Communities of Commerce:

The Legacy of Chinese Immigration on Java

Quoc-Anh Do
Monash University, CEPR, and CESifo

Sebastian Ellingsen
University of Bristol

Gedeon Lim
University of Hong Kong

June 19, 2026

Abstract

This paper studies the long-run impacts of diasporas on local economic development. We do so in the context of one of the most economically influential diasporas: the ethnic Chinese in Southeast Asia. Our instrumental variables strategy exploits 15th-century landing sites on Java that shaped subsequent Chinese settlement but lost maritime access due to 17th-century silting. We find that a 1p.p. higher Chinese share in 1930 leads to 9.4% higher household consumption and 26% higher population today. These effects attenuate sharply during the 1997-1998 crisis, then recover. We find strong effects on firm sales driven by improved financing through local private networks but not through exports. We provide suggestive evidence for the key role of historical ethnic-trading networks within Java. Our findings suggest that despite their small size, ethnic minorities with durable complementarities can play an out-sized, positive role in local economic development.

Keywords: Diaspora; Ethnic minority; Immigration; Ethnic Segregation; Conflict; Discrimination; Private Finance; Networks.

JEL Codes: F63, N35, J15, Z13

*We are grateful for insights and comments from Ying Bai, Sam Bazzi, Sascha Becker, Adam Brzezinski, Konrad Burchardi, Filipe Campante, Zhiwu Chen, Ma Chicheng, Pierre van der Eng, Benjamin Faber, Pauline Grosjean, Maxime Gravoueille, Masyhur Hilmy, Mark Hup, Saumitra Jha, Ruixue Jia, Dana Kassem, James Kung, Horacio Larreguy, Bingjing Li, Weijia Li, Federico Masera, Dilip Mookherjee, Joel Mokyr, Kieu-Trang Nguyen, Nathan Nunn, Laura Panza, Jacopo Ponticelli, Ana Tur Prats, Jens Prüfer, Gianluca Russo, Uta Schönberg, 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 Business School, 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

History abounds with examples of diasporic minority groups that occupy key intermediary roles in commerce and finance. From Jewish communities in Europe to Indian, Lebanese, and Chinese diasporas across much of the developing world, these socially distinctive groups have often played a central role in economic exchange. Yet, a common narrative has portrayed these minorities as "unproductive" or "exploitative", contributing little to the economic well-being of native, majority groups (Bonacich, 1973; Chua, 2004a). In turn, such narratives have fueled historical episodes of exclusion; expulsion; and mass violence (Grosfeld et al., 2020; Anderton and Brauer, 2021; Becker et al., 2022).

This paper directly evaluates these claims by examining whether and how diasporic minorities contribute to long-run economic development in host societies. We do so by focusing on one of the largest and most economically influential diasporas: the ethnic Chinese in Southeast Asia. Leveraging the deep history of Chinese immigration to the island of Java, dating back to the 15th century, we find strong positive causal effects of Chinese communities on the long-run local economic development and welfare of the native population. Importantly, we trace the attenuation and recovery of these economic effects through key historical episodes of discrimination; socioeconomic crises; and recovery. We argue that the recovery of these effects point to the potentially important role of minority, co-ethnic networks in driving long-run development.

Chinese emigration 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 (Macauley, 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 small 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. By 1930, centuries of initially modest Chinese presence had culminated in substantially more settled communities due to large-scale emigration waves in the late 19th and early 20th century. In addition, this data provides us with a snapshot of the ethnic composition across Java before the onset of rapid economic growth in the second half of the 20th century. Lastly, it provides a unique and reasonably truthful snapshot of the ethnic composition of Java: After independence, ethnicity quickly became a politically sensitive topic and was removed from all censuses until 2000.

We begin by documenting a strong link between a district's contemporary economic development and

the size of its Chinese community in 1930. A 1p.p. (0.1p.p.) increase in ethnic Chinese shares is associated with a 3% (0.3%) increase in average household consumption and a 2.6 (0.26) 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, economic development and extraction during the colonial period (Dell and Olken, 2020; De Zwart et al., 2022), and is unlikely to be fully explained 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. To do so, we build on a milestone of Chinese emigration 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). Specifically, we use a district's distance to the expeditions' landing sites in the Muria Strait (a waterway that cut across the then-Northern coast of Java and Mount Muria during the time of Zheng He's expedition, but which no longer exists) as an instrumental variable for Chinese population share in 1930. Importantly, prior to the expedition, these sites were relatively unimportant and did not warrant mention in key historical documents.[1] Hence, our empirical strategy builds on the long tradition of using initial immigration networks as an instrument (Card, 2001; Munshi, 2014). It draws further support from the expeditions' lack of modern navigational technology to precisely choose landing sites (Pascali, 2017; Rivers, 2012), and, most importantly, the complete sedimentation of the Muria Strait since the 17th century that removed all prior waterway advantages of these locations.[2] To further strengthen the causal interpretation of our IV strategy, we implement additional robustness checks to compare our estimates against that of simulated counterfactual landing sites.

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 1p.p. (0.1p.p.) of Chinese share in 1930 to be 9.4% (0.94%) higher household consumption, 26% (2.6%) more district population, and a 14p.p. (1.4p.p.) higher urbanization rate. These long-run impacts 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 are accompanied by strong effects of the Chinese community on the structural transformation of the economy away from agricultural jobs, at the range of 10p.p. (1p.p.) lower employment in agriculture for a 1p.p. (0.1p.p.) higher in Chinese share in 1930.[3]

In light of these sizable effects, we test for heterogeneous effects against the backdrop of social conflicts involving ethnic Chinese. In particular, we focus on the 1998 unrest, which occurred during the Asian

[1] These sites are distinct from the established pre-colonial trade ports (Pekalongan, Gresik, Tuban) that had been documented in Chinese sources since at least the 13th century (e.g. the Zhu Fan Zhi).

[2] Hence, our identification strategy echoes that of Jha (2013) who uses medieval natural harbors that subsequently silted up as an instrument for medieval locuses of Muslim trade. This unique feature of the disruption of a historical intervention due to plausibly exogenous reasons is further reminiscent of the instrument in Valencia Caicedo (2019).

[3] The 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. In addition, these effects are cumulative over a very long duration.

Financial Crisis and the political transition following the fall of President Suharto (Indonesia's longest serving President). This double crisis led to violence against ethnic Chinese throughout Indonesia (see e.g. Varshney et al., 2008). Quantitatively, we find a drop of the positive effects of Chinese presence during the crisis period. Following the recovery and the restoration of equal rights for the Chinese minority under Gus Dur's government (1999-2001), we find a quantitatively strong recovery of the positive effects of Chinese presence.

To better understand these results, we investigate the role of firms.[4] Contemporary firm data does not contain complete names of owners, hence, we are unable to identify the direct impact of Chinese firms per se. Nonetheless, we draw on data from the universe of Indonesian manufacturing firms (1980-2013) and find strong effects of historical Chinese presence on overall 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 region (Suryadinata, ed, 2008; Suryadinata, 2015).

This time-varying pattern is difficult to reconcile with explanations based solely on time-invariant locational fundamentals or persistent physical or institutional infrastructure. Instead, we interpret these findings as likely to be more consistent with the possibility that Chinese communities function as economic intermediaries, leveraging dense social and commercial networks to facilitate investment and stimulate firm growth—much of which survived the 1997-1998 crisis.[5] To test this, we use data from a historical business registry of Chinese firms to provide direct evidence on the structure of these networks: districts with higher Chinese shares have a larger number of middleman firms and are characterized by greater intra-group concentration and district-hub network linkages. Both of which are suggestive of denser social and commercial networks amongst the ethnic Chinese diaspora.

We find limited evidence for leading alternative mechanisms. The main results remain strong after controlling for historical determinants of colonial extraction and early development (see 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 rule out other potential mediators including market access, international trade, and international aid (Rauch, 2001; Rauch and Trindade, 2002), and human capital spillovers (Becker et al., 2020; Arbatlı and Gokmen, 2023).

Taken together, our findings, especially on the time-varying effects of ethnic Chinese shares, 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 activ-

[4]The ethnic Chinese business community on Java has been called one of the most powerful economic forces in the world (Redding, 2013). Despite their relatively small share, ethnic Chinese capital across Thailand, Malaysia, Indonesia, and the Philippines totaled US$11.7 billion relative to total foreign direct investments (mainly from U.S. and Japanese Capital) of US$6.4 billion (Wu and Wu, 1980, p. 34).

[5]Using 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.

ity, 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 (post-)colonial policies, has persisted to the present day. Our findings therefore 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 on the long-term economic and political impacts of ethnic minorities (Acemoglu et al., 2011; Voigtländer and Voth, 2012; Hornung, 2014; Becker and Pascali, 2019) and, in particular, occupationally segregated groups (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: 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 the historical Chinese diaspora, an extremely understudied minority group, on long-term economic development. Our findings on time-varying effects during the 1997 Asian Financial Crisis bridges both strands of literature by examining how the broader political context mediates the economic impact of ethnic minorities in times of crisis. In this aspect, our paper also relates to a recent literature that shows the long-run economic resilience of economic elites through adversity, especially after a reversal of policies or rejection of discriminatory beliefs (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 (2014, 2018) highlights the importance of durable inter-ethnic complementarities combined with mechanisms to share the gains arising from such exchange. In the Indonesian context, ethnic Chinese communities relied on kinship-based networks to organize economic activity, creating high barriers to entry that limited the sharing of rents derived from trade (Jha, 2013). Grosfeld et al. (2020) show that such conditions — where a middleman minority controls economic networks with high barriers to entry — make the minority vulnerable to large political shocks, as these shocks undermine the repeated interactions on which the networks depend. This prediction is borne out in our setting: we document that the economic benefits from Chinese communities collapsed during the 1997–1998 crisis and associated anti-Chinese violence, before partially recovering.

Nevertheless, our results point to a (paradoxical) silver lining—despite the organization of ethnic Chi-

nese trading networks through community and kinship-based ties, both of which might have limited initial entry and ethnic majority gains (Jha, 2013), persistent state-led enforcement of surplus-sharing norms and institutions can lead to economic benefits for the majority. In doing so, such institutions may partially mitigate the tensions arising from minority-specific economic complementarities. Hence, we provide novel 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. (2024) 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 relatively low-skilled migration can promote long-term economic development in initially weak-institutionalized settings, where (the immigration of) groups with strong kinship networks can potentially provide a substitute for state institutions. As such, our findings also build on the impact of social ties on long-term economic development (Alesina and Giuliano, 2014; Chen et al., 2022b).

Finally, our paper contributes to quantitative studies on the determinants of long-term economic development, especially that of colonialism both in general (see e.g. Acemoglu et al., 2001, 2005) and in Indonesia (Pepinsky, 2016; Dell and Olken, 2020; De Zwart et al., 2022; Kuipers, 2022; Lim, 2024).[6] To the best of our knowledge, our paper is one of the first to study the time-varying effects of colonial-era policies that were explicitly designed to strengthen ethnic occupational and social stratification. We show how, as a result of such policies, the occupational specialization of a minority group can continue to have a lasting impact on economic development despite centuries of (de jure and de facto) discrimination and social conflict.

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. Section 6 presents our leading mechanisms of effects on firms and networks. Section 7 explores alternative mechanisms. Section 8 provides a discussion of the main findings. Section 9 concludes.

2 Chinese Communities on Java

In this section, we describe the historical 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.

[6] A closely related study is Kuipers (2022) who studies the political effects of opium concessions on Java.

2.1 Chinese Immigration to Indonesia

Pre 15th Century Chinese Emigration. The historical evidence suggests that there was somewhat limited Chinese emigration to Java before the 15th century.[7] 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 encountered very few Chinese on shore in the ancient port towns of Tuban and Gresik in Eastern Java.

One of the first historical records of major, state-sponsored, Chinese voyages into Java was that of Admiral Zheng He (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, from Southeast Asia to East Africa. These expeditions established persistent diplomatic, commercial, and migrant networks in Southeast Asia—in 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.[8] 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, prior to the expedition, these sites did not appear to have warranted mention in key historical documents such as the Zhu Fan Zhi that was written prior to the expedition.[9] It is further crucial that the entire 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. We further discuss how our IV builds on these historical accounts in Section 5.

Post 15th Century Chinese Emigration. 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

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

[8] Those 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.

[9] This work was written in 1225 by the superintendent of customs in Quanzhou, China (a then-major port on the Southern coast of China), Chao Jukua and provides fairly detailed descriptions of the Kingdom of Srivijaya and specific locations and anchorages in Sunda, Java, Sumatra and Malaya.

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). However, for much of this period, especially up to the mid 19th-century, imperial Chinese edicts nominally prohibited emigration and "few Chinese who went to the South Seas ... had any intention of establishing residence" (Wu and Wu, 1980). Emigration accelerated sharply only from the second half of the 19th century, driven by the opening of treaty ports after the Opium Wars, the Taiping Rebellion (1850–1864), and recurring famines. These waves intensified further in the early 20th century after the fall of the Qing dynasty (1911) (Wu and Wu, 1980).

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 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 (Onghokham, 1989). By 1930, the bulk of the Chinese population was the product of this relatively compressed period of large-scale emigration. 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).

