Ethnic Divisions & Ethno-Nationalism: Evidence from Forced Migration in Malaysia *

Chun Chee Kok
UC Louvain

Gedeon Lim
University of Hong Kong

Danial Shariat
UC Berkeley

Abu Siddique
Royal Holloway & IFS

Shunsuke Tsuda
University of Essex

August 21, 2026

Abstract

We examine the long-run political effects of a Cold War era program that resettled over 500,000 ethnic minority Chinese—many of which were recent economic immigrants—into mono-ethnic villages across Malaysia. Today, areas surrounding these villages remain highly ethnically segregated. Yet ethnic-majority Malay communities in these areas exhibit lower electoral support for the ethnonationalist coalition, greater identification with a more inclusive Malaysian identity, and weaker zero-sum beliefs. We find little evidence of cohort-level heterogeneity, though overall effects appear stronger in regions with historical interethnic economic complementarities. Turning to mechanisms, Malays report greater intergroup contact in the workplace and situational trust towards ethnic minorities—effects are amplified (attenuated) under economic complementarities (competition). Strikingly, effects on measures of broader social integration remain muted. Our findings show how intergroup contact, when supported by economic complementarities, can durably weaken the demand for ethnonationalist politics but may fall short of fostering broader social integration.

***JEL codes**: D72, J15, N45, O15, R23.*
Keywords: Ethnic Diversity, Ethno-Nationalism, Economic Complementarities, Colonial Policy, Forced Resettlement, Interethnic Contact, Social Cohesion.

*We thank Sam Bazzi, Sascha O. Becker, Dilip Mookherjee, Robert A. Margo, Paul Raschky, and everyone at SoDa Labs for advice and support. We thank Yan Yiing Lim, Yuxin (Aaron) Hu, and HKU students for excellent research assistance. We also thank Madhav Aney, Fatima Aqeel, Jeanet Sinding Bentzen, Dan Bogart, Prashant Bharadwaj, Jesse Bruhn, Yiming Cao, Ujjayant Chakravorty, Arnaud Chevalier, Melissa Dell, Quoc-Anh Do, Vitaliia Eliseeva, Sebastian Ellingsen, Marcel Fafchamps, James Feigenbaum, James Fenske, Claudio Ferraz, Oded Galor, Teevrat Garg, Thomas Gautier, Matthew Gudgeon, Thea Howe Choon, Maulik Jagnani, Saumitra Jha, Lawrence Jin, Reka Juhasz, Gaurav Khanna, Nick Khaw, Bryant Kim, Cynthia Kinnan, Masahiro Kubo, Takashi Kurosaki, Claudio Labanca, Yu-Hsiang Lei, Bingjing Li, Matt Lowe, Sara Lowes, Pushkar Maitra, Eoin McGuirk, Harrison Mitchell, Chicheng Ma, Salma Mousa, Gautam Nair, Nathan Nunn, Elvin Ong, Albert Park, Sahar Parsa, Tom Pepinsky, Gianluca Russo, Raul Sanchez de la Sierra, Uta Schoenberg, Mikhail Rosli, Enrico Spolaore, Munir Squires, Chang Sun, Felipe Valencia Caicedo, Pierre-Louis Vezina, Sujata Visaria, Tom Vogl, David Weil, Xiaodong Zhu, for conversations and feedback. We thank audiences at Utrecht, RFBerlin Migration Forum (2026), BÆM Development Workshop (Bristol), Diversity and Inclusion Workshop (Milan-Bicocca), Essex, SWET, Migration and Development Conference (PSE), NEUDC, Leicester, Early Career Workshop QPE (KCL), Development Economics Workshop (Manchester) (all 2025), UBC, Tufts, Brown, PacDev (Stanford), UCSD, UC Irvine (all 2024), ASREC Australasia (Monash), Korean Economic History Association (Seoul), LKYSPP, Warwick CAGE Summer School, Australasian Public Choice Conference (all 2023), Monash, NUS, SMU, Royal Holloway, DWI Berlin (all 2022), Hong Kong Empirical Microeconomics Workshop (HKU-HKUST), Virtual Historical Political Economy Workshop (2021), Boston University DRG (2020). Lim thanks local universities for advice and administrative assistance on pilot visits and survey work in 2018/2019. We thank Danesh Prakash Chacko and Tindak Malaysia for election advice, election (a-)spatial data and support. Kok is grateful for financial support from Monash University's Dean's Postgraduate Research Excellence Award, Donald Cochrane Postgraduate Research Scholarship, Monash Graduate Scholarship, Enhanced International Research Experience Grant Scheme. Lim is grateful for financial support from BU's Initiative on Cities (2018), The University of Hong Kong's Early Career Scheme (2020-2022). This paper was previously circulated as "Interethnic Proximity, (Complementarities), and Politics: Evidence from Malaysia" ;"The Effects of Long-Term Ethnic Segregation in Malaysia" and Lim's Ph.D. thesis chapter (2020): "Ethnic Segregation and Identity Politics: Evidence from Malaysia." All errors are our own. This project received IRB approval from HKU HREAS (EA210013). Contacts: chun.kok@uclouvain.be, gedeonl@hku.hk, danial_shariat@berkeley.edu, abu.siddique@rhul.ac.uk, shunsuke.tsuda@essex.ac.uk.

1 Introduction

In many democracies, authoritarian-populist parties have increasingly mobilized nationalism against religious and ethnic minorities (Alesina and La Ferrara, 2005; Guriev and Papaioannou, 2022).The stakes are especially high because the outcome of electoral politics shapes access not only to public goods and redistributive policies, but also to public employment, land, education, and economic opportunities. Yet, a central puzzle is why ethnic diversity sometimes fuels the rise of ethno-nationalism, but in other cases weakens it. This is especially given that hostility toward minorities is often directed not only at new arrivals, but also at long-settled communities embedded in local economic and social life (Stevenson, 2026).

This paper studies a politically central margin of this process: how sustained exposure to a substantial minority presence across half a decade affects ethno-nationalistic preferences and attitudes. We do so by combining granular voting data with rich data from an original in-person survey. Voting provides a natural revealed measure of the demand for ethno-nationalistic politics. It records how social perceptions, economic interests, and partisan appeals are translated into political choice, and granular electoral data allow us to measure this process at the same local scale at which intergroup contact and changes in beliefs occurs. Aggregated across individuals, these choices have first-order implications for democratic stability, social cohesion, and economic development (Alesina and La Ferrara, 2005; Glaeser, 2005; Guriev and Papaioannou, 2022; Jha, 2025).

We answer these questions in the context of Malaysia—one of the most ethnically polarized countries in the world (Horiwitz, 1989; Horowitz, 2000), that has *not* experienced any large-scale conflict in the past five decades despite (i) large waves of "uncontrolled" economic immigration from China during the early 20th century;[1] (ii) widely held perceptions that ethnic-minorities (the Chinese in particular) control a disproportionate share of economic resources. Specifically, we study how persistent proximity to an ethnic minority community affects ethnic majority group preferences for ethno-nationalistic politics in the long run. To do so, we leverage a 1950s colonial, military resettlement program in British Malaya that forcibly resettled over half a million widely dispersed rural ethnic Chinese into hundreds of compact, fenced, ethnic Chinese New Villages (CNVs hereinafter).[2] The program was implemented to sever the rural support networks of anti-British communist insurgents, who were perceived to be predominantly ethnic Chinese. As a result, this program created a dispersed network of sharply demarcated Chinese enclaves surrounding pre-

[1]During the mid-1920s, the Malayan peninsular experienced annual inflows of more than 200,000 in-migrants. This led to total population increasing from 2.6 million (1921) to 4.4 million (1931) of which the the ethnic Chinese population increased by nearly a third, from 1.2 million to 1.7 million. In fact, by the 1931 census, Malays were a slight minority at 1.6 million.

[2]This was nearly the entire rural Chinese population in British Malaya: roughly 25% of the total Chinese and 10% of the country's population. Selective resettlement based on pro-Communist leanings was thus highly unlikely.

existing Malay villages that significantly reshaped the ethnic geography of Malaysia. Freedom of movement was reinstated in 1960, but most CNVs have persisted until today, largely due to the grant of land titles to the resettled Chinese population, who had previously lacked formal land rights (Nyce, 1973).

For identification, we exploit the plausibly exogenous nature of the resettlement process. We retroactively construct the universe of "second-best" resettlement sites based on plausibly exogenous military criteria (Matranga and Natkhov, 2026) and implement a spatial randomization inference-style approach, analogous to Dell and Olken (2020). This procedure yields a counterfactual, spatial distribution of potential resettlement sites, which allows us to compare the exposure effects of proximity to real CNVs with those of proximity to the counterfactual sites. Importantly, for identification, rich archival documents attest to the plausibly, quasi-random nature of the original site selection process—eventual CNV sites were often sub-optimal due to cost constraints and the lack of concrete data on pre-resettlement locations of rural Chinese (Section 4.2). Counterfactuals achieve balance across a host of pre-treatment geographical and socioeconomic characteristics.

To conduct our analysis, we assemble a rich dataset with two novel features. First, we assemble an exhaustive set of administrative data comprising archival, geospatial, census, and voting data at the finest gradations possible. Second, we conduct an original in-person survey to collect rich measures of interethnic contact, attitudes, and economic outcomes of ethnic majority Malays living near actual and counterfactual CNV sites ($N = 1,990$). Importantly, this allows us to observe pre-resettlement village characteristics through retrospective surveys.[3] In parallell, we conduct a supplementary survey of CNV leaders to pinpoint pre-resettlement, Chinese locations.

We find balance across a range of pre-resettlement characteristics between Malays living near treated and counterfactual CNV sites. In particular, we find balance across Malays' economic status and a host of pre-resettlement interethnic attitudes such as their grandfathers' interactions with Chinese, trust of Chinese and perceived incidence of intergroup conflict. These results suggest that the British did not locate actual CNV sites based on pre-existing sentiments and/or attitudes of the Malay community towards the Chinese.

The Malaysian context offers two advantages. First, the segregation of Malays and Chinese into distinct economic niches allows us to study the long-run effects of complementarities versus competition on the demand for ethno-nationalistic politics. Second, because Malaysian voters rarely change their registered voting addresses (Jomo, 2017), vote shares can plausibly be interpreted as reflecting (i) changes in the political behavior and preferences of Malays who were born

[3]For this primary survey, no existing census of villages in Malaysia exists. Hence, we hand-digitized the exact location of every named Malay village within the universe of polling districts in our sample states. This enabled us to draw a geographically precise sample of Malay respondents based on proximity to nearby CNVs.

there; (ii) the effects of interethnic proximity during a Malay's formative years; and (iii) the effects of interethnic proximity at later-life migration destinations, if any.

Our primary outcome for political behavior is the vote share for the ethnonationalist coalition, the National Front (*Barisan Nasional*, or BN) in 2013 and 2018 federal elections. BN is a ruling coalition of ethnic-based parties that governed Malaysia for nearly six decades since independence, and whose main component party was established to advocate against equal citizenship rights for ethnic minorities, and has consistently advocated and implemented affirmative, Malay-first policies (Jomo, 2017). We interpret BN vote share as a behavioral proxy for the demand for ethnonationalistic politics. We support this interpretation by studying effects on attachment to ethnic vis-a-vis national identity; beliefs regarding the inclusivity of Malaysian identity; and zero-sum beliefs. We find strong convergent patterns.

Vote shares are observed at the polling district level, which is the most disaggregated unit available.[4] Importantly, this allows us to control for parliament seat fixed effects. That is, we estimate the effects of proximity to CNVs on Malay communities located *within* the same parliament seat. This allows us to hold candidate and party identity constant, so as to better interpret vote shares as a clean measure of voter behavior. Throughout, we focus on the exposure effects of CNVs on ethnic majority Malays in surrounding areas by applying a "doughnut-hole approach", excluding all polling districts containing initial CNVs. This isolates the exposure effects of proximity to CNVs while accounting for potential spillovers.[5]

Main results. We find a persistent reduction in BN vote shares near CNVs across both state and federal elections in 2013 and 2018. Polling districts within 0–2 km of CNVs register 3–8 percentage points (pp) (5-18%) lower support for the ethnonationalist coalition relative to polling districts in less proximate distance bins, which is potentially consistent with a decrease in demand for ethnonationalistic politics in areas more proximate to CNVs. Importantly, results are robust to controlling for measures of historical market access (Donaldson and Hornbeck, 2016).

In support of this interpretation, original survey data reveals that Malays residing closer to CNVs report greater identification with a more inclusive Malaysian identity, and weaker zero-sum beliefs (Chinoy et al., 2026). Specifically, treated Malays are: (i) more likely to identify as Malaysian first (national identity) (rather than their ethnic Malay identity); (ii) less likely to agree that having Malay ancestry is important for being Malaysian.

We rule out four leading, alternative explanations: (i) public goods, (ii) differences in turnout, (iii) ethnic composition, and (iv) selective migration. First, we find higher contemporary road

[4]This level of disaggregation is only available from the 2013 general election onwards.

[5]Our treatment estimates reflect the effects of both *increased* exposure to a higher concentration of Chinese populations around CNVs and potentially *decreased* exposure in depopulated origin areas. While we cannot disentangle these effects, the latter is likely minimal, as the original Chinese settlements were largely dispersed and located near rivers and jungle peripheries. Section 4 provides further details.

density near CNVs, but largely insignificant differences in the quantity and quality of educational and health public goods. Given pre-treatment balance on road density, we interpret these results as increases in connectivity-driven development driven by proximity to CNVs, without parallel improvements in service provision or its quality. Second, differences in voter turnout are quantitatively small and likely insufficient to explain observed differences in vote shares. Third, based on full-count electoral rolls, we find small and statistically insignificant differences in the share of registered ethnic Chinese voters between polling districts located within 0-2km of CNVs and less proximate distance bins.[6] Fourth, using original primary survey data on individual-level Malay migration histories, we find little evidence of differential in- or out-migration between Malay villages that are located more or less proximate to CNVs.

Observed differences in vote shares might simply arise from urbanization and/or exposure to new settlers (Balietti et al., 2021; Chetty et al., 2016). To isolate the effects of interethnic exposure, we conduct a placebo test using a contemporaneous, forced resettlement program of ethnic Malays that created Malay New Villages (MNVs). The establishment of MNVs took place (under otherwise similar conditions) but much fewer were established (given British perceptions of the ethnic Chinese as the main supporters of the Communists) (Dobby, 1952; Humphrey, 1971). Reassuringly, we find few differences in vote shares based on proximity to MNVs, strengthening our preferred interpretation that observed effects are likely driven by *interethnic* exposure and contact.

Qualitatively, we argue that our results are unlikely to arise from (i) gerrymandering of polling district borders; (ii) distaste against coercive resettlement policies. First, colonial resettlement of CNVs into close proximity with rural Malay enclaves led to the ethnic Chinese constituting a small number of voters in the heart of each parliament seat. It would have been administratively difficult for the ruling coalition to redraw polling districts borders to selectively isolate Malay voters within 2km of CNVs. Second, resettlement was carried out by the *British* colonial authorities, not by BN—it is unlikely that Malays would vote against BN for a policy they did not implement.

Mechanisms We investigate three potential mechanisms. First, we explore how proximity shaped downstream outcomes of interethnic contact, attitudes and social integration (Bazzi et al., 2019; Lowe, 2021, 2024). Second, we investigate the extent to which proximity might have led to greater economic prosperity and shared economic interests. Third, we leverage proximity to colonial-era rubber estates and tin mines to examine heterogeneous effects of economic competition versus complementarities (Artiles, 2026; Becker and Pascali, 2019; Horowitz, 2000). In addition, we rule out differences in the top-down provision of local public goods and ethnic patronage.

Consistent with *positive contact*, we find that Malays living near CNVs report visiting CNVs

[6] In addition, a back-of-the-envelope calculation shows that, even if all ethnic Chinese voters (unrealistically) voted against the ethnonationalist coalition, implausibly high turnout rates would be required to account for observed differences in vote shares.

more often and more social interactions with Chinese in the labor market and workplaces (but not in schools). They also report significantly higher situational trust (entrusting their child to a Chinese neighbor) but no differences in generalized trust towards Chinese nor attitudes towards inter-ethnic marriage and the number of close Chinese friends. These results suggest that sustained contact in the labor market might help attenuate ethno-nationalistic preferences and attitudes but might not be sufficient for fostering broader social integration among ethnic groups—even after six decades.

We find some evidence of increases in economic prosperity. Treated Malays report slightly higher income and wealth but not increases in educational levels—suggesting that economic interdependencies in existing labor and product markets likely underpin observed economic benefits. Importantly, however, controlling for contemporaneous income and wealth does not eliminate effects on attitudes nor vote shares, suggesting that observed political effects are not driven entirely by economic benefits.

Lastly, we examine the role of interethnic economic complementarities. The attitudinal and political effects of proximity appear to depend crucially on the presence of inter-ethnic complementarities or competition (Alesina and La Ferrara, 2005; Horowitz, 2000; Jha, 2013). We find strong (suggestive) evidence that complementarities (competition) leads to positive (negative) effects. Treated Malays in non-agricultural sectors, are more likely to perceive Chinese as economically beneficial and report that profits would decline in their absence, consistent with complementarities through supply chains and market access. In comparison, Malays employed in agriculture today (the traditionally Malay-dominant sector), report lower trust; empathy; and stronger perceptions of inter-ethnic competition. These opposing patterns suggest that contact alone is insufficient for political moderation; the *nature* of the underlying economic relationship, where certain economic sectors generate non-replicable and non-expropriable complementarities, might play an important role (Jha, 2025).

When did these complementarities first arise and to what extent are these changes in sociopolitical attitudes entirely driven by economic benefits? We further exploit variation in proximity to historical rubber estates and tin mines—labor-intensive sectors where private British employers frequently hired both Malays and Chinese in interdependent roles (Ross, 2014; Siew, 1953). Crucially, the rubber and tin mining industries in Malaysia, once the largest in the world, largely collapsed in the 1980s, leading to the attenuation of economic prosperity in these areas (Shah, 2019). In spite of this, we find that the negative effect on contemporary ethnonationalist vote shares is concentrated in areas located near historical estates and mines. Primary survey results exhibit a similar pattern, with positive attitudes concentrated in these areas despite a lack of positive economic effects. Taken together, strong, *initial* economic dependencies and persistent contact appear to be sufficient for generating persistently positive attitudes and voter demand for less exclusionary

politics over the long-run.

Related literature. This paper speaks to four strands of literature. The first is the literature on identity politics (Guriev and Papaioannou, 2022). Viewed through the lens of Padró i Miquel (2007), ethnonationalist voting may persist because majority-group voters fear that a multi-ethnic coalition would jeopardize group-based access to rents and protection. Relatedly, Saleh and Tirole (2021) show that identity can be shaped by material policy instruments: rulers may use taxes, public goods, or access to state benefits not only to extract revenue, but also to induce identity compliance and expand political support. Our contribution is to identify an empirical condition under which this politics-of-fear constraint potentially weakens. In areas where Malays experienced sustained proximity to ethnic-minority Chinese communities, Malay voters exhibit lower electoral support for the ethnonationalist coalition. Furthermore, given that Malay-first redistribution made ethnonationalist rule materially valuable to the majority group, our results show that sustained, positive-sum intergroup contact can reduce the perceived threat of out-group inclusion and loss of identity-based patronage, by promoting stronger identification with a more inclusive national identity and, ultimately reducing the demand for identity politics.

Second, we contribute to the large literature on the political and economic effects of (forced) resettlement and interethnic coexistence that builds on the seminal work by Allport (1954) (e.g. Bazzi et al., 2016, 2019; Carlitz et al., 2025; Dippel, 2014; Peters, 2022). Much of the literature has documented substantial heterogeneity. Coexistence appears to generate positive economic benefits in some contexts (e.g. Tabellini, 2020) but not in others (e.g. Dippel, 2014; Scott, 2020).[7] Similarly, in terms of political effects, some work finds that co-existence leads to greater support for right-wing parties (e.g. Tabellini, 2020), while others have found evidence in the opposite direction (e.g. Calderon et al., 2023).[8] We reconcile these differences by showing that one possible reason is the presence/absence of economic complementarities. By examining the effects of forced coexistence *within* the same context, we show that local differences in inter-ethnic competition and complementarities can lead to bifurcating long run effects. Furthermore, we build on Tabellini (2020) by providing conditions under which positive effects *can* accrue despite persistent, de-facto residential segregation.

Relatedly, our work builds on Bazzi et al. (2019) and Carlitz et al. (2025). First, we provide evidence for the positive effects of contact even in a setting of extreme ethnic polarization and segregation *across* villages. Second, Carlitz et al. (2025) find positive effects of contact on national

[7] Hence, our paper also relates to studies on the effects of historical immigration (e.g. Long et al., 2022; Sequeira et al., 2020). In particular, Tabellini (2020) finds economic benefits but political backlash of historical immigration in US cities. In contrast, we study rural areas and provide evidence of how persistent, intergroup contact in complementary labor markets can instead foster political moderation.

[8] A notable exception is (Steinmayr, 2021) who finds evidence of both positive *and* negative effects: short-term exposure (sustained contact) leads to an increase (decrease) in the far right vote share.

identity but little effects on democracy or trust. We complement this by showing that market interactions and extensive economic complementarities, stemming from colonial-era occupational niches, *can* drive positive, long-run effects on political preferences and national identity.

Third, we contribute to the literature on the effects of *contact and diversity on socio-political preferences* (Algan et al., 2016; Billings et al., 2021; Bursztyn et al., 2024; Enos, 2014; Lowe, 2021; Rao, 2019; Schindler and Westcott, 2021). We use an original survey to document novel evidence that economic complementarities, even in a setting of extreme ethnic polarization and antagonistic ethnic politics *can* lead to political moderation at the ballot box (e.g. Calderon et al., 2023; Fouka et al., 2022). The closest to ours is Bursztyn et al. (2024), who study the effects of persistent exposure to foreign-origin, ethnic minorities and find that their presence leads to positive attitudes toward these groups through greater intergroup contact and knowledge. Our work complements these results by showing how exposure can lead to weaker zero-sum mindsets and increases in voter demand for less exclusionary politics. We further document a striking asymmetry: despite more than half a century of intergroup exposure, deeper social integration remains muted. Our findings thus provide cautionary, policy-relevant evidence that persistently positive patterns of economic contact alone might not lead to assimilation in the long-run.

Last, our paper connects to the literature on intergroup complementarities (Ghosh, 2025; Jedwab et al., 2019; Jha, 2013, 2025; Montalvo and Reynal-Querol, 2021) and zero-sum thinking (Ali et al., 2025; Carvalho et al., 2023; Chinoy et al., 2026). To the best of our knowledge, we provide novel, quasi-experimental evidence of *how* zero-sum beliefs can be shaped by proximity. Combining unusually rich, original individual-level surveys with granular data on vote shares, we trace how, over a sixty-year period, intergroup contact shaped zero-sum beliefs and, ultimately, voter behavior at the ballot box. We further show how, in a setting where economic sectors generated non- replicable and non-expropriable complementarities (Jha, 2013, 2018, 2025), initially, positive-sum economic dependencies, *can* lead to a lasting divergence in long-run attitudes and political outcomes.

