078)	13.276*** (4.615)	5.569** (2.224)	0.094*** (0.031)
R-squared	0.597	0.674	0.548	0.712	0.535
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	13.753	13.753	13.753	13.753	13.753
Anderson-Rubin, P-value	0.218	0.015	0.042	0.097	0.001

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)
Partial R-squared	0.091	0.091	0.091	0.091	0.091
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: DISTANCE TO THE MURIA STRAIT AND ETHNIC COMPOSITION

<i>Reduced Form</i>	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Chinese Share (1930)	Native Share (1930)	European Share (1930)	Other Asian Share (1930)	European & Oth. Asian Share (1930)	Sugar Factory Presence
Ln(Distance to Muria Strait)	-0.0820*** (0.0261)	0.0865** (0.0327)	-0.00605 (0.00727)	0.00152 (0.00267)	-0.00453 (0.00841)	-0.00928 (0.0147)
R-squared	0.523	0.572	0.524	0.494	0.564	0.475
Mean (Dep. Var.)	0.906	98.749	0.253	0.092	0.345	0.178
Standard Deviation (Dep. Var.)	1.214	1.927	0.690	0.259	0.837	0.383
Fixed Effects	Province	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54	54
Observations	242	242	242	242	242	242

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* include province fixed effects, highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: IV: 1930 CHINESE SHARE, GROWTH, AND LONG-RUN DEVELOPMENT

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	△ Density (1930–2000)	Light Intensity (1992)	Light Intensity (2000)	Light Intensity (2010)	△ Light Intensity (1992–2010)
Chinese Share (percent)	0.131** (0.062)	1.912** (0.750)	2.327** (1.101)	5.569** (2.224)	3.657** (1.584)
R-squared	0.441	0.839	0.824	0.712	0.486
Mean (Dep. Var.)	0.794	4.154	8.504	13.865	9.711
Standard Deviation (Dep. Var.)	0.294	6.181	6.834	10.244	5.749
Kleibergen-Paap WID, F-stat	13.753	13.753	13.753	13.753	13.753
Anderson-Rubin, P-value	0.039	0.076	0.137	0.097	0.119

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)
Partial R-squared	0.091	0.091	0.091	0.091	0.091
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: 1930 CHINESE SHARE AND NIGHTLIGHTS GROWTH

Panel A: IV Specification

Dependent Variable:	(1)	(2)	(3)	(4)
	Light Intensity Growth (1992–1996)	Light Intensity Growth (1996–1997)	Light Intensity Growth (1997–1998)	Light Intensity Growth (1998–2010)
Chinese Share (percent)	0.368** (0.182)	-0.259* (0.146)	-0.280 (0.215)	0.163** (0.082)
R-squared	0.579	0.587	0.646	0.576
Mean (Dep. Var.)	1.194	-0.454	-0.700	0.510
Standard Deviation (Dep. Var.)	0.667	0.867	1.052	0.386
Kleibergen-Paap WID, F-stat	15.921	21.544	19.445	18.574
Anderson-Rubin, P-value	0.139	0.152	0.213	0.118

<i>Panel B: First Stage</i>				
Ln(Distance to Muria Strait)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)	-0.158*** (0.043)
Partial R-squared	0.091	0.091	0.091	0.091
Fixed Effects	Province	Province	Province	Province
Number of Regencies	54	54	54	54
Observations	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: 1930 CHINESE SHARE AND STRUCTURAL TRANSFORMATION

Panel A: IV Specification

Dependent Variable:	(1)	(2)	(3)
	Agriculture Employment Share	Manufacture Employment Share	Services Employment Share
Chinese Share (percent)	-10.315** (4.158)	5.783 (4.706)	4.533 (3.607)
R-squared	0.539	0.377	0.592
Mean (Dep. Var.)	41.655	22.465	35.880
Standard Deviation (Dep. Var.)	17.873	11.432	10.644
Kleibergen-Paap WID, F-stat	3.408	3.408	3.408
Anderson-Rubin, P-value	0.102	0.395	0.022

<i>Panel B: First Stage</i>			
Ln(Distance to Muria Strait)	-0.163* (0.088)	-0.163* (0.088)	-0.163* (0.088)
Partial R-squared	0.097	0.097	0.097
Fixed Effects	Province	Province	Province
Number of Regencies	54	54	54
Observations	230	230	230

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: 1930 CHINESE SHARE AND FIRM OUTCOMES

<i>Panel A: IV Specification</i>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Transformation:	Ln(.)						Arsinh(.)	
Dependent Variable:	Firm Count	Firm Sales	Profits	Estimated Capital	Book Value Capital	Assets	Export Status	Export Share
Chinese Share (percent)	0.099 (0.337)	0.426*** (0.151)	0.232 (0.221)	0.037 (0.173)	-0.494** (0.223)	0.007 (0.186)	-0.087*** (0.032)	-0.093*** (0.034)
Ln(Firm Sales)			1.152*** (0.030)	0.611*** (0.037)	1.241*** (0.057)	0.423*** (0.071)	0.088*** (0.008)	0.111*** (0.009)
Fixed Effects	Province	Province, Sector						
Number of Regencies	54	54	54	54	54	54	54	54
Observations	242	152,345	152,345	152,345	152,345	152,345	152,345	152,345
R-squared	0.405	0.117	0.222	0.087	0.253	0.054	0.264	0.288
Mean (Dep. Var.)	6.056	15.239	12.562	3.727	11.411	5.936	0.236	0.321
Std. Dev. (Dep. Var.)	0.923	2.670	7.083	6.212	6.528	6.890	0.667	0.805
Kleibergen-Paap F-stat	13.753	15.152	15.135	15.135	15.135	15.135	15.135	15.135
Anderson-Rubin p-value	0.781	0.001	0.292	0.835	0.048	0.968	0.000	0.001
<i>Panel B: First Stage</i>								
Ln(Distance to Muria Strait)	-0.158*** (0.043)	-0.191*** (0.049)	-0.192*** (0.049)	-0.192*** (0.049)	-0.192*** (0.049)	-0.192*** (0.049)	-0.192*** (0.049)	-0.192*** (0.049)

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaja) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 9: 1930 CHINESE SHARE AND FIRM INVESTMENTS*Panel A: IV Specification*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable: Investment/Sales	Total	Private Sources	Reinvested Earnings	Stock Market	Domestic Loans	FDI	Government	Foreign Loans
Chinese Share (percent)	14.434* (7.660)	21.049* (11.699)	1.371 (0.958)	-0.293** (0.133)	-4.159 (3.032)	-0.298 (0.239)	-0.020 (0.129)	-0.001 (0.001)
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54	54	54	54	54
Observations	152,345	152,345	152,345	152,345	152,345	152,345	152,345	152,345
R-squared	0.009	0.007	0.002	0.001	0.004	0.001	0.003	0.002
Mean (Dep. Var.)	53.829	57.346	3.189	0.150	6.055	0.072	0.179	0.006
Std. Dev. (Dep. Var.)	1068.391	1362.108	135.678	9.020	200.612	9.392	10.101	0.387
Kleibergen-Paap F-stat	15.135	15.135	15.135	15.135	15.135	15.135	15.135	15.135
Anderson-Rubin p-value	0.039	0.032	0.119	0.010	0.151	0.178	0.874	0.476

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.192*** (0.049)							
-------------------------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------	----------------------

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Supplemental Appendix

(For online publication only)

A.1 Data Sources and Definitions	51
A.2 Robustness empirical strategy using counterfactual landing sites	52
A.3 Robustness empirical strategy using recentered IV	53
A.4 Calculating Market Access at the Regency Level	54
A.5 Additional Results	55

A.1 Data Sources and Definitions

Ethnic composition in 1930: We use ethnic composition data from the 1930 Dutch East Indies Census (*Volkstelling, 1930*) to construct ethnic Chinese shares as our main measure of the historical size of the ethnic Chinese community. The 1930 Census used “social criteria”, broadly defined as language spoken, customs and habits to distinguish between different ethnic groups—the Chinese who had a different religion (“Confucianism in the midst of a sea of Islam”), arranged marriage practices, and “clannish” associations and festivals were clearly set apart ([Van Klinken, 2003](#)). In addition, we digitize the 1930 share of “other foreign Easterners”, 88% of which were Arabs ([Pepinsky, 2016](#)).

Number and Type of Chinese Businesses: We obtain the number and type of Chinese businesses across various economic sectors from a Dutch colonial business registry, the “Directory of Agricultural and Trading companies in Colonial Indonesia” (*Handboek voor Cultuur- en Handelsondernemingen in Nederlandsch-Indië*). This dataset was significantly improved by and described thoroughly in [Van der Eng \(2022\)](#). This directory is *not* a complete listing of all ethnic Chinese business—it lists all businesses that were *voluntarily* incorporated by owners as limited liability companies under company law, which by 1855 applied to firms of ethnic Chinese entrepreneurs in Indonesia ([Liem, 1952](#)). We restrict our analyses to three main variables that are relatively more complete: (i) number of ethnic Chinese firms; (ii) (shareholder) equity; and (iii) economic sectors. Businesses in the Handbook are classified as ethnic Chinese on the basis of their headquarters, name of their company, and, most importantly, whether the names of owners and directors were ethnic Chinese. Shareholder equity refers to funds that shareholders invested in the company at incorporation or at agreed moments after incorporation until full subscription and are based on public announcement of incorporation of venture ([Van der Eng, 2022](#)). Economic sectors are based on names of companies and brief descriptions in their articles of association and, again, must be interpreted with caution give that companies often tried to incorporate as much diversification as possible given prohibitive costs of amending articles of association and hence, many Chinese firms were multipurpose and could actually have straddled multiple sectors.

Historical districts: 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.

Distance to the Muria Strait: First, we trace the locations of landing sites on the Muria Strait, including Demak, Koedoes, Pati, Waroe (Rembang). We exclude Semarang and Binangoen (Lasem) from this calculation, as they are geographically (just) outside the strait. (We have already controlled extensively for the distance to Semarang, as it is the capital of Central Java.) The measure of distance to the Muria Strait is calculated as the smallest geodesic distance from each district (or village)’s centroid to the set of those 4 points.

In order to replicate the Muria Strait into counterfactual locations, we calculate the length of the Strait by pinning down the centroids of the Muria Strait districts, based on modern-day GIS boundaries and historical sources. We then calculate the geodesic distance passing through each and every district centroid, starting from Semarang to Binangoen to obtain the total length of the Muria Straits.