Selection and Composition of Immigrants. While much of the literature has emphasized the unskilled nature of migrant labor from China 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 suggests that there was a strong demand for skilled workers (Bassino and van der Eng, 2021).

This was especially true for ethnic Chinese emigration flows to Java in the early 20th century. Some of these Chinese 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)—with many of them establishing themselves as skilled artisans and businessmen (Booth, 1998; Rush, 1991; Die, 1943).[10] Quantitatively, 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.

With regards to historical capital stock, the high costs of migration suggests that many, if not all Chinese

[10] 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.

emigrants, would have hailed from extended families that had the ability to put together sufficient capital to afford transportation fees to the Dutch East Indies.[11] By the 20th century, qualitative accounts suggest that the Chinese emigrant community (and "other foreign Asians/Easterns") as a whole had established businesses and had higher average incomes than indigenous Indonesians (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 in the region (Pan, 1999, p. 152). Hence, throughout Dutch colonial rule, Chinese communities played an important role as intermediaries between the local population and the Europeans (Houben, 2002, p. 74). The Chinese were active as tax farmers and middlemen in retail trade; entrepreneurship; and informal capital markets with much of this being due to their detailed local knowledge and extensive kinship-based organizations. [12] Hence, the Dutch relied heavily on the Chinese to perform economic functions. This led to a widespread, persistent perception among local Javanese of the Chinese as exploitative middlemen (Suryadinata, ed, 2004).

The economic activities of the Chinese communities thereby played an important role in enabling both the Dutch colonial state and commerce to flourish. However, various restrictions on the movement of Chinese were imposed to maintain ethnic divisions and prevent the Chinese from becoming too economically powerful. These systems remained in place until the late 19th century.

By 1935, the occupational structure of the ethnic Chinese continued to differ substantially from the local Javanese. Ethnic Chinese in Java worked in three main occupations: (i) trading (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%).[13] Importantly, only 5.1% of Chinese involved in trade were involved in banking and finance.[14]

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 to 2 percent of the total Indonesian population of 230 million. Despite their small share, however, the

[11] 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.

[12] A key manner in which Chinese 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 percent of the government revenue in the 19th century (Houben, 2002, p. 74).

[13] Small-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%).

[14] Notably, 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).

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.[15]

The 1997–1998 Crisis and Anti-Chinese Violence As targets of discriminatory policies since independence, the ethnic Chinese have long been vulnerable to periodic episodes of social conflict. In particular, the coincidence of the 1997 economic crisis and the fall of Suharto's regime led to riots and targeted assaults against ethnic Chinese across Indonesia in May 1998, resulting in the deaths of at least a thousand ethnic Chinese (Purdey, 2006). The riots subsided in 1999, and the interim President, Abdurrahman Wahid ("Gus Dur"), reversed most discriminatory policies against the Chinese minorities in 1999–2001. The crisis had lasting consequences for the economic role of ethnic Chinese in Indonesia: many ethnic Chinese entrepreneurs and their capital fled to Singapore, Hong Kong, and other countries in the region (Borsuk, 2014). These events will be central to our empirical analysis in Sections 5 and 6, where we leverage the crisis as a source of temporal variation to better distinguish between time-invariant locational advantages and the role of Chinese social and commercial networks.

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. Table 1 displays summary statistics of key variables and Section A.1 reports additional details.

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

[15] Authors' own calculations.

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).[16]

We do not use ethnic Chinese shares in 2000 as our main explanatory variable, as the identification of ethnic Chinese in 1930 was possibly more accurate than in 2000. First, post-independence Indonesian governments introduced widespread discriminatory policies against ethnic Chinese, leading many to fear openly identifying as ethnic Chinese. 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 undermining the potential veracity of the ethnicity question on the 2000 census.[17]

Beyond measurement, the 1930 census is a potentially meaningful cross-section of Chinese settlement. As described in Section 2, prior to the mid-19th century, the ethnic Chinese population on Java was quantitatively small and largely confined to scattered merchant communities (Wu and Wu, 1980). Instead, the great bulk of Chinese emigration was compressed into a relatively short period driven by successive upheavals in China (e.g. the Opium Wars, Taiping Rebellion, the fall of the Qing dynasty in 1911) and the rapid expansion of the Dutch colonial economy. Much of the spatial distribution of the Chinese population in 1930 was therefore the product of a relatively compressed period of large-scale emigration from the late-19th century, much of which was based on pre-existing trans-maritime kinship networks which never effectively recovered with the onset of World War II.

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.[18]

3.2 Additional Data Sources

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 sim-

[16]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).

[17]Figure A.3 plots ethnic Chinese shares in 2000 against 1930 and shows (i) persistence within districts ($r = 0.655$) but (ii) across districts, a general decline potentially consistent with ethnic Chinese riots in 1998 decreasing individual's willingness to self-identity as ethnic Chinese.

[18]An 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.

ilar 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).[19]

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 7 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 ($\geq$ 20-30 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 (1830-1870s) (Dell and Olken, 2020). We use data on the location of opium concessions granted to the Chinese (1886) (Kuipers, 2022). In both cases, we calculate the distance of each district to these locations. Furthermore, we include information on the number of Chinese officers by district from Cribb (1997). To

[19] Unfortunately, the lack of ethnic identifiers prevents us from computing measures of inter-ethnic inequality.

account for historical market access and infrastructure, we use the digitized colonial railroad network (1925) as well as the Great Postal Road (1808-1811, the main colonial thoroughfare across Java). 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 across the entire Java. In Section 5 we further assess causality using an IV approach that leverages the locations of historical Chinese-expedition landing sites in Central and East Java that subsequently silted up. We begin by splitting all districts on Java 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, consumption, 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

$$Y_i = \gamma_p + \beta C_i + X_i'\delta + Z_i'\gamma + u_i, \tag{1}$$

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, expressed in percentage from 0 to 100. γ_p denotes province fixed effects,[20] and u_i denotes the error term, estimated with adjustment for clustering at the regency level.[21]

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,

[20] Java comprised three provinces during the colonial period: West, Central, and East Java.

[21] There are 54 regencies in the main sample. Clustering by the coarser level of residency (with very few clusters), does not affect the qualitative message. Results available upon request.

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 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 household consumption (averaged across 2000-2011). Panel A presents the estimates without the inclusion of any controls. Panel B includes provincial fixed effects and Panel C includes geographic controls for elevation, longitude, the logarithm of distance to the Northern coast, the logarithm of total area (1930), the logarithm of total dry agricultural land, total wet agricultural land (both in 1963) and foreigner shares (1930).[22] Finally, Panel D includes the quadratic polynomial of logarithm of distance to each major provincial capital in Java

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

(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).[23]

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 0.05 to around 0.15. This means that districts with a one percentage point (0.1p.p.) difference in the Chinese population share (a 100 (10) percent increase from a mean of around 1 percent) differ in population density by between 5 and 15 (0.5 and 1.5) 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 (0.1p.p.) increase in the Chinese population share increases the share of villages designated urban between 4 and 6 (0.4 and 0.6) percentage points. The estimates are again precisely estimated across the different specifications. In Columns (4) and (5), we also consider the impact on nightlight luminosity and the average level of consumption. For both outcomes, we continue to find positive effects, although the precision is more sensitive to the inclusion of controls. The impact on consumption is more modest and ranges between a 3 and 6 (0.3 and 0.6) percent change for a one percentage point (0.1p.p.) difference in the Chinese share. In particular, the results in Panel D indicate that controlling for distances to (historical) urban centers will be crucial for any identification strategy that seeks to identify the effect of ethnic Chinese on contemporary development. 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. Thus far, most of the estimated effects are robust to including the various controls. However, the inclusion of controls in Panel C and D of Table 2 tends to reduce the estimated impact. This is consistent with unobserved confounders explaining the estimated coefficients. To explore this in more detail, we follow Cinelli and Hazlett (2020) and calculate the robustness value. This approach enables us to quantify the explanatory power that 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

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

effects are unlikely to be driven by unobserved confounders. Except for consumption, we find robustness values ranging between 30 to 40 percent. As a result, unobserved confounders that explain between 30 and 40 percent of the variation in the Chinese population shares and the various outcomes can account for a true effect size of zero. We believe the omission of such confounders from Equation 1 is unlikely, as it 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.

In addition, an important potential 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 colonial presence, which could have affected economic development in a myriad of ways (see, e.g., Dell and Olken, 2020). In Table A.1, we therefore test for robustness to the addition of key controls such as whether a district contains (i) a 19th century sugar factory; (ii) distance to colonial railroads (1925); (iv) distance to the Great Postal Road (1808-1811); (v) distance to opium concessions (1886). All in all, we do not find strong evidence that these factors explain the association between economic development and the location of Chinese communities. Importantly, our IV results, remain robust to the iterative inclusion of these additional controls (discussed in Section 5).

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 1p.p. change in the Chinese population share is a relatively large difference, accounting for an approximately 100 percent increase from the sample mean. Moving forward, we continue to report estimated effects in terms of a 0.1p.p. increase in Chinese shares. In sum, 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 even without

Chinese immigrants. While the robustness checks in the previous section suggest that sorting is unlikely to account for the entire effect, we cannot rule out the importance of sorting in driving these patterns. 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 documented landing sites of Admiral Zheng He's expeditions in the 15th century—one of the largest maritime expeditions launched by Imperial China—located along the Muria Straits in Central and East Java. Specifically, we use distance to the expedition's landing sites along the Muria Straits as an instrumental variable (IV) for Chinese population share in 1930. Crucially, our IV leverages on the fact that these landing sites along the Muria Strait were transient anchorages on a now-vanished waterway, distinct from the established pre-colonial trade ports (e.g. Gresik, Tuban).

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, based on historical records, we use distance to landing sites along the Muria Straits as our main IV. We do so for a few reasons. However, before we discuss these, it is worth mentioning that we do not claim that the choice of these landing sites arose out of Zheng He's lack of prior knowledge of where to land his crew. Rather, the validity of our IV rests on the claim that even if these sites were known to Zheng He's expedition—a definite possibility given the history of trade and commerce between China and Indonesia—they were unlikely to have been chosen for pre-determined geographic fundamentals that might be correlated with contemporary economic development.

First, these sites along the Muria Straits appear to have been relatively underdeveloped and historical records, to the best of our knowledge, do not mention a strong Chinese presence in these areas, prior to the arrival of Zheng He's crews. This is important given that Zheng He had similarly visited major pre-industrial cities of the Majapahit Kingdom (Mills, 1971) which did not appear to have resulted in a persistent, substantive presence of ethnic Chinese.[24] Furthermore, a key historical document, the Zhu Fan Ji (a 13th century pre-Zheng He, imperial Chinese record of known ports and landing sites in Southeast Asia, including on Java) does not contain, to the best of our knowledge, any detailed record of known landing sites or Chinese presence along the Muria Straits. Last, the size of Zheng He's fleet;[25] local legends of post-

[24] Distance to these cities might capture potential confounders in terms of pre-industrial state capacity or kingdom centralization. Specifically, these cities include the current-day towns of Tuban, Gresik, Surabaya, and Trowulan (the site of the ancient Majapahit capital itself). We control for polynomials of distance to each of these cities.

[25] 27,000–28,000 officers and soldiers, 62 large ships and nearly 200 medium and small vessels (Huang, 2005)

Zheng He Chinese settlement at Semarang,[26] a key node of Chinese settlement in our baseline sample; and a sizeable Chinese presence in these areas by 1930, suggests that some part of Zheng He's crew were likely to have settled in these areas as they passed through.

Second, that there might exist a potential arbitrariness with which Zheng He picked these sites given the lack of sufficient technology to determine longitudes in maritime navigation until the 19th century (Pascali, 2017; Rivers, 2012). This possibility is 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 Tuban and Gresik. We leverage this possibility in an alternative IV strategy (Section 5.3.1) where we compare the effect of proximity to landing sites on the Muria Strait with proximity to hypothetical landing sites near Cirebon, the best deep-water port in Java that Zheng He might have alternatively landed at.[27]

Third, and most importantly, the Muria Straits silted up, turned into land in the 17th century, and has remained inaccessible to maritime traffic for nearly 3 centuries. These initial landing sites therefore are unlikely to have retained locational fundamentals particularly amenable to trade in subsequent centuries, strengthening the case of the exclusion restriction.[28] In this aspect, our identification strategy echoes Jha (2013) who uses medieval natural harbors that subsequently silted up as an instrument for medieval locuses of Muslim trade. 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. We further probe relevant instrumental variable assumptions in Section 5.3.

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 following model,

$$Y_i = \alpha_p + \rho C_i + X_i'\omega + \varepsilon_i, \tag{2}$$

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 fixed effects for Java's five provinces. We control for an extensive set of covariates X_i that may have affected Chinese settlement locations as well as long-term development. They

[26] Huang (2005) (pp.62) further writes of a detailed, local account (based on newspaper sources) of Wang Jinghong, Zheng He's deputy, who on one of their return voyages from East to West Java, suddenly became severely ill. As a result, Zheng He ordered his crew to stop at a river mouth located in Semarang where he found a cave, outside of which he constructed a small hut for Wang's recuperation. After ten days, Zheng He resumed his voyage and left Wang behind together with a small crew; provisions; and a small vessel. Wang eventually recovered but never returned to China; instead leading his crew members to open the largely unsettled frontier in Semarang and settling down as residents. Etymology-wise, Semarang and its largest temple are named after Zheng He in the Chinese language. Separately, Abdurrahman Wahid, the first post-Suharto Indonesian president, has a mixed native and Chinese ancestry, and, according to newspaper accounts, is said to have some ancestors that can be traced back to Zheng He's crews.