2 Institutional Background and Conceptual Framework

2.1 Institutional Background

Ethnic Diversity and Interethnic Relations. British rule over Peninsular Malaysia lasted from 1786 to 1957. In 2010, the population comprised of 64.5% ethnic Malays, 25.9% ethnic Chinese, and 8.9% ethnic Indians. Ethnic Malays and Chinese differ substantially in terms of language, religion and culture and interethnic marriages are rare (Nagaraj, 2009).

The in-migration of ethnic Chinese to Malaysia dates back to the 15^{th} century, but many that lived in rural areas prior to resettlement, arrived mostly from the early 20^{th} century, driven by

political upheaval in China and drawn to British Malaya's labor demands in rubber and tin (Kim, 1998; Strauch, 1981; Wang, 1959). Colonial policy further entrenched ethnic divisions via land-use policies that incentivized the Malays to remain in subsistence farming, and the Chinese to enter lucrative sectors like mining and trade (Kratoska, 1982). This institutionalized economic disparity laid the groundwork for persistent post-independence tensions, with many Malays perceiving the Chinese as disproportionately rich and economically advantaged.

Electoral Politics in Malaysia. In 1969, the economic gap between Malays and Chinese culminated in widespread racial riots, which were largely driven by economic grievances from ethnic Malays.[9] In response, the ruling coalition, the National Front (*Barisan Nasional*), led by the United Malays National Organization (UMNO), launched the 1970 New Economic Policy (NEP) to reduce Malay poverty and interethnic inequalities. The NEP granted ethnic Malays preferential access to land, education, public-sector jobs, housing, and corporate equity (Jomo, 2017).[10] These policies, together with the persistent political narrative of Malay ethno-nationalism (Gungwu, 2021), enabled the National Front to maintain political dominance and govern without interruption up till 2018.

In 2013, the opposition People's Alliance (*Pakatan Rakyat*)—a multi-ethnic coalition of People's Justice Party (PKR), Democratic Action Party (DAP), and Parti Islam Malaysia (PAS)—made major gains, campaigning on political reform and a more inclusive, multi-ethnic form of governance. In particular, a key policy agenda was that of Malaysia as a multicultural nation, which should grant equal rights and representation to all, regardless of ethnicity or religion.[11] Though the National Front retained power, it was increasingly dominated by UMNO, which pushed a more assertive pro-Malay agenda, including "Ketuanan Melayu" (Malay Sovereignty) (Ostwald and Oliver, 2020).[12] In 2018, PAS exited the opposition over ideological splits, mainly over the implementation of Islamic sharia law, and formed the conservative *Gagasan Sejahtera* (GS) bloc. The opposition rebranded itself as the Alliance of Hope (*Pakatan Harapan*), still led by PKR and Anwar Ibrahim. The election became a three-way race: reformist (Alliance of Hope), conservative Islamist (GS), and ethno-nationalist (National Front). In a historic result, and amid a wave of corruption episodes, the Alliance of Hope won the popular vote and formed Malaysia's first non–National Front federal government.

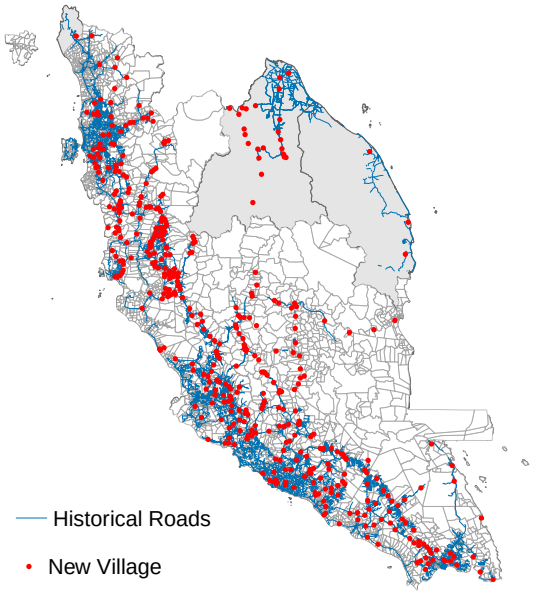
[9] We do not have data on 1969 violence but note that qualitative fieldwork suggests that very few of our surveyed villages reported experiencing any violent conflicts. In addition, proximity to CNVs does not appear to predict any differences in pre-resettlement, small-scale arguments and unhappiness between Malays and Chinese. See Table C.4.

[10] Ravallion (2020a) finds that post-NEP, all major ethnic groups saw income growth, with Malays gaining most. Ravallion (2020b) later shows that while the NEP reduced poverty, its effect weakened after the initial years.

[11] PKR appealed to urban Malays; DAP to Chinese and Indian voters; and PAS to rural Malays.

[12] The Malaysian Chinese Association and Malaysian Indian Congress—UMNO's coalition partners—had become politically marginal by 2013.

Figure 1: Location of New Villages and Historical Roads



Notes: Locations of 452 New Villages are represented by red dots, and the pre-resettlement road network is represented by dark blue lines. The white polygons indicate polling district boundaries in 2013. Polygons shaded in grey are the states of Kelantan and Trengganu, excluded from our analysis (See Appendix B.1). Source: Lee (2012); Lim and Song (2002), authors' own geo-referencing and HIND1035-series maps from 1947 (National Library of Australia).

The Electoral System in Malaysia. Malaysia operates as a first-past-the-post parliamentary democracy. Since 1963, elections have been held every five years at both the federal and state levels. Voters elect representatives to the federal House of Representatives and to State Legislative Assemblies from single-member constituencies, with each federal parliament seat typically comprising two to six state constituencies. To that end, all our analyses control for federal parliament seat fixed effects and state-party constituency match-ups, where applicable.

All Malaysians aged 21 and above were eligible to vote in both elections.[13] Voters are generally assigned to polling stations near their registered voting address, which almost always maps directly to a single polling district.[14] Notably, Malaysians rarely update their registered voting addresses despite later-life migration, often preferring to vote in their hometowns.[15] In Section 3.2, we explain how this informs our interpretation of vote share differences.

The Colonial Resettlement of Ethnic Chinese to CNVs. To study the impact of geographic proximity to ethnic Chinese, we exploit a top-down colonial resettlement program implemented by the British military during the Malayan Emergency (1948-1960). The program resulted in the forced resettlement of about 573,000 rural ethnic Chinese (nearly the entire rural Chinese

[13]The voting age was lowered to 18 prior to the 2022 elections.

[14]In nearly all cases, there is a 1-to-1 mapping between polling stations and districts.

[15]This may be due to the high cost of updating registration, opportunities to visit family, or the perception that rural votes carry more weight due to malapportionment (Jomo, 2017). For example, Kuala Lumpur, despite its large population, has no state legislature and relatively fewer federal seats.

population, which is about 25% of the Chinese population and 10% of the total population of Peninsular Malaysia in 1947) to fenced-up New Villages across Peninsular Malaysia (Sandhu, 1964). Figure 1 shows the distribution of all New Villages that we successfully geolocated (We describe this process in Section 3). Following military site selection criteria, nearly all CNVs lie along a historical main road for ease of reinforcement. In Section 4.2, we describe how we exploit this feature and other plausibly exogenous characteristics of site selection for identification.

This resettlement was part of the British response to a rising communist insurgency, led by the Malayan Communist Party (MCP), which the British perceived as being chiefly supported by the ethnic Chinese community. As the MCP turned against British rule, communist insurgents began targeting strategic economic assets like rubber estates and tin mines. The British military viewed mass resettlement as essential to severing support to communists and restoring political stability.[16]

The program rapidly resettled rural Chinese (mostly engaged in farming, pig-rearing, and smallholder mining and estate work) into fenced and heavily surveilled compounds located 3.2 to 9.6 km from their original settlements (Nyce, 1973). The resettlement was rapid and often unexpected, with the bulk of resettlement taking place within the first 3 years of the Emergency. Humphrey (1971) described it as a "laboratory experiment", because the resettlement imposed strict movement restrictions as CNVs were heavily surveilled, often enclosed by barbed wire, and largely isolated from neighboring communities from 1948-1960. Notably, the British aimed to minimize *labor dislocation* from pre-existing work sites and, despite nightly curfews, villagers were allowed to commute to their workplaces by foot or bicycle during the day.[17]

Post-Resettlement, Persistence of CNVs. Freedom of movement returned post-Emergency in 1960, yet most CNVs remained intact. This is largely attributed to the formalization of land tenure: many Chinese had been squatters without legal claims to land. Land titles awarded in CNVs provided newfound security and economic stakes, encouraging settlers to remain despite urban migration opportunities (Nyce, 1973; Strauch, 1981). Many Chinese have continued to remain in these villages due to the strong social and economic foundations they built over time (Strauch, 1981).

2.2 Conceptual Framework

We present a simple conceptual framework to clarify the channels through which the colonial resettlement program might have shaped the political identity of the ethnic majority Malays. We build

[16]It is worth noting that political stability in Malaya was important to the British because Malaya was one of the largest colonial revenue sources due to high demand and international prices of tin and rubber, the main exports of Malaya throughout the early and mid-20^{th} century.

[17]In the case of farmers, the British actively sought to purchase and convert land that was located adjacent to CNVs, to varying degrees of success.

on Shayo (2009), who models political preferences (which are shaped by one's political identity) as jointly determined by economic interests and social identity, and on Jha (2018), where interethnic complementarities can sustain peaceful coexistence and competition can intensify conflict. We extend these ideas to a spatial setting in which forced resettlement creates persistent variation in interethnic proximity between the ethnic majority, M, and ethnic minority, m.

The setup: Consider an M individual i in polling district d at distance r from the nearest cluster of resettled m. Individual i's support for the ethnonationalist coalition, v_i, is modeled as:

$$v_i = V(\tau_i,\, \pi_i,\, \eta_i) \tag{1}$$

where τ_i captures i's interethnic attitudes (e.g., higher trust, lower zero-sum beliefs/threat), π_i captures economic well-being (e.g., better income and wealth), and η_i denotes other determinants of political behavior (e.g., religiosity, candidate characteristics) that we treat as noise but can often be absorbed by constituency fixed effects and controls.

We assume $\frac{\partial V}{\partial \tau} < 0$: more positive interethnic attitudes reduce ethnonationalist support. In Shayo's (2009) framework, interethnic attitudes shape political preferences by modulating the salience of ethnic identity in political choice. This assumption is based on evidence that interpersonal trust shapes political and institutional preferences (Algan and Cahuc, 2010; Tabellini, 2008) and that zero-sum thinking strongly predicts support for exclusionary policies (Chinoy et al., 2026). The sign of $\frac{\partial V}{\partial \pi}$ is ambiguous a priori, as higher economic well-being may reduce the demand for ethnic patronage networks or strengthen the in-group's perceived stake in distributive politics.

Therefore, the resettlement program may affect v_i by changing both τ_i and π_i:

$$\tau_i = T(c(r),\, \bar{\tau}(r, \lambda)) \;; \qquad \pi_i = \Pi(a(r),\, \lambda) \tag{2}$$

Here $c(r)$ is routine interethnic contact (in markets, workplaces, schools), decreasing in r ($c' < 0$); $\bar{\tau}(r, \lambda)$ captures inherited attitudes, such as norms of trust and cooperation that were shaped by the post-resettlement history of interethnic interaction and transmitted across generations (Bisin and Verdier, 2001; Tabellini, 2008). These inherited norms vary with both distance and historical economic structure: areas closer to CNVs experienced more intensive interethnic contact, and areas with historical complementarities developed more positive interethnic norms. The inclusion of $\bar{\tau}(r, \lambda)$ allows attitudes to persist through intergenerational transmission even if contemporary contact has declined (Schindler and Westcott, 2021). $a(r)$ captures agglomeration and connectivity spillovers from m clusters, also decreasing in r ($a' < 0$). Finally, $\lambda \in \{\lambda^L, \lambda^H\}$ indexes the local economic structure, distinguishing areas with historical interethnic complementarities (λ^H: e.g., estate and mine economies with occupational specialization between M and m) from areas where

M and m competed for similar resources (λ^L: e.g., subsistence agriculture).

Role of economic complementarities vs competition: A key implication is that the effect of intergroup contact on attitudes depends on whether interactions are historically structured by complementarities or competition. In our framework, this heterogeneity operates through inherited norms: historical economic structure shapes $\bar{\tau}(r, \lambda)$, which in turn governs how present-day contact translates into attitudes. Formally, the marginal effect of contact on attitudes, $\frac{\partial \tau_i}{\partial c} = T_c(c(r), \bar{\tau}(r, \lambda))$, depends on both present-day contact and inherited norms.

We assume that contact is more likely to improve attitudes when inherited norms are more cooperative, i.e., $\frac{\partial^2 T}{\partial c \, \partial \bar{\tau}} > 0$, and that historical complementarities generate higher inherited pro-outgroup norms: $\bar{\tau}(r, \lambda^H) > \bar{\tau}(r, \lambda^L)$. Overall, these assumptions imply that the marginal attitudinal return to contact is more positive in complementarity areas:

$$\left.\frac{\partial \tau_i}{\partial c}\right|_{\lambda=\lambda^H} > \left.\frac{\partial \tau_i}{\partial c}\right|_{\lambda=\lambda^L} \tag{3}$$

Intuitively, complementarities can make repeated interethnic exchange relatively more mutually beneficial and raise the returns to maintaining cooperative norms (Jha, 2013, 2018). In such settings ($\lambda = \lambda^H$), inherited norms $\bar{\tau}(r, \lambda)$ are more likely to be pro-outgroup, so contact is more likely (relative to λ^L) to increase trust and reduce zero-sum beliefs. In our context, rubber estates and tin mines provide a natural motivation for this. In more competitive settings ($\lambda = \lambda^L$), contact may be less effective or even increase perceived threat (Caselli and Coleman, 2013). This heterogeneity motivates separating attitudinal from economic channels.

Persistence and disentangling channels: Substituting equation (2) into equation (1) and differentiating with respect to r by the chain rule, treating λ and η_i as fixed, gives:

$$\frac{dv_i}{dr} = \underbrace{V_\tau \Big[T_c \, c'(r) + T_{\bar{\tau}} \, \bar{\tau}_r(r, \lambda) \Big]}_{\text{Attitudinal channel}} + \underbrace{V_\pi \, \Pi_a \, a'(r)}_{\text{Economic channel}} \tag{4}$$

The first term is the *attitudinal* channel: proximity shapes interethnic attitudes both through present-day contact and through inherited norms shaped by the post-resettlement history of interethnic interaction, which in turn affects ethnonationalist support. The second is the *economic* channel: proximity generates agglomeration benefits that may affect political behavior through economic gains. Since both $c(r)$ and $a(r)$ are functions of r, variation in r alone cannot separate channels. Heterogeneity by λ helps interpretation because λ shifts inherited norms and thereby changes how contact maps into attitudes.

If historical complementarities produced persistent cooperative norms, transmitted across gen-

erations through families, local institutions, and social networks, then $\bar{\tau}(r, \lambda)$ can sustain political effects even when current economic benefits are small. Nunn and Wantchekon (2011) provide direct evidence that historical shocks to interethnic trust persist over long horizons through cultural transmission.

Empirical implications: The framework leads to three testable implications. First, ethnonationalist support should increase with distance from CNVs. Second, proximity effects on attitudes and political behavior should be stronger (attenuated or reversed) in regions of economic complementarities (competition). Third, political and attitudinal effects should persist even when contemporaneous economic effects are weak.

3 Data

3.1 Sampling Frame

Chinese New Villages. We focus on initially *rural* CNVs, given that the resettlement program largely involved rural-to-rural resettlement. To that end, we construct a baseline sample of 208 CNVs from a historical survey by the Malayan Christian Union (1958). This sample is obtained as follows: We first geolocated 452 New Villages using maps published by the Ministry of Housing and Government (Lee, 2012), manually verified using Google Earth. We then imposed four sample criteria to include only those (i) whose primary language was Mandarin or a Mandarin dialect; (ii) located in seven West Coast states; (iii) located in historically rural census districts;[18] and (iv) with at least 200 Chinese voters in 2013 electoral rolls. Figure 2 shows the geographical distribution of our baseline sample of CNVs. Table C.1 compares in-sample (208) and out-of-sample (244) New Villages and suggests that our in-sample villages are fairly representative.[19] The only statistically significant differences are that in-sample villages are more likely to be located along main highways and to speak a Chinese dialect. The latter is directly reflective of our first constraint; the former possibly reflects our fourth constraint—villages situated along main thoroughfares were more likely to persist till today.[20] Hence, our estimates are particularly informative about the long-run effects of CNVs that have persisted to the present. See Appendix B.1 for full construction

[18] Qualitative fieldwork revealed that resettlement patterns between CNVs in historically rural and urban areas differed markedly. In particular, many "urban" CNV villagers had never experienced resettlement—village authorities simply had contiguous existing house lots fenced up so as to obtain federal government subsidies. See Appendix B.1 for further details.

[19] This data was collected shortly after resettlement (1956-1958) but given (i) movement restrictions remained in place and (ii) no data was collected given security issues in the immediate post-resettlement years, they represent, to the best of our knowledge, the only source of CNV characteristics in the early post-resettlement period.

[20] This is consistent with qualitative fieldwork where we came across largely empty New Villages that had become depopulated over time due to a lack of economic viability and/or (natural) disasters that rendered them inhospitable. We describe one such case of the latter in Appendix B.1.

details and the detailed rationale behind our sample criteria.

Polling districts. We use polling-district-level electoral data from the 2013 and 2018 Malaysian General Elections. This is the most disaggregated level at which data on vote shares are available. In our sample, there are 939 districts in 2013 and 1,004 in 2018, each averaging 10.9 km^2 and 1,519 registered voters.[21] On average, a federal parliamentary constituency seat contains 18 polling districts (Figure 2). To obtain this sample, we imposed four restrictions. First, to ensure that we compare outcomes only across polling districts that could been possible candidates for siting a CNV, we restrict our analysis to districts within 10km of a CNV. Second, to study exposure effects on Malays, we exclude districts that contain an initial CNV. Third, in line with our focus on initially rural CNVs, we restrict our analysis to historically rural districts. Last, given our focus on interactions between ethnic Malays and Chinese, we exclude districts where Indians are the dominant ethnic group.[22] See Appendix B.2 for full construction details.

3.2 Treatment and Primary Outcomes

Treatment: Interethnic proximity. We define interethnic proximity as the straight-line (fly-by-crow) distance from the centroid of each polling district to the nearest CNV (Figure 2). We use fly-by-crow distances instead of historical road distances because a substantial number of pre-existing, non-resettlement villages in the pre-resettlement period were (i) not accessible by main roads and (ii) likely connected through informal dirt paths rather than paved roads. As a result, straight-line distance would more accurately capture potential inter-village movement and contact during this period.

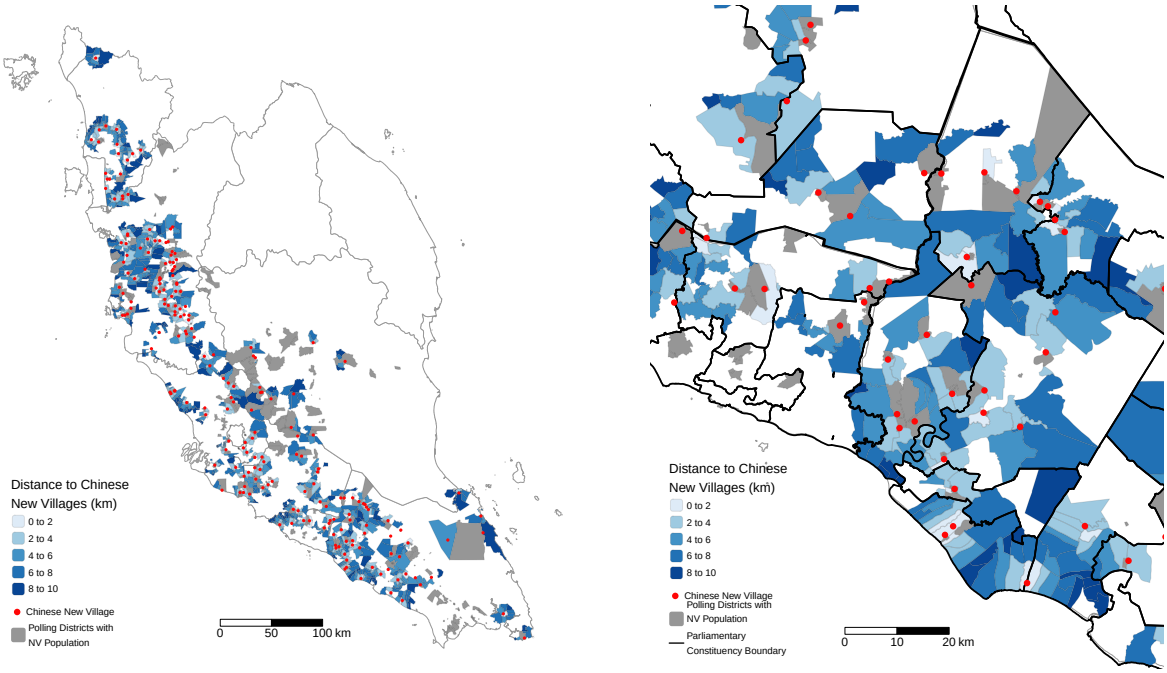
Primary outcome: Vote shares. Our primary political outcome is the National Front (BN) vote share in the 2013 and 2018 General Elections, measured at the polling district level. We interpret BN vote share as a behavioral proxy for Malay ethnonationalist identity.[23] Since Malaysians rarely update registered voting addresses, vote shares primarily reflect preferences shaped by growing up near CNVs, though they may also reflect later-life exposure at migration destinations. All estimations control for proximity to the nearest 1947 urban center to account for this. Figure C.1 maps polling-district BN vote shares in the 2013 federal election. Appendix B.2 provides additional background on party platforms, our choice of BN vote share, and the advantages of polling-district election data for identification.

[21] The number of registered voters is roughly equivalent to an electoral precinct in the US (Longuet-Marx, 2024).

[22] The political economy of these polling districts continues to be markedly different—many were historical plantation lines, where (descendants of) marginalized Indian coolies and rubber tappers, brought into colonial Malaysia as indentured labor, continue to reside today (Kratoska, 1982).

[23] Albeit, there are limitations to this interpretation in 2018 given that PAS—which promoted a conservative Malay-Muslim identity——ran independently.

Figure 2: Proximity to the Nearest CNV at the Polling District-Level



Panel A **Panel B**

Notes: Panel A shows the location of CNVs in red dots and the distances to the nearest CNVs in the estimating sample. Panel B is an enlarged area for illustrative purposes. Polygons depict polling district boundaries. Polygons shaded in darker colors indicate greater distances to the nearest CNV. Black borders represent parliamentary constituency boundaries. In grey, are polling districts that contain a CNV. In white, are polling-districts (i) further than 10km from a CNV; (ii) are historically urban; or (iii) have an ethnic-Indian majority and are hence excluded from our regression sample. We discuss these restrictions in Section 3.