Sectoral Employment Shares: We construct measures of sectoral 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.

Socio-Economic Surveys: We are interested in estimating the impact of Chinese presence, be it through business establishments, credit, loans, or other mechanisms, 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 and 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.

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 (minimum 20 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*).

Other Data We also include a variety of variables capturing differences in the intensity of colonial extraction across districts. To this end, we include geocoded information on the location of sugar mills from [Dell and Olken \(2020\)](#). We use data on the location of opium concessions granted to the Chinese by district from [Kuijpers \(2022\)](#). In both these cases we calculate the distance of each district to these locations. To this end, we also include information on the number of Chinese officers by district using information from [Cribb \(1997\)](#). To account for historical market access and infrastructure, we include information on the historical railroad as well as the great postal road. Lastly, we include a range of information on location fundamentals such as elevation, agricultural potential, and distance to the coastline. For measures of agricultural productivity we use data from [Bazzi et al. \(2020a\)](#). As a robustness check, we also include data on caloric potential from [Galor and Özak \(2015\)](#) and [Galor and Özak \(2016\)](#). Data on communist vote shares in the 1950s legislative elections are from [Bazzi et al. \(2020b\)](#). Data on social conflict between 1993 and 2003 is from UNSFIR ([Varshney et al., 2008](#)).

A.2 Robustness empirical strategy using counterfactual landing sites

Equation 2 compares differences in economic development across locations with differing sizes of Chinese communities, where the Chinese share is instrumented by the distance to Zheng He's landing sites in the Muria Strait. A different approach to take advantage of the arbitrary landing sites of Zheng He's expeditions is to consider a set of counterfactual landing sites that could have been picked by the expeditions, and then compare the hypothetical consequences of those landing sites on the construction of the IV with the actual IV constructed from the distance to historical landing sites.

For this purpose, we choose the natural counterfactual landing site of Cirebon, right at the limit between Central Java and West Java. Cirebon is the key deep-water port in Central and East Java, and has also attracted a lot of Chinese immigrants by 1930, as shown in Figure 1. However, it was unlikely a landing site by Zheng He's expeditions (see [Salmon and Lombard \(1981\)](#) on the "Chinese chronicles of Cirebon and Samarang").

Based on Cirebon, we recreate four hypothetical landing sites on the Northern coastline to the East of Cirebon, in such a way that their coastal distances to Cirebon replicate respectively the distances from the actual landing sites of Demak, Koedoes, Pati, and Waroe (Rembang) to the city of Semarang. In other words, we translate the landing sites of Semarang, Demak, Koedoes, Pati, and Waroe along the Northern coastline to the West so that Semarang is brought onto Cirebon, and then we use the translated locations of the other four landing sites as the counterfactual landing sites. We then calculate the counterfactual geodesic distance from each district in our sample to those counterfactual landing sites.

Next, we make use of those counterfactual distances in stratified comparisons with the actual distances to the historical landing sites in the Muria Strait. To do that, we stratify each set of historical and counterfactual distances into ten strata each, and then match the i -th stratum of historical distances with the i -th stratum of counterfactual distances, $i \in \{1, \dots, 10\}$. Last, we run the main specifications in equations 1 and 2 in this matched sample, including the stratum fixed effects. This means we are comparing locations where the historical distances are very close to the counterfactual distances, thus focusing only on the variation that Zheng He's crews stopped in the Muria Strait but not at the counterfactual locations near Cirebon.

The OLS results reported in Table A.6 are much comparable to the main OLS results from Table 2. Similarly, the IV results reported in Table A.7 closely trace the main IV results from Table 3. They thus lend more credibility to the main interpretation of the long-run effects of the Chinese communities.

A.3 Robustness empirical strategy using recentered IV

Equation 2 compares differences in economic development across locations with differing sizes of Chinese communities, where the Chinese share is instrumented by the distance to Zheng He's landing sites in the Muria Strait. Even if the location of those landing sites and initial settlements is randomly assigned, those sites are still constrained to be on a specific segment of the Northern coastline of Java, and that restriction would carry its limits over the calculation of distances from all districts in our sample to the landing site. Consequently, districts that are closer to the Northern coast tend to have a shorter distance to the Muria Strait, and using this variation in our empirical analysis may entail confounding factors that could have produced the main results without a real impact of the Chinese communities.

This concern has been treated in a generalized setting in Borusyak and Hull (2023)'s examination of instruments as nonlinear functions of an arbitrary treatment. We follow Borusyak and Hull (2023)'s method by calculating the *expected* treatment of each district, denoted $\mathbb{E}[\ln(D_{pd})]$ where the expectation is taken over alternative treatment assignment procedures over all the potential initial settlement sites, and then extracting it from the instrument to compute the recentered IV as $D_{pd}^r = D_{pd} - \mathbb{E}[D_{pd}]$.

To address this concern, we follow Borusyak and Hull (2023) and calculate the *expected* treatment of each district that we denote $\mathbb{E}[\ln(D_{pd})]$ where the expectation is taken over alternative treatment assignment procedures over all the potential initial landing sites. Specifically, we randomly draw each potential set of initial landing sites by first picking randomly a main counterfactual landing point P on the Northern coastline that spans from Cirebon to Surabaya. We choose this range, as it covers the historical passage of Zheng He's expeditions to Java, throughout Central Java all the way to Surabaya, where the Brantas river that runs through the Majapahit capital in Trowulan pours into the sea at Surabaya. For each main counterfactual landing point P , we translate the set of historical landing sites along the Northern coastline so that Semarang is translated onto P , and the other four hypothetical landing points would be at the same coastal distances from P as the actual distances from Demak, Koedoes, Pati, and Waroe (Rembang) to the city of Semarang. Those four translated locations of the four landing sites will be used as the counterfactual landing sites. We then calculate the counterfactual geodesic distance from each district in our sample to those counterfactual landing sites. Overall, in this exercise we replicate the same procedure as we did for Cirebon in section 5.2.1, but this time for each random draw of a main landing point on this particular segment on the Northern coastline.

By repeating this procedure for a large number of draws, we can thus calculate the expected distance to any counterfactual landing strip $\mathbb{E}[\ln(D_{pd})]$, from which we calculate the new recentered IV D_{pd}^r . This IV is then applied in the same specification as in equation 2.

The results are shown in Table A.8. They are slightly smaller but still quantitatively close to the results in Table 3. The first stage's strength is still reasonably good, with a Kleibergen-Paap weak-instrument-robust F-statistic of 12.59. This robustness exercise thus lends additional credibility to our main interpretations of the long-run effects of the Chinese communities.

A.4 Calculating Market Access at the Regency Level

In this section, we elaborate on the construction of the market access measure. We start by constructing the bilateral travel times between all regencies using information from IRMS on the Indonesian highway network. The network is depicted in Figure A.1. Following Baum-Snow et al. (2020), we assume that the average travel speed on and off a highway is 95 kph and 25 kph respectively. We furthermore assume that routes between locations are chosen to minimize travel time. Again following Baum-Snow et al. (2020), we assume trade costs between locations i and j are given by

$$\tau_{ij} = 1 + 0.004T_{it}^{0.8}, \quad (3)$$

where T_{it} denote the travel time between locations i and j . The market access of a location can then be expressed by the following recursive formula,

$$MA_i = \sum_{j=1}^R \tau_{ij}^{1-\sigma} \frac{Y_j}{MA_j}, \quad (4)$$

where $1 - \sigma$ is the trade elasticity, R the number of regencies, and Y_j the total production/expenditure in



Figure A.1 : This figure depicts the Indonesian highway network used to construct the trade cost matrix.

region j (see e.g. Allen and Arkolakis, 2023). We set $1 - \sigma = 4$ as a baseline case. From the formula, one can see that the market access of region i is higher when the income of surrounding regions is higher, the region has a higher geographical centrality as captured by the trade costs, or the market access of surrounding regencies is low (which is the case when these regions have fewer alternative trading partners). We then use data on bilateral trade costs for τ_{ij} and data on employment times average wages to calculate Y_j for each regency. We then solve the system of equations for MA_i . The solution to this equation is the market access measure used in Section 7 of the paper. Due to data limitations, we can only calculate the market access at the regency level.

A.5 Additional Results

**Table A.1 : OLS: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
ALTERNATIVE STANDARD ERROR CORRECTION**

<i>Residency Clustering</i>		(1)	(2)	(3)	(4)	(5)
Dependent Variable:		Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)		0.085** (0.036)	0.170** (0.060)	7.234* (3.563)	2.648** (0.917)	0.032** (0.014)
R-squared		0.600	0.684	0.580	0.749	0.595
Mean (Dep. Var.)		5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)		0.585	0.616	22.725	10.244	0.172
Number of Residencies		11	11	11	11	11
Observations		242	242	242	242	242

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* include province fixed effects, highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area, quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by residency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.2 : OLS: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
ADDITIONAL CONTROLS**

Panel A: Sugar Factory Control

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.068 (0.041)	0.155*** (0.056)	6.496* (3.346)	2.487** (1.169)	0.031** (0.015)
Sugar Factory Presence	0.436*** (0.151)	0.377*** (0.101)	18.426*** (4.293)	4.024** (1.638)	0.037 (0.027)
R-squared	0.642	0.713	0.631	0.761	0.598
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Observations	242	242	242	242	242

Panel B: Communist Vote Control

Chinese Share (percent)	0.084839*** (0.031452)	0.169248*** (0.049393)	6.882602** (2.988075)	2.629229** (1.112811)	0.030091** (0.014022)
Communist Votes (1955/1957)	0.072566 (0.360839)	0.183189 (0.355282)	28.898123* (15.956491)	5.904238 (5.603295)	0.270002** (0.119964)
R-squared	0.585	0.659	0.575	0.731	0.600
Mean (Dep. Var.)	5.854	6.644	26.121	13.665	11.558
Standard Deviation (Dep. Var.)	0.576	0.593	22.270	9.779	0.168
Observations	241	241	241	241	241

Panel C: Colonial Railroad Control

Chinese Share (percent)	0.034824 (0.037027)	0.127875*** (0.042127)	6.007648** (2.862311)	2.404521** (1.136112)	0.029891* (0.015254)
Ln(Distance to Colonial Railroads) (1925)	-0.115650** (0.046823)	-0.097384*** (0.028369)	-2.811626** (1.248899)	-0.559210 (0.411333)	-0.005444 (0.007090)
R-squared	0.643	0.711	0.597	0.752	0.596
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Observations	242	242	242	242	242