[27] In the Annnals of Cirebon and Semarang, Cirebon was highlighted as an equally important site vis-a-vis Semarang.

[28] A case in point is Semarang, Central Java, a key ethnic Chinese settlement in our sample. While an important port in the 18th century, it was not suited for steamships, which had deeper drafts. Its importance therefore diminished sharply in the second half of the 19th century (Onghokham, 1989).

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 an approximation of 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 $\mathbb{E}\left(\varepsilon_i \mid D_i, X_i\right) = 0$.

These assumptions rest on two key features of the context. First, there is likely to be a high degree of persistence in the settlement patterns of Chinese across Java. We hypothesize that early settlement in a given area reduced the cost of settlement for subsequent waves of migrants. This mechanism is consistent with research emphasizing the importance of kinship ties and pre-existing clan networks in Southeast Asia, which were important determinants of where Chinese immigrants chose to settle (Setijadi, 2023; Hup and De Zwart, 2024). Second, our identification strategy assumes that the locations in which Chinese initially settled were not systematically more conducive to economic growth in the era of rapid economic growth in the second half of the 20th century. While the locations were more accessible from the sea in the past, sedimentation and changes in the location of the coastline have made this less likely. Moreover, the early settlement decisions were made under conditions of limited information about the economic potential of the inland areas. Lastly, we control for a range of observable location fundamentals today, such as agricultural productivity.

Throughout, we exclude West Java (which includes present-day Jakarta, the capital of Indonesia) for a number of reasons. First, there is little record of Zheng He's expedition stopping along the coast of West Java—much of West Java including Batavia (present-day Jakarta) was settled only in the early 17th century (Ju-kua, 1911; Huang, 2005). Second, selective sorting of emigrants into West Java are likely more endogenous given that the Dutch awarded monetary incentives to encourage ethnic Chinese and other trading communities to settle in and around Batavia (Huang, 2005). Third, Batavia today is a key port through which the majority of Indonesia's shipping traffic passes through. Hence, it is arguably more difficult to identify the effects of Chinese settlements on contemporary development in much of West Java.[29]

[29]Nonetheless, Table A.2 shows IV results on key outcomes of contemporary development. Overall, while the IV is, expectedly, much weaker in this extended sample, the estimates remain qualitatively similar.

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).

As a preliminary test of the exclusion restriction, we re-estimate our first-stage regressions using a host of alternative observable outcomes. A lack of statistical significance between our instrument and these outcomes would strengthen the plausible exogeneity of our IV. To that end, we regress the shares of other ethnic groups, represented in the 1930 census as Native, European, and other foreigners, and sugar factory presence—a potentially key, deep-rooted colonial-era determinant of contemporary development (Dell and Olken, 2020)—on our instrument. Table A.3 shows that our instrument does not predict the respective shares of all other minority groups nor colonial-era sugar factory presence (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), suggesting that any violation of the exclusion restriction is unlikely to operate through these channels. Subsection 5.3 further assesses the exclusion restriction and other leading concerns in detail.

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 1%, and a 1 standard deviation increase in the ethnic Chinese share is around 1.5%. As such, a one percentage point increase in the Chinese share is substantial. A more appropriate unit of comparison is a 0.1 percentage point increase, representing roughly a 10 percent change relative to the mean. Second, the magnitude corresponds to a low difference in annual growth rates, as these differences likely have accrued over many decades. To illustrate, our IV estimate for population density (0.261, Table 3, Column 2) implies that a 0.1 percentage point higher Chinese share is associated with approximately 2.6% higher population density in 2000. Over a 70-year-period (1930– 2000), this level difference corresponds to an annualized growth differential of about 0.04 percentage points.

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.

Table A.4 tests for the sensitivity of our estimates to controlling for important historical determinants of colonial-era development. We do so by iteratively controlling for the presence of 19th century sugar factories; distance to colonial railroads; distance to the Great Postal Road (a main thoroughfare during the Dutch Colonial era); and distance to opium concession boundaries. Throughout, results remain largely similar, strengthening the causal interpretation of our estimated effects.

5.3 Assessing the IV Exclusion Restriction

A central concern with our IV is that distance to the Muria Straits may capture persistent geographical advantages of coastal areas that would have independently promoted development regardless of subsequent Chinese settlements. We address this through a series of tests that collectively strengthen the credibility of our exclusion restriction. Specifically, while these sites have been inaccessible for 3 centuries (and might possibly have been arbitrarily chosen by Zheng He's crew), distance to these sites may incorporate information that correlates with other potential determinants of development. It is thus important to try to isolate the plausibly exogenous source of variation in Chinese shares, as much as possible.

To that end, in addition to the placebo tests in 5.1.1 we show that our results are robust to: (i) a counterfactual approach where, informed by the historical literature, we compare outcomes of districts that are equidistant from the Muria Strait landing sites with that of counterfactual sites where Zheng He's expeditions could plausibly have landed; (ii) a generalized re-centered IV approach using the universe of possible counterfactual "Muria Straits" along the northern-most coast of Central and East Java (Borusyak and Hull, 2023); (iii) restricting our sample to districts located along a narrow segment of the north coast.

5.3.1 Counterfactual Landing Sites

A concern with our IV is that distance to the Muria Strait may capture proximity to the Northern coast more broadly, rather than ethnic Chinese settlements arising from Zheng He's landing sites. Specifically, while these sites have been inaccessible for 3 centuries, distance to these 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. To do so, we implement a counterfactual approach that considers various sets of landing sites that could have been picked by the expeditions (but which were not) and compare the hypothetical consequences of distance to those landing sites vis-a-vis distance to Zheng He's actual landing sites.

Cirebon Counterfactual. We construct counterfactual landing sites around Cirebon—a major deep-water port located at the westernmost edge of Central Java that has, due to it's (continued) port status, historically attracted a number of Chinese immigrants but was not a Zheng He landing site (Salmon and Lombard, 1981).[30] We then use stratified comparisons to absorb variation common to any Northern-coast location (See Appendix A.2 for estimation details). Table A.5 shows that our results hold within matched strata of distance to actual landing sites along the Muria Straits vis-a-vis distance to fake landing sites centered around the Cirebon coastline, suggesting that observed positive effects on development in our original IV specification is possibly driven by distance to the actual landing sites (and subsequent ethnic Chinese settlement) rather than other unobserved determinants of development.

Systemic Counterfactual: Riverine access to Majapahit. We complement the Cirebon approach with a systemic approach that algorithmically identifies counterfactual landing sites that were equally accessible to the centre of the Majapahit empire—one of the largest pre-industrial empires on Java. While coastal developments related to the empire were mostly centered on the Eastern coast of Java and a key historical document, the Zhu Fan Zhi (a 13th century pre-Zheng He, imperial Chinese record of known ports and landing sites in Southeast Asia, including on Java) has no record of known landing sites along the Muria Straits, concerns might remain that Zheng He chose to land at these sites based on riverine access to a known hinterland. Using the universe of Java's river network, we identify all north-coast river mouths that are connected to the Majapahit centre via navigable river routes (See Appendix A.3 for further details). Next, for each district, we calculate the distance to the nearest such counterfactual landing site and construct a continuous IV which takes the value of the difference between the log distance to the Muria Strait (our main IV) and the log distance to the nearest counterfactual site. Hence, this instrument captures variation in relative proximity to Zheng He's actual landing sites versus plausible alternative sites.

Table A.6 presents results. Panel A (B) restricts our sample to districts within 25km (50km) of the northern coastline (i.e. for which subsequent Chinese settlement from initial landing sites was most plausible). The estimated effects on population density in 2000 and household consumption remain positive and statistically significant across both panels albeit we lose precision on urbanization and light intensity in some specifications due to the smaller number of observations.

Recentered IV. Another potential concern is that, even if the initial location of landing sites along the Muria Straits is plausibly exogenous, those sites are still constrained to be on a specific segment of the Northern coastline of Java. 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 similar confounding factors as mentioned above, that could have generated observably similar developmental effects even in the absence of Chinese communities.

[30] See Salmon and Lombard (1981) on the "Chinese chronicles of Cirebon and Samarang".

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 $\mathbb{E}\left[\ln\left(\mathrm{D}_{pd}\right)\right]$ 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 $\mathrm{D}^{r}_{pd} = \mathrm{D}_{pd} - \mathbb{E}\left[\mathrm{D}_{pd}\right]$. This IV is then applied in the same specification as in Equation 2. To compute $\mathbb{E}\left[\ln\left(\mathrm{D}_{pd}\right)\right]$, we construct alternative sets of landing sites in a generalization of our construction of counterfactual landing sites around Cirebon. To do so, we start from any arbitrary point on the segment of the Northern coastline spanning from Cirebon to Surabaya,[31] then taking the average of the distance between each district and the universe of all such counterfactual landing sites, over all counterfactual segments. For intuition, Figure A.2 provides a graphical depiction and Section A.4 provides a detailed discussion of construction and estimation.

The results are shown in Table A.7. Coefficient magnitudes are slightly smaller but quantitatively close to those in Table 3, with a Kleibergen-Paap weak-instrument-robust F-statistic of 12.59. This exercise lends additional credibility to our main interpretations of observed positive development effects as stemming primarily from the greater share of ethnic Chinese communities.

5.3.2 Restricting to Coastal Districts.

In a further test of the exclusion restriction, we restrict the sample to the subset of districts that could plausibly have been landing sites—i.e. districts that would have been equally "at risk of treatment." If persistent coastal advantages drive our results, then restricting our sample only to these districts should attenuate our estimates. Table A.8, Panel A presents our baseline IV results for ease of comparison. Panel B (Panel C) presents (recentered) IV results restricting our sample to districts within 25km of Java's northern coastline. The first-stage Kleibergen-Paap F-statistic is ranges from 25.6-33.4; the estimated effects on population density in 2000, nighttime light intensity, and household consumption all remain positive and statistically significant.

5.4 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 reasons. 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 4 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 pos-

[31] This 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.

itive 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 $\frac{1.91}{8.71} \times 0.277 = 0.0607$). Per a 0.1p.p. higher Chinese share, this translates into approximately 0.6% higher GDP per capita.

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. (0.1p.p.) increase in Chinese shares is associated with a 3.657 (0.37) unit 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).

5.4.1 Chinese Communities through Social Conflicts

The estimated positive effects of Chinese shares over the long-run (1992–2010) potentially mask fluctuations from intermediate episodes of social conflict.[32] As described in Section 2, the 1997–1998 Asian Financial Crisis culminated in large-scale anti-Chinese riots and targeted assaults, causing a sharp disruption to the economic role of Chinese communities across Indonesia.[33]

To examine variation across those periods, we separately regress the annualized growth of nighttime light intensity (i.e., the change in light intensity divided by the number of elapsed years) on 1930 ethnic Chinese shares across three periods of: (i) 1992-1997 (anti-Chinese discriminatory policies under the height of the Suharto regime); (ii) 1997-1999 (the Asian Financial Crisis and targeted violence against the ethnic Chinese); and (iii) 1999-2010 (the recovery period with the abolishment of most discriminatory policies). Throughout, to better ensure that we are comparing areas with similar levels of economic development, we control for the level of nighttime light intensity in 1992; 1997; and 1999 respectively.

Table 5 shows that a higher ethnic Chinese share in 1930 leads to (i) a positive, albeit statistically insignificant effect on the growth of nighttime light intensity from 1992 to 1996; (ii) a substantially negative

[32]The frequency of social conflicts over these different periods are highlighted in Table A.9. 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 (Column 2: 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, potentially because naming those conflicts as such was a taboo before 2000 (Column 4).

[33]While anti-Chinese violence in 1998 was more frequent in Central and East Java—areas with historical Chinese settlements—the intensity of violence was far greater in more modern commercial centers such as Jakarta and its surrounding environs (e.g. Bekasi). Jakarta's single episode (May 13) resulted in 1,188 deaths and 4,123 shops destroyed, accounting for 97% of all fatalities (UNSFIR). Bekasi, another rapidly urbanizing area where the Chinese had relatively shallower roots, saw 3 killed and 567 shops destroyed. By contrast, the most severe episode in a key historic settlement of Central Java (Surakarta, 33 killed, 280 shops) was an order of magnitude smaller than Jakarta. Although data limitations prevent us from probing these further, these descriptive patterns appear to be consistent with the view that relatively deeper historical integration and complementary economic roles of the ethnic Chinese in Central/East Java may have moderated the intensity of anti-Chinese violence (Jha, 2013). Textual analysis of early 20th century Chinese firm registration data supports this: In Central and East Java, the Chinese production sector was diversified across sugar processing (18.3%), manufacturing (16.2%), and ice/mineral water (18.3%) — activities with potential local employment spillovers. In contrast, in West Java (including present-day Jakarta), Chinese production was dominated by plantations (54.5%); and Chinese non-trade middlemen sectors had a relatively higher proportion of firms in private estate management (24.6% in West Java vs 4.5% in Central and East Java) that potentially extracted corvée obligations and harvest taxes from the local population.

growth effect from 1997 to 1999; and (iii) a *positive* growth effect from 1999 to 2010. The magnitudes of the corresponding estimates on annual growth are larger than the estimate in column (5) of Table 4, showing the importance of accounting for fluctuations in economic growth across intermediate periods of social conflict. Table A.10 further controls for pre-crisis population density (1996) and shows that negative effects in 1997-1999 remain robust and, in fact, become even larger in magnitude. This suggests that observed negative effects are unlikely to be a result of areas with a higher share of ethnic Chinese being more economically developed in scale and/or intensity to begin with—before the onset of the Asian Financial Crisis—and hence, suffering disproportionately negative effects from the crisis.