Voter turnout and registered voters. We measure turnout at the polling-district level using data from the Malaysian Electoral Commission, focusing on the period when federal and state elections were held concurrently. We also construct polling-district ethnic shares from full-count 2013 voter rolls, which uniquely record voters' self-identified ethnicity. We are unable to do so using voter rolls from the 2018 General Election, as the self-reported ethnicity variable is missing for a large number of voters.

Local economic development, urbanization, and public goods. Census data are too coarse for our setting. We proxy local development and urbanization using 2014 satellite nighttime lights and 2010 population from the Global Human Settlement Layer, aggregated to 1km$\times$1km grid cells and polling districts. We measure public goods using geocoded administrative lists of schools in 2010 from the Ministry of Education, and health facilities in 2022 from the Ministry of Health, and construct polling-district road density from Open Street Map by computing road kilometers per square kilometer. See Appendix B.3 for data sources and construction details.

Primary individual-level survey data. To investigate the underlying mechanisms behind the political effects, we use our original primary survey data on attitudes and behaviors of Malays toward Chinese. We describe these at length when presenting results in Section 6.1.

4 Empirical Strategy

4.1 Ordinary Least Squares Estimation

Our baseline empirical specification is:

$$Y_{d,p} = \alpha + \sum_{k=1}^{4} \beta_k distCNV_d^k + \gamma \mathbf{X}_d + \theta_p + \epsilon_{d,p} \tag{5}$$

where $Y_{d,p}$ is an outcome of interest in polling district d of federal parliamentary constituency p. $distCNV_d^k$ are indicators equal to 1 if the geodesic distance from the centroid of polling district d to the nearest CNV is 2-4km (k=1), 4-6km (k=2), 6-8km (k=3), and 8-10km (k=4). The omitted bin is 0-2km.[24] $\mathbf{X}_d$ is a vector of predetermined topographic, geo-climatic, and demographic controls defined at the polling-district level.[25] Importantly, demographic controls include distance to the nearest urban center, population density, and ethnic Chinese shares in 1947. These controls ensure that we compare areas with similar levels of pre-treatment diversity and economic prosperity.[26] θ_p is federal constituency fixed effects. We exclude polling districts that contain CNVs ("doughnut hole" approach) and cluster standard errors at the federal parliament seat level.

4.2 Counterfactual Site-Based Identification Strategy

The potentially endogenous placement of CNVs raises concerns that OLS estimates of β_k may be biased. For instance, since British authorities prioritized road access for military logistics (dhu Renick, 1965), proximity to CNVs may correlate with proximity to roads and long-run market access. Similarly, if CNVs were placed in areas where Malays were already more accepting of Chinese settlers, observed voting patterns may reflect pre-existing attitudes rather than the effects of interethnic proximity.[27] To address this, we implement counterfactual analyses inspired by Dell and Olken (2020). We outline our approach here and provide further details in Appendix A.2.

[24] Alternatively, we could extend our analysis to include districts that are greater than 10km and designate all such districts as the omitted control group. Unfortunately, the small number of polling districts that lie within the same parliamentary constituency *and* whose centroids lie more than 10km from a CNV prevent us from doing so.

[25] See "Other Controls" in Appendix B.3 for the full list.

[26] Regressions using state legislative seat election results further control for indicators for every possible combination of party match-ups at the state constituency level.

[27] Balance tests using retrospective survey data provides some evidence against this (See Section 6.1, Table C.4).

This approach leverages the fact that CNV sites were selected rapidly under military duress—with over half a million Chinese being resettled in just three years. This urgency meant that site selection largely followed a set of over-arching military criteria (see Appendix A.1), rather than pre-existing economic or political conditions. No systematic records of counterfactual sites exist. Instead, we retrospectively reconstruct feasible counterfactual CNV locations based on actual CNV locations and British planning criteria. Specifically, we generate 1,000 sets of counterfactual CNV sites that satisfy (1) road-access, (2) topographic-suitability, (3) Malay Reservation exclusions, and (4) spatial-balance.

The key identifying assumption is that realized CNV sites were not uniquely optimal and that similarly suitable alternative sites possibly existed.[28] Rich archival documents support this, with British officials frequently mentioning deviations from optimal sites due to two, key local constraints of: (i) limited prior knowledge of Chinese squatter locations; and (ii) unexpectedly high land acquisition costs. Importantly, it is worth mentioning that our identification strategy does *not* rely on the claim that the British did not have a good *general* knowledge of these regions. Rather, we argue that the British, in many cases, were unlikely to have had sufficiently *local* knowledge nor capacity to pin down optimal CNV sites.

First, the British lacked sufficiently granular, local knowledge of scattered, rural Chinese squatter settlements and small Malay villages. 1947 census enumerators were told explicitly to exclude small, rural villages where: "the value of information was thought to be so slight that it will not justify the time and money spent (Del Tufo, 1949)." There were two implications: (i) officials frequently 'discovered' previously unknown settlements. For example, in one instance, an officer remarked that: "The total strength of the squatters was estimated at 8,000 ... (but) subsequent investigation in the area revealed ... the strength of the squatter was 26,000 (Markandan, 1954)." (ii) The British could largely rely only on data from hasty, pre-resettlement 'squatter' censuses. For example, a senior British official wrote: "The number of illegal squatters are unknown. In all 118 ... only the approx. position of 13 huts are known ... regarding illegal squatters and huts I have no information. This has been a bad area for a long time and there were limited opportunities for inspection."[29]

Second, sub-optimal CNV sites were often chosen due to unexpectedly high costs of acquisitioning private land from non-Malay landowners.[30] This was because a key aim, in addition to military purposes, was economic— British planners sought to minimize labor dislocation of Chi-

[28] This discussion is based on primary documents from the Johor State Archives, the National Archives of Malaysia, the Imperial War Museum in London, and the National Archives of Singapore.

[29] Source: District Forest Officer, Lumut, to Senior Forest Officer, Perak, 13 March 1950, File No. 12 in F.O.D. B. 35/48, National Archives of Malaysia.

[30] Revenue from Korean War-driven rubber sales quickly became insufficient, as estimated projections for the resettlement of each subsequent New Village rapidly exceeded what the British had initially planned for.

nese squatters from rubber estates and tin mines (Humphrey, 1971).[31] Hence, suitable CNV sites were often in private hands and site choice was largely determined by (i) who was willing to sell their land at short notice and (ii) at a low enough price. For instance, in the resettlement of Sungei Way, the eventual CNV site was chosen because it was found to contain mostly old rubber trees, which authorities could reacquisition at a low enough cost: "I regret that I can find no other suitable land. The area further to the west and south is under mining title ... the main road (there) was recently diverted by the company at a cost of over half a million dollars ... the land to the east bordering Petaling Jaya is also under mining title and consists mostly of mining pools."[32] This choice of suboptimal CNV sites would, if anything, suggest that estimated economic effects on proximate Malay villages might be a lower bound.

Data and procedure. To measure road accessibility and exclude sites within Malay Reservations, we digitize the full universe of 1947 pre-resettlement roads and Malay Reservation boundaries using the highest-resolution maps available. To our knowledge, this is the most detailed mapping source available for capturing local-level variation in the pre-resettlement period. Figure C.2 presents an example of these historical maps.

Importantly, despite the British military's goal of minimizing squatter dislocation, not all CNVs were located at sites that had the highest levels of pre-resettlement market access.[33] This is consistent with the historical accounts cited above and further underscores how the lack of sufficiently *local* information led to deviations from optimal site locations.

Figure 3 illustrates the procedure for counstructing counterfactual sites. Panel A shows an actual CNV (Buloh Kasap, purple dot), pre-resettlement roads (red lines), and suitable polygons shaded in green. Panel B shows a suitable counterfactual site (blue dot), and Panel C displays a full set of counterfactual sites that meet all four conditions, including spatial balancing.

Estimating equation. We estimate the polling district-level counterfactual specification using the same controls, $\mathbf{X}_d$, and parliamentary constituency fixed effects, θ_p, as in Equation (5):[34]

$$Y_{d,p} = \alpha + \sum_{k=1}^{4} \beta_{k,\text{real (fake)}}\, distCNV_d^{k,\text{real (fake)}} + \gamma \mathbf{X}_d + \theta_p + \epsilon_{d,p}. \tag{6}$$

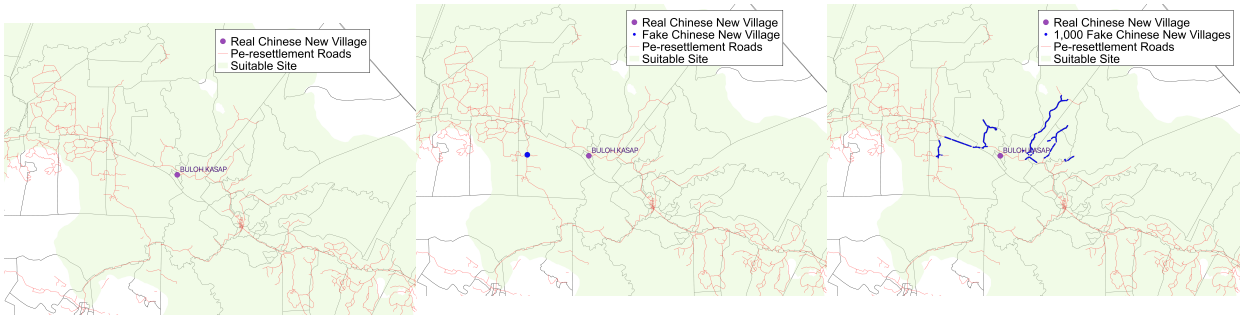
[31] It is important to note that not all villagers in CNVs worked in rubber estates or tin mines, we leverage this variation to study interethnic complementarities in Section 6.

[32] Source: Acquisition of Agricultural Land for Sungei Way New Village, Klang, 1952–1953, File No. 1957/0302416W, Sel.Sec. 2771/52, National Archives of Malaysia. Similarly, in the state of Negeri Sembilan, a mining company outright refused to sell its land despite the pending expiry of its mining lease.

[33] Specifically, for our survey states of Johor and Perak, we can use the near-universe of geolocated, historical villages and the complete road network to construct market access measures (Donaldson and Hornbeck, 2016). We find that market access does not differ significantly between real and fake CNVs, and that the counterfactual results are robust to controlling for the market access of both.

[34] We also estimate economic outcomes at the grid-cell level; the grid-cell specification is analogous.

Figure 3: Counterfactual Chinese New Villages Example



Panel A: Actual CNV | **Panel B: Counterfactual CNV** | **Panel C: 1,000 Counterfactual CNVs**

Notes: This figure illustrates the construction of the counterfactual CNVs, as described in Section 4.2.

For each outcome, we estimate (6) once using realized CNVs and again using each of 1,000 counterfactual CNV sets. The effect for distance bin k is $\hat{\beta}_{k,\text{real}} - \frac{1}{1000}\sum_{s=1}^{1000}\hat{\beta}_{k,\text{fake}}^{(s)}$, which differences out location advantages shared by real and counterfactual sites (e.g. road proximity) while retaining within-constituency comparisons via θ_p and all pre-determined controls as in Equation (5) (including 1947 population density and ethnic Chinese share). We exclude polling districts that contain CNVs (i.e., "doughnut hole") throughout.

To illustrate this, consider a stylized example with only one criterion: proximity to roads. Given this criterion and that nearly all CNVs are located along main roads (Figure 1), the coefficients in Equation 5 implicitly compare outcomes of places that are progressively located further away from main roads to those of CNV sites located right next to a main road. In comparison, the estimation of the coefficients in Equation 6 is analogous to taking a double difference. Specifically, the subtraction allows us to compare the effects of a real CNV, located next to a main road, on surrounding areas located 0-2km away vis-a-vis the effects of a fake CNV, that is similarly located next to a main road, on surrounding areas located 0-2km away: allowing us to directly purge the effects of proximity to roads (i.e., any pre-existing locational advantages).

Figure C.3 provides a graphical illustration of the key steps we take for inference. To compute statistical significance, we follow the randomized inference literature to compare the actual coefficients in Equation 6 to the empirical distribution of the coefficients of 1,000 counterfactual regressions. Specifically, we compute p-values by comparing the position of the $distCNV_d^{k,real}$ coefficient to that of the distribution of absolute values of the 1,000 counterfactual $distCNV_d^{k,fake}$ coefficients. A small p-value implies that patterns near the actual CNVs would have been unlikely to arise in the absence of resettlement. Panel A of Figure C.3 illustrates this procedure.

Balance checks. We check whether polling districts located at varying distances to real and counterfactual CNV sites are balanced on key geographic, topographical, and pre-treatment demo-

graphic characteristics. Ideally, if our approach is valid, we should observe few differences across these variables. Figure C.4 plots the difference between distance-to-real CNV distance coefficients and the average of 1,000 distance-to-counterfactual CNV coefficients relative to the 0-2km distance bin. Significance is denoted by crosses (above 95^{th} percentile), solid dots (above 90^{th}), and hollow dots (below 90^{th}).

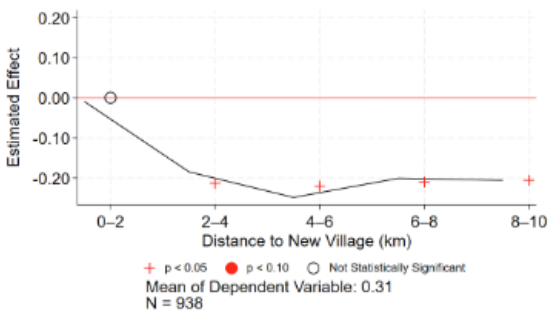
Reassuringly, relative to the 0-2km distance bin, we find no statistically significant differences across all variables except for a slight difference in distance to urban centers in 1947 (Panel J: polling districts located in the 4-6km and 6-8km distance bins are 2-2.5km further from an urban center, relative to the sample mean of 26km) and the percentage share of Chinese in 1947 (Panel M: polling districts located in further distance bins had 3.5-6p.p. lower share of ethnic Chinese, relative to the sample mean of 39%). To that end, in all our regressions, we control for these two variables and all other covariates in Figure C.4. Hence, all regression estimates presented are *conditional on the pre-resettlement distribution of ethnic Chinese in 1947.*[35]

Pre-resettlement exposure to Chinese communities. Our treatment estimates may capture both *increased* exposure to Chinese populations around CNVs and *decreased* exposure in depopulated origin areas. However, precisely geo-located squatter locations from primary survey data indicate this latter confound is likely minimal (See Appendix B.4). First, pre-resettlement Chinese squatter communities were geographically dispersed. Second, due to their remote nature and limited connectivity, Malay communities likely experienced significantly less baseline exposure to the Chinese, compared to post-resettlement patterns of sustained contact subsequent to the lifting of movement restrictions.

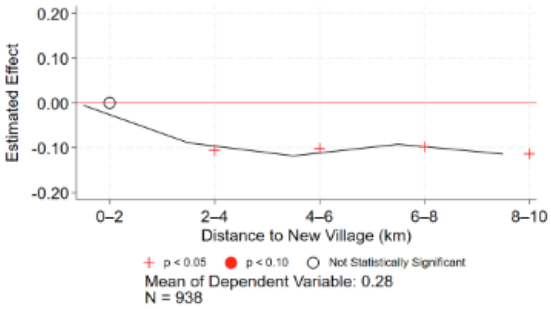
Contemporary Ethnic Segregation We use voter rolls from 7.7 million voters in the 2013 federal elections to compute and regress measures of ethnic segregation on distance to CNVs (Equation (5)). This is the only regression in which we compute our outcome variable by including ethnic shares of CNV polling districts (i.e. of ethnic Chinese themselves). Figure 4 shows that, across typical measures of ethnic polarization and fractionalization (Montalvo and Reynal-Querol, 2005), the mean level of polarization in our estimation sample ranges from 0.31-0.47 and polling districts located within 0-2km of CNVs have around 0.20 points higher levels of polarization (approximately 42-64% higher). In sum, CNV resettlement continues to have a persistent effect on ethnic segregation patterns across Malaysia with areas proximate to CNVs exhibiting substantively higher levels of segregation.

[35]Interethnic contact over the 20-30 years prior to resettlement, including in 1947, might have been conceivably lower relative to the 50-60 year post-Emergency period (starting from 1962) when movement restrictions were lifted and the Malaysian economy became rapidly industrialized. Regardless, given that the bulk of ethnic Chinese migrants arrived in the early 1920s and largely lived in poorly connected, highly dispersed settlements, our treatment estimates might also be interpreted as the exposure effects on Malays of both resettlement and *re-concentration* of the rural

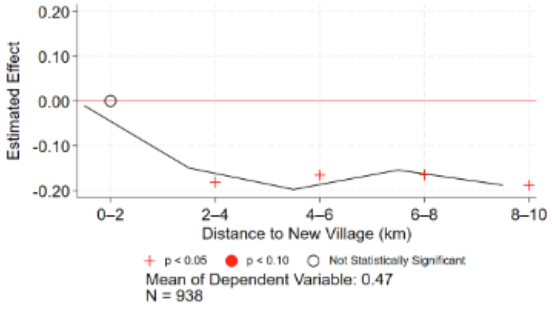
Figure 4: Effects of Proximity on Ethnic Diversity (2013)



RQ Malay and Chinese



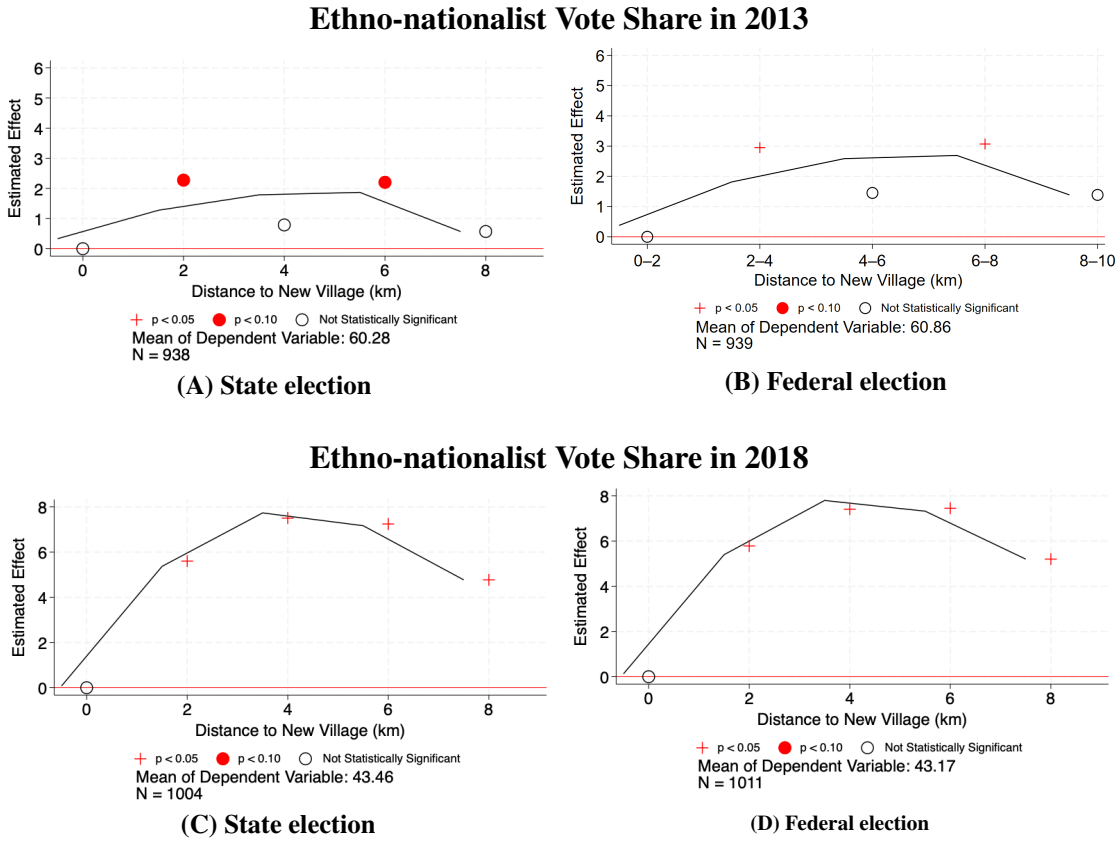
Fractionalization



RQ 3 groups

Notes: Figures plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects, geographical and pre-treatment controls. In addition, regressions of state election results include indicators for all possible combinations of party match-ups at the state constituency level. The means of analogous estimates computed from 1,000 counterfactual New Village configurations are subtracted from each actual coefficient. The points are fit with a linear spline. P-values compare the effect of distance to the nearest actual New Village to the effects of distance to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations. RQ measures are measures of ethnic polarization, constructed following Montalvo and Reynal-Querol (2005). Fractionalization and RQ 3 Groups includes ethnic Indian shares.

Figure 5: Effects of Proximity to CNVs on Ethno-nationalist Electoral Support



Notes: Figures plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects, geographical and pre-treatment controls. In addition, regressions of state election results include indicators for all possible combinations of party match-ups at the state constituency level. The means of analogous estimates computed from 1,000 counterfactual New Village configurations are subtracted from each actual coefficient. The points are fit with a linear spline. P-values compare the effect of distance to the nearest actual New Village to the effects of distance to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

5 Effects on Ethno-nationalist Preferences and Attitudes

5.1 Main Results

Ethno-Nationalist Vote Shares To account for potential biases in the OLS arising from the possibly endogenous placement of CNVs, we primarily present results from our counterfactual analyses.[36] As outlined in Section 4.2, we closely follow British military site-selection criteria

ethnic Chinese population.

[36]Table C.2 presents the OLS results. In the 2013 election (Panel A), polling districts 2–10 km from CNVs had 3–6pp higher support for the ethnonationalist coalition than those within 0–2km, representing a 5–10% increase relative to the 60% mean. The 2018 election (Panel B) shows similar but much stronger spatial patterns, despite an overall national-level drop in support for the ethnonationalist coalition (from 60% to 43%): vote shares in 2–10 km bins were still 7–12pp (or 16–28%) higher. Results are similar using Conley standard errors and population-weighted exposure measures using GHSL. Replicating this analysis at the 1947 Census District level ($N = 77$), we find effects of similar, though imprecisely estimated, magnitudes. All results available upon request.

to generate 1,000 sets of hypothetical resettlement sites such that the point estimate of the effect of being distance k away from a CNV is now given by the difference between the coefficient of $distCNV_d^{k,real}$ and the average of the coefficients of 1,000 $distCNV_d^{k,fake}$. We run these regressions at the polling district level for both state and federal legislative seat elections.

Because inference is based on randomization rather than asymptotic standard errors, we do not report conventional confidence intervals. Instead, p-values are computed by comparing the magnitude of each real-CNV coefficient to the empirical distribution of the corresponding 1,000 fake-CNV coefficients; a p-value of, say, $p < 0.05$ indicates that less than 5% of counterfactual coefficients exceed the real coefficient in absolute value.