Panel D: Great Postal Road Control

Chinese Share (percent)	0.076111** (0.037344)	0.166940*** (0.050791)	6.909520** (3.111516)	2.586005** (1.125832)	0.032287** (0.014406)
Ln(Distance to Great Postal Roads)	-0.100003 (0.082606)	-0.037181 (0.036899)	-3.545676** (1.402533)	-0.682141 (0.566976)	0.000246 (0.011041)
R-squared	0.615	0.686	0.592	0.751	0.595
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Observations	242	242	242	242	242

Panel E: Opium Boundary Control

Chinese Share (percent)	0.086747*** (0.030177)	0.168244*** (0.049441)	6.968560** (3.062680)	2.599629** (1.123733)	0.030465** (0.014618)
Ln(Distance to Opium Boundaries) (1886)	-0.027579 (0.060961)	0.038659 (0.043845)	4.892794*** (1.727404)	0.899488 (0.726411)	0.033213** (0.013210)
R-squared	0.601	0.686	0.598	0.752	0.609
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Observations	242	242	242	242	242

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* include province fixed effects, highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area, quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaija) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.3 : OLS: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
INCLUDING WEST JAVA**

<i>Panel A: OLS without Control</i>		(1)	(2)	(3)	(4)	(5)
Dependent Variable:		Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)		0.133*** (0.035)	0.209*** (0.036)	6.236*** (1.347)	3.633*** (0.674)	0.060*** (0.007)
R-squared		0.163	0.334	0.261	0.348	0.265
Mean (Dep. Var.)		5.782	6.743	28.650	15.372	11.639
Standard Deviation (Dep. Var.)		0.684	0.746	25.162	12.690	0.239

<i>Panel B: Province Fixed Effects</i>					
	(1)	(2)	(3)	(4)	(5)
Dependent Variable:		Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)
Chinese Share (percent)		0.146*** (0.033)	0.205*** (0.036)	6.193*** (1.370)	3.588*** (0.692)
R-squared		0.237	0.346	0.268	0.370
Mean (Dep. Var.)		5.782	6.743	28.650	15.372
Standard Deviation (Dep. Var.)		0.684	0.746	25.162	12.690

<i>Panel C: Geographical Controls & Foreign Shares</i>					
	(1)	(2)	(3)	(4)	(5)
Dependent Variable:		Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)
Chinese Share (percent)		0.063*** (0.017)	0.099*** (0.029)	3.959*** (1.418)	1.888** (0.762)
R-squared		0.660	0.613	0.427	0.529
Mean (Dep. Var.)		5.778	6.722	28.079	14.956
Standard Deviation (Dep. Var.)		0.679	0.716	24.409	11.936

<i>Panel D: Urban Center Controls</i>					
	(1)	(2)	(3)	(4)	(5)
Dependent Variable:		Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)
Chinese Share (percent)		0.045** (0.019)	0.069** (0.026)	3.579*** (1.311)	1.120* (0.609)
R-squared		0.685	0.677	0.520	0.693
Mean (Dep. Var.)		5.778	6.722	28.079	14.956
Standard Deviation (Dep. Var.)		0.679	0.716	24.409	11.936
Observations		340	340	340	340

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* include province fixed effects, highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area, quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.4 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
INCLUDING WEST JAVA**

IV Specification

	(1)	(2)	(3)
Dependent Variable:	Ln(Consumption) (2000–2010)	Ln(Pop. Density) (2000)	Δ Light Intensity (1992–2010)
Chinese Share (percent)	0.058** (0.028)	0.116 (0.097)	3.062* (1.651)
R-squared	0.930	0.651	0.338
Mean (Dep. Var.)	11.66	-7.09	9.76
Kleibergen-Paap WID, F-stat	7.46	7.01	8.41
Anderson-Rubin, P-value	0.058	0.307	0.198
Fixed Effects	Province	Province	Province
Observations	47,872	344	344

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, the natural logarithm of total area (1930), total dry agricultural land (1963), total wet agricultural land (1963), longitude, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). Observations are at the village-level for column (1) and district-level for columns (2) and (3). Standard errors are clustered by regency.
 * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5 : STRUCTURAL TRANSFORMATION: INCLUDING WEST JAVA

IV Specification

	(1)	(2)	(3)
Dependent Variable:	Agriculture Employment Share	Manufacturing Employment Share	Service Employment Share
Chinese Share (percent)	-7.074** (2.972)	4.129 (4.847)	2.945 (3.754)
R-squared	0.234	0.110	0.230
Mean (Dep. Var.)	39.65	22.35	38.01
Kleibergen-Paap WD, F-stat	1.50	1.50	1.50
Anderson-Rubin, P-value	0.354	0.596	0.301
Fixed Effects	Province	Province	Province
Observations	19,706	19,706	19,706

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, the natural logarithm of total area (1930), total dry agricultural land (1963), total wet agricultural land (1963), longitude, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). Observations are at the village-level for column (1) and district-level for columns (2) and (3). Standard errors are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.6 : OLS: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
COUNTERFACTUAL LANDING SITES AROUND CIREBON**

OLS Specification

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.0895** (0.0348)	0.166*** (0.0512)	6.532** (2.991)	2.558** (1.143)	0.0323** (0.0149)
R-squared	0.619	0.707	0.618	0.771	0.656
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Observations	242	242	242	242	242

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.7 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
COUNTERFACTUAL LANDING SITES AROUND CIREBON**

Panel A: IV Specification

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	Ln(Pop. Density) (1930)	Ln (Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.150** (0.0720)	0.203*** (0.0594)	8.867*** (2.922)	5.154** (2.028)	0.0706*** (0.0261)
R-squared	0.615	0.705	0.613	0.743	0.634
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	9.033	9.033	9.033	9.033	9.033
Anderson-Rubin, P-value	0.145	0.050	0.082	0.142	0.003

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.152*** (0.0506)	-0.152*** (0.0506)	-0.152*** (0.0506)	-0.152*** (0.0506)	-0.152*** (0.0506)
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.8 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
RECENTERED IV**

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.0805 (0.0760)	0.208*** (0.0681)	9.951** (4.238)	4.632** (2.230)	0.0802*** (0.0274)
R-squared	0.600	0.682	0.574	0.732	0.559
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	12.585	12.585	12.585	12.585	12.585
Anderson-Rubin, P-value	0.349	0.024	0.078	0.145	0.003

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0865*** (0.0244)	-0.0865*** (0.0244)	-0.0865*** (0.0244)	-0.0865*** (0.0244)	-0.0865*** (0.0244)
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. **p* < 0.1, ***p* < 0.05, ****p* < 0.01.

**Table A.9 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
SUGAR FACTORY CONTROL**

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.0918 (0.0606)	0.248*** (0.0852)	12.18*** (3.556)	5.572*** (2.052)	0.0960*** (0.0301)
Sugar Factory presence	0.430*** (0.137)	0.354*** (0.110)	17.06*** (5.148)	3.284 (2.016)	0.0209 (0.0362)
R-squared	0.642	0.702	0.602	0.719	0.533
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	14.479	14.479	14.479	14.479	14.479
Anderson-Rubin, P-value	0.174	0.014	0.024	0.085	0.002

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0759*** (0.0200)	-0.0759*** (0.0200)	-0.0759*** (0.0200)	-0.0759*** (0.0200)	-0.0759*** (0.0200)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. **p* < 0.1, ***p* < 0.05, ****p* < 0.01.

**Table A.10 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
COMMUNIST VOTE CONTROL**

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.150** (0.0761)	0.305*** (0.0772)	14.38*** (4.413)	6.542*** (2.194)	0.110*** (0.0297)
Communist Votes (1955/1957)	0.0304 (0.363)	0.0951 (0.335)	24.03 (15.38)	3.364 (5.552)	0.218* (0.123)
R-squared	0.579	0.635	0.523	0.657	0.496
Mean (Dep. Var.)	5.854	6.644	26.121	13.665	11.556
Standard Deviation (Dep. Var.)	0.576	0.593	22.270	9.779	0.168
Kleibergen-Paap WID, F-stat	13.998	13.998	13.998	13.998	13.998
Anderson-Rubin, P-value	0.123	0.007	0.028	0.068	0.000

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0775*** (0.0207)	-0.0775*** (0.0207)	-0.0775*** (0.0207)	-0.0775*** (0.0207)	-0.0775*** (0.0207)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	241	241	241	241	241

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. **p* < 0.1, ***p* < 0.05, ****p* < 0.01.

**Table A.11 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
COLONIAL RAILROAD CONTROL**

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.0939 (0.0860)	0.256*** (0.0817)	13.53*** (5.121)	5.995** (2.503)	0.102*** (0.0321)
Ln(Distance to Colonial Railroads) (1925)	-0.106*** (0.0403)	-0.0757*** (0.0282)	-1.535 (1.412)	0.0501 (0.557)	0.00676 (0.00846)
R-squared	0.639	0.693	0.550	0.700	0.522
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	13.515	13.515	13.515	13.515	13.515
Anderson-Rubin, P-value	0.334	0.019	0.061	0.111	0.004

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0719*** (0.0196)	-0.0719*** (0.0196)	-0.0719*** (0.0196)	-0.0719*** (0.0196)	-0.0719*** (0.0196)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. **p* < 0.1, ***p* < 0.05, ****p* < 0.01.