Taken together, these results provide novel evidence of how changes in district-level economic growth across Indonesia appear to have been closely tied to a sharp discontinuity in the political and social position of the ethnic Chinese. More broadly, these results provide evidence of how discriminatory periods, where a minority is socially persecuted, can lead to strongly negative overall economic outcomes, but effects can reverse and became substantially positive upon the abolishment of such policies. Such reversals are broadly in line with the extant literature on Jewish firms in France (Do et al., 2024) and rich families in China (Alesina et al., 2020).

More importantly, they hint at the potentially important role of ethnic minority Chinese social and economic networks as a key mechanism of persistence for observed positive development effects. Looting and violence against Chinese shops and trading houses might have temporarily reversed these effects, but, in the medium to long-run, some of these networks might have reconstituted in the post-conflict period, leading to some recovery in economic benefits. We further probe this possibility using contemporary firm-level data and historical Chinese business registry data in Section 6.

5.4.2 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 the Chinese communities on the structural transformation of the Indonesian economy. We explore this possibility in Table 6, 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. (0.1p.p.) increase in ethnic Chinese shares leads to a 10.315 (1) 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. (0.1p.p.) higher ethnic Chinese shares in 1930, has on average, 5.8 p.p. (0.58p.p.) more individuals working in the manufacturing sector, and 4.53 p.p. (0.45p.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 Anderson-Rubin test indicates

positive 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.

6 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).[34] 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 manufacturing firms—key drivers of industrialization—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: 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 2013 (the last year of available data).

Last, we use the Dutch Registry of Businesses to identify historical Chinese firm presence and suggestive evidence of ethnic Chinese network effects from the early 20th century—both of which might explain observed contemporary positive effects.

6.1 Chinese Population Shares and Firm Performance

Table 7 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

[34]Case 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 parat 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).

sale shown in column 2 is both statistically and economically strong. 1p.p. (0.1p.p.) increase in the share of Chinese is estimated to imply an increase of 53% $(= \exp(0.426) - 1)$ (4.4%) 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 firms in Chinese communities do not 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. Echoing the results on growth in nighttime light intensity (Table 5), firms in Chinese communities are larger (Table A.11) and grow faster (Table A.12) during the pre- and post- crisis period (Columns 1, 3 and 4). However, during the crisis period from 1997 to 1999, they suffer more than firms elsewhere, and experience little positive and even somewhat negative effects (Columns 2 and 4). Strikingly, however, after the crisis, firms in Chinese communities appear to grow faster and regain their cross-sectional advantage (Column 3).

6.2 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 8. There appears to be little evidence of favorable state-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 substantial: private-source investment is 21 p.p. higher (as a share of sales) (2.1 p.p. per 0.1p.p. of Chinese share). 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 ac-

counts of the Chinese in Southeast Asia, especially in comparison with the Jews in Europe (Suryadinata, ed, 2008; Suryadinata, 2015). Hence this result is potentially 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.

Given increases in firm-level private financing, do we also observe general increases in the financial access of households? The latter might be another possible mechanism of persistence akin to much of the literature on Jewish-minority presence and banking access (see e.g. Pascali, 2016). Table A.18 shows that households in places with more Chinese presence have no advantage in financial access (as measured across all categories from the household survey *Susenas*).[35] There thus appears to be a clear distinction between the positive effects of Chinese communities on private, firm-level financing networks and that on household-level financial access.

Investment Dynamics. Table A.13 further focuses on the variation of firms' private sources of investment over the three periods centered around anti-Chinese violence in 1998. The advantages in terms of private investment created in Chinese communities were particularly strong in the first period before the 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.[36]

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, some of which appear to have been disrupted by targeted ethnic violence and institutional shocks.

Firm Employment, Wages, and Labor Productivity Dyanamics. Stronger investment among firms in Chinese communities should also produce advantages on other dimensions. We focus on labor advantages in Tables A.14, A.15, and A.16. Chinese share in 1930 generates positive effects on total employment; total wage bill; and average wages during the pre- and much of the post-crisis period. We find similar flunctuations in the effect of Chinese shares on labor productivity (Table A.17). The concentration of positive effects in terms of higher wages and labor productivity suggests that firms in areas with larger Chinese commu-

[35] If anything, households from those districts are at a small disadvantage: being 2.5% less likely to have access to business loans; and 2.6% more likely to sell their belongings in financial distress (both numbers significant at 5% by the AR p-value) (0.25% and 0.26% respectively per 0.1p.p. of Chinese share).

[36] This is consistent with two possibilities: (i) measurement-wise, post-crisis we only observe one more year of data on investment sources in 2006; (ii) after the crisis, 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).

nities might have been more likely to pursue capital-intensive, skill-biased production strategies vis-a-vis labor-intensive growth.

Strikingly, we again see a largely negative effect of Chinese shares during the crisis period of 1997-1999 which appear to have wiped away much of the positive effects accrued to districts with larger Chinese communities (Columns 4, Tables A.14–A.17). However, apart from employment (Table A.14) which, like firm sales, does not appear to have exhibited much of a recovery, we see a recovery in positive effects of Chinese shares for all other variables in the post-crisis period.

Interpretation of Results. 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 (the private financing of) 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 contemporary 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.

To further probe the potential importance of these networks, we next turn to examine early 19th century business registration data from which we can explicitly identify firms under ethnic Chinese ownership and (clan) networks based on textual analysis of director names.

6.3 Historical Chinese Firms: Social and Economic Networks

To trace the roots of these contemporary firm advantages, we use three waves of post-1930 registration data of historical Chinese firms from the Dutch Registry of Businesses (1891-1940). This was a period of rapid economic liberalization, during which Chinese communities and entrepreneurs across Java diversified their economic activities (Zanden and Marks, 2012). The key advantage of this dataset is that we are able to identify Chinese-owned firms and the full names of ethnic Chinese directors to construct proxies for social and economic networks. We describe this dataset in detail in Appendix A.1.

Importantly, a number of founders and predecessors of prominent, contemporary Indonesian Chinese

conglomerates appear in this dataset. For example, firms registered under *Oei Tiong Ham*, the sugar king of Java are present (Ham, 1989). The parent firm, *Kian Gwan Handel* was eventually nationalized and became a food production, state-holding conglomerate that continues to exist today. Similarly, *Sampoerna Handel* (a single 1935 firm that manufactured hand-rolled cigarettes) is today a $7.4 billion Philip Morris subsidiary in Indonesia that remains under the control of the initial founder's descendants.

To measure historical Chinese firm development, we construct the number of Chinese firms; average firm equity; the number of middlemen vis-a-vis production firms at the district-year level. This decomposition allows us to test whether the instrument's variation operates specifically through an increase in middleman commercial networks. To measure networks, we use textual analysis to construct two measures of: (i) *Dialect HHI*: a Herfindahl-Hirschman Index that measures dialect (clan-based) concentration of Chinese firm directors at the district-year level (Chen et al., 2022b; Greif and Tabellini, 2017); (ii) *Top-3 Share (Strengh of Hub-Surname Clan Links)*: the share of firms whose director shares the same surname as one of the top-3-surname-clans in the three largest hub cities on Java (Batavia, Semarang, and Surabaya) (Macauley, 2021). We further describe variable construction in Appendix A.6.[37] We interpret a higher *Dialect HHI* as capturing a greater concentration and network strength of Chinese dialect groups within a district; and a higher *Top-3* share as possibly suggestive of stronger commercial networks between district-level firms and same-surnamed director-owned firms in hub cities.

Table 9 presents results. Panel A shows that a higher Chinese share predicts significantly more Chinese firms (Column 1) and higher average equity (Column 2), with the effect operating entirely through middleman firms (Column 3) and not production firms (Column 4). Panel B examines the impact of Chinese shares on networks. Higher Chinese shares predicts a lower diversity in dialect group ownership of firms (Column 1) and higher hub-clan concentration among all firms (Column 2) which appear to be concentrated only in middleman firms (Column 3), but not among production firms (Column 4). These results suggest that the observed positive effects on both contemporary firms and overall development might, in part, stem from the greater (historical) specialization of Chinese communities in middlemen firms and commercial intermediation that are more well-connected to cross-regional (dialect and surname) clan networks. An alternative interpretation is that a higher *Dialect HHI* might capture lower competition among dialect groups and/or greater monopolistic power. This interpretation, however, would appear to be inconsistent with observed positive long-run economic development effects.

Taken together, positive effects on Chinese firm development would appear to precede the period of accelerated growth in the 1960s. The positive effects on contemporary economic development could therefore be driven by both the Chinese communities themselves since 1930, and the increased economic development and economic networks they had generated even before independence.

[37] Surname clans are typically a subset of dialect-group clans and both are likely to have operated in colonial Java.

7 Alternative Mechanisms

Next, we examine several leading 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.

7.1 Negative Spillovers to Neighboring Districts

In Table A.19, 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.

7.2 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).[38] Table A.20 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.[39]

7.3 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 international 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

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

[39] These 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.

(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 7 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 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.5). We do not find evidence of this channel (Table A.21).

Third, we consider whether Chinese communities receive more favorable investment and aid, 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 6), international investments and aids from China were unlikely to have been substantial, since China was still quite underdeveloped (Dreher et al., 2019).

8 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).[40] While Javanese kinship ties were

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

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 (Onghokham, 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 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.

9 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.

Importantly, these positive effects are not static. We document that the economic benefits of Chinese presence collapsed during the 1997–1998 crisis, when anti-Chinese violence disrupted the co-ethnic commercial and financial networks through which these effects operate, but substantially recovered following the restoration of equal rights. This time-varying pattern—echoed in both district-level nightlight growth and firm-level data—is difficult to reconcile with explanations based solely on time-invariant locational fundamentals, and instead point to the resilience of kinship-based minority networks as a key mechanism of long-run persistence.

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.

Akbulut-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.

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 Liong'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 A. Hassan, Lisa Tarquinio, and Stephen J. Terry, "Immigration, Innovation, and Growth," *American Economic Review*, 2024, *114* (7), 2118–2158.

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, Fanghao, Ruichi Xiong, and Xiaobo Zhang, "Familiar strangers: the role of diaspora networks in foreign investment and long-run development," *Rotman School of Management Working Paper*, 2022, (4004159).

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," *Cliometrica*, 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.

Do, Quoc-Anh, Roberto Galbiati, Benjamin Marx, and Miguel A. Ortiz Serrano, "J'Accuse! Antisemitism 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 Prokopczuk, 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.

Ham, Ong Hok, "Chinese Capitalism in Dutch Java," *Southeast Asian Studies*, 1989, *27* (2), 156–175.

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.

Huang, Kunzhang, *Yinni huaqiao huaren shi:(gudai zhi 1949)*, Guangdong Gaodeng Jiaoyu Chubanshe, 2005.

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*.

Ju-kua, Chau, *Chau Ju-kua: His Work on the Chinese and Arab Trade in the Twelfth and Thirteenth Centuries, Entitled Chu Fan Chi*, St. Petersburg: Imperial Academy of Sciences, 1911.

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 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 Distribueerende Tusschenhandel der Chineezen 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://gedeon89.github.io/gedeonlim.com/Lim_Bengkok_edited_pre_sub.pdf.

Macauley, 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, "Banks and Development: Jewish Communities in the Italian Renaissance and Current Economic Performance," *The Review of Economics and Statistics*, 2016, *98* (1), 140–158.

_____ , "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.

Redding, Gordon, *The spirit of Chinese capitalism*, New York, NY, USA: De Gruyter, 2013.

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 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.

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.