Figure 5 presents results for vote shares, plotting estimated coefficients for each distance bin, fitted with a linear spline.[37] Graphs A and B (C and D) show results for the 2013 (2018) elections. These results attest to the possibility that greater proximity to CNVs led to a reduction in support for the ethnonationalist coalition. Throughout, the magnitude of the effects appear to be consistently smaller than OLS estimates. This is to be expected since OLS results likely capture upward bias from unobservables behind original siting decisions, such as better market access, economic fundamentals, or pre-existing political attitudes in areas located further away from CNVs.

Specifically, in the 2013 election, polling districts 2–4km from CNVs show about 2-3pp higher support for the ethnonationalist coalition relative to the 0–2km bin, with effects generally attenuating with distance. For both state and federal elections, the 2–4km and 6–8km bins are statistically different from 0–2km at conventional levels, while effects in the 4–6km bins are positive but not statistically significant. These attenuated effects might reflect the slightly higher prevalence of pre-resettlement Chinese squatter sites in areas 4–6 km from actual CNV sites (Figure B.1).

Panels C and D presents results for vote shares in the 2018 elections. Relative to 0–2km, support is 5–8pp higher across the 2–10km bins, and all estimates are significant at the 5% level.[38] Relative to the 0-2km bin, areas in 4-6km and 6-8km bins are about 8pp *more* likely to vote for BN; whereas areas in 2-4km and 8-10km bins are about 6pp *more* likely to do so, showing a hump-shaped relationship in ethnonationalist support by distance to CNVs. Taken together, these findings reveal the localized nature of political preference shifts induced by proximity to CNVs. The attenuation of effects at greater distances suggests that the mechanisms driving changes in political preferences—whether economic and/or social—are strongest in areas immediately adjacent to CNVs.

[37] Table C.3 reports the same results with detailed quantitative information.

[38] There may be contamination bias if a fake CNV lies in close enough proximity to an actual CNV. If so, some "control" polling districts in the fake regressions (i.e., polling districts within 2-10 km from a fake CNV) would be located within 0-2km of a real CNV, leading to an *underestimation* of the true treatment effects. Therefore, we viw our reported estimates as a conservative lower bound, and these results are robust to excluding any "control" polling district located within 0-2 km from a real CNV, in all 1,000 fake regressions. Results are available upon request.

Zero-Sum, Ethnic Attitudes, and Identity. Observed lower electoral support for the ethno-nationalist coalition could potentially reflect other voter-side preferences. To further understand the extent to which electoral effects capture changes in ethno-nationalist preferences and identity, we turn to original survey data (described in Section 6).

We estimate a simple OLS equation where we regress our outcome variables of interest on a dummy for whether a Malay respondent lives near a real CNV (vis-a-vis a counterfactual, fake CNV) (See Equation 7). We first examine effects on zero-sum attitudes: that is the belief that the gains experienced by one group comes at the expense of another (Foster, 1965). The left panel of Figure 6 shows that treated Malays are less likely to agree with the statements that "Chinese take away Malay jobs" and "Chinese take away business/trade". Interestingly, the effect on "If the Chinese get richer, others get poorer" is not statistically significant.[39] These shifts in zero-sum perceptions might be an important explanation for the decline in support for the ethnonationalist coalition we observe, as zero-sum beliefs have been shown to influence political preferences and support for exclusionary policies (Chinoy et al., 2026).

Turning to effects on ethnic identity and interethnic trust, treated Malays are more likely to primarily identify as Malaysian first (i.e. national identity) and are less likely to agree that having (ethnic) Malay ancestry is important for being Malaysian (first two rows, right panel of Figure 6). In terms of trust, treated Malays are more willing to entrust their kids to their Chinese neighbors (situational trust) but we find no effects on generalized trust towards ethnic Chinese (last two rows, right panel of Figure 6). The latter underlines the potential difficulties of proximity in generating broader, more generalized positive effects (Lowe, 2024).

Taken together, these results suggest that sustained, interethnic proximity over six decades has, concomitant with weakening ethno-nationalist electoral support, led to a weakening of zero-sum beliefs and ethno-nationalistic attitudes. Despite the widely-held perception that the ethnic Chinese control the vast majority of economic resources in Malaysia, treated Malays are *less* likely to believe that the Chinese took away their jobs, business, and trading opportunities. Similarly, interethnic proximity appears to have led treated Malays to hold a more inclusive conception of the national Malaysian identity. This is striking given that the idea of a "Malaysian Malaysia" (racial equality and a decrease in the preponderance of Malay affirmative action) has been a constant clarion call that the multi-ethnic opposition coalition has been striking since the opposition resurgence in 2008. It is also the primary ideology that UMNO, the key ethnic Malay component party of BN was established (even during pre-independence) to campaign and advocate against (Ng, 2010).

[39] Fieldwork suggests that this is possibly because our rural-area respondents found the question to be slightly too abstract and hence difficult to understand.

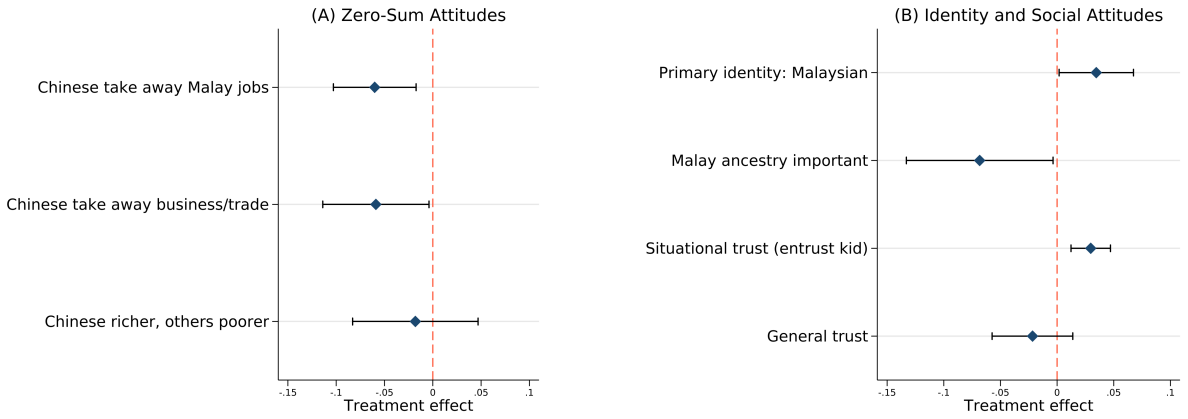


Figure 6: Effects of Proximity on Zero-Sum, Ethnic Attitudes & Identity

Notes: This figure reports OLS estimates using the primary survey data, where dots represent point estimates and blue lines represent 90% confidence intervals. The sample consists of individuals who were born in the surveyed village. "Primary identity: Malaysian" is an indicator variable that takes the value of 1 if a respondent ranks Malaysian identity as his first choice in response to a question that asks respondents to rank Malaysian, (ethnic) Malay, religious, and district identity. "Malay ancestry important" is a standardized outcome variable that asks respondents, from a scale of 1-4, to what extent they agree that having Malay ancestry is important for being (defined as) a Malaysian. "General trust" refers to a standardized outcome variable of the question: "From a scale of 1-4, to what extent do you trust Chinese (in general)". All regressions include nearest CNV fixed effects and two-way clustered standard errors at the cohort and nearest CNV-level.

5.2 Alternative Explanations and Robustness Checks

Our results suggest that Malays living closer to CNVs are less likely to vote for the ethnonationalist coalition. We interpret this as possibly reflecting a shift in political identity and preferences among the ethnic Malay majority. Here, we address several leading, alternative explanations.

Public goods. Political preferences could depend on the availability of public goods where voters reside (Calabrese et al., 2006). In Malaysia, public goods are largely funded by the federal government with limited availability of state and local government funds (Aspinall et al., 2022). Fieldwork further suggests that public goods are seldom funded by bottom-up, within-village sources. Figure C.5 presents results on educational, infrastructure, and health public goods. Panel A, B and D suggest that, aside from a small difference in the number of schools in the 2-4km bin, there are few differences in educational and health public goods across distance bins.[40] Null effects on educational public goods are particularly important given that observed differences in inclusive political behavior could simply have been a result of improved human capital from better access to educational public goods (Glaeser et al., 2007). Our results suggest that this is unlikely.[41]

In comparison, Panel C indicates significantly higher road density near CNVs (the 0-2 km bin) compared to surrounding areas ($p < 0.05$ for the 2-10 km bins). Increased connectivity could have led to greater economic prosperity which, in turn, could be partially responsible for the observed

[40] Data on health public goods is from 2022, well after observed vote shares. Hence, we interpret these results as exploratory. To the best of our knowledge, geolocated data in earlier time periods is unavailable.

[41] Section 7 provides evidence of few differences in the education of Malays close to and further from CNVs.

moderation in political preferences.[42] We further discuss this possibility when discussing other mechanisms in Section 6.

Turnout. Lower vote shares near CNVs could simply reflect higher turnout driven by racial threat (Enos, 2014). We examine the effects of distances to CNVs on voter turnout in the 2013 elections. In Panel A of Figure 7, effects are statistically significant but quantitatively small, ranging from 0.5-0.8pp across the 2-6km bins. Given that we observe differences in vote shares of 3pp, it is unlikely that our results are driven solely by differential turnout rates.

Ethnic composition. Differences in vote shares might reflect differences in ethnic composition if, over time, a greater share of ethnic Chinese chose to reside around CNVs.[43] If this was the case, observed lower vote shares near CNVs might simply reflect higher ethnic Chinese shares. Reassuringly, Panels B and C of Figure 7 show that there are few differences in the ethnic composition of registered voters across distance bins.[44] To further quantify the potential role of ethnic Malay voters in driving the decreases in ethnonationalist vote share, we also perform a back-of-the-envelope calculation. The results in Figure C.7 show that, even under the extreme (assumption) that all ethnic Chinese voters voted *against BN*, observed vote share differences persist unless Chinese turnout exceeds 50–60%. Given that typical Chinese turnout rates rarely surpass 50% (Malay Mail, 2024), this suggests that neither changes in ethnic Chinese turnout nor ethnic composition are sufficient to explain the full decline in *BN* support. See Appendix D.1 for details.

Selective migration. We use primary survey data on the migration histories of about 2,000 Malays residing in villages near CNVs to examine the possibility of selective in- or out-migration of Malays in response to CNV resettlement.Figure C.8 shows that only around 1% of villagers have ever moved out of (and very few have sold their houses and permanently left) Malay villages since 1960, and there are little to no differences in in-/out-migration rates across Malay villages located closer to (treated) vis-a-vis further from (control) CNVs (Panels A and B).[45] Moreover, 78% of respondents were themselves born in their village of residence (Panel C). These figures are

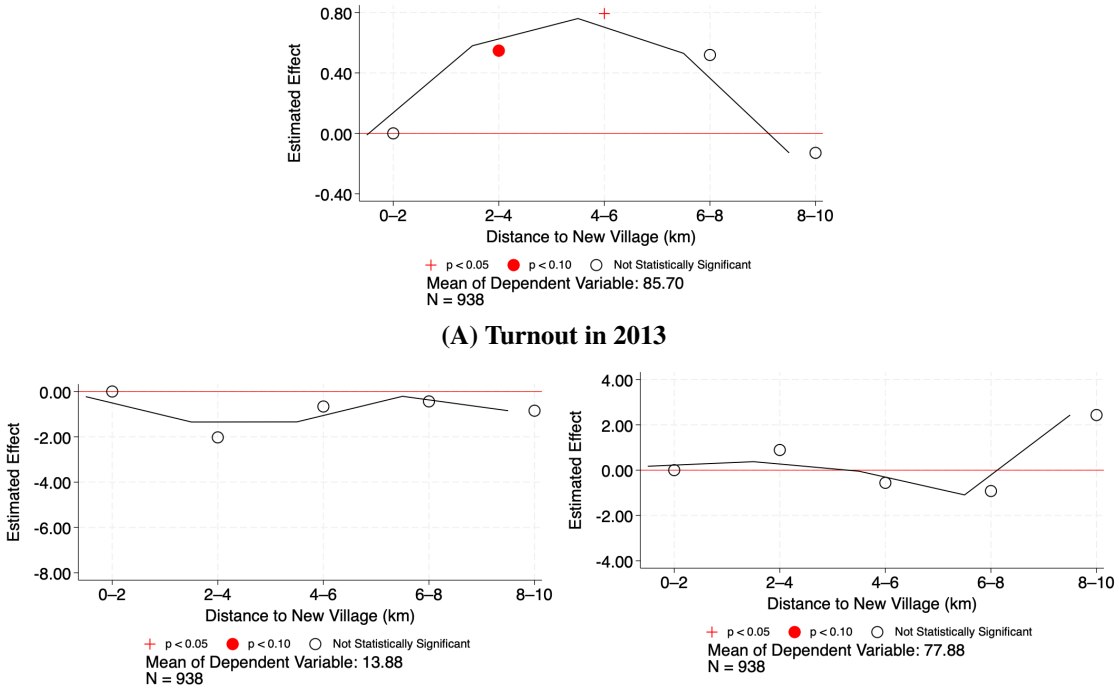
[42] Differences in roads are unlikely to have been driven by the strategic targeting of public goods. For strategic targeting to explain our results, either political parties would have had to differentially target roads towards Malay villages within 0-2km of CNVs and/or Malay leaders residing in these villages would have had been more successful at obtaining public goods. Using primary survey data, we show that Malay village leaders in greater proximity to CNVs are not more likely to perceive that public goods allocations to CNVs eases lobbying for public goods. Results available upon request.

[43] That is, while our doughnut-hole design explicitly excludes polling districts containing initial CNVs, fieldwork indicates that the children of initial Chinese New Villagers tend to settle in adjacent areas after marriage.

[44] Figure C.6 provides a further robustness check by excluding polling districts where the ethnic Chinese share in 1947 exceeded 80%. The estimated coefficients remain unchanged in magnitude and become somewhat more precise.

[45] We chose 1960 as a cutoff date as pilot surveys found that respondents were more likely to recall events from the 1960s (as opposed to the 1949/1950s). Furthermore, free movement was severely curtailed up till the end of the Emergency in 1962. Hence, substantive migration was likely to have occurred only from the 1960s.

Figure 7: Effects of Proximity to CNVs on Voter Turnout and Ethnic Composition



(A) Turnout in 2013

(B) Share of Registered Ethnic Chinese Voters in 2013

(C) Share of Registered Ethnic Malay Voters in 2013

Notes: Panel B plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects, geographical and pre-treatment controls. Each coefficient is computed by subtracting the mean of the analogous counterfactual estimates (computed from 1,000 counterfactual New Village configurations, the distribution of which is displayed in each subpanel of Panel A) from each actual coefficient (the value of which is displayed as a red vertical line in each subpanel of Panel A). The points in Panel B are fitted with a linear spline. p-values compare the effect of distance to the nearest actual New Village to the effects of distance to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

similar across both treatment and control Malay villages (all $p > 0.10$), implying near identical in-/out-migration of Malays in response to Chinese resettlement.[46]

Distaste for coercive resettlement policies? Finally, lower BN vote shares may instead reflect a greater distaste for coercive policies, given that Malay communities living closer to CNVs might have had greater indirect exposure to coercive resettlement policies. This is unlikely for two reasons. First, resettlement was carried out by the British colonial government and not by BN. Hence, it would be unlikely for Malay voters to punish BN at the ballot box for a policy they did not implement. Second, as we show below, a placebo test exploiting a nearly identical forced resettlement program targeting Malays (resulting in the creation of Malay New Villages (MNVs)) finds little impact on exposed Malays' political behavior. If aversion to coercion were driving our results, these effects should be stronger where Malays witnessed the coercive resettlement of co-ethnics.

[46]Given our primary interest in studying exposure effects on Malays, we are less concerned about selective in- and out-migration of *Chinese* villagers. In fact, this is why we chose *not* to study effects on the ethnic Chinese.

6 Potential Mechanisms for Political Behavior

We find a positive effect of proximity to CNVs on the moderation of Malay political behavior. This section uses rich, individual-level, primary survey data to investigate three potential mechanisms. First, we explore how proximity shaped downstream outcomes of interethnic contact, attitudes and social integration (Bazzi et al., 2019; Lowe, 2021, 2024). Second, we investigate the extent to which proximity might have led to greater economic prosperity, weaker zero-sum beliefs, and shared economic interests (Chinoy et al., 2026). Third, we leverage the rich spatial setting of colonial-era rubber estates and tin mines to examine heterogenous effects in relation to the presence of (historical) economic competition versus complementarities (Becker and Pascali, 2019; Horowitz, 2000).

We find that intergroup proximity led to greater contact and positive economic outcomes. The presence of economic competition (complementarities) largely reverses (amplifies) these results. Throughout, we find little evidence of social integration. We further document that, in areas of *historically high* complementarities, proximate Malays continue to exhibit positive interethnic attitudes; weaker zero-sum beliefs; and lower vote shares today *despite* the lack of observed economic effects. We hypothesize that sustained intergroup proximity, through initial economic interdependence, could have resulted in persistent changes in attitudes and zero-sum beliefs. In turn, these changes translated into lasting, positive effects at the ballot box—even as initial economic gains faded away.

6.1 Interethnic Contact: Attitudes, Behaviors, and Economic Outcomes

We begin with micro-level mechanisms, drawing on our novel primary survey data to examine how proximity to CNVs affected interethnic attitudes, behaviors, and economic outcomes of Malays. Understanding these mechanisms is important for interpreting the observed political effects, which may reflect differences in interethnic exposure across locations.

Primary survey on micro-level contact and attitudes: Sampling and variable description. We conducted a novel, in-person survey in partnership with one of Malaysia's leading survey firms, *Ilham*.[47] We do so as there are no existing datasets that capture micro-level measures of interethnic contact or attitudes in this context. We collected individual-level data from December 2024 to June 2025 in both 'treated' and 'control' Malay villages (defined as being located within 2km of a real and fake CNV site, respectively) across the states of Johor and Perak—the two states with the largest number of resettled ethnic Chinese. We selected CNVs (and their associated Malay villages) based on a feasibility criterion: a CNV was included if it had at least one Malay

[47] All fieldwork, data collection, and data cleaning was supervised by our co-author, Gedeon Lim.

village within 2 km, and if its corresponding counterfactual site also had at least one Malay village within 2 km (We further describe our methodology in our *Empirical Strategy* below.) Our survey targeted 1,990 randomly selected Malay males, aged 18 and above, from 75 Malay villages.[48] The sample was stratified by age to ensure representation of both older (60+) and younger (under 60) Malays.[49] To minimize concerns around selection (Section 5.2), our analysis focuses only on respondents who were born in the surveyed villages (about 78% or 1,562 respondents).

The survey captured several dimensions; we focus on seven key dimensions that are likely to be the most relevant to our hypothesized mechanisms—(A) *Contact*: frequency of contact with Chinese peers in villages, schools, and workplaces. (B) *Social integration*: presence of close Chinese friends (strong ties), number of ethnic Chinese phone contacts (weak ties), and acceptance of intermarriage within the family. (C) *Trust*: general trust in Chinese, and willingness to entrust a Chinese person with childcare. (D) *Emotions*: seeing Chinese being treated disrespectfully makes Malay feel bad, and seeing Chinese wealth makes Malay feel jealous. (E) *Economic status & education*: self-reported monthly income and educational attainment, and enumerators' visual assessment of respondents' wealth. (F) *Zero-sum beliefs*: whether Chinese take away jobs and business opportunities from Malays, and the Chinese become richer at the expense of the Malays. (G) *Business and labor market outcomes*: whether the presence of Chinese affects Malays' wage and business profits, and whether both ethnic groups work in the same occupation.

Empirical strategy. We take three key steps. First, for each actual CNV resettlement site, we use 1:1 nearest-neighbor propensity score matching on four key military criteria of elevation, slope, market access to pre-existing villages, and distance to the nearest pre-existing Malay village, to select a single, counterfactual resettlement site.[50] These four variables are our best available proxies for CNV site selection (and closely follows the military criteria underlying the construction of counterfactual sites in Section 5. Also see Section 2.1).[51] In addition, just as in our counterfactual exercise, we continue to restrict the set of possible counterfactual sites only to grid-cells that were located along (historical) roads. Second, we select a maximum of two treated and two control Malay villages based on proximity to, respectively, an actual or counterfactual CNV resettlement site. Specifically, we define a treated (control) Malay village as a village that is located within

[48] We do not have female respondents because the pilot revealed major logistical challenges in recruiting female respondents, given our predominantly Muslim survey sample.

[49] Specifically, enumerators were instructed to conduct stratified random sampling at the village-level. Before beginning surveys, enumerators would visit the village chief to understand the various lanes and neighborhoods in each village. Each enumerator pair would then begin from a representative neighborhood of the village, from a randomly chosen house, and select survey participants by skipping every 3 doors.

[50] We perform matching at the 1 km $\times$ 1 km grid-cell level, the most disaggregated level possible.

[51] Elevation and slope measure defensibility; market access measures ease of sending military reinforcements; and distance to the nearest Malay village takes into account the possibility that the colonial government was concerned with local Malay majority sentiments towards ethnic Chinese settlements. See Appendix A for details.

2km of an actual (counterfactual) CNV resettlement site. Third, we collected individual-level data using in-person, door-to-door surveys in these treatment and control villages.

Balance on pre-resettlement characteristics. Table C.4 presents balance tests to assess whether surveyed Malay villages located within 2km of an *actual* CNV site (treated) differ systematically from those located within 2km of a *counterfactual* CNV site (control). Panel A compares respondents' individual-level characteristics and pre-resettlement, interethnic attitudes based on recollections of their grandparents' attitudes and behavior when they were young.[52] We find balance across individual demographics of age, marital status and the pre-resettlement economic status of a respondent's family ("Family was rich before Chinese resettlement"). Importantly, we largely find balance across a host of pre-resettlement interethnic attitudes such as grandfathers interactions with Chinese, trust of Chinese and perceived incidence of intergroup conflict, across treatment and control villages. Hence, it does not appear that the British located actual CNV sites in areas where Malays had pre-existing, positive sentiments towards the Chinese.[53]

Panel B presents balance tests on village-level geographical characteristics, including those that we did not explicitly match on, such as measures of pre-resettlement population and year of village establishment from primary survey data. Reassuringly, we do not find any significant differences. Overall, the treated and control groups appear well balanced.

To estimate the effects of proximity on the seven key dimensions listed above, we estimate the following equation using OLS:

$$Y_{iv} = \alpha + \beta_1 Treat_{iv} + \theta_{enum} + \theta_c + \epsilon_{iv} \tag{7}$$

where Y_{iv} is the outcome of Malay i in village v; $Treat_{iv}$ is the treatment variable that takes the value of 1 if a Malay lives in a Malay village whose centroid lies within 2km of a *real* CNV, otherwise the treatment variable takes the value of 0 when a Malay lives within 2km of a *fake* CNV; θ_{enum} takes the value of 1 if an enumerator was from East Malaysia; θ_c denotes nearest CNV fixed effects. We report two-way clustered standard errors, clustering by age cohort (above versus below age 60) and nearest CNV.[54] We also report p-values adjusted for multiple hypothesis testing using Romano and Wolf (2005) stepdown procedure.