**Table A.12 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
GREAT POSTAL ROAD CONTROL**

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.00371 (0.127)	0.252*** (0.0813)	9.909** (4.536)	5.384** (2.253)	0.107*** (0.0364)
Ln(Distance to Great Postal Roads)	-0.106 (0.0812)	-0.0299 (0.0397)	-3.291** (1.341)	-0.444 (0.643)	0.00663 (0.0118)
R-squared	0.607	0.677	0.584	0.717	0.508
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	9.150	9.150	9.150	9.150	9.150
Anderson-Rubin, P-value	0.977	0.033	0.139	0.149	0.003

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0788*** (0.0260)	-0.0788*** (0.0260)	-0.0788*** (0.0260)	-0.0788*** (0.0260)	-0.0788*** (0.0260)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.13 : IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT:
OPIUM BOUNDARY CONTROL**

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.147* (0.0853)	0.289*** (0.0825)	13.89*** (5.018)	5.914** (2.391)	0.0962*** (0.0270)
Ln(Distance to Opium Boundaries) (1886)	-0.0314 (0.0613)	0.0309 (0.0446)	4.446** (1.761)	0.686 (0.801)	0.0290** (0.0134)
R-squared	0.596	0.668	0.556	0.704	0.542
Mean (Dep. Var.)	5.861	6.655	26.426	13.865	11.558
Standard Deviation (Dep. Var.)	0.585	0.616	22.725	10.244	0.172
Kleibergen-Paap WID, F-stat	13.211	13.211	13.211	13.211	13.211
Anderson-Rubin, P-value	0.162	0.017	0.062	0.109	0.002

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0773*** (0.0213)	-0.0773*** (0.0213)	-0.0773*** (0.0213)	-0.0773*** (0.0213)	-0.0773*** (0.0213)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14 : 1930 CHINESE SHARE AND NIGHTLIGHTS GROWTH - VILLAGE
IV Specification

Dependent Variable:	(1)	(2)	(3)	(4)
	Light Intensity Growth (1992–1996)	Light Intensity Growth (1996–1997)	Light Intensity Growth (1997–1998)	Light Intensity Growth (1998–2010)
Chinese Share (percent)	0.351** (0.171)	-0.068 (0.142)	-0.316* (0.177)	0.181* (0.094)
R-squared	0.231	0.390	0.372	0.359
Mean (Dep. Var.)	1.253	-0.525	-0.774	0.550
Standard Deviation (Dep. Var.)	1.141	1.621	1.793	0.552
Kleibergen-Paap WID, F-stat	17.430	19.014	18.728	18.172
Anderson-Rubin, P-value	0.144	0.634	0.137	0.147

Notes: The table reports IV estimates. The unit of analysis is at the village level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15 : 1930 CHINESE SHARE AND STRUCTURAL TRANSFORMATION - VILLAGE

Panel A: IV Specification

Dependent Variable:	(1)	(2)	(3)
	Agriculture Employment Share	Manufacture Employment Share	Services Employment Share
Chinese Share (percent)	-9.762** (4.021)	5.554 (4.662)	4.208 (3.574)
R-squared	0.561	0.414	0.600
Mean (Dep. Var.)	39.887	22.909	37.204
Standard Deviation (Dep. Var.)	17.887	11.530	10.468
Kleibergen-Paap WID, F-stat	4.125	4.125	4.125
Anderson-Rubin, P-value	0.118	0.413	0.041

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.159** (0.078)	-0.159** (0.078)	-0.159** (0.078)
Partial R-squared	0.096	0.096	0.096
Fixed Effects	Province	Province	Province
Number of Regencies	54	54	54
Observations	12,778	12,778	12,778

Notes: The table reports IV estimates. The unit of analysis is at the village level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16 : FIRM SALES AND 1930 CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Sales) (1980–1996)	Ln(Sales) (1997–1999)	Ln(Sales) (2000–2013)	Ln(Sales) All Years	Ln(Sales) All Years
Chinese Share (percent)	0.220* (0.118)	0.119 (0.0889)	0.187* (0.0978)	0.212* (0.109)	0.130 (0.0896)
Chinese Share (percent) # Period 2				-0.0934** (0.0412)	
Chinese Share (percent) # Period 1					0.0861** (0.0380)
Chinese Share (percent) # Period 3					0.0569 (0.0412)
R-squared	0.243	0.204	0.219	0.407	0.414
Mean (Dep. Var.)	12.228	13.351	14.676	13.614	13.614
Standard Deviation (Dep. Var.)	2.021	2.005	2.012	2.316	2.316
Kleibergen-Paap WID, F-stat	15.878	15.476	14.204	7.495	5.264
Anderson-Rubin, P-value	0.008	0.089	0.015	0.019	0.020

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.115*** (0.0289)	-0.121*** (0.0307)	-0.113*** (0.0301)	-0.114*** (0.0295)	-0.120*** (0.0302)
Ln(Distance to Muria Strait) # Period 2				-0.00653*** (0.00251)	
Ln(Distance to Muria Strait) # Period 1					0.00536* (0.00274)
Ln(Distance to Muria Strait) # Period 3					0.00712*** (0.00190)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year				
Number of Regencies	54	54	54	54	54
Observations	460,422	142,375	657,209	1,260,006	1,260,006

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17 : 1930 CHINESE SHARE AND CONFLICT FREQUENCY

Panel A: IV Specification

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	UNSFIR Data				ACLED Data
	Conflicts Before 1998	Conflicts 1998	Conflicts After 1998	Anti Chinese Conflicts (1990–2003)	
Chinese Share (percent)	1.534** (0.601)	4.338 (2.889)	4.416 (3.094)	-0.0408 (0.0345)	37.39** (15.97)
R-squared	0.694	-1.611	0.097	0.473	0.801
Mean (Dep. Var.)	0.868	0.636	2.364	0.054	32.814
Standard Deviation (Dep. Var.)	3.588	2.014	5.739	0.290	111.331
Kleibergen-Paap WID, F-stat	13.736	13.736	13.736	13.736	13.736
Anderson-Rubin, P-value	0.019	0.027	0.052	0.235	0.018

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0777*** (0.0210)	-0.0777*** (0.0210)	-0.0777*** (0.0210)	-0.0777*** (0.0210)	-0.0777*** (0.0210)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.18 : FIRM SALES GROWTH AND 1930 CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Sales Growth (1980–1996)	Sales Growth (1997–1999)	Sales Growth (2000–2013)	Sales Growth All Years	Sales Growth All Years
Chinese Share (percent)	0.00971** (0.00465)	0.00430 (0.00616)	0.0136* (0.00711)	0.0127** (0.00591)	0.00273 (0.00613)
Chinese Share (percent) # Period 2				-0.00838 (0.00670)	
Chinese Share (percent) # Period 1					0.00712 (0.00676)
Chinese Share (percent) # Period 3					0.0110 (0.00738)
R-squared	0.008	0.030	0.008	0.010	0.011
Mean (Dep. Var.)	0.117	0.185	0.129	0.131	0.131
Standard Deviation (Dep. Var.)	0.676	0.679	0.895	0.800	0.800
Kleibergen-Paap WID, F-stat	15.789	15.797	13.960	7.383	5.263
Anderson-Rubin, P-value	0.009	0.434	0.023	0.016	0.020

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.115*** (0.0290)	-0.120*** (0.0303)	-0.113*** (0.0303)	-0.114*** (0.0297)	-0.120*** (0.0302)
Ln(Distance to Muria Strait) # Period 2				-0.00640** (0.00272)	
Ln(Distance to Muria Strait) # Period 1					0.00509 (0.00342)
Ln(Distance to Muria Strait) # Period 3					0.00715*** (0.00249)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year				
Number of Regencies	54	54	54	54	54
Observations	377,390	132,093	596,559	1,106,042	1,106,042

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include natural logarithm distance and squared natural logarithm of distance to major colonial cities (Batavia, Semarang, Soerabaia) and ports (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.19 : FIRM PRIVATE INVESTMENT/SALES AND CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable: Investment/Sales	Private Investment (1980–1996)	Private Investment (1997–1999)	Private Investment (2000–2013)	Private Investment All Years	Private Investment All Years
Chinese Share (percent)	0.352** (0.141)	-0.00997 (0.00748)	-0.00315 (0.00666)	0.237** (0.093)	-0.00997 (0.01961)
Chinese Share (percent) # Period 2				-0.247** (0.098)	
Chinese Share (percent) # Period 1					0.362*** (0.130)
Chinese Share (percent) # Period 3					0.00682 (0.02493)
Kleibergen-Paap WID, F-stat	14.189			7.657	5.231
Anderson-Rubin, P-value	0.661	0.783			

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.117*** (0.0295)	-0.121*** (0.0307)	-0.116*** (0.0310)	-0.117*** (0.0298)	-0.121*** (0.0307)
Ln(Distance to Muria Strait) # Period 2				-0.00377 (0.00240)	
Ln(Distance to Muria Strait) # Period 1					0.00362 (0.00305)
Ln(Distance to Muria Strait) # Period 3					0.00408** (0.00167)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year				
Number of Regencies	52	52	52	52	52
Observations	447,457	142,690	152,656	742,803	742,803

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.20 : FIRM LABOR AND CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Employment) (1980–1996)	Ln(Employment) (1997–1999)	Ln(Employment) (2000–2013)	All Years	All Years
Chinese Share (percent)	0.124* (0.0719)	0.0384 (0.0417)	0.0413 (0.0385)	0.106* (0.0640)	0.0438 (0.0423)
Chinese Share (percent) # Period 2				-0.0676* (0.0350)	
Chinese Share (percent) # Period 1					0.0785** (0.0378)
Chinese Share (percent) # Period 3					0.00816 (0.0132)
R-squared	0.097	0.107	0.103	0.098	0.100
Mean (Dep. Var.)	3.966	3.966	3.927	3.960	3.960
Standard Deviation (Dep. Var.)	1.141	1.127	1.089	1.131	1.131
Kleibergen-Paap WID, F-stat	15.784	15.451	14.211	7.770	5.231
Anderson-Rubin, P-value	0.013	0.292	0.201	0.020	0.026

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.117*** (0.0295)	-0.121*** (0.0307)	-0.118*** (0.0312)	-0.117*** (0.0297)	-0.120*** (0.0302)
Ln(Distance to Muria Strait) # Period 2				-0.00343 (0.00259)	
Ln(Distance to Muria Strait) # Period 1					0.00345 (0.00240)
Ln(Distance to Muria Strait) # Period 3					0.00281 (0.00177)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year				
Number of Regencies	54	54	54	54	54
Observations	462,530	142,645	108,610	713,785	713,785

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.21 : FIRM WAGE BILL AND CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Wages) (1980–1996)	Ln(Wage Bill) (1997–1999)	Ln(Wage Bill) (2000–2013)	Ln(Wage Bill) All Years	Ln(Wage Bill) All Years
Chinese Share (percent)	0.265* (0.137)	0.0792 (0.0712)	0.135* (0.0744)	0.196* (0.101)	0.0924 (0.0737)
Chinese Share (percent) # Period 2				-0.117** (0.0497)	
Chinese Share (percent) # Period 1					0.173** (0.0741)
Chinese Share (percent) # Period 3					0.0427 (0.0480)
R-squared	0.278	0.366	0.354	0.505	0.512
Mean (Dep. Var.)	10.280	9.736	12.495	11.367	11.367
Standard Deviation (Dep. Var.)	1.778	2.215	1.982	2.264	2.264
Kleibergen-Paap WID, F-stat	15.804	15.363	14.162	7.457	4.720
Anderson-Rubin, P-value	0.005	0.177	0.043	0.025	0.008