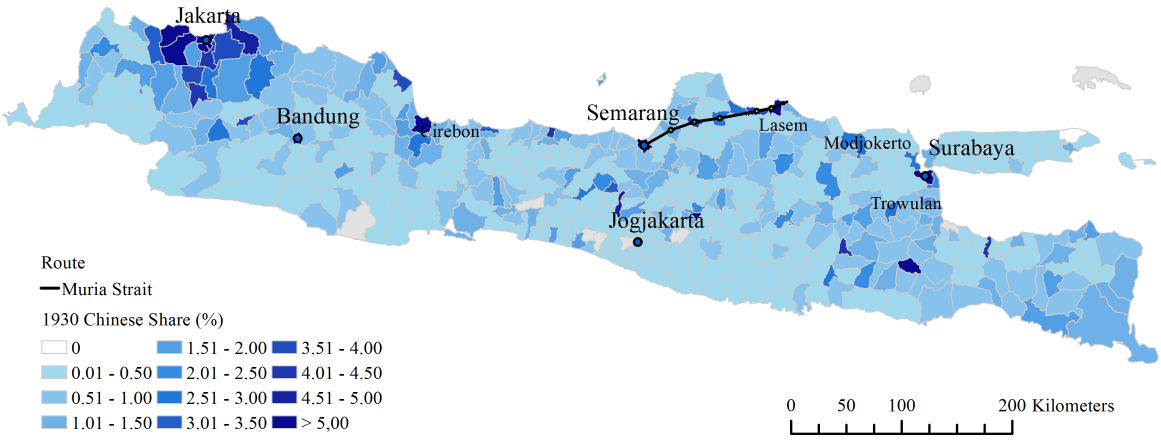
Wu, Yuan-Li, "Chinese Entrepreneurs in Southeast Asia," *American Economic Review: Papers and Proceedings of the Ninety-Fifth Annual Meeting of the American Economic Association*, 1983, *73* (2), 112–117.

Wu, Yuan-li and Chung-Hsi Wu, *Economic Development in Southeast Asia; the Chinese Dimension*, Stanford University, Stanford, California, USA: Stanford University Press, 1980.

Xi, Beijing Daxue Zhongguo Yuyan Wenxue and Yuyanxue Jiaoyanshi, "Hanyu Fangyin Zihui," *Phonetic dictionary of Chinese dialects.) Beijing: Wenni GaigeChubanshe*, 1962.

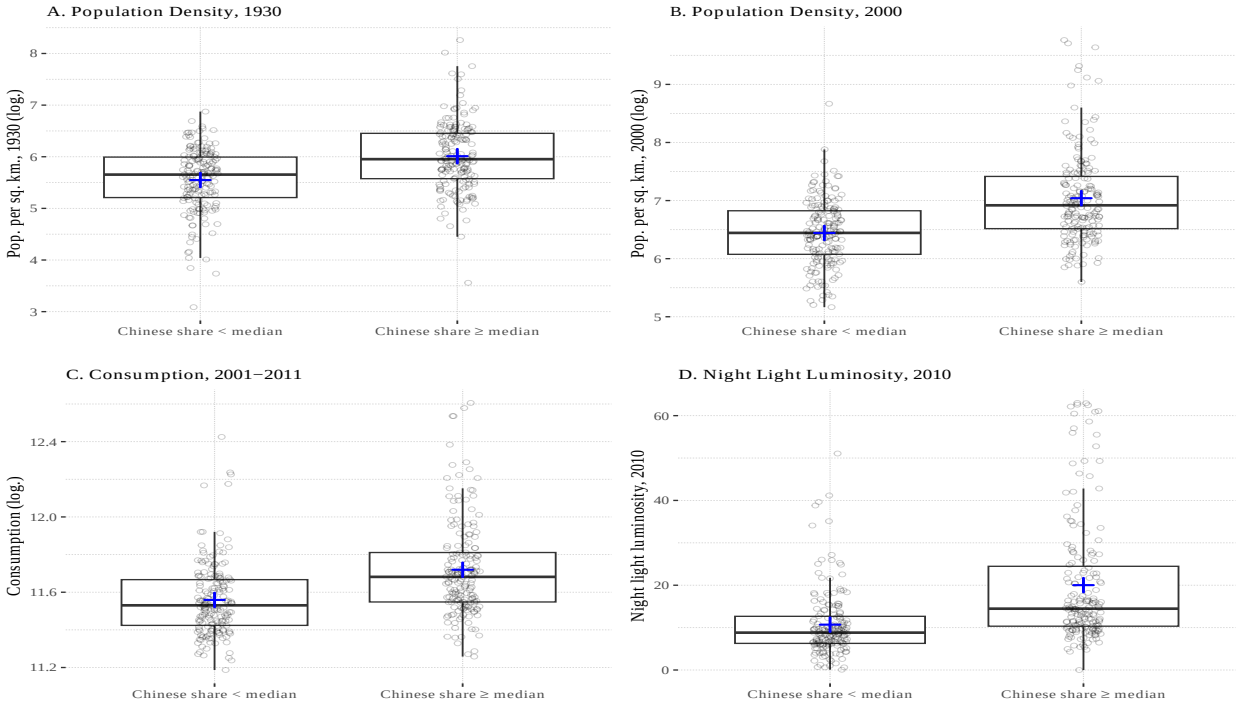
Zhongguo, Fei Xiaotong's Xiangtu, Gary G. Hamilton, and Wang Zheng, *From the Soil: The Foundations of Chinese Society*, Berkeley, California, USA: University of California Press, 1992.

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



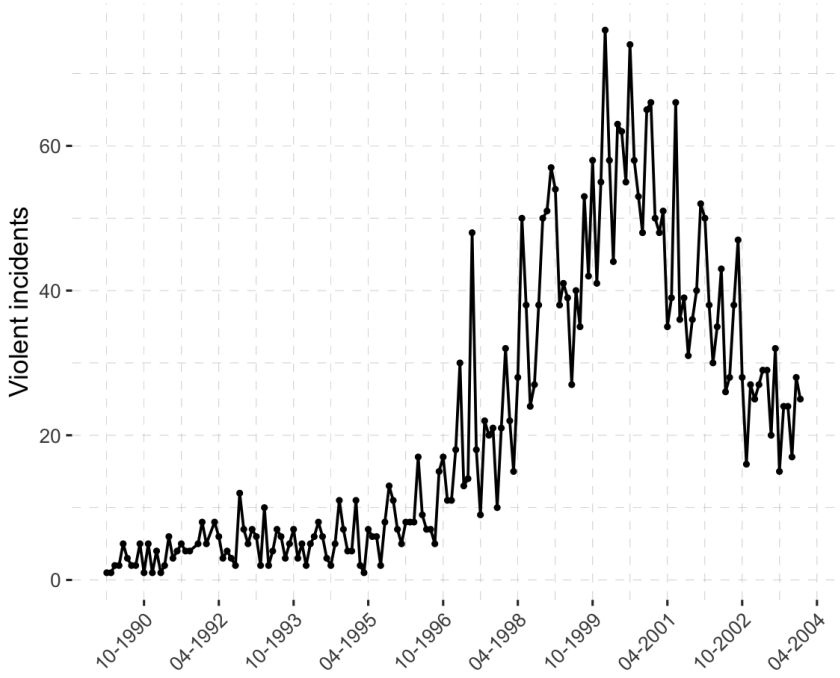
Notes: This figure depicts the underlying variation in our treatment variable: the share of Chinese by district-level across 1930 districts. The solid black line traces the Muria Strait across historical landing sites of Zheng He's expeditions in the early 15th-century. Source: Indonesian Population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

Figure 2: Outcomes for districts with above/below median Chinese population shares in 1930.



Notes: The figure depicts the distribution of the main district-level development 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 household-level consumption averaged across 2000-2011 from annual rounds of the Indonesian Household Consumption Survey (*SUSENAS*). Panel D. depicts the night light luminosity in 2010. The scatter plots are jittered.

Figure 3: 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
Panel A: Predetermined Geographic Variables								
Ln(Distance to Muria Strait)	340	11.7	2.51	-10.1	13.2	11.6	12.0	12.6
Ln(Distance to Coast)	340	6.86	4.67	0	11.2	0	9.53	10.3
Altitude	340	746.7	811.9	1	6000	161	600	1000
Panel B: 1930 Ethnic Shares & Development								
Chinese Share (1930)	340	1.03	1.45	0.0085	11.9	0.25	0.61	1.18
Native Share (1930)	340	98.6	2.32	79.3	100.0	98.5	99.2	99.7
European & American Share (1930)	340	0.29	0.92	0	9.00	0.020	0.085	0.22
Ln(Population Density) (1930)	340	5.78	0.68	3.09	8.26	5.36	5.78	6.21
Ln(Total District Area) (1930)	340	5.62	0.61	3.56	7.84	5.20	5.63	5.95
Panel C: Intermediate Development								
Ln(Dry Land Area) (1963)	340	10.4	0.51	8.64	11.7	10.3	10.6	10.8
Ln(Wet Rice Area) (1963)	340	10.4	0.48	7.92	11.3	10.1	10.4	10.7
Ln(Distance to Colonial Railroads)	340	8.46	1.27	2.06	10.8	7.78	8.65	9.46
Panel D: Contemporary Development								
Ln(Population Density) (2000)	340	6.72	0.72	5.17	9.71	6.28	6.67	7.08
% Urbanization (2010)	340	28.1	24.4	0	100	9.09	21.8	40
Light Intensity (2010)	340	15.0	11.9	0	63	7.82	11.0	16.9
Consumption (2001–2011)	340	11.6	0.23	11.2	12.6	11.5	11.6	11.7
Nightlights Growth (1992–2010)	340	9.79	5.93	-0.87	37.6	6.26	8.35	11.8

Table 2: OLS: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT: (INCLUDING WEST JAVA)

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)	0.133***	0.209***	6.236***	3.633***	0.060***
	(0.035)	(0.036)	(1.347)	(0.674)	(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

Panel B: Province Fixed Effects

Chinese Share (percent)	0.146***	0.205***	6.193***	3.588***	0.052***
	(0.033)	(0.036)	(1.370)	(0.692)	(0.007)
R-squared	0.237	0.346	0.268	0.370	0.452
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

Panel C: Geographical Controls & Foreign Shares

Chinese Share (percent)	0.063***	0.099***	3.959***	1.888**	0.027**
	(0.017)	(0.029)	(1.418)	(0.762)	(0.011)
R-squared	0.660	0.613	0.427	0.529	0.546
Mean (Dep. Var.)	5.778	6.722	28.079	14.956	11.631
Standard Deviation (Dep. Var.)	0.679	0.716	24.409	11.936	0.229

Panel D: Urban Center Controls

Chinese Share (percent)	0.045**	0.069**	3.579***	1.120*	0.015
	(0.019)	(0.026)	(1.311)	(0.609)	(0.009)
R-squared	0.685	0.677	0.520	0.693	0.688
Mean (Dep. Var.)	5.778	6.722	28.079	14.956	11.631
Standard Deviation (Dep. Var.)	0.679	0.716	24.409	11.936	0.229
Observations	340	340	340	340	340

Notes: The table reports OLS estimates. The unit of analysis is at district level. *Controls* are as described in Section 4.2. Specifically: *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	0.261***	13.276***	5.569**	0.094***
	(0.093)	(0.078)	(4.615)	(2.224)	(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***	-0.158***	-0.158***	-0.158***	-0.158***
	(0.043)	(0.043)	(0.043)	(0.043)	(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: 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**	1.912**	2.327**	5.569**	3.657**
	(0.062)	(0.750)	(1.101)	(2.224)	(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***	-0.158***	-0.158***	-0.158***	-0.158***
	(0.043)	(0.043)	(0.043)	(0.043)	(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 5: 1930 CHINESE SHARE AND FLUCTUATIONS IN NIGHTLIGHTS GROWTH

Panel A: IV Specification

	(1)	(2)	(3)
Dependent Variable:	Light Intensity Growth (1992–1997)	Light Intensity Growth (1997–1999)	Light Intensity Growth (1999–2010)
Chinese share (percent)	0.206	-0.308**	0.213**
	(0.128)	(0.135)	(0.0925)
R-squared	0.521	0.505	0.569
Mean (Dep. Var.)	0.864	-0.020	0.496
Standard Deviation (Dep. Var.)	0.439	0.515	0.414
Kleibergen-Paap WID, F-stat	15.921	19.445	17.397
Anderson-Rubin, P-value	0.202	0.056	0.090

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.139***	-0.131***	-0.140***
	(0.0349)	(0.0296)	(0.0337)
Partial R-squared	0.075	0.037	0.079
Fixed Effects	Province	Province	Province
Number of Regencies	54	54	54
Observations	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 STRUCTURAL TRANSFORMATION

Panel A: IV Specification

	(1)	(2)	(3)
Dependent Variable:	Agriculture Employment Share	Manufacture Employment Share	Services Employment Share
Chinese Share (percent)	-10.315**	5.783	4.533
	(4.158)	(4.706)	(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

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.163*	-0.163*	-0.163*
	(0.088)	(0.088)	(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 7: 1930 CHINESE SHARE AND FIRM OUTCOMES

Panel A: IV Specification

	(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	0.426***	0.232	0.037	-0.494**	0.007	-0.087***	-0.093***
	(0.337)	(0.151)	(0.221)	(0.173)	(0.223)	(0.186)	(0.032)	(0.034)
Ln(Firm Sales)			1.152***	0.611***	1.241***	0.423***	0.088***	0.111***
			(0.030)	(0.037)	(0.057)	(0.071)	(0.008)	(0.009)
Fixed Effects	Province	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	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

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.158***	-0.191***	-0.192***	-0.192***	-0.192***	-0.192***	-0.192***	-0.192***
	(0.043)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(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, 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 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*	21.049*	1.371	-0.293**	-4.159	-0.298	-0.020	-0.001
	(7.660)	(11.699)	(0.958)	(0.133)	(3.032)	(0.239)	(0.129)	(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***	-0.192***	-0.192***	-0.192***	-0.192***	-0.192***	-0.192***	-0.192***
	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(0.049)	(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. $^{*}p < 0.1,\ ^{**}p < 0.05,\ ^{***}p < 0.01$.