[52] Specifically, we asked these questions only to respondents aged 35 and above. Respondents were asked to answer these questions based on what they knew about their grandfather and any recollections of stories that their grandfather might have told them about the pre-resettlement period, while he was still alive.

[53] The only statistically significant imbalance is that grandfathers in control villages were more willing to accept marriage with Chinese than those in treated villages. Hence, given control areas appear slightly more open to Chinese at baseline, any observed treatment effect on treated Malays today in terms of positive attitudes would likely be a conservative estimate of the true effect.

[54] Results are largely similar when we control for pre-determined Malay village-level characteristics.

Table 1: The Effects of Proximity on Malay Attitudes and Behaviors toward Chinese

	Contact			Social Capital & Trust	
	(1) $= 1$ if $\geq$ some Chinese prim. or sec. school	(2) $= 1$ if $\geq$ some Chinese colleagues	(3) $= 1$ if ever visited CNV	(4) $= 1$ if have good Chinese friend	(5) $= 1$ if $\geq$ somewhat willing Chinese marriage
Treat	-0.018	0.048***	0.080***	-0.023	0.020
	(0.018)	(0.017)	(0.026)	(0.017)	(0.018)
R^2	0.099	0.036	0.052	0.134	0.070
Mean Dep. Var. (Control)	0.140	0.097	0.161	0.158	0.699
Std. Dev. Var. (Control)	0.347	0.296	0.368	0.365	0.459
Observations	1083	1045	1335	1354	1357
Romano-Wolf p-value	0.186	0.003	0.003	0.236	0.316
Cluster	Cohort $\times$ Nearest Chinese New Village				
FE	Nearest Chinese New Village				

Notes: This table reports OLS estimates using the primary survey data. The sample consists of individuals who were born in the surveyed village. The dependent variables in columns 1 through 5 are indicator variables that take the value of 1 when a Malay: had at least some (>= 10% to more than half) Chinese primary or secondary school classmates (col 1); has at least some (>= 10% to more than half) Chinese colleagues (col 2); has ever visited the nearest Chinese New Village (col 3); has at least one good friend who is Chinese (col 4); is somewhat willing or very willing to have a family member marry a Chinese (col 5). All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level. The Romano-Wolf p-value is computed based on 1000 resamples. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results: Contact, attitudes and behaviors towards Chinese. Table 1 reports our main survey results on Malays' interactions with, and attitudes and behaviors toward the Chinese. Note that the final sample size is reduced by about 200 observations due to missing data on specific control variables. We find that Malays living near a real CNV report significantly more interethnic interactions compared to those in control areas: in terms of visiting the CNVs (column 1, $p < 0.01$) and in workplaces (column 3, $p < 0.01$). Interactions in schools are, however, not statistically significant at conventional levels, possibly because Chinese children primarily study in Chinese-medium schools.

Increased social interactions do not, however, appear to translate into deeper social integration. We find no significant differences in the number and quality of friendships, nor attitudes towards interethnic marriage (columns 4-6, $p > 0.10$).

Results: Economic outcomes and education. Next, we report micro-level economic impacts in Table 2. Malays living near a real CNV are not significantly more likely to engage in non-agricultural jobs compared to those living farther away (column 1). Treated Malays, however, report significantly higher monthly income (column 2, $p < 0.05$).[55] This finding is corroborated by enumerator assessments of respondents' visible wealth, which is based on house size, quality, and plot size (column 3, $p < 0.05$). Taken together, Malays living closer to real CNVs appear

[55]This result is robust to using different thresholds of RM3,000, RM4,000, RM4,500, and RM5,000. We use RM3,500 as it reflects the average monthly earnings in rural Malaysia.

Table 2: The Effects of Proximity on Economic Outcomes and Education of Malays

	Economic Status & Education			
	(1) $= 1$ if non-agricultural job	(2) $= 1$ if high income ($\geq$ RM3,500)	(3) Wealth class (1-4)	(4) Years of education
Treat	0.041	0.045**	0.053**	0.077
	(0.030)	(0.018)	(0.025)	(0.155)
R^2	0.073	0.038	0.076	0.053
Mean Dep. Var. (Control)	0.679	0.134	1.608	10.066
Std. Dev. Var. (Control)	0.467	0.341	0.641	2.610
Observations	1197	1237	1342	1293
Romano-Wolf p-value	0.115	0.012	0.024	0.519
Cluster	Cohort $\times$ Nearest Chinese New Village			
FE	Nearest Chinese New Village			

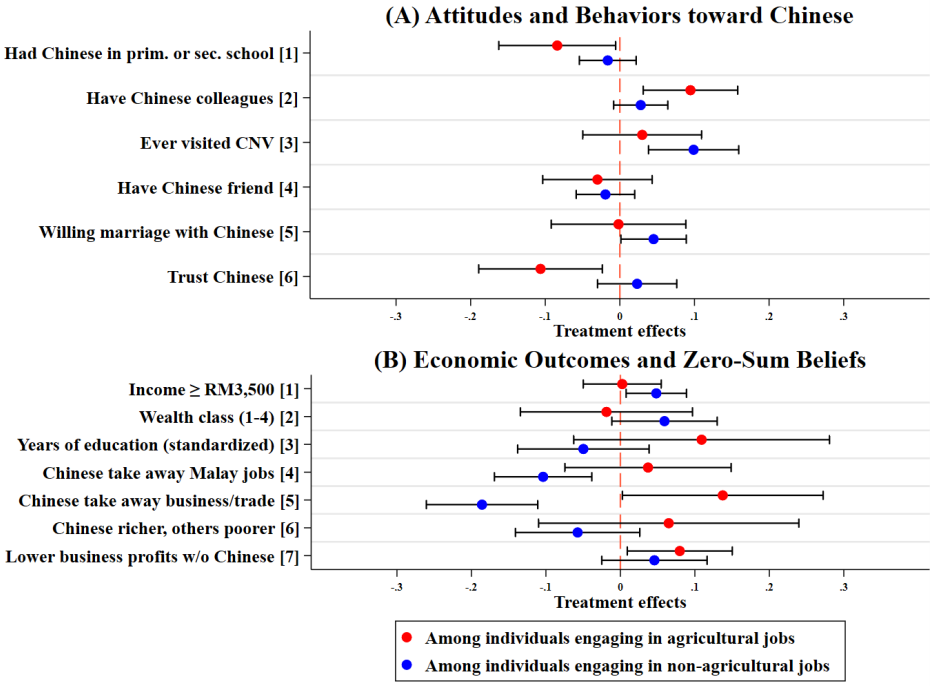
Notes: This table reports OLS estimates using the primary survey data. The sample consists of individuals who were born in the surveyed village. The dependent variables are defined as follows: a non-agricultural job is the current (or most recent) job (col 1); monthly income higher than or equal to RM3,001–RM3,500 (col 2); the categorical wealth class (1=lower, 2=lower middle, 3=upper middle, 4=top) based on enumerators' assessment (col 3); years of education (col 4). The sample size varies across specifications because each outcome variable has a different number of missing observations. All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level. The Romano-Wolf p-value is computed based on 1000 resamples. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

to be economically better off, possibly due to higher value-added jobs *within* sectors, rather than a greater transition towards non-agricultural sectors. This would be consistent with qualitative evidence on product linkages between Malays and Chinese within the agricultural sector (Strauch, 1981). We find no statistically significant difference in years of education (column 4), suggesting that economic gains are unlikely to be driven by human capital accumulation.

Heterogeneity: contemporary economic competition vs. complementarities. Naturally, the presence of economic competition and complementarities might mediate the effects of proximity and contact. For instance, economic competition might lead to negative effects on attitudes and outcomes. Malaysia presents a rich setting to study these effects given that the agriculture sector has long been associated with Malays given their constitutional rights as "Sons of the Soil" (Horowitz, 2000). Figure 8 presents evidence on heterogeneity in attitudes and economic outcomes among Malays working in contemporary agricultural (red dots) versus non-agricultural jobs (blue dots). The motivation is that interethnic competition (complementarities) might have been more (less) salient for Malays in agricultural work, given that the agricultural sector is typically perceived as the main source of employment for Malays.

We note, however, that contemporary occupational choice might itself be an endogenous out-

Figure 8: Contemporary Interethnic Competition vs Complementarities: Heterogeneity based on Malay Agricultural vs Non-Agricultural Jobs



Notes: This figure reports OLS estimates using the primary survey data, where colored dots represent point estimates and black lines represent 90% confidence intervals. The sample consists of individuals who were born in the surveyed village. The sample split is based on whether a respondent's current primary job is agricultural or non-agricultural ($n = 360$ and 848, respectively). For respondents who are currently unemployed, we classify occupations based on their most recent job. The dependent variables in Panel A (B) are the same as those reported in Table 1 (Table 2) and variable definitions are found in the respective table notes. All regressions include nearest CNV fixed effects and two-way clustered standard errors at the cohort and nearest CNV-level.

come of past differences in attitudes and economic prosperity. For instance, Malays with more negative attitudes towards the Chinese might self-select into jobs that minimize interactions with Chinese. To that end, we view these results as exploratory, and further test for heterogenous effects of proximity and contact based on *historical* occupational shares (See Section 6.2).

Agricultural sector Within Malays in the agricultural sector (red dots), our results are largely suggestive of economic competition. Treated Malays are *more* likely to report having Chinese colleagues relative to control Malays (row 3 of Panel A). Yet, greater proximity here translates into significantly *lower* generalized trust and empathy (rows 7 and 9, Panel A) and a heightened perception that the Chinese pose a competitive threat to businesses (rows 5, Panel B).

Non-agricultural sector Conversely, within Malays in the non-agricultural sector (blue dots), the evidence points towards the presence of economic complementarities. Despite not having more Chinese colleagues (row 3 of Panel A), treated Malays report higher situational trust and empathy (rows 8 and 9, Panel A). Economically, these individuals report higher incomes. In terms of zero-sum beliefs, they are significantly *less* likely to view the economic landscape through a zero-sum lens, perceiving less competition for jobs and business opportunities (rows 4 and 5, Panel B).

Taken together, these findings, though admittedly suggestive, point to interethnic economic competition and complementarities as a potentially important mechanism of persistence for translating proximity into improved social cohesion and shared economic sentiments. We provide further evidence below.

6.2 Historical Competition vs. Complementarities

Interethnic proximity can promote negative political attitudes through economic competition (Becker and Pascali, 2019; Horowitz, 2000) or positive attitudes through complementarities and specialization. This could, in turn, affect productivity, social cohesion, and tolerance toward out-groups (Alesina and La Ferrara, 2005; Jha, 2013). At the time of CNV resettlement, nearly 70% of Malays were employed in (non-cash-crop) agriculture (Del Tufo, 1949). If proximity to (historically) agricultural CNVs, led to greater competition over agricultural resources (such as land and water) and economic strain for local Malays, we would expect higher support for ethnonationalist parties in areas immediately surrounding agricultural CNVs relative to more distant areas. Conversely, (historically) non-agricultural CNVs (particularly, those engaged in rubber tapping or tin mining) may have generated economic complementarities, potentially promoting positive, downstream economic interactions. In this case, we would expect lower ethnonationalist vote shares near CNVs associated with these sectors, as economic benefits and/or intergroup contact facilitated better interethnic relations and more moderate political attitudes.

To test these hypotheses, we use two separate datasets. First, we conduct a macro-level analysis (polling district-level) based on heterogeneity in CNV-distance to historical rubber estates and tin mines. We focus on rubber estates and tin mines as these were two of the largest non-farm economic sectors, both for CNV and Malay villagers (Del Tufo, 1949).[56] We compute fly-by-crow distances from each CNV to the nearest rubber estate or tin mine and test for heterogeneous exposure effects between polling districts located at varying distances from CNVs $<=$ 5km of a rubber estate or tin mine, and those that are located at varying distances from CNVs $>$ 5km of a rubber estate or tin mine. Second, using primary survey data, we conduct a heterogeneity analysis at the individual-level. Specifically, we split our sample into: (i) villages where the rubber and tin sector accounted for the largest employment shares in 1960 (*historically high* complementarities); (ii) villages where the two sectors did not account for the largest employment shares in 1960 (*historically low* complementarities).[57]

The (historical) presence of rubber estates and tin mines in Malaysia offer a rich setting to study the effects of interethnic economic complementarities stemming from extensive labor and product

[56] Specifically, we digitize pre-resettlement rubber estate polygons and tin mines using high-resolution HIND1035 maps (ANU Library archives) and the Map of the Malayan Peninsula 1891 (NLA).

[57] For each *Malay* village in our survey sample, we asked *Malay* village leaders to list 2-3 job sectors with the highest employment shares in 1960. We classify villages using this data.

market linkages. First, rubber and tin fostered complementarities through task and occupational specialization. Rubber production, in particular, required multiple complementary tasks within estates. For example, British estate owners frequently employed both Chinese and Malays but assigned them to distinct, complementary roles (Lees, 2017).[58] Separately, the Chinese presence, particularly around tin mines, increased the demand for agricultural products from Malays (Hirschman and Yeoh, 1979).[59] Second, these sectors generated non-replicable and non-expropriable complementarities: Chinese commercial networks and access to international trade were costly for Malays to build independently (Gungwu, 2021; Macauley, 2021); while Malays supplied abundant labor that the Chinese alone could not provide and mobilize. In turn, the mobility of the Chinese potentially deterred Malays from pursuing exclusionary actions so as to retain mutually beneficial economic benefits from the Chinese presence.[60] These features align with the conditions for social cohesion highlighted by (Jha, 2013, 2018).

Importantly, there has been a marked change in the contemporary participation of ethnic Chinese New Villagers across economic sectors. Today, the tin mining industry has collapsed and there has been a significant reduction in the number of rubber plantations (Shah, 2019). Hence, most rural ethnic Chinese New Villagers today are engaged in the agricultural production of vegetables, fruits and palm oil, with a size-able minority engaged in commerce and trade.[61] There is also a well-documented trend of ethnic Chinese New Villagers out-migrating to towns and cities given limited economic opportunities in and around CNVs (Strauch, 1981, 2013).[62] This will be important for interpreting results.

Macro-level results: Proximity of CNVs to historical rubber estates and tin mines. Figure 9 presents results from our counterfactual analyses. Throughout, we control for the identical set of controls as in Equation 5, including important pre-resettlement characteristics such as distance to urban centers and ethnic Chinese shares in 1947. Graphs A-C displays results for polling districts located at varying distances from CNVs that *did not* have a historical rubber estate or tin min located within 5km (*low* complementarities). Graphs D-F displays results for polling districts at varying distances from CNVs that *had* a historical rubber estate or tin min located within 5km (*high* complementarities).

[58] For instance, Malays were often hired for clearing of scrub land and construction. Subsequently, the ban on indentured Indian and Chinese laborers in 1910 led estate owners to hire more Malay workers as rubber tappers. By 1940, there were 45,000 Malay workers vis-a-vis 77,000 Chinese workers on rubber estates (pp.183 of Lees (2017). These numbers likely increased during the Korean War-driven expansion in rubber estates (Sandhu, 1964).

[59] Due to data limitations, we are unable to disentangle these channels but note that complementarities would be more likely to have arisen through rubber given higher Chinese *and* Malay employment on estates.

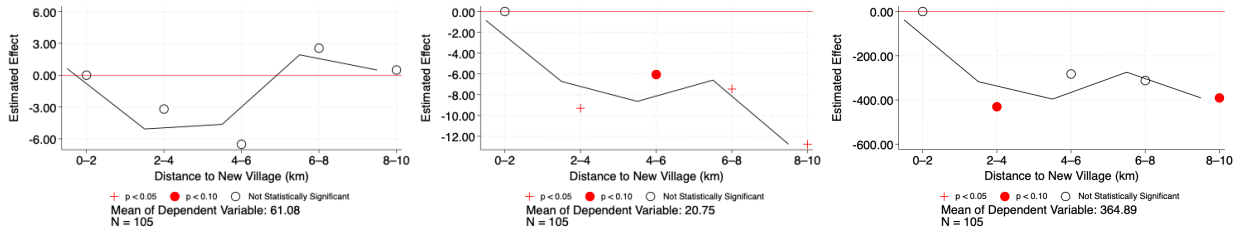
[60] Even though the rural Chinese might have lacked the global mobility of their urban counterparts, many possessed viable outside options for domestic, rural-to-urban migration through extensive Chinese networks (Strauch, 1981).

[61] This is reflected in our surveys where 83% (14%) of CNVs' main occupation today belong to the former (latter). This breakdown does not differ between historically agriculture CNVs vs historically tin and rubber CNVs.

[62] Hence, and given the lack of a comparable control group, we do not focus on ethnic Chinese in our analyses.

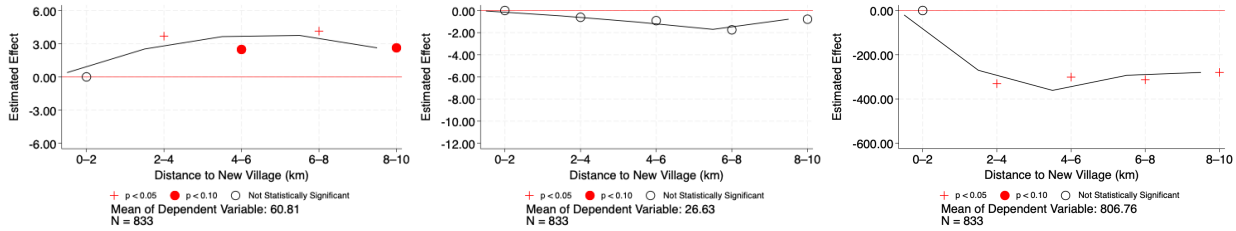
Figure 9: Effects of Proximity to CNVs by Interethnic Complementarities

Historically *Low* Complementarities: No Rubber Estate or Tin Mine $\leq$ 5km



(A) Vote Share **(B) Mean Luminosity** **(C) Population Density**

Historically *High* Complementarities: Rubber Estate or Tin Mine $\leq$ 5km



(D) Vote Share **(E) Mean Luminosity** **(F) Population Density**

Notes: The outcome variables for (A) and (D) are the vote share for the ethno-nationalistic coalition at the federal elections in 2013. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects and geographical and pre-treatment controls. The means of analogous estimates computed from 1,000 counterfactual New Village configurations are subtracted from each actual coefficient. The points are fit with a linear spline. P-values compare the effect of distance to the nearest actual New Village to the effects of distance to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

Political effects Graph A (D) displays results for vote shares in the *low (high)* complementarities sample. We find statistically insignificant differences across all distance bins in the *low* complementarities sample (Graph A). In contrast, in the *high* complementarities sample, we find a negative and statistically significant impact on ethno-nationalist vote shares in polling districts immediately adjacent to CNVs ($p < 0.05$, Graph D). Specifically, we observe a 4p.p. lower vote share for the ethnonationalist coalition in polling districts located within the 0-2 km distance bin, relative to the 2-4km bin (and comparable effects relative to the 4-10km bins). Notably, this effect size is larger than that of the pooled sample (3p.p. in Graph B of Figure 5).

Economic effects (low complementarities) Graphs B-C displays results for economic development in the *low* complementarities sample. In contrast to null effects on vote shares, we find largely positive effects on nighttime light intensity and population density for polling districts located within the 0-2km distance bin vis-a-vis those located in less proximate distance bins.[63] One explanation could be the shift of ethnic Chinese from small to medium and large-scale agricultural

[63] Albeit, effects in the 4-8km bin for population density are statistically insignificant.

production. Such businesses today employ a large number of international migrant workers. In turn, labor expansion may have generated demand for non-agricultural services of Malays, contributing to the observed increases in nighttime luminosity.

Why did contemporary economic benefits in areas of *historically low* complementarities *not* lead to positive political effects? One possible explanation is that changes in political behavior and attitudes may respond more slowly than changes in recent economic activity. Another is that, in this subsample, Malay villages in areas close to CNVs were initially more underdeveloped (than those in high complementarities areas) to begin with. In turn, highly localized agglomeration forces from the contemporary economic transition of CNVs to medium/large-scale agriculture, led to these Malay villages becoming significantly more developed than less proximate Malay villages.

Economic effects (high complementarities) Graphs E-F displays results for economic development in the *high* complementarities sample. In contrast to negative effects on vote shares, effects on nightlight light intensity are insignificant across all distance bins (Graph E). Effects on population density are positive and statistically significant (Graph F) for polling districts located within 0-2km. Together, these results suggest that contemporary economic prosperity has attenuated in these areas over time. This is consistent with the decline of rubber and tin industries (Shah, 2019), which, in turn, led to the gradual out-migration among ethnic Chinese from former rubber tapping and tin mining CNVs (Strauch, 1981, 2013).[64]

Negative effects on vote shares together with muted effects on economic prosperity suggest a potential role for *historically high* interethnic complementarities in driving observed results on voting behavior. Historical complementarities might have led to past, shared economic agglomeration effects and prosperity with proximate Malay villages. Although these gains have largely faded (as reflected in null effects on nighttime light intensity), positive effects on political attitudes appear to have persisted over time. These results, however, might not fully capture individual-level differences in Malay attitudes. We next turn to heterogeneity analyses of primary survey data, based on historical occupational shares of Malay villages.

Micro-level results: Malay villages with high vs low employment in rubber and tin. Figure C.9 plots survey outcomes separately for Malay villages with *low* (red) versus *high* (blue) historical employment in rubber estates and tin mines (i.e. *low* versus *high* complementarities). Overall, we do not observe any statistically significant differences in effects *across* both sub-samples. For a subset of outcomes, the small sample size in the *low* sub-sample further prevents the estimation

[64]The nature of historical economic spillovers from rubber and tin might also explain the lack of observed effects of proximity on contemporary economic development. In these areas, initially positive economic benefits from CNVs might have been larger and/or more geographically dispersed across Malay villages within a wider radius. Hence, the collapse of these industries might have led to an overall observed decline in the economic prosperity of *all* Malay villages (and little contemporary differences in economic prosperity across distance bins). The substantively lower mean of nightlights and population density in Panels B-C vis-a-vis Panels E-F supports this interpretation.

of standard errors.[65] We therefore treat these results as exploratory and describe specific patterns *within* each subsample of treated and control villages.

Within areas of *low complementarities*, proximity to a real CNV leads to higher income and wealth (Panel B: Rows 1-2)—-potentially consistent with higher demand effects from proximate CNVs (as noted in our macro-level results). Effects on zero-sum beliefs, however, appear to be small and are largely statistically insignificant (Panel B: Rows 4 & 6).[66] In contrast, within areas of *high* complementarities, proximity to real CNVs leads to consistently *weaker* zero-sum beliefs about jobs, business, and in-group gains (Panel B: Rows 4-6); higher situational trust (Panel A: Row 8); and envy (Panel A: Row 10). Notably, though suggestive, treated Malays do not have higher levels of income nor wealth today. Taken together, these results are consistent with a role for intergroup production links in shaping attitudes (Ghosh, 2025) and, importantly, attests to how initial economic gains might have faded in areas of *historically high* complementarities.