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.117*** (0.0294)	-0.122*** (0.0311)	-0.113*** (0.0301)	-0.115*** (0.0297)	-0.122*** (0.0306)
Ln(Distance to Muria Strait) # Period 2				-0.00702*** (0.00258)	
Ln(Distance to Muria Strait) # Period 1					0.00465* (0.00247)
Ln(Distance to Muria Strait) # Period 3					0.00818*** (0.00215)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year				
Number of Regencies	54	54	54	54	54
Observations	462,801	131,849	653,750	1,248,400	1,248,400

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.22 : FIRM AVERAGE WAGE AND CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Avg. Wage) (1980–1996)	Ln(Avg. Wage) (1997–1999)	Ln(Avg. Wage) (2000–2013)	Ln(Avg. Wage) All Years	Ln(Avg. Wage) All Years
Chinese Share (percent)	0.141** (0.0674)	0.0581 (0.0400)	0.0492* (0.0297)	0.122** (0.0605)	0.0614 (0.0401)
Chinese Share (percent) # Period 2				-0.0638** (0.0274)	
Chinese Share (percent) # Period 1					0.0803** (0.0346)
Chinese Share (percent) # Period 3					-0.0119 (0.0206)
R-squared	0.540	0.544	0.757	0.645	0.648
Mean (Dep. Var.)	6.313	5.741	7.317	6.352	6.352
Standard Deviation (Dep. Var.)	1.075	1.674	2.068	1.455	1.455
Kleibergen-Paap WID, F-stat	15.807	15.358	14.020	7.769	5.239
Anderson-Rubin, P-value	0.003	0.068	0.053	0.007	0.013
<hr/>					
<i>Panel B: First Stage</i>					
Ln(Distance to Muria Strait)	-0.117*** (0.0294)	-0.122*** (0.0311)	-0.118*** (0.0315)	-0.117*** (0.0297)	-0.122*** (0.0306)
Ln(Distance to Muria Strait) # Period 2				-0.00468* (0.00278)	
Ln(Distance to Muria Strait) # Period 1					0.00463* (0.00246)
Ln(Distance to Muria Strait) # Period 3					0.00367* (0.00194)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year	Province, Sector, Year	Province, Sector, Year	Province, Sector, Year	Province, Sector, Year
Number of Regencies	54	54	54	54	54
Observations	462,382	131,820	103,624	697,826	697,826

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.23 : FIRM LABOR COUNT PRODUCTIVITY AND CHINESE SHARE

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(L Prod.) (1980-1996)	Ln(L Prod.) (1997-1999)	Ln(L Prod.) (2000-2013)	Ln(L Prod.) All Years	Ln(L Prod.) All Years
Chinese Share (percent)	0.122** (0.0610)	0.0796 (0.0516)	0.0864** (0.0430)	0.116** (0.0578)	0.0852* (0.0506)
Chinese Share (percent) # Period 2				-0.0366** (0.0180)	
Chinese Share (percent) # Period 1					0.0347* (0.0194)
Chinese Share (percent) # Period 3					0.00291 (0.0196)
R-squared	0.344	0.257	0.271	0.473	0.475
Mean (Dep. Var.)	8.271	9.386	10.237	8.782	8.782
Standard Deviation (Dep. Var.)	1.351	1.351	1.354	1.544	1.544
Kleibergen-Paap WID, F-stat	15.880	15.469	14.239	7.812	5.264
Anderson-Rubin, P-value	0.005	0.039	0.004	0.009	0.011

<i>Panel B: First Stage</i>					
Ln(Distance to Muria Strait)	-0.115*** (0.0289)	-0.121*** (0.0307)	-0.118*** (0.0312)	-0.116*** (0.0293)	-0.120*** (0.0302)
Ln(Distance to Muria Strait) # Period 2				-0.00486* (0.00290)	
Ln(Distance to Muria Strait) # Period 1					0.00533* (0.00274)
Ln(Distance to Muria Strait) # Period 3					0.00281 (0.00176)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector, Year				
Number of Regencies	54	54	54	54	54
Observations	459,993	142,330	108,360	710,683	7106,83

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.24 : 1930 CHINESE SHARE AND FINANCIAL ACCESS

<i>Panel A: IV Specification</i>										
Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Chinese Share (percent)	-0.022 (0.024)	-0.090 (0.337)	-0.025* (0.013)	0.016 (0.016)	-0.003 (0.004)	0.007 (0.007)	0.026 (0.016)	0.003 (0.013)	0.006 (0.020)	-0.043 (0.026)
R-squared	0.181	0.335	0.307	0.231	0.413	0.464	0.364	0.475	0.503	0.412
Mean (Dep. Var.)	0.045	4.912	0.126	0.155	0.030	0.045	0.131	0.115	0.645	0.539
Standard Deviation (Dep. Var.)	0.100	2.332	0.050	0.067	0.016	0.021	0.061	0.061	0.096	0.104
Kleibergen-Paap WID, F-stat	13.753	13.753	13.753	13.753	13.753	13.753	13.753	13.753	13.753	13.753
Anderson-Rubin, P-value	0.349	0.796	0.049	0.314	0.454	0.210	0.026	0.786	0.778	0.065

<i>Panel B: First Stage</i>									
Ln(Distance to Muria Strait)	-0.158*** (0.043)								
Partial R-squared	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091
Fixed Effects	Province, Sector								
Number of Regencies	54	54	54	54	54	54	54	54	54
Observations	242	242	242	242	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaja) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.25 : HISTORICAL FIRM PERFORMANCE

Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Firm Count	Ln(Deflated Equity)	Ln(Average Equity)	Chinese Firm Count	Chinese Ln(Deflated Equity)	Chinese Ln(Average Equity)
Ln(Distance to Muria Strait)	-0.053*** (0.013)	-0.246*** (0.062)	-0.224*** (0.057)	-0.034** (0.014)	-0.199** (0.087)	-0.190** (0.084)
R-squared	0.581	0.405	0.308	0.549	0.369	0.291
Mean (Dep. Var.)	0.268	1.095	0.920	0.172	0.752	0.650
Standard Deviation (Dep. Var.)	1	2	2	1	2	2
Fixed Effects	Province	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54	54
Observations	2,178	2,178	2,178	2,178	2,178	2,178

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.26 : 1930 CHINESE SHARE AND CHINESE OFFICERS

Reduced Form

	(1)	(2)	(3)	(4)	(5)
Period:	1867		1930		
Dependent Variable:	1st, 2nd Ranks	1st, 2nd, 3rd Ranks	1st, 2nd Ranks	1st, 2nd, 3rd Ranks	Non-Chinese Any Rank
Ln(Distance to Muria Strait)	-0.00708 (0.00957)	-0.0267** (0.0115)	-0.0200 (0.0133)	-0.0368*** (0.00470)	-0.00896 (0.0104)
R-squared	0.341	0.370	0.337	0.301	0.343
Mean (Dep. Var.)	0.058	0.140	0.054	0.182	0.029
Standard Deviation (Dep. Var.)	0.234	0.348	0.226	0.386	0.168
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.27 : 1930 CHINESE SHARE AND PROMINENT CHINESE

Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)
Period:	1941–1984			1941–1925		
Dependent Variable:	# Prominent	# Prominent /Total Pop.	# Prominent /Chinese Pop.	#Prominent	# Prominent /Total Pop.	# Prominent /Chinese Pop.
Ln(Distance to Muria Strait)	-0.0671* (0.0368)	-0.000000662* (0.000000362)	-0.0000910 (0.0000750)	0.00484 (0.00843)	2.90e-08 (9.73e-08)	-0.0000399 (0.0000424)
R-squared	0.835	0.392	0.057	0.682	0.261	0.060
Mean (Dep. Var.)	0.558	0.000	0.001	0.269	0.000	0.000
Standard Deviation (Dep. Var.)	2.559	0.000	0.007	1.141	0.000	0.004
Fixed Effects	Province	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54	54
Observations	242	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.28 : 1930 CHINESE SHARE AND PROMINENT CHINESE IN BUSINESS
Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)
Period:	1941–1984				1941–1925	
Dependent Variable:	# Prominent	# Prominent /Total Pop.	# Prominent /Chinese Pop.	#Prominent	# Prominent /Total Pop.	# Prominent /Chinese Pop.
Ln(Distance to Muria Strait)	-0.0219 (0.0193)	-0.000000206 (0.000000179)	-0.0000241 (0.0000220)	0.000538 (0.00204)	-1.78e-09 (2.34e-08)	-0.00000985 (0.0000107)
R-squared	0.645	0.283	0.055	0.398	0.144	0.059
Mean (Dep. Var.)	0.136	0.000	0.000	0.045	0.000	0.000
Standard Deviation (Dep. Var.)	0.592	0.000	0.002	0.245	0.000	0.001
Fixed Effects	Province	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54	54
Observations	242	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.29 : 1930 CHINESE SHARE AND PROMINENT CHINESE IN CULTURE
Reduced Form

	(1)	(2)	(3)	(4)	(5)	(6)
Period:	1941–1984				1941–1925	
Dependent Variable:	# Prominent	# Prominent /Total Pop.	# Prominent /Chinese Pop.	#Prominent	# Prominent /Total Pop.	# Prominent /Chinese Pop.
Ln(Distance to Muria Strait)	-0.0452* (0.0234)	-0.000000457* (0.000000251)	-0.0000669 (0.0000533)	0.00430 (0.00701)	3.07e-08 (7.78e-08)	-0.0000301 (0.0000318)
R-squared	0.858	0.420	0.058	0.682	0.283	0.061
Mean (Dep. Var.)	0.421	0.000	0.000	0.223	0.000	0.000
Standard Deviation (Dep. Var.)	2.026	0.000	0.005	0.960	0.000	0.003
Fixed Effects	Province	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54	54
Observations	242	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.30 : 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT SPILLOVERS