Table 9: 1930 CHINESE SHARE, HISTORICAL CHINESE FIRMS AND NETWORKS

Panel A: Firm Scale and Sector

	(1)	(2)	(3)	(4)
Dependent Variable:	Chinese Firm Count	Chinese Ln(Avg Equity)	Ln(# Mid. Firms+1)	Ln(# Prod. Firms+1)
Chinese Share (percent)	0.463***	1.765**	0.462***	0.029
	(0.157)	(0.894)	(0.138)	(0.027)
R-squared	0.497	0.224	0.401	0.567
Mean (Dep. Var.)	0.172	0.641	0.135	0.033
S.D. (Dep. Var.)	0.498	1.601	0.444	0.191
Kleibergen-Paap F	14.587	14.587	14.587	14.587
Anderson-Rubin p	0.029	0.013	0.014	0.358

Panel B: Clan Composition

	(1)	(2)	(3)	(4)
Dependent Variable:	Dialect HHI	Top-3 All Firms	Top-3 Middleman	Top-3 Production
Chinese Share (percent)	0.217***	0.161***	0.157***	0.050
	(0.073)	(0.043)	(0.041)	(0.031)
R-squared	0.258	0.090	0.046	0.289
Mean (Dep. Var.)	0.090	0.040	0.036	0.015
S.D. (Dep. Var.)	0.238	0.174	0.163	0.116
Kleibergen-Paap F	14.587	14.587	14.587	14.587
Anderson-Rubin p	0.001	0.025	0.021	0.214
Number of Regencies	54	54	54	54
Observations	726	726	726	726

Notes: IV estimates. First-stage output surpressed for presentational purposes. District-year panel using post-1929 CBI firm registrations. Panel A: Chinese Firm Count = Ln(1 + number of Chinese firms); Avg Equity = Ln(average deflated equity); Middleman (trade, brokerage, finance etc) and Production (manufacturing, plantation etc) counts = Ln(1 + sector count). Panel B: Dialect HHI = $\sum_d s_d^2$ across Hokkien, Hakka, Cantonese, Teochew; Top-3 = share of firms directed by TAN, OEI, or LIEM clans, decomposed into middleman and production firms. See Appendix A.6 for further details. Number of regencies and observations are identical across all columns. All specifications include year FE, province FE, and geographical and urban center controls. 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 52

A.2 Robustness: Identifying Counterfactual Land Sites around Cirebon 53

A.3 Robustness: Identifying Counterfactual Landing Sites Based on Riverine Access to Majapahit Kingdom 54

A.4 Robustness: Empirical Strategy using Re-centered IV 54

A.5 Calculating Market Access at the Regency Level 55

A.6 Calculating Measures of Ethnic Chinese Networks from Historical Firm Data (CBI) 56

A.7 Additional Figures and Tables 58

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 (classified by middlemen vs production). 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 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 Kuipers (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: Identifying Counterfactual Land Sites around Cirebon

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 consider a set of counterfactual landing sites that were equally likely to have been picked by the expeditions, and then compare the hypothetical consequences of those landing sites (and the counterfactual 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 Java, and has also attracted a lot of Chinese immigrants by 1930 (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, $i \in \{1, \cdots, 10\}$. 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 IV results reported in Table A.5 closely trace the main IV results from Table 3. They thus lend credibility to the main interpretation of the long-run effects of the Chinese communities.

A.3 Robustness: Identifying Counterfactual Landing Sites Based on Riverine Access to Majapahit Kingdom

To systematically identify potential landing sites, we first constructed a river network using HydroRIVERS data from HydroSHEDS. In preventing shorter main rivers from being spuriously connected under a larger tolerance, main river systems were classified as short or long based on a 7.5 km shape length threshold. Segments exceeding this threshold were first connected using a 2 km tolerance. The resulting network was then merged with segments below the threshold and reconnected using a 0.5 km tolerance. This two-stage procedure allowed larger gaps in major river systems to be bridged while minimizing false connections among shorter rivers. The resulting networks were then used to identify river mouths connected to the Majapahit river system through an Origin-Destination (OD) cost matrix analysis in ArcGIS. Majapahit's centroid was specified as the origin, while destinations were limited to river mouths along Java's northern coast, from Cirebon (Western-most coastal city at the border of Central Java) to Surabaya (East Java).

A.4 Robustness: Empirical Strategy using Re-centered 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 $\mathbb{E}\left[\ln\left(\mathrm{D}_{pd}\right)\right]$ 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 $\mathrm{D}^{r}_{pd} = \mathrm{D}_{pd} - \mathbb{E}\left[\mathrm{D}_{pd}\right]$.

To address this concern, we follow Borusyak and Hull (2023) and calculate the *expected* treatment of each district that we denote $\mathbb{E}\left[\ln\left(\mathrm{D}_{pd}\right)\right]$ 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 to 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 5.3.1, 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 $\mathbb{E}\left[\ln\left(\mathrm{D}_{pd}\right)\right]$, from which we calculate the new recentered IV D^{r}_{pd}. This IV is then applied in the same specification as in equation 2.

The results are shown in Table A.7. 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.5 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

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

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,

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

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

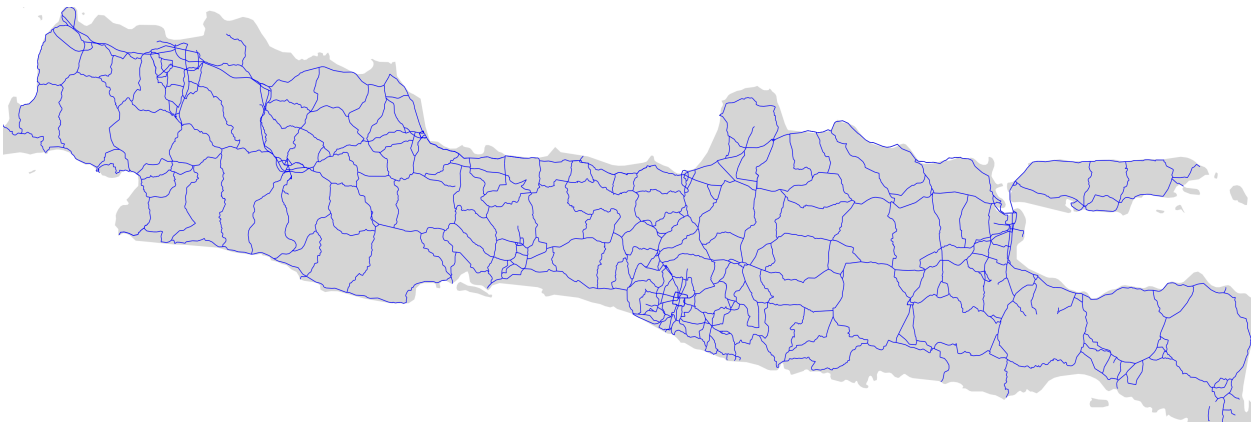


Figure A.1: This figure depicts the Indonesian highway network used to construct the trade cost matrix.

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.6 Calculating Measures of Ethnic Chinese Networks from Historical Firm Data (CBI)

Middlemen Firms We further distinguish Chinese *middleman* firms (trade, brokerage, commission agency, transport, financial intermediation) from that of *production* firms (manufacturing, plantations, sugar factories, mining). Firm classification is based on Dutch-language activity descriptors in the CBI firm name field (*handel* = trade, *fabriek* = factory, *cultuur* = plantation, *toko* = shop, etc.); 1,872 unique firms are classified, of which 73% are middleman and 15% production, with the remainder hybrid or ambiguous. This decomposition allows us to test whether the instrument's variation operates specifically through middleman commercial networks, as predicted by the historical accounts of Chinese as intermediaries between the indigenous population and European colonial firms.

Dialect-HHI. We assign dialect probabilities to each of the 5,063 director records in the Colonial Business Indonesia database using a surname-to-dialect crosswalk constructed from six sources. The primary linguistic reference is the *Hanyu Fangyin Zihui* (HYFYZH, Xi and Jiaoyanshi, 1962), which provides IPA pronunciations of 84 surname characters across 20 Chinese dialect points; we extract five relevant to Southeast Asian Chinese: Xiamen (Hokkien), Chaozhou (Teochew), Meixian (Hakka), Guangzhou (Cantonese), and Fuzhou (Hokchia, grouped with Hokkien). We supplement this with attested Dutch-Indonesian spellings from Suryadinata (2015); romanization variants from online genealogical databases; a probabilistic spelling-to-character crosswalk from Chen et al. (2022a); and the *Daftar Marga Tionghoa di Jakarta* (PMTSI Library, Tangerang), which provides Chinese-character identifications for 168 Indonesian Chinese surnames. Together, these sources yield character identifications and dialect likelihoods for 141 standardized surnames covering 4,776 of 5,063 director records (94.3%). The remaining 287 unmatched records—mostly low-frequency spellings appearing in 1–3 records—are assigned national dialect priors from the 1930 *Volkstelling* (Hokkien 56.6%, Hakka 20.5%, Cantonese 13.9%, Teochew 9.0%).

Dialect assignment is probabilistic rather than deterministic because Dutch colonial romanizations of Chinese surnames are often compatible with multiple dialect origins. We apply Bayes' rule: $P(\text{dialect} \mid \text{spelling}) \propto P(\text{spelling} \mid \text{dialect}) \times P(\text{dialect})$, where $P(\text{dialect})$ is the *Volkstelling* population prior and $P(\text{spelling} \mid \text{dialect})$ is a binary indicator for whether the dialect's phonology can produce the observed spelling, determined by manual review of HYFYZH IPA data. For example, the surname TAN maps to the character *Chén*, whose Hokkien colloquial reading is *tan* and Teochew colloquial reading is *ta*; Hakka and Cantonese produce *ts'in* and *t'n* respectively, which cannot yield "TAN." The posterior is therefore 86% Hokkien and 14% Teochew. A fundamental limitation is that Hokkien and Teochew share identical pronunciations for 48% of surname characters, and Hakka and Cantonese overlap for 24%, imposing irreducible ambiguity on any classification from romanized spellings alone.

From these director-level dialect probabilities, we construct district-level measures by averaging across firms within each district (or district-year cell for panel specifications). Our main measure is the *Dialect HHI*, defined as $\sum_{d \in \{\text{Hok, Hak, Can, Teo}\}} s_d^2$ where s_d is the expected share of firms whose directors belong to dialect group d. Higher values indicate more linguistically homogeneous Chinese communities.

Top-3 Share: Strength of Hub-Surname Clan Links. Chinese business communities in colonial Java were also organized around surname-based clan (*marga*) networks that facilitated access to credit, trading connections, and commercial information across districts. We measure the concentration of these networks using the *top-3 clan share*: the fraction of a district's registered Chinese firms whose director bears the surname TAN, OEI, or LIEM/LIM. These three surname clans in the Colonial Business Indonesia database rank among the top three clans in each of the three major commercial hubs of Surabaya, Semarang, and Batavia.

A firm is classified as "top-3" if *any* of its listed directors carries one of these surnames. We compute the measure at the district-year level.

The correlation between dialect HHI and top-3 clan share is near zero ($r = 0.02$), confirming that dialect composition captures a genuinely distinct dimension of community network structure rather than trading links with same-surnamed commercial families.