Takeaways. We interpret these results as evidence that Malays living in close proximity to CNVs might have experienced *initial* economic benefits from employment in rubber estates and tin mines. These economic benefits led to positive interethnic attitudes that persisted over time and which are reflected in differences in contemporary voting behavior (see Section 5). Economic benefits, however, might have faded over time given that many ethnic Chinese New Villagers have out-migrated to towns and cities due to perceived, limited economic opportunities around CNVs (Strauch, 1981, 2013). Taken together, we hypothesize that sustained intergroup proximity could have, through initial economic interdependence, resulted in persistent changes in interethnic attitudes and zero-sum beliefs. In turn, these changes appear to have translated into concrete, long-lasting, positive effects at the ballot box, even as initial economic gains have faded away.

7 Conclusion

In this paper, we leverage a forced resettlement program to study how persistent interethnic proximity shapes long-run political and economic development in Malaysia. More than seven decades later, Malays living closer to Chinese resettlement sites exhibit lower support for the ethnonational-

[65] Panel A, Rows 2, 6, 10; Panel B, Row 7. This is expected, given that the empirical strategy for our primary survey focused on selecting treated and control Malay villages based on *proximity to CNVs* (See *Empirical strategy*, Section 6). Given the extensive complementarities between Chinese and proximate Malays that we document, a large proportion of Malay village(r)s in our primary survey sample were historically involved in non-agricultural jobs.

[66] Notably, results on the indicator for "Chinese take away business/trade" are statistically significantly negative in the low complementarities (Panel B: Row 5). This, however, is possibly consistent with the possibility that, in the low complementarities subsample, Malays living near real CNVs specialized in agriculture, might have had little contact with ethnic Chinese that were, historically, largely engaged in individual small-scale agriculture and family farms. Hence, these Malays might be less likely to hold *any* prior stereotypes of ethnic Chinese. This is partially consistent with qualitative fieldwork.

ist coalition, a more inclusive conception of national identity; weaker attachment to ethnic identity; and weaker zero-sum beliefs. Economic gains alone do not appear to account for the entirety of observed political effects. To shed light on mechanisms, we conducted a large-scale, in-person, retrospective mixed-methods survey of ethnic Malays to measure interethnic contact and political attitudes across generations.

Malays residing closer to CNVs show increased interethnic contact, higher trust, lower zero-sum beliefs, and better economic outcomes. Many of these effects are attenuated (amplified) by economic competition (complementarities). However, effects on broader social integration and attitudes remain muted, underscoring the limitations of interethnic contact in fostering broader social cohesion. The shift in observed political behavior appears to have emerged from routine exposure—especially in the labor market—and historical interethnic economic complementarities. Naturally, this raises the question of *which* contexts might lend themselves more readily to such positive effects of proximity. We cannot answer this conclusively, but these conditions (in particular, the opportunity for repeated, equal status interactions (Allport, 1954)) may be replicable, policy-wise, in highly segregated urban environs. This relates to, for instance, Wong (2013), who studies an ethnic integration policy that mandates ethnic quotas in Singaporean public housing.

As conflict-driven, forced displacement continues to rise, our findings carry renewed relevance. Understanding how spatial integration, even under adverse beginnings, can moderate identity and political behavior is important for designing policies aimed at promoting social cohesion. Further research, such as that on the role of local governance and cohort-level exposure across an individual's life cycle might be particularly fruitful.

References

Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800.

Algan, Y. and Cahuc, P. (2010). Inherited trust and growth. *American Economic Review*, 100(5):2060–2092.

Algan, Y., Hémet, C., and Laitin, D. D. (2016). The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation. *Journal of Political Economy*, 124(3):696–733.

Ali, S. N., Mihm, M., and Siga, L. (2025). The political economy of zero-sum thinking. *Econometrica*, 93(1):41–70.

Allport, G. W. (1954). The nature of prejudice. *Addison-Wesley google schola*, 2:59–82.

Artiles, M. (2026). Ethnic diversity, historical economic exchange, and development: Evidence from andean peru. *American Economic Journal: Applied Economics*, 18(2):229–265.

Aspinall, E., Weiss, M. L., Hicken, A., and Hutchcroft, P. D. (2022). *Mobilizing for elections: Patronage and political machines in Southeast Asia*. Cambridge University Press.

Balietti, S., Getoor, L., Goldstein, D. G., and Watts, D. J. (2021). Reducing opinion polarization:

Effects of exposure to similar people with differing political views. *Proceedings of the National Academy of Sciences*, 118(52):e2112552118.

Bazzi, S., Gaduh, A., Rothenberg, A. D., and Wong, M. (2016). Skill transferability, migration, and development: Evidence from population resettlement in indonesia. *American Economic Review*, 106(9):2658–2698.

Bazzi, S., Gaduh, A., Rothenberg, A. D., and Wong, M. (2019). Unity in diversity? how intergroup contact can foster nation building. *American Economic Review*, 109(11):3978–4025.

Becker, S. O. and Pascali, L. (2019). Religion, division of labor, and conflict: Anti-semitism in germany over 600 years. *American Economic Review*, 109(5):1764–1804.

Billings, S. B., Chyn, E., and Haggag, K. (2021). The long-run effects of school racial diversity on political identity. *American Economic Review: Insights*, 3(3):267–284.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, 97(2):298–319.

Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2024). The immigrant next door. *American Economic Review*, 114(2):348–384.

Calabrese, S., Epple, D., Romer, T., and Sieg, H. (2006). Local public good provision: Voting, peer effects, and mobility. *Journal of Public Economics*, 90(6-7):959–981.

Calderon, A., Fouka, V., and Tabellini, M. (2023). Racial diversity and racial policy preferences: the great migration and civil rights. *The Review of Economic Studies*, 90(1):165–200.

Carlitz, R., Morjaria, A., Mueller, J., and Osafo-Kwaako, P. (2025). State building in a diverse society. *Review of Economic Studies*, 92(6):3704–3740.

Carvalho, J.-P., Bergeron, A., Henrich, J., Nunn, N., and Weigel, J. (2023). *Zero-sum thinking, the evolution of effort-suppressing beliefs, and economic development*. Number w31663. National Bureau of Economic Research.

Caselli, F. and Coleman, W. J. (2013). On the theory of ethnic conflict. *Journal of the European Economic Association*, 11(suppl_1):161–192.

Chetty, R., Hendren, N., and Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902.

Chinoy, S., Nunn, N., Sequeira, S., and Stantcheva, S. (2026). Zero-sum thinking and the roots of us political differences. *American Economic Review*, 116(3):1052–1096.

Del Tufo, M. V. (1949). *Malaya, comprising the Federation of Malaya and the Colony of Singapore: A report on the 1947 census of population*. Crown Agents for the Colonies.

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

dhu Renick, R. (1965). The emergency regulations of Malaya: Causes and effect. *Journal of Southeast Asian History*, 6(2):1–39.

Dippel, C. (2014). Forced coexistence and economic development: Evidence from native american reservations. *Econometrica*, 82(6):2131–2165.

Dobby, E. H. G. (1952). Resettlement transforms malaya: A case-history of relocating the population of an asian plural society. *Economic Development and Cultural Change*, 1(3):163–189.

Donaldson, D. and Hornbeck, R. (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Enos, R. D. (2014). Causal effect of intergroup contact on exclusionary attitudes. *Proceedings of the National Academy of Sciences*, 111(10):3699–3704.

Foster, G. M. (1965). Peasant society and the image of limited good. *American Anthropologist*, 67(2):293–315.

Fouka, V., Mazumder, S., and Tabellini, M. (2022). From immigrants to americans: Race and assimilation during the great migration. *The Review of Economic Studies*, 89(2):811–842.

Ghosh, A. (2025). Religious divisions and production technology: Experimental evidence from india. *Journal of Political Economy*, 133(10):3249–3304.

Glaeser, E. L. (2005). The political economy of hatred. *The Quarterly Journal of Economics*, 120(1):45–86.

Glaeser, E. L., Ponzetto, G. A., and Shleifer, A. (2007). Why does democracy need education? *Journal of Economic Growth*, 12:77–99.

Gungwu, W. (2021). *Home is Where We Are*. NUS Press.

Guriev, S. and Papaioannou, E. (2022). The political economy of populism. *Journal of Economic Literature*, 60(3):753–832.

Hirschman, C. and Yeoh, S.-P. (1979). Ethnic patterns of urbanization in peninsular malaysia, 1947-1970. *Asian Journal of Social Science*, 7(1):1–19.

Horiwitz, D. L. (1989). Incentives and behaviour in the ethnic politics of sri lanka and malaysia. *Third World Quarterly*, 11(4):18–35.

Horowitz, D. L. (2000). *Ethnic groups in conflict*. University of California Press.

Humphrey, J. W. (1971). *Population resettlement in Malaya*. Northwestern University.

Jedwab, R., Johnson, N. D., and Koyama, M. (2019). Negative shocks and mass persecutions: evidence from the black death. *Journal of Economic Growth*, 24(4):345–395.

Jha, S. (2013). Trade, institutions, and ethnic tolerance: Evidence from south asia. *American Political Science Review*, 107(4):806–832.

Jha, S. (2018). Trading for peace. *Economic Policy*, 33(95):485–526.

Jha, S. (2025). Breaking through the ethnic growth trap. *Economic Policy*, 40(123):765–802.

Jomo, K. (2017). The new economic policy and interethnic relations in malaysia. In *Global Minority Rights*, pages 239–266. Routledge.

Kim, K. K. (1998). Malaysia: Immigration and the growth of a plural society. *Journal of the Malaysian Branch of the Royal Asiatic Society*, 71(1 (274):1–25.

Kratoska, P. H. (1982). Rice cultivation and the ethnic division of labor in British Malaya. *Comparative Studies in Society and History*, 24(2):280–314.

Lee, T. H. (2012). *Chinese schools in Peninsular Malaysia: The struggle for survival*. Institute of Southeast Asian Studies.

Lees, L. H. (2017). *Planting empire, cultivating subjects: British Malaya, 1786–1941*. Cambridge University Press.

Lim, T.-H. and Song, W.-Y. (2002). 50 years of chinese new villages in malaysia. *(Revised), Kuala Lumpur: Malaysian Chinese Research Centre*.

Long, J., Medici, C., Qian, N., and Tabellini, M. (2022). The impact of the chinese exclusion act on the us economy. *Harvard Bus. Sch. Work. Pap*, pages 23–008.

Longuet-Marx, N. (2024). Party lines or voter preferences? explaining political realignment. In *2024 APPAM Fall Research Conference*. APPAM.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, 111(6):1807–44.

Lowe, M. (2024). Has intergroup contact delivered? *Annual Review of Economics (forthcoming)*.

Macauley, M. (2021). *Distant shores: colonial encounters on China's maritime frontier*. Princeton University Press.

Malay Mail (2024). Chinese voter turnout surges to nearly 50pc in Mahkota polls, says DAP's Chin Tong. Accessed: 2024-11-04.

Malayan Christian Union (1958). *A survey of the New Villages in Malaya*. International Missionary Council Archives.

Markandan, P. (1954). *The Problem of the New Villages in Malaya*. Number 1. D. Moore.

Matranga, A. and Natkhov, T. (2026). All along the watchtower: Military landholders and serfdom consolidation in early modern russia. *The Review of Economic Studies*, 93(4):2784–2818.

Montalvo, J. G. and Reynal-Querol, M. (2005). Ethnic polarization, potential conflict, and civil wars. *American Economic Review*, 95(3):796–816.

Montalvo, J. G. and Reynal-Querol, M. (2021). Ethnic diversity and growth: Revisiting the evidence. *Review of Economics and Statistics*, 103(3):521–532.

Nagaraj, S. (2009). Intermarriage in malaysia. *Malaysian Journal of Economic Studies*, 46(1):75.

Ng, I. (2010). *The Singapore lion: A biography of S. Rajaratnam*. Institute of Southeast Asian Studies.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7):3221–3252.

Nyce, R. (1973). *Chinese New Villages in Malaya: A community study*. Malaysian Sociological Research Institute.

Ostwald, K. (2017). Malaysia's electoral process: The methods and costs of perpetuating umno rule. *Available at SSRN 3048551*.

Ostwald, K. and Oliver, S. (2020). Four arenas: Malaysia's 2018 election, reform, and democratization. *Democratization*, 27(4):662–680.

Padró i Miquel, G. (2007). The control of politicians in divided societies: The politics of fear. *The Review of Economic Studies*, 74(4):1259–1274.

Peters, M. (2022). Market size and spatial growth—evidence from germany's post-war population expulsions. *Econometrica*, 90(5):2357–2396.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in delhi schools. *American Economic Review*, 109(3):774–809.

Ravallion, M. (2020a). Ethnic inequality and poverty in malaysia since may 1969. part 1: Inequality. *World Development*, 134:105040.

Ravallion, M. (2020b). Ethnic inequality and poverty in malaysia since may 1969. part 2: Poverty. *World development*, 134:105039.

Romano, J. P. and Wolf, M. (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469):94–108.

Ross, C. (2014). The tin frontier: Mining, empire, and environment in southeast asia, 1870s–1930s. *Environmental History*.

Saleh, M. and Tirole, J. (2021). Taxing identity: Theory and evidence from early islam. *Econometrica*, 89(4):1881–1919.

Sandhu, K. S. (1964). Emergency resettlement in malaya. *The Journal of Tropical Geography*.

Schindler, D. and Westcott, M. (2021). Shocking racial attitudes: black gis in europe. *The Review of Economic Studies*, 88(1):489–520.

Scott, J. C. (2020). *Seeing like a state: How certain schemes to improve the human condition have failed*. yale university Press.

Sequeira, S., Nunn, N., and Qian, N. (2020). Immigrants and the making of america. *The Review of Economic Studies*, 87(1):382–419.

Shah, N. (2019). *Striving for inclusive development: From Pangkor to a modern Malaysian State*. Oxford Fajar Sdn. Bhd. under licence from Oxford University Press.

Shayo, M. (2009). A model of social identity with an application to political economy: Nation, class, and redistribution. *American Political Science Review*, 103(2):147–174.

Siew, N. C. (1953). Labour and tin mining in malaya.

Steinmayr, A. (2021). Contact versus exposure: Refugee presence and voting for the far right. *Review of Economics and Statistics*, 103(2):310–327.

Stevenson, B. (2026). When democracy falters: A multidisciplinary, multi-book review essay on polarization, populism, and authoritarianism. *Journal of Economic Literature*, 64(2):679–697.

Strauch, J. (1981). Chinese new villages of the malayan emergency, a generation later: A case study. *Contemporary Southeast Asia*, 3(2):126–139.

Strauch, J. (2013). Chinese village politics in the malaysian state. In *Chinese Village Politics in the Malaysian State*. Harvard University Press.

Tabellini, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, 123(3):905–950.

Tabellini, M. (2020). Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration. *The Review of Economic Studies*, 87(1):454–486.

Wang, G. (1959). *A short history of the Nanyang Chinese*. Number 13. D. Moore.

Wong, M. (2013). Estimating ethnic preferences using ethnic housing quotas in singapore. *Review of Economic Studies*, 80(3):1178–1214.

ONLINE APPENDIX

Chun Chee Kok, Gedeon Lim, Danial Shariat, Abu Siddique, and Shunsuke Tsuda

A Site Selection Criteria

A.1 Actual Military site selection criteria

Based on the Report on Squatter Resettlement in Various States, File No: B.A. Selangor 119/50, the British produced a set of plans and procedures for site selection. These criteria are:

1. Resettlement villages were to be located on a main road or other major transportation artery.

2. Villages were to be relocated, wherever possible, on rolling terrain to promote drainage.

3. Squatters were to be concentrated into compact villages that were fenced in and protected by a police post capable of commanding the entire village, most importantly the village gate.

4. Villages were to be sited in such a manner as to minimize squatter dislocation.

5. Sufficient water was to be supplied, either from adjacent towns or from wells within the village. Health and fire regulations were expected to be observed.

6. Amenities such as schools, dispensaries, and community centers had to be provided as quickly as possible.

7. Sufficient agricultural land of good quality was to be provided for all agriculturalists forced to abandon their previous holdings.

The site selection criteria were driven primarily by military expediency rather than the economic and social well-being of the resettled population. Hence, some criteria were followed through completely but not all. The criteria that were followed through are the following. Chinese New Villages (CNVs) were located close to a main road or transportation artery to increase accessibility by the British military to these villages in case of communist attacks.[67] Second, CNVs were on high ground to improve defensibility from the communists. The resettled often lived under the surveillance regime during the Malayan Emergency.[68] In theory, according to the full site selection criteria, a New Village was to possess basic amenities and sufficient agricultural land. In practice, however, the rapid strategic demands of resettlement resulted in the non-compliance of many criteria that were more tangential to military objectives Phee (2012).

[67] In Figure 1, we plot the location of CNVs and historical roads. We observe that there is a high correlation between the location of 452 CNVs and historical roads in our sample.

[68] This entailed curfews, body searches at checkpoints, communal kitchen arrangements, food restrictions, and identity certificate registration. There were fortified sentry boxes and watch towers with floodlights to "guard" the CNVs. A police station was located either near the main gate of the village or placed at a high point for surveillance.

A.2 Reconstructing feasible counterfactual CNV locations

Our counterfactual analyses would be valid only if eventual, real CNV sites were not necessarily located at the most optimal locations—especially with regards to pre-existing interethnic proximity. I.e. it would have to be the case that the relocation process involved a certain degree of quasi-randomness and equally suitable candidate sites existed. Below, we summarize key insights from rich archival documents suggesting that, absent intervening factors, the British might indeed have placed CNVs at alternative locations. Furthermore, proximity to pre-existing Malay villages was unlikely to have been a priority nor possibility.[69]

While it is broadly agreed that the British chose sites to adhere to a set of over-arching military criteria, it is also clear, however, that optimal site selection decisions were often hindered by on-the-ground realities that the British were unable to foresee. Specifically, for each CNV, the British would choose 3-4 candidate sites but the eventual, real CNV site was often sub-optimal due to (i) private land acquisition costs (due to the need to minimize dislocation of Chinese from work sites); and (ii) minimal, prior local knowledge (the lack of granular data on the location and number of pre-resettlement Chinese squatter settlements and rural Malay villages).[70]

First, the British possessed extremely limited knowledge of Chinese squatter locations (and rural Malay villagers), forcing them to conduct fresh "squatter" censuses under severe time pressure and at substantial time and manpower cost. Officials frequently "discovered" previously unknown settlements during the resettlement process.[71] Furthermore, the 1947 Census excluded the enumeration of many small Malay villages in rural environs. Given military expediencies and cost constraints, this made it unlikely that the British could have strategically chosen sites based on

[69] Given the Malayan Census of 1947 Del Tufo (1949), one might be worried that the British must have had rich, granular data on the population and location of ethnic Chinese (and Malay) settlements that would have allowed them to fine-tune CNV site locations. Three historical facts suggest that this was unlikely: (i) the 1947 Census was conducted immediately after the British had regained control of Malaya after the end of World War II and 3 years of Japanese Occupation in 1945. Anthropological and historical accounts suggest that, both before and after World War II, large-scale in/out-migration flows of the ethnic Chinese, which the British were entirely unable to keep track of, were exceedingly common given Japanese discrimination and (mass-)killings of ethnic Chinese as retaliation for their (perceived) support of China (Gungwu, 2021; Sandhu, 1964); (ii) 1947 Census documents explicitly note that, given limited post-war state capacity, the census only enumerated (semi-)urban towns and large villages. Hence, most rural villages and squatter settlements were left out due to their dispersed character and small population size; (iii) This is underscored by various federal and state "Squatter Committee" reports that repeatedly emphasized the need to conduct squatter censuses to more clearly ascertain the exact number of individuals to be resettled (See e.g. "The Squatter Problem in the Federation in 1950", National Archives of Malaysia.)

[70] "Federation of Malaya: Report on the emergency in Malaya from April 1950 to November 1951 by Lieut-General Sir Harold Briggs, Director of Operations. Kuala Lumpur Government Press 1951 (National Archives of Singapore, pp3)" writes: "These Chinese squatters were (a relic of the Japanese occupation) *numbering anything* up to 500,000 spread over the country and *beyond effective administration by the Government*" (Appendix A)". These reports, and various archival documents documenting a myriad of pre-resettlement squatter censuses, strongly suggest that the British often did not know the location nor population sizes of squatter settlements.

[71] For instance, Markandan (1954) (pp7) writes that a resettlement officer was told to expect 8,000 squatters. Subsequent investigations, however, revealed nearly 29,000 squatters in an area of nine square miles.

pre-existing locations and sentiments of local Malay communities.[72]

Second, sub-optimal CNV sites were often chosen due to unexpectedly high costs of acquisitioning private land.[73] This was a consequence of the need to minimize the distance of CNVs from rubber estates and tin mines in which Chinese squatters worked.[74] Hence, the only suitable candidate CNV sites, were almost always under private ownership/lease and site choice was often determined by (i) which private entity was willing to sell their land at short notice and (ii) the land-sale price. For instance, in the resettlement of Sungei Way New Village (Selangor), the eventual CNV site was chosen because it was found to contain mostly old rubber trees (which authorities could reacquisition at a low enough cost). In comparison, two other candidate sites that were strictly preferable in terms of minimizing labor dislocation, were under mining titles and had "just been diverted by the (mining) Company at a cost of over half a million dollars".[75]

Together, the historical record suggests that eventual CNV locations were largely determined by the extent of low-cost, private land that happened to be available at short notice, and the little local knowledge that authorities could muster (within months) in a hostile, foreign environment (in which they did not speak the local language and had to rely largely on local intermediaries), rather than by targeted, strategic concerns.

Hence, using declassified British military planning documents (see Appendix A.1), we reconstruct the set of feasible counterfactual CNV locations. Specifically, we generate 1,000 sets of counterfactual CNVs, each satisfying the following plausibly exogenous site selection rules:[76]

1. *Road accessibility and livelihood continuity:* Counterfactual resettlement sites must be located along a main road and within 2.5–10 km of the original CNV site via the pre-resettlement road network.

2. *Topographical suitability:* Sites must have elevation and slope below the 90^{th} percentile of actual CNVs and contain, at least as much topographically-suitable land within a 2.5km radius (the average size of a CNV Nyce (1973)) as the 10^{th} percentile of actual CNVs.

In addition, we impose two additional constraints that closely adhere to local conditions that the British faced:

[72]This lack of prior knowledge is detailed in various state-level "Squatter Committee" reports as well as "The Squatter Problem in the Federation in 1950" (Johor State Archives and National Archives of Malaysia.) In fact, the British had to rapidly train hundreds of local Chinese to serve as District Liaison officers at extremely short notice and there was an extreme shortage of such officers in the early years (Source: District Liaison Office Accounts, National War Museum, London, United Kingdom).