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Chinese Share (percent)	0.0972 (0.0766)	0.266*** (0.0868)	14.73*** (4.962)	6.366*** (2.148)	0.106*** (0.0335)
Neighbors' Chinese Share (percent)	-0.0308 (0.0333)	0.0462 (0.0520)	3.166 (3.399)	0.120 (0.922)	0.0293 (0.0218)
R-squared	0.707	0.679	0.513	0.669	0.463
Mean (Dep. Var.)	5.824	6.612	24.798	13.236	11.545
Standard Deviation (Dep. Var.)	0.558	0.596	21.783	9.453	0.168
Kleibergen-Paap WID, F-stat	11.862	11.862	11.862	11.862	11.862
Anderson-Rubin, P-value	0.157	0.061	0.028	0.148	0.000

Panel B: First Stage for Chinese Share

Ln(Distance to Muria Strait)	-0.0838*** (0.0204)	-0.0838*** (0.0204)	-0.0838*** (0.0204)	-0.0838*** (0.0204)	-0.0838*** (0.0204)
Neighbors' Ln(Distance to Muria Strait)	0.00281 (0.0153)	0.00281 (0.0153)	0.00281 (0.0153)	0.00281 (0.0153)	0.00281 (0.0153)

Panel C: First Stage for Neighbors' Chinese Share

Ln(Distance to Muria Strait)	0.0121 (0.0203)	0.0121 (0.0203)	0.0121 (0.0203)	0.0121 (0.0203)	0.0121 (0.0203)
Neighbors' Ln(Distance to Muria Strait)	-0.102*** (0.0209)	-0.102*** (0.0209)	-0.102*** (0.0209)	-0.102*** (0.0209)	-0.102*** (0.0209)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	1,131	1,131	1,131	1,131	1,131

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* include province fixed effects, altitude, the natural logarithm of total area (1930), total dry agricultural land (1963), total wet agricultural land (1963), longitude, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.31 : 1930 CHINESE SHARE AND HUMAN CAPITAL

Panel A: IV Specification

Dependent Variable:	(1)	(2)	(3)	(4)	(5)
	High School Graduates (2010)	Bachelor's Degree (2010)	Master's Degree (2010)	# High School per 1,000 (2011)	# University per 1,000 (2011)
Chinese Share (percent)	-0.887 (3.397)	0.320 (0.492)	0.0363 (0.0469)	0.0107 (0.00882)	0.00537** (0.00227)
R-squared	0.565	0.578	0.541	0.337	0.381
Mean (Dep. Var.)	18.214	3.879	0.206	0.086	0.008
Standard Deviation (Dep. Var.)	8.658	2.502	0.214	0.039	0.011
Kleibergen-Paap WID, F-stat	3.268	3.268	3.268	13.736	13.736
Anderson-Rubin, P-value	0.767	0.539	0.390	0.178	0.002

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0757* (0.0419)	-0.0757* (0.0419)	-0.0757* (0.0419)	-0.0777*** (0.0210)	-0.0777*** (0.0210)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	231	231	231	242	242

Panel C: IV Specification

Dependent Variable:	(1)	(2)	(3)
	Chinese High School Graduates (2010)	Chinese Bachelor's Degree (2010)	Chinese Mater's Degree (2010)
Chinese Share (percent)	0.0406 (0.0556)	0.0341 (0.0340)	0.00352 (0.00293)
R-squared	0.745	0.720	0.721
Mean (Dep. Var.)	0.134	0.086	0.008
Standard Deviation (Dep. Var.)	0.273	0.240	0.018
Kleibergen-Paap WID, F-stat	3.327	2.920	2.429
Anderson-Rubin, P-value	0.585	0.446	0.322

Panel D: First Stage

Ln(Distance to Muria Strait)	-0.0765* (0.0419)	-0.0718* (0.0420)	-0.0625 (0.0401)
Baseline Controls	Yes	Yes	Yes
Fixed Effects	Province	Province	Province
Number of Regencies	52	50	48
Observations	205	182	100

Notes: Columns (1) to (3) report the Shares of individuals whose highest educational attainment is within each category. Specifically, "High School" includes high school and any equivalent secondary education qualification. "Undergraduate" includes bachelor's degrees and all equivalent tertiary education credentials. "Master's" includes master's degrees and all equivalent postgraduate qualifications. All regressions include controls of altitude, the natural logarithm of total area (1930), total dry agricultural land (1963), total wet agricultural land (1963), longitude, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). Standard errors are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.32 : CHINESE COMMUNITIES AND MARKET ACCESS

Reduced Form

Dependent Variable:	(1)	(2)	(3)
	Market Access (2010)	Market Access (2010)	Market Access (2010)
Ln(Distance to Murait Strait)	-0.003 (0.001)	-0.004 (0.001)	0.001 (0.001)
R-squared	0.162	0.686	0.942
Mean (Dep. Var.)	14.78	14.78	14.78
Geographic/Agroclimatic Controls		Yes	Yes
Urban Centers Controls			Yes
Number of Regencies	68	68	68
Observations	68	68	68

Notes: The table reports OLS estimates. Chinese share 1930 denotes the share of the logarithm of the total population with Chinese ethnicity in 1930 measured from 0-100. The unit of observation is at the regency-level. Regressions are based on a cross-section of 95 regencies. *Dependent variable:* The natural logarithm of market access in 2010 (see Appendix for details). *Controls:* Altitude, the natural logarithm of total area (1930), total dry agricultural land (1963), total wet agricultural land (1963), longitude, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.33 : 1930 CHINESE SHARE AND 2003–2009 CHINESE AID

Panel A: IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Infrastructure # Aid Project	Infrastructure Aid Amount	Infrastructure # Aid Project	Infrastructure Aid Amount
Chinese Share (percent)	0.227** (0.112)	23.78** (10.95)	0.259** (0.127)	26.29** (12.34)
Light Intensity (1998)			-0.0120 (0.0106)	-0.950 (0.891)
R-squared	0.378	0.269	0.357	0.251
Mean (Dep. Var.)	0.157	11.723	0.157	11.723
Standard Deviation (Dep. Var.)	0.456	45.138	0.456	45.138
Kleibergen-Paap WID, F-stat	13.736	13.736	19.505	19.505
Anderson-Rubin, P-value	0.034	0.041	0.037	0.037

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0777*** (0.0210)	-0.0777*** (0.0210)	-0.0663*** (0.0150)	-0.0663*** (0.0150)
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province
Number of Regencies	54	54	54	54
Observations	242	242	242	242

Panel C: IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Non-Infrastructure # Aid Project	Non-Infrastructure Aid Amount	Non-Infrastructure # Aid Project	Non-Infrastructure Aid Amount
Chinese Share (percent)	-0.627** (0.304)	-0.0308** (0.0149)	-0.694** (0.348)	-0.0344** (0.0168)
Light Intensity (1998)			0.0257 (0.0198)	0.00135 (0.000948)
R-squared	0.450	0.238	0.429	0.205
Mean (Dep. Var.)	1.826	0.023	1.826	0.023
Standard Deviation (Dep. Var.)	1.106	0.045	1.106	0.045
Kleibergen-Paap WID, F-stat	13.736	13.736	19.505	19.505
Anderson-Rubin, P-value	0.038	0.036	0.041	0.037

Panel D: First Stage

Ln(Distance to Muria Strait)	-0.0777*** (0.0210)	-0.0777*** (0.0210)	-0.0663*** (0.0150)	-0.0663*** (0.0150)
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province
Number of Regencies	54	54	54	54
Observations	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the firm level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.34 : 1930 CHINESE SHARE AND SOCIAL ATTITUDES

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Trust	Community	Toleration for Other Ethnicities	Toleration for Other Religions	Trust towards Religious Leader
Chinese Share (percent)	0.0502 (0.0325)	-0.00305 (0.0452)	0.117 (0.167)	0.103 (0.163)	-0.0314 (0.0216)
R-squared	0.080	0.292	0.219	0.198	0.068
Mean (Dep. Var.)	2.985	2.524	2.210	2.174	3.087
Standard Deviation (Dep. Var.)	0.114	0.168	0.754	0.780	0.120
Kleibergen-Paap WID, F-stat	13.768	13.768	13.768	13.768	13.768
Anderson-Rubin, P-value	0.069	0.947	0.503	0.558	0.046

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.0782*** (0.0211)	-0.0782*** (0.0211)	-0.0782*** (0.0211)	-0.0782*** (0.0211)	-0.0782** (0.0211)
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Province	Province	Province	Province	Province
Number of Regencies	54	54	54	54	54
Observations	235	235	235	235	235

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* include province fixed effects, altitude, the natural logarithm of total area (1930), total dry agricultural land (1963), total wet agricultural land (1963), longitude, and quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* are clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.35 : 1930 CHINESE SHARE AND STATE CAPACITY

<i>Panel A: IV Specification</i>							
Dependent Variable:	(1) Special Allocation Grant	(2) General Allocation Grant	(3) Natural Resource Revenue Sharing	(4) Own Source Revenue	(5) Other Revenue	(6) Tax Revenue Sharing	(7) Total Revenue
Chinese share (percent)	-579.4 (426.4)	-6662.3** (2693.1)	236.2 (285.4)	-990.7 (901.6)	768.3 (692.3)	-451.7 (613.7)	-7038.1 (4402.4)
R-squared	0.407	0.283	0.300	0.486	0.351	0.394	0.258
Mean (Dep. Var.)	5209.674	65478.849	884.225	9398.596	12625.183	5261.843	1.01e+05
Standard Deviation (Dep. Var.)	2356.091	20934.790	1415.482	5151.328	4043.647	2460.710	30191.587
Kleibergen-Paap WID, F-stat	13.736	13.736	13.736	13.736	13.736	13.736	13.736
Anderson-Rubin, P-value	0.248	0.011	0.453	0.181	0.270	0.384	0.078

<i>Panel B: First Stage</i>							
Ln(Distance to Muria Strait)	-0.0777*** (0.0210)						
Baseline Controls	Yes						
Fixed Effects	Province						
Number of Regencies	54	54	54	54	54	54	54
Observations	242	242	242	242	242	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. *Geographical controls* include highest altitude level, longitude, natural logarithm of distance to nearest coastline, total area, total dry rice field area, total wet rice field area. *Urban center controls* include quadratic polynomials of the natural logarithm of distances to major provincial capitals (Batavia, Semarang, Soerabaia) and historical Majapahit coastal cities (Tuban, Gresik, Trowulan). *Standard errors* clustered by regency. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