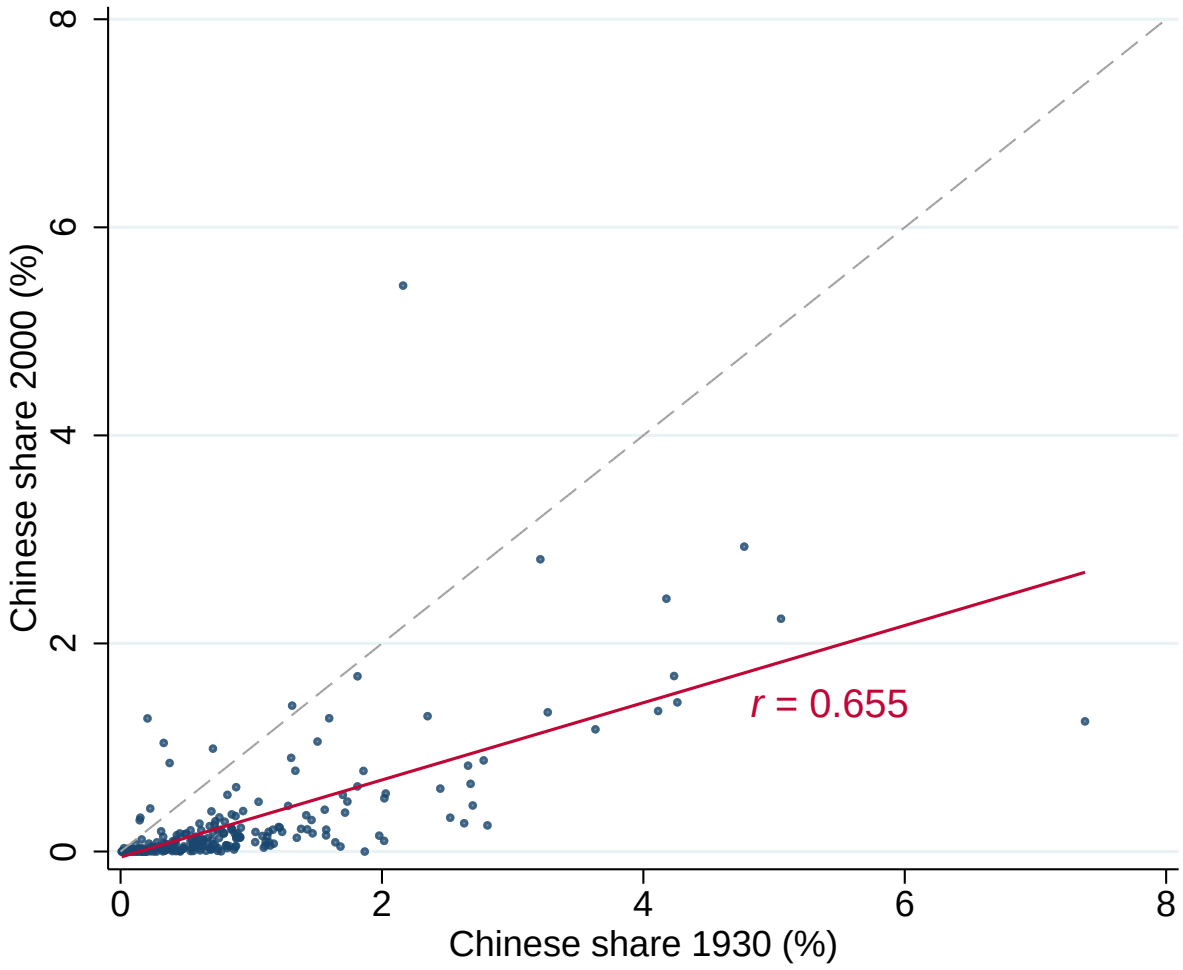
A.7 Additional Figures and Tables

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



Notes: This figure illustrates our recentered instrumental variable strategy (Borusyak and Hull, 2023). 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 A.3: Persistence of Ethnic Chinese Shares, 1930-2000



Notes: Each dot is a 1930 district in Central and East Java. The y-axis plots the Chinese population share from the 2000 Indonesian Population Census; the x-axis plots the Chinese share from the 1930 Census (Volkstelling). We exclude 2 observations for presentational purposes. Dashed line indicates one-for-one persistence. Solid line is the OLS fit (r = 0.655, N = 246).

Table A.1: OLS: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT: ADDITIONAL CONTROLS (INCLUDING WEST JAVA)

Panel A: Sugar Factory 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)	0.041**	0.064**	3.294**	1.088*	0.015
	(0.020)	(0.027)	(1.397)	(0.610)	(0.009)
Sugar Factory presence	0.276***	0.297***	17.739***	1.995	0.013
	(0.090)	(0.084)	(3.595)	(1.534)	(0.028)
R-squared	0.698	0.690	0.562	0.695	0.688
Mean (Dep. Var.)	5.778	6.722	28.079	14.956	11.631
Standard Deviation (Dep. Var.)	0.679	0.716	24.409	11.936	0.229
Observations	340	340	340	340	340

Panel B: Dist. to Col. Railroads Control

Chinese share (percent)	0.032*	0.054**	3.212**	1.024*	0.014
	(0.019)	(0.024)	(1.272)	(0.606)	(0.009)
Ln Distance to Colonial Railroads (1925)	-0.109***	-0.122***	-2.914***	-0.764**	-0.009*
	(0.037)	(0.025)	(1.008)	(0.360)	(0.005)
R-squared	0.716	0.711	0.537	0.698	0.689
Mean (Dep. Var.)	5.778	6.722	28.079	14.956	11.631
Standard Deviation (Dep. Var.)	0.679	0.716	24.409	11.936	0.229
Observations	340	340	340	340	340

Panel C: Dist. to Postal Roads Control

Chinese share (percent)	0.047**	0.073***	3.801***	1.168*	0.016*
	(0.019)	(0.026)	(1.364)	(0.609)	(0.009)
Ln Dist. to Great Postal Roads (1808-1811)	-0.037	-0.074**	-5.054***	-1.082*	-0.018*
	(0.044)	(0.034)	(1.274)	(0.610)	(0.010)
R-squared	0.688	0.685	0.553	0.699	0.692
Mean (Dep. Var.)	5.778	6.722	28.079	14.956	11.631
Standard Deviation (Dep. Var.)	0.679	0.716	24.409	11.936	0.229
Observations	340	340	340	340	340

Panel D: Dist. to Opium Control

Chinese share (percent)	0.049***	0.069***	3.531***	1.109*	0.015
	(0.018)	(0.026)	(1.299)	(0.608)	(0.009)
Ln Distance to Opium Boundaries (1886)	-0.055	0.000	0.701	0.158	0.003
	(0.036)	(0.034)	(1.547)	(0.480)	(0.010)
R-squared	0.691	0.677	0.520	0.693	0.688
Mean (Dep. Var.)	5.778	6.722	28.079	14.956	11.631
Standard Deviation (Dep. Var.)	0.679	0.716	24.409	11.936	0.229
Observations	340	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.2: 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**	0.116	3.062*
	(0.028)	(0.097)	(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.3: PLACEBO: DISTANCE TO THE MURIA STRAIT, ETHNIC COMPOSITION, AND SUGAR FACTORIES

Reduced Form

	(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***	0.0865**	-0.00605	0.00152	-0.00453	-0.00928
	(0.0261)	(0.0327)	(0.00727)	(0.00267)	(0.00841)	(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 A.4: IV: 1930 CHINESE SHARE AND CONTEMPORARY DEVELOPMENT — ADDITIONAL CONTROLS

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ln(Pop. Density) (1930)	Ln(Pop. Density) (2000)	% Urban (2010)	Light Intensity (2010)	Consumption (2001–2011)
Panel A: Sugar Factory Control					
Chinese Share (percent)	0.0918	0.248***	12.18***	5.572***	0.0960***
	(0.0606)	(0.0852)	(3.557)	(2.050)	(0.0301)
Kleibergen-Paap F	14.48	14.48	14.48	14.48	14.48
Anderson-Rubin p	0.174	0.014	0.024	0.085	0.002
Panel B: Colonial Railroad Control					
Chinese Share (percent)	0.0939	0.256***	13.53***	5.995**	0.102***
	(0.0860)	(0.0817)	(5.130)	(2.501)	(0.0321)
Kleibergen-Paap F	13.52	13.52	13.52	13.52	13.52
Anderson-Rubin p	0.334	0.019	0.061	0.111	0.004
Panel C: Great Postal Road Control					
Chinese Share (percent)	0.00371	0.252***	9.909**	5.384**	0.107***
	(0.127)	(0.0813)	(4.529)	(2.255)	(0.0364)
Kleibergen-Paap F	9.15	9.15	9.15	9.15	9.15
Anderson-Rubin p	0.977	0.033	0.139	0.149	0.003
Panel D: Opium Boundary Control					
Chinese Share (percent)	0.147*	0.289***	13.89***	5.914**	0.0962***
	(0.0854)	(0.0826)	(5.024)	(2.393)	(0.0270)
Kleibergen-Paap F	13.21	13.21	13.21	13.21	13.21
Anderson-Rubin p	0.162	0.017	0.062	0.109	0.002
Observations	242	242	242	242	242

Notes: Each panel reports IV estimates from a separate regression that adds the indicated control to the baseline specification. The unit of analysis is the district. Only the coefficient on the endogenous variable (Chinese Share) is shown; coefficients on each additional control and first-stage results are omitted for brevity. *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.5: IV: 1930 Chinese Share and Contemporary Development: Counterfactual Landing Sites around Cirebon

Panel A: IV Specification

Dependent Variable:	(1) Ln(Pop. Density) (1930)	(2) Ln (Pop. Density) (2000)	(3) % Urban (2010)	(4) Light Intensity (2010)	(5) Consumption (2001–2011)
Chinese Share (percent)	0.150**	0.203***	8.867***	5.154**	0.0706***
	(0.0720)	(0.0594)	(2.922)	(2.028)	(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***	-0.152***	-0.152***	-0.152***	-0.152***
	(0.0506)	(0.0506)	(0.0506)	(0.0506)	(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.6: IV: 1930 Chinese Share and Contemporary Development: Counterfactual Landing Sites around Majapahit

	(1) Ln Pop. Density (1930)	(2) Ln Pop. Density (2000)	(3) % Urban (2010)	(4) Light Intensity (2010)	(5) Consumption (2001–2011)
Panel A: ≤25km from north coastline					
Chinese share (percent)	-0.041	0.114*	3.649	1.140	0.096***
	(0.053)	(0.066)	(5.574)	(0.783)	(0.028)
KP Wald F-stat	29.775	29.775	29.775	29.775	29.775
AR p-value	0.632	0.001	0.224	0.005	0.000
Regencies	16	16	16	16	16
Observations	57	57	57	57	57
Panel B: ≤50km from north coastline					
Chinese share (percent)	0.034	0.139***	2.721	3.470**	0.053**
	(0.105)	(0.054)	(3.080)	(1.764)	(0.026)
KP Wald F-stat	9.242	9.242	9.242	9.242	9.242
AR p-value	0.755	0.070	0.443	0.187	0.009
Regencies	26	26	26	26	26
Observations	100	100	100	100	100

Notes: The table reports IV estimates. The unit of analysis is at the district level. The excluded instrument is the difference between log distance to the Muria Strait and log distance to the nearest counterfactual landing site that shared similar riverine access to the centre of the Majapahit Kingdom (See Section A.3 for details). All regressions include baseline controls (distances to 6 historical cities with quadratic terms) and province fixed effects. Standard errors clustered at the regency level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: 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	0.208***	9.951**	4.632**	0.0802***
	(0.0760)	(0.0681)	(4.238)	(2.230)	(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***	-0.0865***	-0.0865***	-0.0865***	-0.0865***
	(0.0244)	(0.0244)	(0.0244)	(0.0244)	(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.8: IV: 1930 Chinese Share and Contemporary Development: Restricting to Coastal Districts

	(1) Ln Pop. Density (1930)	(2) Ln Pop. Density (2000)	(3) % Urban (2010)	(4) Light Intensity (2010)	(5) Consumption (2001–2011)
Panel A: Baseline IV Sample					
Chinese share (%)	0.130	0.261***	13.276***	5.569**	0.094***
	(0.093)	(0.078)	(4.617)	(2.220)	(0.031)
KP Wald F-stat	13.753	13.753	13.753	13.753	13.753
AR p-value	0.218	0.015	0.042	0.097	0.001
Regencies	54	54	54	54	54
Observations	242	242	242	242	242
Panel B: $\leq 25km$ from north coastline, baseline IV					
Chinese share (%)	0.016	0.176**	5.891	2.219***	0.110***
	(0.043)	(0.081)	(5.920)	(0.683)	(0.028)
KP Wald F-stat	25.649	25.649	25.649	25.649	25.649
AR p-value	0.065	0.000	0.116	0.000	0.000
Regencies	16	16	16	16	16
Observations	57	57	57	57	57
Panel C: $\leq 25km$ from north coastline, Recentered IV					
Chinese share (%)	0.014	0.167**	5.390	1.828***	0.101***
	(0.042)	(0.077)	(5.841)	(0.639)	(0.028)
KP Wald F-stat	33.370	33.370	33.370	33.370	33.370
AR p-value	0.044	0.000	0.114	0.000	0.000
Regencies	16	16	16	16	16
Observations	57	57	57	57	57

Notes: IV estimates using 2SLS. Panels A–B use log distance to the Muria Strait as the excluded instrument. Panel C uses the Borusyak-Hull recentered instrument. Panel A uses the full baseline sample of Central and East Java districts. Panels B and C restrict our sample to north-coast districts (within 25km of the north coastline. All regressions include province fixed effects and baseline controls. Standard errors clustered at the regency level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: 1930 Chinese Share and Conflict Frequency

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
	UNSFIR Data				ACLED Data
Dependent Variable:	Conflicts Before 1998	Conflicts 1998	Conflicts After 1998	Anti Chinese Conflicts (1990–2003)	Contemporary Conflicts
Chinese Share (percent)	1.534**	4.338	4.416	-0.0408	37.39**
	(0.601)	(2.889)	(3.094)	(0.0345)	(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***	-0.0777***	-0.0777***	-0.0777***	-0.0777***
	(0.0210)	(0.0210)	(0.0210)	(0.0210)	(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.10: 1930 Chinese Share and Nightlights Growth (1997-1999): Controlling for Population Density