[73]Revenue from rubber sales owing to the Korean War enabled the British to finance much of the resettlement process, but estimated projections for the resettlement of each New Village eventually amounted to hundreds of thousands of British pounds each, far exceeding what the British had planned for.

[74]It is important to note that not all villagers in CNVs worked in rubber estates or tin mines, we leverage this variation to study interethnic complementarities in Section 7.

[75]Similarly, in Seremban district (Negeri Sembilan) a mining company outright refused to sell their land despite their mining dredge having been closed down and the pending expiry of their mining lease.

[76]See Appendix A: Site Selection Criteria for more details on the military rationale for the first two criteria.

3. *Exclusion of Malay reservations:* Malay Reservations are large tracts of federal and state lands created by the British to "safeguard" the heritage and land of the native, ethnic Malays. Accordingly, the British could not resettle Chinese communities on these lands, and we constrain all counterfactual sites to lie outside designated Malay reservations (Kratoska, 1983, 1982).

4. *Spatial balancing:* To ensure geographic comparability, each set of counterfactual CNVs is constructed by randomly shifting actual CNV locations while maintaining balance in the north–south and east–west directions. We implement a simulated annealing procedure to minimize discrepancies in the distribution of distances between actual CNVs and each set of counterfactual CNVs (following Dell and Olken (2020)). This procedure ensures that the relative distances between counterfactual CNV sites mirror those between actual CNVs. This is important for two reasons. First, the British military likely maintained minimum distances between CNVs to prevent coordination among resettled Chinese villagers that might support communist insurgency. Second, although granular data on pre-resettlement Chinese settlements is largely unavailable, spatial balancing helps replicate the distribution of actual CNVs, which were likely sited to minimize disruption to villagers' economic activities dhu Renick (1965).

B Data Appendix

B.1 Chinese New Villages

The main source of information on the location of Chinese New Villages is *A Survey of the New Villages in Malaya* published by the Malayan Christian Union (1958). The census contains information on the names of the New Villages, their prevailing Chinese dialect spoken, their estimated population, whether there was evangelistic work performed in the village, whether medical facilities and amenities were available. To identify the exact location of these New Villages, we manually matched the village names listed by the Malayan Christian Union (1958) with the maps from the Ministry of Housing and Local Government, Malaysia in 2012 Lee (2012). We successfully identified and geolocated a total of 452 CNVs. We further imposed four sample criteria that gave us 208 CNVs, which we use in our analysis:

1. We kept all CNVs whose primary medium of language was recorded as Mandarin or a Mandarin dialect in 1958.[77] In cases where language information was missing, we supplemented the Malayan Christian Union (1958) with data on the name and location of Chinese-medium schools Lim and Song (2002), the presence of which is highly correlated with the presence of a CNV.

[77]Note that a small number of ethnic Malays were resettled into Malay New Villages. This took place largely in areas where the Communists were deemed to be extremely active ("Black" areas) and was meant to protect the local Malay population from Communist attacks dhu Renick (1965). We use exposure to Malay New Villages as a test for alternative mechanisms in Section 6.

2. We focused on seven West Coast states and excluded CNVs in two East Coast states, where interethnic contact patterns continue to be markedly different—chiefly due to a much less-developed rubber and non-existent tin industry which led to a much smaller wave of early 20th century in-migration from China (Sidhu, 1976).[78] In particular, these two East Coast states were the last to come under British control and had a much smaller rubber industry (nearly non-existent in the case of Trengganu). Accordingly, the share of ethnic Chinese both in urban and rural areas was much lower in both states. This was due to a number of factors: (i) their traditional economies being dominated by rice; (ii) the absence of any known tin reserves; and (iii) extensive land reservations for Malays that made it difficult for large-scale land purchases and conversion of land (by non-Malays) to take place (Sidhu, 1976).

3. We excluded all CNVs located in 1947 urban census districts.[79] We did so given (i) the patterns of resettlement between CNVs in historically rural and urban areas drastically differed. In particular, our fieldwork suggested that many "urban" CNV villagers had never experienced resettlement but simply had their existing house lots fenced up; (ii) the possibility that patterns of economic and political development might have been very different from rural districts, especially due to earlier waves of selective in-migration into towns pre-CNV resettlement. Furthermore, the small number of urban districts makes it difficult for our research design to adequately account for differential levels of selective in/out-migration given the lack of individual-level data that is geolocated at a level that is lower than census districts.

4. We excluded all CNVs that have fewer than 200 Chinese voters in the 2013 electoral rolls. Results are qualitatively similar if we use lower thresholds. The most likely reason for the depopulation of these CNVs is a lack of economic viability and/or (natural) disasters that rendered them inhospitable. For example, one such CNV in the state of Johor that we visited no longer had any viable Chinese presence in the old village quarters due to a fire that broke out in the 1970s and burnt down nearly the entire village.

B.2 Voting Variables: Malaysian General Elections in 2013 and 2018

Polling districts. Our sample includes 939 polling districts (located in 81 parliamentary constituencies) in 2013 and 1,004 polling districts in 2018. To obtain this sample, we impose four sampling restrictions.[80] First, to ensure that we are comparing outcomes only across polling dis-

[78] We focus on Johor, Perak, Negeri Sembilan, Selangor, Kedah, Malacca, Pahang, and exclude Kelantan and Trengganu.

[79] Specifically, we digitize census district polygons from the 1947 Malayan Population Census and classify a census district as urban if it contains a major town.

[80] This compares to 2,371 excluded polling districts in 2013. In line with our focus on historically rural polling districts, excluded polling districts have lower average BN vote shares; lower 2013 ethnic Chinese shares; marginally lower turnout rates; lower 1947 ethnic Chinese shares and population density; lower elevation; and largely more

tricts that could have possibly been candidates for siting a CNV, we restrict our sample to all polling districts located 10 kilometers from a CNV. This is informed by military site selection criteria (see Section 4.2) where Chinese were largely resettled between 3.2 and 9.6 kilometers from their original locations dhu Renick (1965).[81] Second, to better isolate the effects of exposure to CNVs, we conduct a "doughnut hole" analysis by excluding all polling districts that contain a CNV. Third, we restrict our sample to historically rural polling districts. In this way, we can interpret our results as the effects on areas that started from a similar level of development.[82] Last, given our focus on studying the effects of interethnic proximity between ethnic Chinese and Malays, which are the two largest ethnic groups in Malaysia, we exclude all polling districts where ethnic Indians are the dominant group.[83]

Voting outcome. We use federal and state election vote shares for the National Front (BN) in both the 2013 and 2018 General Elections as our main measure of political preferences for ethnonationalist policies, and as a behavioral proxy for Malay political identity. BN ran on a Malay-first platform, in contrast to the more inclusive, multi-ethnic coalitions of the opposition, the People's Alliance (rechristened the Alliance of Hope in 2018).

We do not separately analyze contests involving the Islamic party (PAS) for three reasons. First, ethnicity and religion are tightly linked in Malaysia: nearly all Malays are Muslims by constitutional definition. Second, BN's main Malay party (UMNO) has historically promoted a conservative Malay-Muslim identity, making its platform broadly similar to PAS; in 2013, PAS ran within a multi-ethnic coalition and our results are robust to excluding those contests. Third, in 2018 PAS moved further right within a separate conservative coalition (GS) but won only 16.89% of the popular vote, limiting its aggregate influence.[84]

We use polling-district data for three reasons: (i) polling districts are smaller and more demographically homogeneous (an average federal constituency contains 18 polling districts); (ii) unlike

favorable agro-climatic conditions. Results available upon request.

[81] Ideally, we would measure the distance between each polling district and the centroid of the closest original ethnic Chinese squatter camp location, but to the best of our knowledge, such granular data does not exist at a systematic level due to the hastiness of the resettlement process.

[82] Specifically, we digitize census district polygons from the 1947 Malayan Population Census and classify a census district as urban if it contains a major town. A major town is defined by the Census as any town containing a population of more than 10,000 inhabitants. We then overlay contemporary polling district polygons over 1947 census district polygons and define a contemporary polling district as historically rural if more than 50% of its geographical area lies outside a 1947 urban census district polygon.

[83] Specifically, given the three main ethnic groups, we exclude all polling districts where ethnic Indian shares in 2013 were greater than 66.6%. As mentioned in Section 2, Indians comprise 8.9% of Malaysia's population. Furthermore, the political economy of Indian-dominant polling districts is markedly different—many were historical plantation lines, where plantation owners constructed purpose-built dormitories and amenities, and where (many descendants of) marginalized Indian coolies and rubber tappers who were brought into colonial Malaysia as indentured labor continue to reside (Kratoska, 1982).

[84] We can further test for differences in religious identity and/or religiosity using primary survey data.

federal and state constituencies—where malapportionment favoring BN is well documented Ostwald (2017)—polling districts contain roughly equal numbers of voters and are less subject to boundary manipulation; and (iii) the fine geography lets us measure distance to CNVs precisely and include parliamentary seat fixed effects.[85] We therefore use federal-election vote shares as the primary outcome unless otherwise noted.

Finally, we interpret vote-share differences as reflecting both formative and later-life exposure to CNVs. Because Malaysians rarely change voting addresses, vote shares mainly reflect preferences shaped by growing up near CNVs; in areas with out-migrants who return to vote, they may also reflect exposure at migration destinations. All specifications, therefore, control for proximity to the nearest 1947 urban center.

Our data include polling district identifiers, vote counts of each coalition in 2013 and 2018, and the number of registered voters by ethnic groups in 2013. We construct measures of the vote share of *Barisan Nasional*, which is the total votes received by Barisan Nasional over the total number of votes cast in each polling district in both 2013 and 2018. We construct the ethnic share by dividing the number of registered Chinese or Malay voters by the total number of registered voters in each polling district in 2013. We construct voter turnout, which is the total number of votes cast over the total number of registered voters in each polling district in 2013.

B.3 Spatial Variables

Nighttime luminosity. We use nighttime luminosity data from satellite images in 2014 as a proxy for local economic activity and development (Hodler and Raschky, 2014; Michalopoulos and Papaioannou, 2013). We use remotely sensed nighttime lights data from NASA's Black Marble product (VNP46A3), which provides monthly cloud-free radiance composites. The data are downloaded using the `blackmarbler` R package, which accesses NASA's Black Marble archive through authenticated API queries. We specify Malaysia as the region of interest using Level-2 administrative boundaries from the GADM database and extract monthly raster layers for the year 2014. The downloaded rasters represent All-Angle Snow-Free Composites, filtered to remove low-quality observations based on quality flags. To compute average luminosity for each polling district in both 2013 and 2018 boundaries, as well as grid cells, for each month, we calculate the mean radiance within the spatial boundary of each polling district, then average across all twelve months of 2014 to obtain a single measure of mean annual luminosity in 2014.

Population density. We measure local population density using gridded population estimates from the 2010 Global Human Settlement Layer (GHSL), which provides population counts at approximately 1 km resolution. For each polling district, we calculate the total population by

[85] Pre-2013 electoral results are only available at the parliamentary or state-seat level, which is too coarse to exploit within-seat distance variation to CNVs.

summing the raster values that fall within its boundaries using the `exactextractr` package. To compute density, we divide this population sum by the land area of the polling district, calculated in square kilometres. The resulting measure represents the number of persons per square kilometre in each polling district in 2010. This variable serves as a proxy for local settlement intensity.

Public goods. We measure educational public goods by geo-referencing a complete list of all primary and secondary schools in Malaysia in 2010 from the Ministry of Education. These records contain the point coordinates and the number of teachers and students for each school. We measure health public goods by geo-referencing a complete list of all clinics and hospitals in Malaysia in 2022 from the Ministry of Health. Due to the lack of geocoded administrative records prior to 2013, we use the 2022 distribution as a proxy, noting that major public infrastructure investment typically varies little over a single decade. We aggregate both data up to the polling district level. Lastly, we measure contemporary road density at the polling-district level using data from Open Street Maps (OSM). We do so by overlaying the contemporary OSM road network over polling district boundaries and computing the total length of roads within each polling district. We then divide the total road length by the area (in square kilometers) of each polling district.

Other controls. We include a wide range of geographical variables to construct the controls in the regressions. These include measures of: (i) topography (elevation, slope, aspect, coastlines, (ii) soil quality (% of topsoil carbon, % of topsoil sodicity, type of soil, and class of drainage), and (iii) pre-resettlement variables (nearest urban centre in 1947, population density in 1947, Chinese share in 1947). Below, we briefly discuss the construction and sources of these variables.

Topography. We create topographical variables using the *Harmonized World Soil Database* (HWSD). The raster files are compiled from high-resolution source data and aggregated to 30-arc-second grids. We compute elevation for each polling district as the average elevation over the entire polling district polygon, using raster data from HWSD. Slope and aspect data were also computed for each polling district similarly. For aspect data, the variables equal to the average share of 30-arc-second grids that are north-, south-, east-, and west-facing grids of each polling district. We compute fly-by-crow distances from the polling district centroids to the nearest coastlines.

Soil quality measures. We make use of the FAO GAEZ V4 data for soil quality measures. HWSD provides detailed information on different soil types across the world. We compute the average raster values within each polling district polygon for continuous variables, including the % of topsoil organic carbon and % of topsoil sodicity. These variables proxy soil fertility and salinity. For categorical soil attributes including drainage class and soil texture, we extract the majority pixel value within each polygon using nearest-neighbor resampling to preserve class integrity.

Pre-resettlement variables. We use the population census in 1947 to construct pre-treatment demographic variables Del Tufo (1949). We digitized the list of urban centers with at least 10,000 inhabitants and geolocated each of them. We compute fly-by-crow distances from the polling

district centroids to the nearest urban centers. Moreover, we digitized the count of population by ethnic groups at the Census District level (*mukim*). We then assign population statistics of subdistricts to the polling districts (which are more disaggregated than subdistricts) based on the share of intersected areas between a subdistrict and a polling district.

B.4 Pre-Resettlement Chinese Squatter Locations (Original Survey Data)

Our primary survey data includes a retrospective survey administered to a total of approximately ninety ethnic Chinese community leaders from about 30 CNVs (we targeted three per village). Through semi-structured interviews, we identify (i) the geographic location of each Chinese pre-resettlement squatter location and (ii) distance from the CNV.[86]

Figure B.1 displays the spatial distribution of pre-resettlement locations of Chinese squatters relative to the locations of CNVs. Panels A and B show maps for two districts of Batu Pahat and Kluang (Johor state) and Kinta (Perak state). These maps plot squatter locations (red dots) alongside CNVs (crosses) and illustrate that the pre-resettlement squatter settlements were largely (i) geographically dispersed and (ii) lacked systematic patterns of concentration.

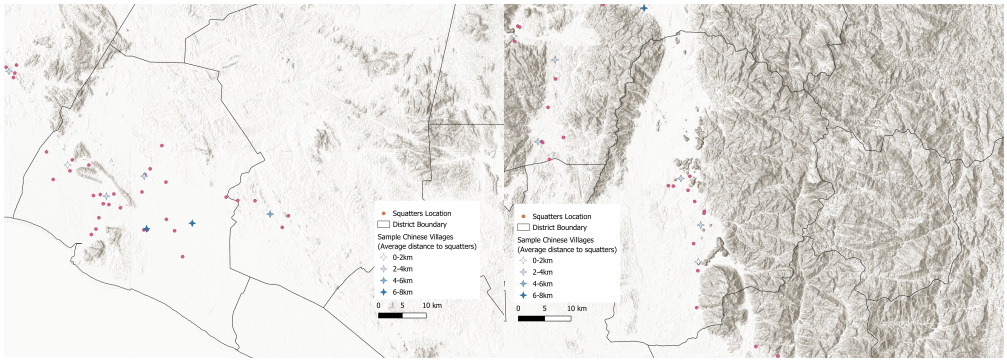
The spatial dispersion of pre-resettlement squatter locations corroborates historical accounts (dhu Renick, 1965), and suggests that any potential effects from decreased exposure or depopulation in the origin areas on existing Malay settlements were likely to have been minimal. Specifically, Malay communities residing across different distance bins would have had some degree of pre-resettlement exposure to Chinese squatters, but the remote, dispersed nature of these squatter locations is likely to have limited inter-settlement connectivity and resulted in a much lower intensity of pre-resettlement exposure to ethnic Chinese, vis-à-vis sustained differences in post-resettlement contact.

Panel C plots the distance between each real CNV and initial Chinese squatter locations from which CNV villagers were relocated from.[87] The highest pre-resettlement exposure appears to be for Malay communities located within the 4-6km distance bin. This will be important for interpreting potential differences in treatment estimates across distance bins in Section 5.

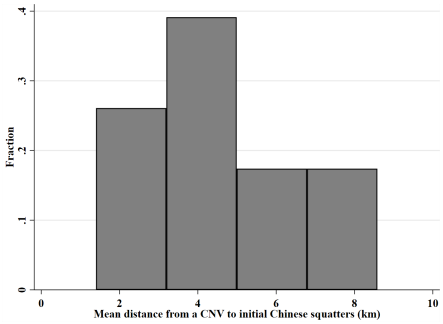
[86] To minimize recall bias, trained enumerators from the local community engaged leaders using high-resolution satellite imagery (Google Maps). Respondents identified landmarks associated with former settlements, which were then geolocated to the highest degree of precision possible.

[87] For more than 80% of the observed CNVs, ethnic Chinese squatters were relocated from distances within 10km—we omit the remaining 20% for presentation purposes.

Figure B.1: The Spatial Distribution of Pre-Resettlement Chinese Squatters



Panel A: Batu Pahat and Kluang **Panel B: Kinta**

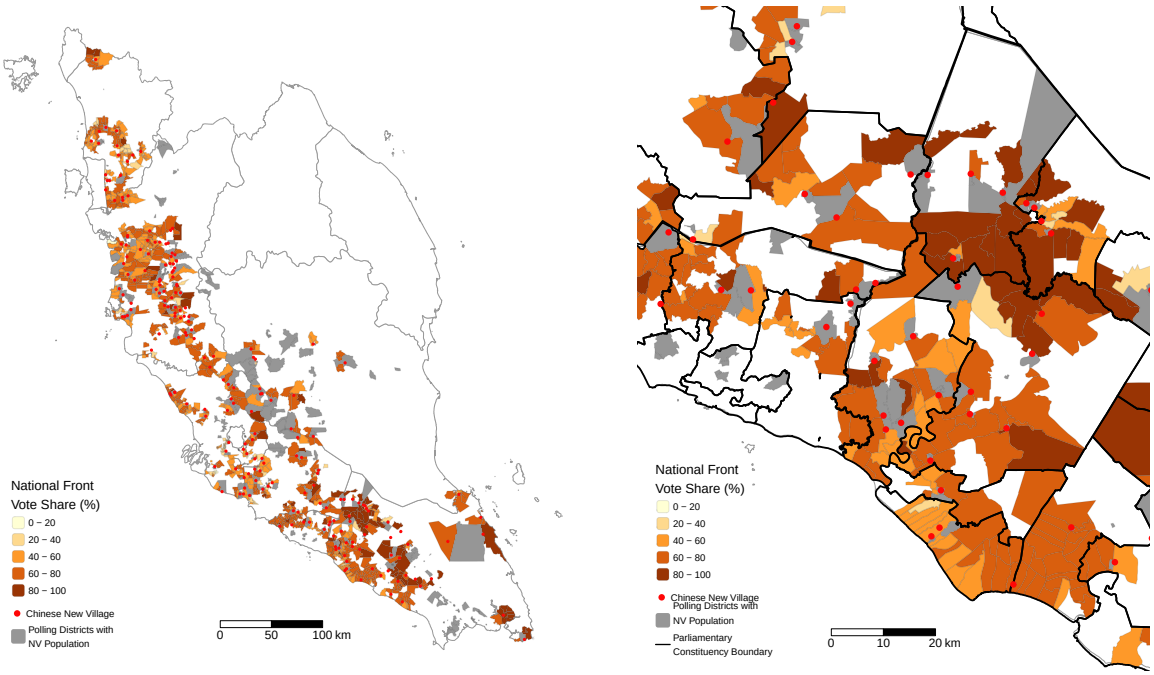


Panel C: CNV-level distance

Notes: Panels A and B plot the pre-resettlement distribution of Chinese squatter locations (red dots) in relation to the actual Chinese New Village (CNV) locations to which ethnic Chinese minorities were resettled (blue crosses). The shade of the blue crosses correspond to the average distance between each CNV location and all-known squatter locations from which the CNV population was resettled from. Darker shades correspond to greater average distances. District boundaries (black lines) and topographical features are mapped in the background. Panel C plots, at the CNV level, the average distance from each CNV to initial (pre-resettlement) Chinese squatter locations. Source: Authors' primary survey data from retrospective surveys with Chinese New Village leaders.

C Additional Figures and Tables

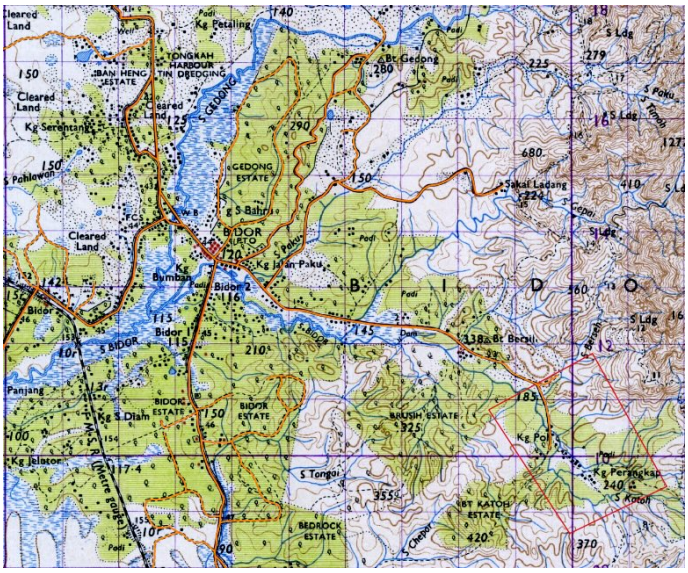
Figure C.1: Vote Shares for the National Front in 2013, Polling District-Level



Panel A **Panel B**

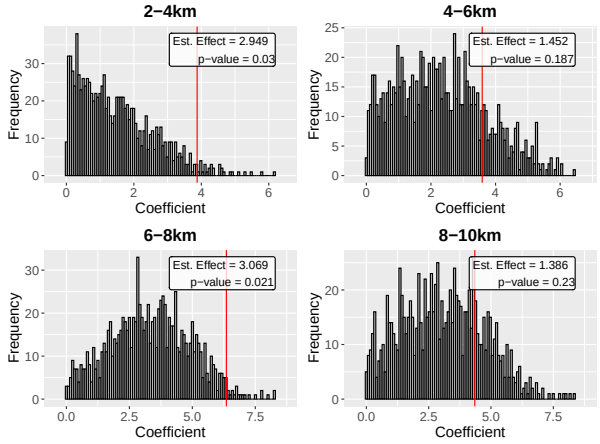
Notes: This figure displays vote shares for the National Front at the polling district level in 2013. Polygons shaded in darker colors indicate greater vote shares for the National Front in 2013. CNVs are represented as red dots. Each parliamentary constituency contains an average of 18 polling districts. Black borders represent parliamentary constituency boundaries. In grey, are polling districts that contain a CNV. In white, are polling-districts (i) further than 10km from a CNV; (ii) are historically urban; or (iii) have an ethnic-Indian majority and are hence excluded from our regression sample. We discuss these restrictions in Section 3.