References for the Appendix

- Acemoglu, Daron, Simon Johnson, and James A. Robinson**, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001, 91 (5), 229.
- _____, _____, and **James A. Robinson**, "Institutions as a Fundamental Cause of Long-Run Growth," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Elsevier, 2005, pp. 385–472.
- _____, **Tarek A. Hassan, and James A. Robinson**, "Social Structure and Development: A Legacy of the Holocaust in Russia," *The Quarterly Journal of Economics*, May 2011, 126 (2), 895–946.
- Ager, Philipp, Leah Platt Boustan, and Katherine Eriksson**, "The Intergenerational Effects of a Large Wealth Shock: White Southerners after the Civil War," *American Economic Review*, 2021, 111 (11), 3767–3794.
- Akulut-Yuksel, Mevlude and Mutlu Yuksel**, "The Long-Term Direct and External Effects of Jewish Expulsions in Nazi Germany," *American Economic Journal: Economic Policy*, 2015, 7 (3), 58–85.
- Alesina, Alberto and Eliana La Ferrara**, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, 43 (3), 762–800.
- _____, and **Paola Giuliano**, "Family Ties," in "Handbook of Economic Growth," Vol. 2, Elsevier, 2014, pp. 177–215.
- Alesina, Alberto F., Marlon Seror, David Y. Yang, Yang You, and Weihong Zeng**, "Persistence Despite Revolutions," Working Paper 27053, National Bureau of Economic Research 2020. Revised March 2021.
- Allen, Treb and Costas Arkolakis**, "Economic Activity across Space: A Supply and Demand Approach," *Journal of Economic Perspectives*, May 2023, 37 (2), 3–28.
- Ananta, Aris, Evi Nurvidya Arifin, M Sairi Hasbullah, Nur Budi Handayani, and Agus Pramono**, *Demography of Indonesia's Ethnicity*, Institute of Southeast Asian Studies, 2015.
- Anderson, Robert Warren, Noel D. Johnson, and Mark Koyama**, "Jewish Persecutions and Weather Shocks: 1100–1800," *The Economic Journal*, 2017, 127 (602), 924–958.
- Anderton, Charles H. and Jürgen Brauer**, "Mass Atrocities and Their Prevention," *Journal of Economic Literature*, 2021, 59 (4), 1240–1292.
- Arbatlı, Cemal Eren and Gunes Gokmen**, "Human Capital Transfers and Sub-national Development: Armenian and Greek Legacy in Post-expulsion Turkey," *Journal of Economic Growth*, March 2023, 28 (1), 1–43.
- Auwalin, Ilmiawan**, "Ethnic Identity and Internal Migration Decision in Indonesia," *Journal of Ethnic and Migration Studies*, 2020, 46 (13), 2841–2861.
- Bai, Ying and Ruixue Jia**, "The Economic Consequences of Political Hierarchy: Evidence from Regime Changes in China, 1000–2000 CE," *Review of Economics and Statistics*, 2023, 105 (3), 626–645.
- Bassino, Jean-Pascal and Pierre van der Eng**, *The First East Asian Economic Miracle: Wages, Living Standards and Foundations of Modern Economic Growth in Southeast Asia, 1880-1938*, Australian National University, 2021.
- Bates, Robert H.**, "Ethnic Competition and Modernization in Contemporary Africa," *Comparative Political Studies*, 1974, 6 (4), 457–484.
- Baum-Snow, Nathaniel, J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt**, "Does Investment in National Highways Help or Hurt Hinterland City Growth?," *Journal of Urban Economics*, January 2020, 115, 103124.

- Bazzi, Samuel, Arya Gaduh, Alexander D Rothenberg, and Maisy Wong**, "Skill transferability, migration, and development: Evidence from population resettlement in Indonesia," *American Economic Review*, 2016, 106 (9), 2658–2698.
- _____, _____, _____, and _____, "Unity in diversity? How intergroup contact can foster nation building," *American Economic Review*, 2019, 109 (11), 3978–4025.
- _____, **Gabriel Koehler-Derrick, and Benjamin Marx**, "The Institutional Foundations of Religious Politics: Evidence from Indonesia," *The Quarterly Journal of Economics*, 2020, 135 (2), 845–911.
- _____, **Masyhur Hilmy, and Benjamin Marx**, "Religion, Education, and the State," *NBER Working Paper* 27073, 2020.
- Becker, Sascha O. and Luigi Pascali**, "Religion, Division of Labor, and Conflict: Anti-semitism in Germany over 600 Years," *American Economic Review*, May 2019, 109 (5), 1764–1804.
- _____, **Irena Grosfeld, Pauline Grosjean, Nico Voigtländer, and Ekaterina Zhuravskaya**, "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review*, May 2020, 110 (5), 1430–1463.
- _____, **Sharun Mukand, and Ivan Yotzov**, "Persecution, Pogroms and Genocide: A Conceptual Framework and New Evidence," *Explorations in Economic History*, 2022, 86, 101471.
- Bestari, Rossi Rizki, Yusuf Reza Kurniawan, and Muhammad Halley Yudhistira**, "The Long-term Effects of Early European Settlement on Local Development: Evidence from Indonesia," *World Development*, 2022, 158, 105977.
- Bonacich, Edna**, "A Theory of Middleman Minorities," *American Sociological Review*, 1973, 38 (5), 583–594.
- Booth, Anne**, "Living Standards and the Distribution of Income in Colonial Indonesia: A Review of the Evidence," *Journal of Southeast Asian Studies*, 1988, 19 (2), 310–334.
- _____, *The Indonesian Economy in the Nineteenth and Twentieth Centuries: A History of Missed Opportunities* A Modern Economic History of Southeast Asia (MEHSA), London, UK: Macmillian Press, 1998.
- Borsuk, Richard**, *Liem Sioe Liang's Salim Group*, Flipside Digital Content Company Inc., 2014.
- 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 2021.
- Caicedo, Felipe Valencia**, "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America," *The Quarterly Journal of Economics*, 2019, 134 (1), 507–556.
- Campante, Filipe R. and Quoc-Anh Do**, "Isolated Capital Cities, Accountability, and Corruption: Evidence from US States," *American Economic Review*, 2014, 104 (8), 2456–2481.
- Card, David**, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 2001, 19 (1), 22–64.
- Carey, Peter**, "Changing Javanese Perceptions of the Chinese Communities in Central Java, 1755–1825," *Indonesia*, 1984, (37).
- Caselli, Francesco and Wilbur John Coleman**, "On the Theory of Ethnic Conflict," *Journal of the European Economic Association*, 2013, 11, 161–192.

- Chen, Zhiwu, Chicheng Ma, and Andrew J Sinclair**, "Banking on the Confucian Clan: Why China Developed Financial Markets so Late," *The Economic Journal*, May 2022, 132 (644), 1378–1413.
- Chua, Amy**, *World on Fire: How Exporting Free Market Democracy Breeds Ethnic Hatred and Global Instability*, New York: Anchor Books, 2004.
- Chua, Christian**, "Defining Indonesian Chineseness under the New Order," *Journal of Contemporary Asia*, 2004, 34 (4), 465–479.
- Cinelli, Carlos and Chad Hazlett**, "Making Sense of Sensitivity: Extending Omitted Variable Bias," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2020, 82 (1), 39–67.
- Claver, Alexander**, *Dutch Commerce and Chinese Merchants in Java: Colonial Relationships in Trade and Finance, 1800-1942*, Vol. 291 of *Verhandelingen van het Koninklijk Instituut voor Taal-, Land en Volkenkunde*, Leiden and Boston: Brill, 2014.
- Cribb, Robert**, *Historical Atlas of Indonesia*, London: Curzon Press, 1997.
- De Zwart, Pim**, "Inequality in Late Colonial Indonesia: New Evidence on Regional Differences," *Clio-metrica*, 2022, 16 (1), 175–211.
- , **Daniel Gallardo-Albarrán, and Auke Rijpma**, "The Demographic Effects of Colonialism: Forced Labor and Mortality in Java, 1834–1879," *The Journal of Economic History*, 2022, 82 (1), 211–249.
- Deaton, Angus**, *The analysis of household surveys: a microeconometric approach to development policy*, World Bank Publications, 1997.
- Dell, Melissa**, "The persistent effects of Peru's mining mita," *Econometrica*, 2010, 78 (6), 1863–1903.
- 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.
- der Eng, Pierre Van**, "Chinese Entrepreneurship in Indonesia: A Business Demography Approach," *Business History*, 2022, 64 (4), 682–703.
- Die, Ong Eng**, *Chineezen in Nederlandsch-Indie: Sociografie van een Indonesische Bevolkingsgroep*, Van Gorcum, 1943.
- Dippel, Christian and Stephan Hebllich**, "Leadership in Social Movements: Evidence from the "Forty-Eighters" in the Civil War," *American Economic Review*, 2021, 111 (2), 472–505.
- Do, Quoc-Anh, Roberto Galbiati, Benjamin Marx, and Miguel A. Ortiz Serrano**, "J'Accuse! Anti-semitism and financial markets in the time of the Dreyfus Affair," *Journal of Financial Economics*, 2024, 154, 103809.
- Dreher, Axel, Andreas Fuchs, Bradley Parks, Austin Strange, and Michael J. Tierney**, "Aid, China, and Growth: Evidence from a New Global Development Finance Dataset," *American Economic Journal: Economic Policy*, 2021, 13 (2), 135–174.
- , ———, **Roland Hodler, Bradley C. Parks, Paul A. Raschky, and Michael J. Tierney**, "African Leaders and the Geography of China's Foreign Assistance," *Journal of Development Economics*, 2019, 140, 44–71.
- Dreyer, Edward L.**, *Zheng He: China and the Oceans in the Early Ming Dynasty, 1405–1433*, New York: Pearson Longman, 2006.
- Droller, Federico**, "Migration, Population Composition and Long Run Economic Development: Evidence from Settlements in the Pampas," *The Economic Journal*, 2018, 128 (614), 2321–2352.

D'Acunto, Francesco, Marcel Prokopcuk, and Michael Weber, "Historical Antisemitism, Ethnic Specialization, and Financial Development," *The Review of Economic Studies*, May 2019, 86 (3), 1170–1206.