Panel A: IV Specification

	(1)	(2)
Dependent Variable:	Light Intensity Growth (1997–1999)	Light Intensity Growth (1997–1999)
Chinese share (percent)	-0.308**	-0.350**
	(0.135)	(0.156)
R-squared	0.505	0.512
Mean (Dep. Var.)	-0.020	-0.020
Standard Deviation (Dep. Var.)	0.515	0.515
Kleibergen-Paap WID, F-stat	19.445	19.884
Anderson-Rubin, P-value	0.056	0.056
Control for 1996 Pop. Density	No	Yes

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.131***	-0.131***
	(0.0296)	(0.0293)
Partial R-squared	0.037	0.072
Fixed Effects	Province	Province
Number of Regencies	54	54
Observations	242	242

Notes: The table reports IV estimates. The unit of analysis is at the district level. Column (2) adds 1996 population density as a control calculated from the 1996 Village Census. *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: FIRM SALES AND 1930 CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Ln(Sales) (1980-1996)	Ln(Sales) (1997-1999)	Ln(Sales) (2000-2013)	Ln(Sales) All Years
Chinese share (percent)	0.220*	0.119	0.187*	0.212*
	(0.118)	(0.0888)	(0.0977)	(0.110)
Chinese share (percent) X Period 2				-0.0934**
				(0.0414)
R-squared	0.243	0.204	0.219	0.407
Mean (Dep. Var.)	12.228	13.351	14.676	13.614
Standard Deviation (Dep. Var.)	2.021	2.005	2.012	2.316
Kleibergen-Paap WID, F-stat	15.878	15.476	14.204	7.495
Anderson-Rubin, P-value	0.008	0.089	0.015	0.019
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	460422	142375	657209	1260006
F-stat (Chinese Share)	15.878	15.476	14.204	7.992
F-stat (Chinese Share X Period 2)				15.480

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.12: FIRM SALES GROWTH AND 1930 CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Sales Growth (1980-1996)	Sales Growth (1997-1999)	Sales Growth (2000-2013)	Sales Growth All Years
Chinese share (percent)	0.00971**	0.00430	0.0136*	0.0127**
	(0.00466)	(0.00616)	(0.00711)	(0.00593)
Chinese share (percent) X Period 2				-0.00838
				(0.00669)
R-squared	0.008	0.030	0.008	0.010
Mean (Dep. Var.)	0.117	0.185	0.129	0.131
Standard Deviation (Dep. Var.)	0.676	0.679	0.895	0.800
Kleibergen-Paap WID, F-stat	15.789	15.797	13.960	7.383
Anderson-Rubin, P-value	0.009	0.434	0.023	0.016
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	377390	132093	596559	1106042
F-stat (Chinese Share)	15.789	15.797	13.960	8.783
F-stat (Chinese Share X Period 2)				7.901

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.13: FIRM PRIVATE INVESTMENT/SALES AND CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable: Investment/Sales	Private Investment (1980-1996)	Private Investment (1997-1999)	Private Investment (2000-2013)	Private Investment All Years
Chinese share (percent)	0.352**	-0.00997	-0.00315	0.237**
	(0.141)	(0.00744)	(0.00667)	(0.0932)
Chinese share (percent) X Period 2				-0.247**
				(0.0986)
R-squared	0.001	0.041	0.002	0.001
Mean (Dep. Var.)	0.286	0.083	0.064	0.203
Standard Deviation (Dep. Var.)	47.876	0.276	3.580	37.433
Kleibergen-Paap WID, F-stat	15.693	15.457	14.077	7.657
Anderson-Rubin, P-value	0.000	0.140	0.610	0.001
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	447457	142690	152656	742803
F-stat (Chinese Share)	15.693	15.457	14.077	7.736
F-stat (Chinese Share X Period 2)				15.461

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1,\ ^{**}p < 0.05,\ ^{***}p < 0.01$.

Table A.14: FIRM LABOR AND CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Ln(Employment) (1980-1996)	Ln(Employment) (1997-1999)	Ln(Employment) (2000-2013)	Ln(Employment) All Years
Chinese share (percent)	0.124*	0.0384	0.0413	0.106*
	(0.0719)	(0.0417)	(0.0385)	(0.0642)
Chinese share (percent) X Period 2				-0.0676*
				(0.0352)
R-squared	0.097	0.107	0.103	0.098
Mean (Dep. Var.)	3.966	3.966	3.927	3.960
Standard Deviation (Dep. Var.)	1.141	1.127	1.089	1.131
Kleibergen-Paap WID, F-stat	15.784	15.451	14.211	7.770
Anderson-Rubin, P-value	0.013	0.292	0.201	0.020
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	462530	142645	108610	713785
F-stat (Chinese Share)	15.784	15.451	14.211	7.771
F-stat (Chinese Share X Period 2)				15.454

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1,\ ^{**}p < 0.05,\ ^{***}p < 0.01$.

Table A.15: FIRM WAGE BILL AND CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Ln(Wages) (1980-1996)	Ln(Wages) (1997-1999)	Ln(Wages) (2000-2013)	Ln(Wages) All Years
Chinese share (percent)	0.265*	0.0792	0.135*	0.196*
	(0.137)	(0.0714)	(0.0744)	(0.101)
Chinese share (percent) X Period 2				-0.117**
				(0.0496)
R-squared	0.278	0.366	0.354	0.505
Mean (Dep. Var.)	10.280	9.736	12.495	11.367
Standard Deviation (Dep. Var.)	1.778	2.215	1.982	2.264
Kleibergen-Paap WID, F-stat	15.804	15.363	14.162	7.457
Anderson-Rubin, P-value	0.005	0.177	0.043	0.025
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	462801	131849	653750	1248400
F-stat (Chinese Share)	15.804	15.363	14.162	7.904
F-stat (Chinese Share X Period 2)				15.368

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.16: FIRM AVERAGE WAGE AND CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Ln(Avg. Wage) (1980-1996)	Ln(Avg. Wage) (1997-1999)	Ln(Avg. Wage) (2000-2013)	Ln(Avg. Wage) All Years
Chinese share (percent)	0.141**	0.0581	0.0492*	0.122**
	(0.0675)	(0.0401)	(0.0297)	(0.0605)
Chinese share (percent) X Period 2				-0.0638**
				(0.0276)
R-squared	0.540	0.544	0.757	0.645
Mean (Dep. Var.)	6.313	5.741	7.317	6.352
Standard Deviation (Dep. Var.)	1.075	1.674	2.068	1.455
Kleibergen-Paap WID, F-stat	15.807	15.358	14.020	7.769
Anderson-Rubin, P-value	0.003	0.068	0.053	0.007
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	462382	131820	103624	697826
F-stat (Chinese Share)	15.807	15.358	14.020	7.777
F-stat (Chinese Share X Period 2)				15.361

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.17: FIRM LABOR COUNT PRODUCTIVITY AND CHINESE SHARE

IV Specification

	(1)	(2)	(3)	(4)
Dependent Variable:	Ln(L Prod.) (1980-1996)	Ln(L Prod.) (1997-1999)	Ln(L Prod.) (2000-2013)	Ln(L Prod.) All Years
Chinese share (percent)	0.122**	0.0796	0.0864**	0.116**
	(0.0610)	(0.0515)	(0.0429)	(0.0575)
Chinese share (percent) X Period 2				-0.0366**
				(0.0177)
R-squared	0.344	0.257	0.271	0.473
Mean (Dep. Var.)	8.271	9.386	10.237	8.782
Standard Deviation (Dep. Var.)	1.351	1.351	1.354	1.544
Kleibergen-Paap WID, F-stat	15.880	15.469	14.239	7.812
Anderson-Rubin, P-value	0.005	0.039	0.004	0.009
Baseline Controls	Yes	Yes	Yes	Yes
Fixed Effects	Province, Sector	Province, Sector	Province, Sector	Province, Sector
Number of Regencies	54	54	54	54
Observations	459993	142330	108360	710683
F-stat (Chinese Share)	15.880	15.469	14.239	7.814
F-stat (Chinese Share X Period 2)				7.736

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. F-stat (Chinese Share) and F-stat (Chinese Share X Period 2) refer to the partial F-stats for joint significance of the instruments in their separate first-stage regressions. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.18: 1930 Chinese Share and Financial Access

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependent Variable:	Bank Count	Distance to Banks	Any Business Loans	Deficit: Savings	Deficit: Borrow Banks	Deficit: Borrow Rentenir	Deficit: Sell Belongings	Deficit: Sell Assets	Deficit: Borrow Family	Deficit: Borrow Friends
Chinese Share (percent)	-0.022	-0.090	-0.025*	0.016	-0.003	0.007	0.026	0.003	0.006	-0.043
	(0.024)	(0.337)	(0.013)	(0.016)	(0.004)	(0.007)	(0.016)	(0.013)	(0.020)	(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

Panel B: First Stage

Ln(Distance to Muria Strait)	-0.158***	-0.158***	-0.158***	-0.158***	-0.158***	-0.158***	-0.158***	-0.158***	-0.158***	-0.158***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Partial R-squared	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091	0.091
Fixed Effects	Province, Sector	Province, Sector	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	54	54
Observations	242	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, 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.19: 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	0.266***	14.73***	6.366***	0.106***
	(0.0766)	(0.0868)	(4.962)	(2.148)	(0.0335)
Neighbors' Chinese Share (percent)	-0.0308	0.0462	3.166	0.120	0.0293
	(0.0333)	(0.0520)	(3.399)	(0.922)	(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***	-0.0838***	-0.0838***	-0.0838***	-0.0838***
	(0.0204)	(0.0204)	(0.0204)	(0.0204)	(0.0204)
Neighbors' Ln(Distance to Muria Strait)	0.00281	0.00281	0.00281	0.00281	0.00281
	(0.0153)	(0.0153)	(0.0153)	(0.0153)	(0.0153)

Panel C: First Stage for Neighbors' Chinese Share

Ln(Distance to Muria Strait)	0.0121	0.0121	0.0121	0.0121	0.0121
	(0.0203)	(0.0203)	(0.0203)	(0.0203)	(0.0203)
Neighbors' Ln(Distance to Muria Strait)	-0.102***	-0.102***	-0.102***	-0.102***	-0.102***
	(0.0209)	(0.0209)	(0.0209)	(0.0209)	(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.20: 1930 CHINESE SHARE AND HUMAN CAPITAL

Panel A: IV Specification

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	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	0.320	0.0363	0.0107	0.00537**
	(3.397)	(0.492)	(0.0469)	(0.00882)	(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*	-0.0757*	-0.0757*	-0.0777***	-0.0777***
	(0.0419)	(0.0419)	(0.0419)	(0.0210)	(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

	(1)	(2)	(3)
Dependent Variable:	Chinese High School Graduates (2010)	Chinese Bachelor's Degree (2010)	Chinese Mater's Degree (2010)
Chinese Share (percent)	0.0406	0.0341	0.00352
	(0.0556)	(0.0340)	(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*	-0.0718*	-0.0625
	(0.0419)	(0.0420)	(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.21: CHINESE COMMUNITIES AND MARKET ACCESS

Reduced Form

	(1)	(2)	(3)
Dependent Variable:	Market Access (2010)	Market Access (2010)	Market Access (2010)
Ln(Distance to Murait Strait)	−0.003	−0.004	0.001
	(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. $^{*}p < 0.1,\ ^{**}p < 0.05,\ ^{***}p < 0.01$.

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.

Akbulut-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.

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 Liong'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 A. Hassan, Lisa Tarquinio, and Stephen J. Terry, "Immigration, Innovation, and Growth," *American Economic Review*, 2024, *114* (7), 2118–2158.

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, Fanghao, Ruichi Xiong, and Xiaobo Zhang, "Familiar strangers: the role of diaspora networks in foreign investment and long-run development," *Rotman School of Management Working Paper*, 2022, (4004159).

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," *Cliometrica*, 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.

Do, Quoc-Anh, Roberto Galbiati, Benjamin Marx, and Miguel A. Ortiz Serrano, "J'Accuse! Antisemitism 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 Prokopczuk, 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.

Ham, Ong Hok, "Chinese Capitalism in Dutch Java," *Southeast Asian Studies*, 1989, *27* (2), 156–175.

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.

Huang, Kunzhang, *Yinni huaqiao huaren shi:(gudai zhi 1949)*, Guangdong Gaodeng Jiaoyu Chubanshe, 2005.

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*.

Ju-kua, Chau, *Chau Ju-kua: His Work on the Chinese and Arab Trade in the Twelfth and Thirteenth Centuries, Entitled Chu Fan Chi*, St. Petersburg: Imperial Academy of Sciences, 1911.

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 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 Distribueerende Tusschenhandel der Chineezen 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://gedeon89.github.io/gedeonlim.com/Lim_Bengkok_edited_pre_sub.pdf.

Macauley, 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, "Banks and Development: Jewish Communities in the Italian Renaissance and Current Economic Performance," *The Review of Economics and Statistics*, 2016, *98* (1), 140–158.

_____ , "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.

Redding, Gordon, *The spirit of Chinese capitalism*, New York, NY, USA: De Gruyter, 2013.

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 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.

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.

Xi, Beijing Daxue Zhongguo Yuyan Wenxue and Yuyanxue Jiaoyanshi, "Hanyu Fangyin Zihui," *Phonetic dictionary of Chinese dialects.) Beijing: Wenni GaigeChubanshe*, 1962.

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.