Easterly, William and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 1997, 112 (4), 1203–1250.

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.

Enke, Benjamin, "Kinship, Cooperation, and the Evolution of Moral Systems," *The Quarterly Journal of Economics*, May 2019, 134 (2), 953–1019.

Galor, Oded and Ömer Özak, "Land Productivity and Economic Development: Caloric Suitability vs. Agricultural Suitability," *Manuscript*, 2015.

_____ and _____, "The Agricultural Origins of Time Preference," *American Economic Review*, 2016, 106 (10), 3064–3103.

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

Greif, Avner, "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders," *The Journal of Economic History*, December 1989, 49 (4), 857–882.

_____, "Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition," *The American Economic Review*, 1993, 83 (3), 525–548. Publisher: American Economic Association.

_____, and Guido Tabellini, "The Clan and the Corporation: Sustaining Cooperation in China and Europe," *Journal of Comparative Economics*, February 2017, 45 (1), 1–35.

Grosfeld, Irena, Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya, "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, January 2020, 87 (1), 289–342.

Henderson, J Vernon, Adam Storeygard, and David N Weil, "Measuring Economic Growth from Outer Space," *American Economic Review*, 2012, 102 (2), 994–1028.

Heryanto, Ariel, "Ethnic Identities and Erasure," *Southeast Asian identities: Culture and the politics of representation in Indonesia, Malaysia, Singapore, and Thailand*, 1998, pp. 95–114.

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.

Houben, Vincent J.H., "Java in the 19th Century: Consolidation of a Territorial State," in V.J.H. Houben and J.Th. Lindblad, eds., *The Emergence of a National Economy*, Leiden, Netherlands: KITLV Press, 2002, pp. 56–81.

Huber, Kilian, Volker Lindenthal, and Fabian Waldinger, "Discrimination, Managers, and Firm Performance: Evidence from "Aryanizations" in Nazi Germany," *Journal of Political Economy*, 2021, 129 (9), 2455–2503.

Huff, Gregg and Giovanni Caggiano, "Globalization, Immigration, and Lewisian Elastic Labor in Pre-World War II Southeast Asia," *The Journal of Economic History*, 2007, 67 (1), 33–68.

- 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.
- _____, ""Unfinished business": Historic Complementarities, Political Competition and Ethnic Violence in Gujarat," *Journal of Economic Behavior & Organization*, 2014, 104, 18–36.
- _____, "Trading for Peace," *Economic Policy*, 2018, 33 (95), 485–526.
- Jia, Ruixue and Torsten Persson**, "Choosing Ethnicity: The Interplay Between Individual and Social Motives," *Journal of the European Economic Association*, April 2021, 19 (2), 1203–1248.
- Johnson, Noel D and Mark Koyama**, "Jewish Communities and City Growth in Preindustrial Europe," *Journal of Development Economics*, 2017, 127.
- Kaur, Amarjit**, "Wage Labour in Southeast Asia since 1840," *Neg Yobk: Palgbafe Macmillan*, 2004.
- Keane, Michael P. and Timothy Neal**, "Instrument Strength in Structural Econometric Models," *Journal of Econometrics*, 2023, 235 (2), 344–368.
- Klinken, Gerry Van**, "Ethnicity in Indonesia," in "Ethnicity in Asia," New York, NY, USA: Routledge, 2003, pp. 88–111.
- Kok, Chun Chee, Gedeon J. Lim, Danial Shariat, Abu Siddique, and Shunsuke Tsuda**, "Interethnic Proximity and Political Development," 2025. Available as a working paper on SSRN and IZA.
- Kuipers, Nicholas**, "The Long-Run Consequences of The Opium Concessions for Out-Group Animosity on Java," *World Politics*, 2022, 74 (3), 405–442.
- Landa, Janet Tai**, *Trust, Ethnicity, and Identity: Beyond the New Institutional Economics of Ethnic Trading Networks, Contract Law, and Gift-exchange*, Ann Arbor, Michigan, USA: University of Michigan Press, 1998.
- Leeuwen, Marco HD Van and Ineke Maas**, *HISCLASS: A Historical International Social Class Scheme*, Universitaire Pers Leuven, 2011.
- Liem, Thian Joe**, *Riwajat Semarang*, Semarang: Ho Kim Yoe, 1933.
- Liem, Twan Djie**, "De Distribuerende Tusschenhandel der Chinezen op Java [The Distributive Intermediate Trade of the Chinese in Java]," *Nijhoff*, 1952.
- Lim, Gedeon J.**, "Political Elites and Land Rents: Evidence from Indonesia," Technical Report, The University of Hong Kong 2024. Pre-submission manuscript, available at https://github.com/gedeon89.github.io/gedeonlim.com/Lim_Bengkok_edited_pre_sub.pdf.
- Macaulay, Melissa**, *Distant Shores: Colonial Encounters on China's Maritime Frontier*, Princeton, New Jersey, USA: Princeton University Press, 2021.
- McKeown, Adam**, "Global Migration, 1846–1940," *Journal of World History*, 2004, 15 (2), 155–189.
- Michelman, Valerie, Joseph Price, and Seth D. Zimmerman**, "Old Boys' Clubs and Upward Mobility Among the Educational Elite," *The Quarterly Journal of Economics*, 2022, 137 (2), 845–909.
- Mills, J.V.G.**, *Ying-yai sheng-lan: 'The Overall Survey of the Ocean's Shores'* [1453], Cambridge, UK: Cambridge University Press, 1971.
- Mitra, Anirban and Debraj Ray**, "Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India," *Journal of Political Economy*, 2014, 22 (4).

- Montalvo, José G. and Marta Reynal-Querol**, "Ethnic Diversity and Growth: Revisiting the Evidence," *The Review of Economics and Statistics*, 2021, 103 (3), 521–532.
- Munshi, Kaivan**, "Community Networks and the Process of Development," *Journal of Economic Perspectives*, 2014, 28 (4), 49–76.
- _____, "Social Networks and Migration," *Annual Review of Economics*, 2020, 12, 503–24.
- Onghokham**, "Chinese Capitalism in Dutch Java," *Southeast Asian Studies*, 1989, 27 (2), 156–176.
- Pan, Lynn**, *The Encyclopedia of the Chinese Overseas*, Cambridge, Massachusetts, USA: Harvard University Press, 1999.
- Pascali, Luigi**, "The wind of change: Maritime technology, trade, and economic development," *American Economic Review*, 2017, 107 (9).
- Pepinsky, Thomas B**, "Colonial Migration and the Origins of Governance: Theory and Evidence from Java," *Comparative Political Studies*, 2016, 49 (9), 1201–1237.
- Purdey, Jemma**, *Anti-Chinese Violence in Indonesia, 1996–1999*, Honolulu: University of Hawai'i Press, 2006.
- Rauch, James E**, "Business and Social Networks in International Trade," *Journal of Economic Literature*, December 2001, 39 (4), 1177–1203.
- Rauch, James E. and Vitor Trindade**, "Ethnic Chinese Networks in International Trade," *The Review of Economics and Statistics*, February 2002, 84 (1), 116–130.
- Rivers, Philip J.**, "New Lamps for Old: Modern Nautical Terms for Ancient Marine Practices and the Navigation of the Zheng He Voyages," *Journal of the Malaysian Branch of the Royal Asiatic Society*, June 2012, 85 (1), 85–98.
- Rocha, Rudi, Claudio Ferraz, and Rodrigo R Soares**, "Human Capital Persistence and Development," *American Economic Journal: Applied Economics*, 2017, 9 (4), 105–136.
- Rush, James**, "Placing the Chinese in Java on the Eve of the Twentieth Century," *Indonesia*, 1991, pp. 13–24.
- Salmon, Claudine and Denys Lombard**, *Les chroniques chinoises de Cirebon et de Semarang*, Paris: École Française d'Extrême-Orient, 1981.
- Schwarz, Adam**, *A nation in waiting: Indonesia's search for stability*, Routledge, 2018.
- 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.
- Suryadinata, Leo**, *Admiral Zheng He and Southeast Asia*, Vol. 286, Institute of Southeast Asian Studies, 2005.
- _____, *Understanding the Ethnic Chinese in Southeast Asia*, Singapore: Institute of Southeast Asian Studies, 2007.
- _____, *Prominent Indonesian Chinese: Biographical Sketches*, Institute of Southeast Asian Studies, 2015.
- _____, ed., *Ethnic Chinese as Southeast Asians*, Singapore: Institute of Southeast Asian Studies, 1997.
- _____, ed., *Ethnic Relations and Nation-Building in Southeast Asia: The Case of the Ethnic Chinese*, Singapore: Institute of Southeast Asian Studies, 2004.

- _____, ed., *Ethnic Chinese in Contemporary Indonesia*, Singapore: ISEAS–Yusof Ishak Institute, 2008.
- 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* Routledge Studies in the Growth Economies of Asia (2005), Hoboken, NY, USA: Taylor and Francis, 2012.
- Varshney, Ashutosh, Mohammad Zulfan Tadjoeddin, and Rizal Panggabean**, "Creating Datasets in Information-Poor Environments: Patterns of Collective Violence in Indonesia, 1990–2003," *Journal of East Asian Studies*, December 2008, 8 (3), 361–394.
- Vleming, JL**, *Het Chineesche Zakenleven in Nederlandsch-Indië* number 730, Landsdrukkerij, 1926.
- Voigtländer, Nico and Hans-Joachim Voth**, "Persecution Perpetuated: The Medieval Origins of Anti-Semitic Violence in Nazi Germany," *The Quarterly Journal of Economics*, 2012, 127 (3), 1339–1392.
- Volkstelling**, *Census of 1930 in the Netherlands Indies*. Vol. VII *Chinese and Other Non-indigenous Orientals in the Netherlands Indies*, Batavia: Landsdrukkerij, 1930.
- Wu, Yuan-Li**, "Chinese Entrepreneurs in Southeast Asia," *American Economic Review: Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association*, 1983, 73 (2), 112–117.
- Wu, Yuan-li and Chung-Hsi Wu**, *Economic Development in Southeast Asia; the Chinese Dimension*, Stanford University, Stanford, California, USA: Stanford University Press, 1980.
- Zhongguo, Fei Xiaotong's Xiangtu, Gary G. Hamilton, and Wang Zheng**, *From the Soil: The Foundations of Chinese Society*, Berkeley, California, USA: University of California Press, 1992.