Figure C.2: Digitizing Historical Roads and Malay Reservation Polygons: Extract from HIND1035 1947 Map (State of Perak)

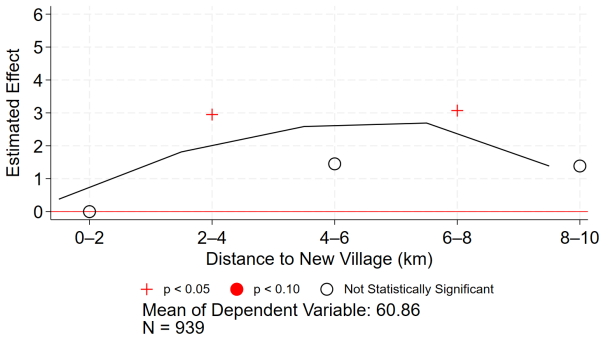


Notes: Scale: 1:63,360 or 1 inch to a mile. The orange lines indicate main roads and the red box indicates an example of a Malay Reservation Area.
Source: HIND 1035, Sheet 2N/14. National Library of Australia.

Figure C.3: Illustration of Methodology: Effects of Proximity to CNVs on Ethno-nationalist Vote Share at the Federal Election (2013)



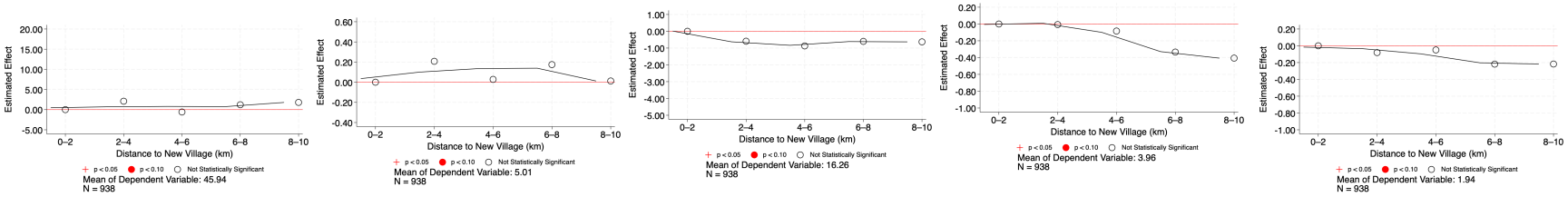
Panel A: Independent Shifts: Counterfactuals



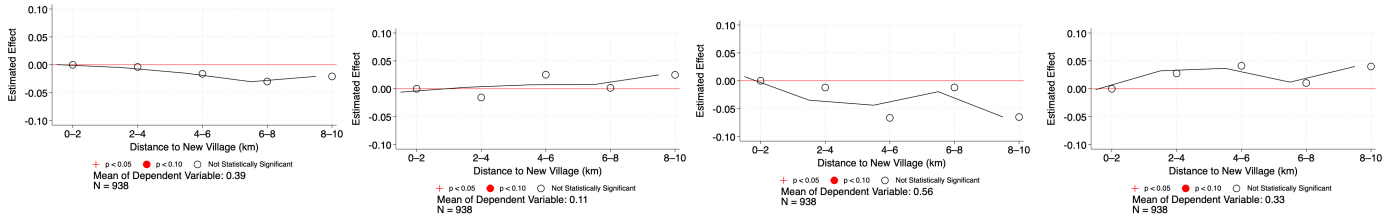
Panel B: Independent Shifts: Plotted Coefficients

Notes: Panel B plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects, geographical and pre-treatment controls. Each coefficient is computed by subtracting the mean of the analogous counterfactual estimates (computed from 1,000 counterfactual New Village configurations, the distribution of which is displayed in each subpanel of Panel A) from each actual coefficient (the value of which is displayed as a red vertical line in each subpanel of Panel A). The points in Panel B are fitted with a linear spline. p-values compare the effect of distance to the nearest actual New Village to the effects of distance to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

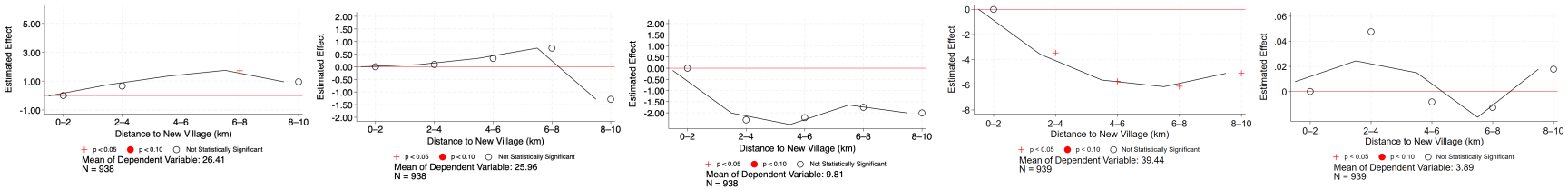
Figure C.4: Geographic, Soil, and Pre-Resettlement Balance



(A) Elevation | **(B) Slope** | **(C) % of East-facing Grids** | **(D) % of Topsoil Organic Carbon** | **(E) % of Topsoil Sodicity**



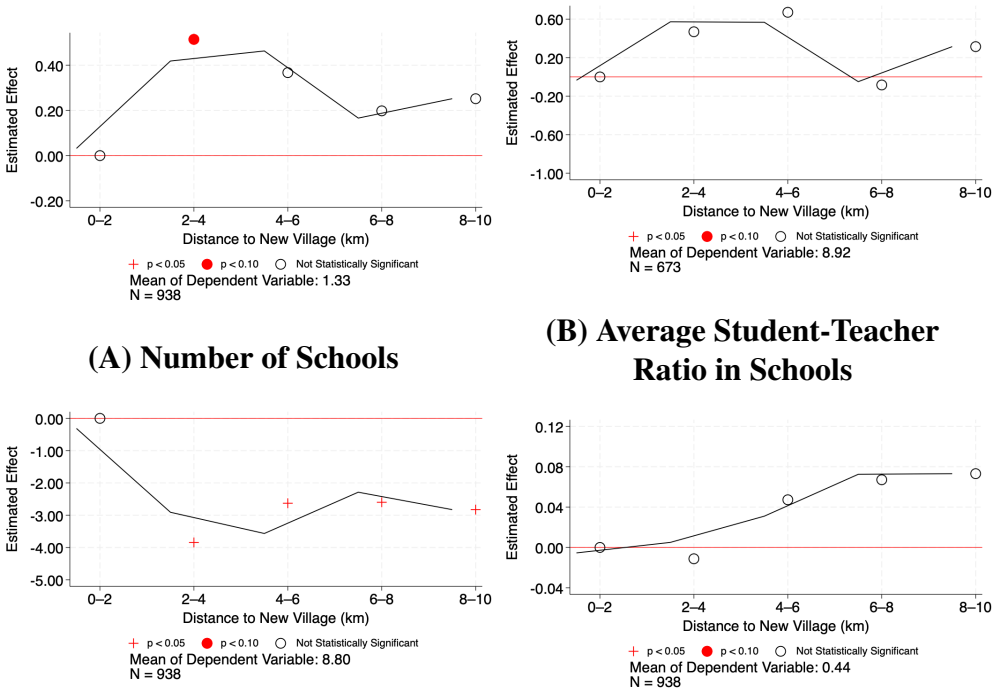
(F) Drainage is very poor (=1) | **(G) Soil is Coarse (=1)** | **(H) Soil is Medium (=1)** | **(I) Soil is Fine (=1)**



(J) Distance to urban centre in 1947 (km) | **(K) Distance to coast (km)** | **(L) Road Density in 1947** | **(M) % Share of Chinese in 1947** | **(N) Log Population Density in 1947**

Notes: We normalized each dependent variable so that it always ranges from -1 to 1. Points plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest New Village, controlling for federal parliamentary constituency fixed effects. The means of analogous estimates computed from 1,000 counterfactual New Village configurations are subtracted from each coefficient. The points are fit with a linear spline. P-values compare the effect of proximity to the nearest actual New Village to the effects of proximity to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

Figure C.5: Effects of Chinese New Villages on Public Goods Provision



(A) Number of Schools

(B) Average Student-Teacher Ratio in Schools

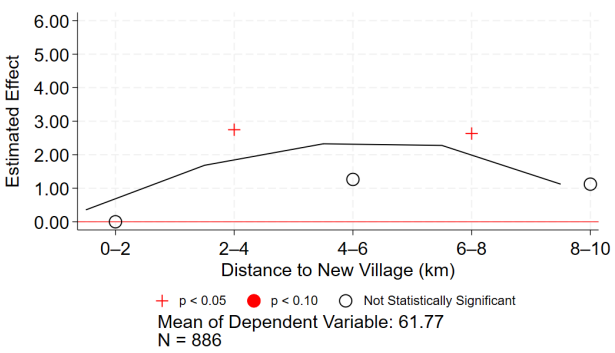
(C) Road Density

(D) Clinics (=1)

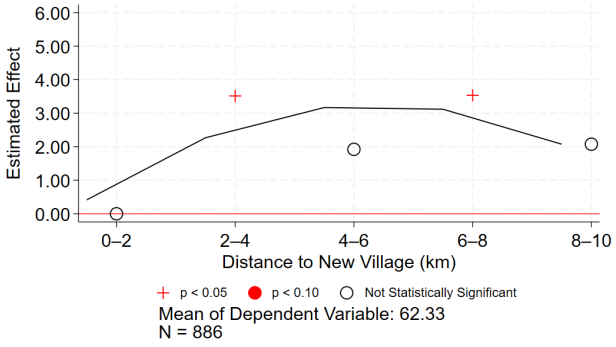
Notes: Points plot coefficients estimated from regression the outcome variable on 2-km bins of distance to the nearest New Village, controlling for federal parliamentary constituency fixed effects and geographical and pre-treatment controls. The means of analogous estimates computed from 1,000 counterfactual New Village configurations are subtracted from each coefficient. The points are fit with a linear spline. p-values compare the effect of proximity to the nearest actual New Village to the effects of proximity to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

Figure C.6: Effects of Proximity to CNVs on Ethno-nationalist Electoral Support (Excluding Polling Districts where 1947 Chinese Share >=80%)

Ethno-nationalist Vote Share in 2013

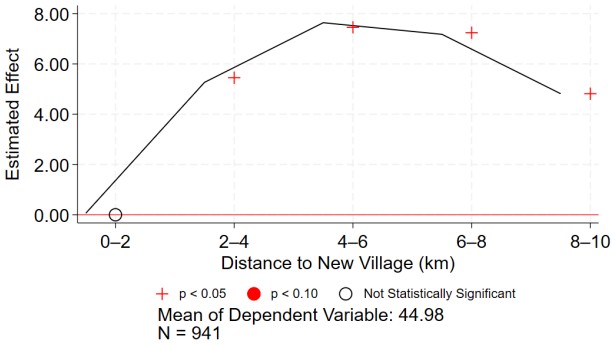


(A) State election

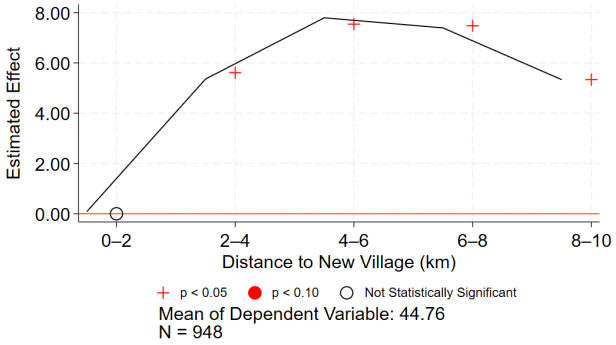


(B) Federal election

Ethno-nationalist Vote Share in 2018



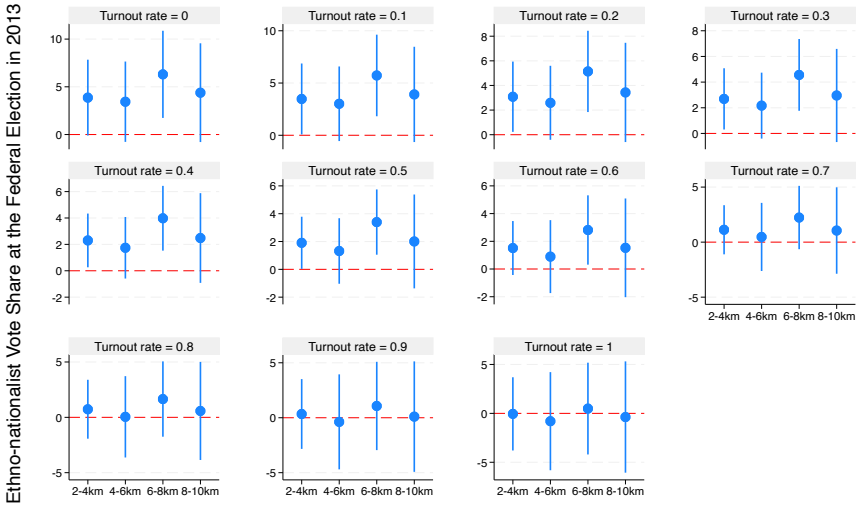
(C) State election



(D) Federal election

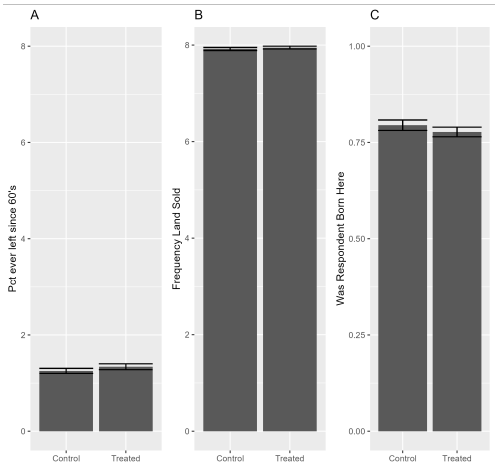
Notes: Figures plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects, geographical and pre-treatment controls. In addition, regressions of state election results include indicators for all possible combinations of party match-ups at the state constituency level. The means of analogous estimates computed from 1,000 counterfactual New Village configurations are subtracted from each actual coefficient. The points are fit with a linear spline. P-values compare the effect of distance to the nearest actual New Village to the effects of distance to the nearest counterfactual New Village, computed from 1,000 counterfactual New Village configurations.

Figure C.7: The Effects of Varying Ethnic Chinese Voter Turnout Rates



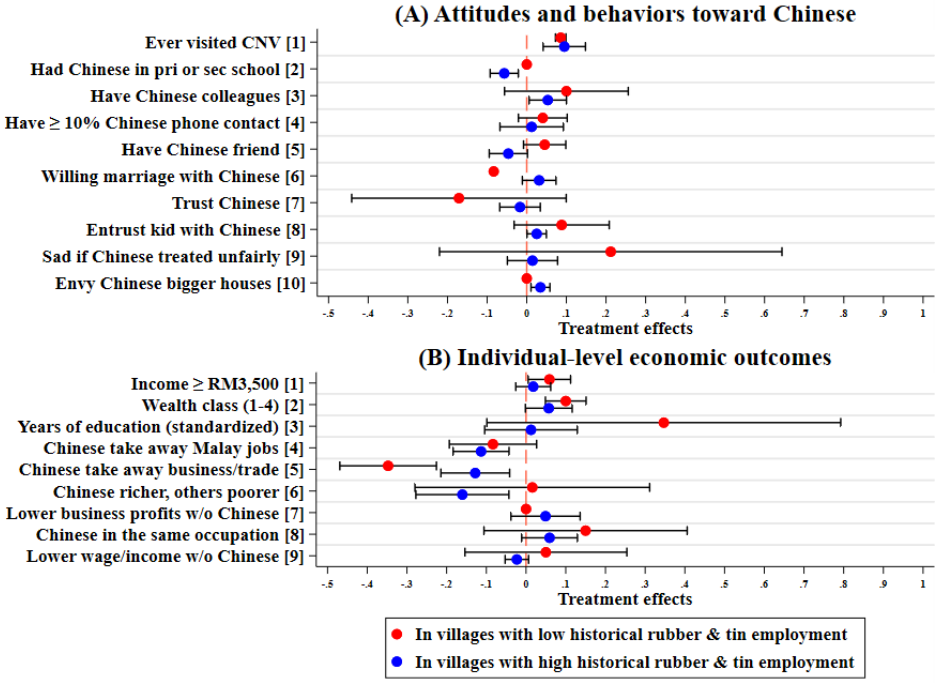
Notes: This figure plots OLS estimates of Equation 5. The outcome variable is vote share for ethno-nationalistic coalition at the polling district level in federal parliament constituency contests. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects and geographical and pre-treatment controls. Each subfigure represents different turnout rates applied to estimate the number of ethnic Chinese who cast their votes, ranging from 0 to 1, under the (implausible) assumption that all ethnic Chinese voters voted against the ethno-nationalist coalition. The sample comprises polling districts in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding polling districts that contain New Village population, polling districts in historically urban areas and polling districts beyond 10km of a New Village.

Figure C.8: Selective In- and Out-Migration Near CNVs



Notes: *Treated (Control)* Malay villages are located 0-2km from a real (fake) CNV. See Section 6 for details on our empirical strategy and matching procedure. Sub-figure A plots categorical responses to the question: "What percentage of residents since the 1960s have sold their house and never returned?" 1 = $<$ 1%; 2 = 1-5%; 3 = 5-10%; 4 = 11-20%; 5 = 21-30%; 6 = 31-40%; 7 = 41-50%; 8 = 51-100%. Sub-figure B plots the frequency with which villagers have ever sold their land in the village, to someone not from the village. 1 = everyday; 2 = weekly; 3 = every two weeks; 4 = every month; 5 = every 3 months; 6 = every 6 months; 7 = every year; 8 = not even yearly. Sub-figure C plots individual-level responses to: "Were you born in the village". 1 = Yes, 0 = No. Responses in A and B are from village leader surveys. Responses in C are from non-leader surveys. P-values of difference-in-means are 0.272, 0.492, and 0.342 for A, B, and C, respectively. $N = 288$ (A and B) and $N = 1990$ (C).

Figure C.9: Historical Interethnic Complementarities: Heterogeneity of Survey Results using Malay Occupational Shares in Rubber/Tin (1960)



Notes: This figure reports OLS estimates using the primary survey data, where colored dots represent point estimates and black lines represent 90% confidence intervals. The sample consists of individuals who were born in the surveyed village. The sample split is based on whether rubber or tin mining had the highest employment share in a village in 1960 or not ($n = 945$ and 247, respectively). We classify a village as "rubber & tin" if more than half of the village leaders report it as having the largest share. The dependent variables in Panel A (B) are the same as those reported in Table 1 (Table 2) and their definitions are found in the table notes. Note that the 'Non-agricultural job' variable has been omitted from Panel B. The absence of confidence intervals for some coefficients in the low historical rubber & tin sample reflects extremely small standard errors or insufficient variation due to limited sample size (See *Micro-level results* in Section 6.2). All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level.

Table C.1: Summary Statistics of New Village Characteristics shortly after Resettlement (1956–1958)

Variable	(1) In-Sample	(2) Out-of-Sample	(3) Diff (1)-(2)	(4) P-value
Transportation Access				
Main highway (=1)	0.692	0.566	0.127	0.005
Railroad (=1)	0.005	0.012	-0.007	0.398
Distance to Urban Center				
Distance within 5 miles (=1)	0.125	0.115	0.010	0.739
Distance within 5-10 miles (=1)	0.144	0.148	-0.003	0.921
Distance within 10-15 miles (=1)	0.144	0.119	0.025	0.426
Type of Government				
District council (=1)	0.466	0.434	0.032	0.498
Municipal council (=1)	0.457	0.484	-0.027	0.569
Self governance (=1)	0.058	0.037	0.021	0.296
Medical Facilities				
None (=1)	0.072	0.090	-0.018	0.487
Settlement Type				
Assimilated to existing and attached settlement (=1)	0.433	0.398	0.035	0.451
New settlement (=1)	0.413	0.381	0.032	0.485
Dominant Language Spoken				
Chinese dialects (=1)	0.832	0.643	0.188	0.000
Observations	208	244	452	

Notes: Source: (Humphrey, 1971; Malayan Christian Union, 1958). Data compiled from historical surveys conducted by the Malayan Christian Union in 1956–1958. In-Sample villages refer to the 208 CNVs included in our analysis out of the 452 CNVs that we were able to geolocate and verify. Out-of-Sample villages are the remaining 244 CNVs that we exclude after applying our sample selection criteria (Section 3). 83/452 (18.4%) villages are missing settlement type. For transport, urban centre, medical facilities and linguistic variables, 85/452 (18.8%) observations have missing values. For type of government, 12/452 (2.65%) observations have missing values.

Table C.2: OLS: Chinese New Villages and Ethno-nationalist Support

Panel (A)	**Dep. Var.: Ethno-nationalist Coalition Vote Share 2013**					
	State Election			Federal Election		
	(1)	(2)	(3)	(4)	(5)	(6)
2-4km	4.611**	3.873**	3.451*	4.706**	4.379**	3.868*
	(1.880)	(1.914)	(1.935)	(2.098)	(2.066)	(1.996)
4-6km	6.866***	3.944**	3.273*	6.800***	4.234*	3.439
	(1.975)	(1.957)	(1.869)	(2.457)	(2.263)	(2.114)
6-8km	11.725***	7.135***	5.893***	12.118***	7.747***	6.305***
	(1.867)	(2.127)	(2.174)	(2.305)	(2.355)	(2.294)
8-10km	7.356***	4.637*	3.692	8.176***	5.467**	4.386*
	(2.175)	(2.502)	(2.551)	(2.556)	(2.660)	(2.599)
Observations	939	939	938	939	939	938
Adjusted R^2	0.088	0.462	0.474	0.046	0.461	0.474
Mean (Dep. Var.)	60.301	60.301	60.283	60.863	60.863	60.843
SD (Dep. Var.)	16.338	16.338	16.338	16.358	16.358	16.355
Parliamentary Constituency FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Panel (B)	**Dep. Var.: Ethno-nationalist Coalition Vote Share 2018**					
	State Election			Federal Election		
	(1)	(2)	(3)	(4)	(5)	(6)
2-4km	8.641***	7.281***	6.587***	8.080***	7.128***	6.520***
	(1.776)	(1.904)	(1.829)	(2.093)	(1.861)	(1.779)
4-6km	13.558***	10.746***	9.958***	13.380***	10.399***	9.678***
	(2.258)	(1.796)	(1.823)	(2.801)	(1.720)	(1.722)
6-8km	16.486***	12.559***	11.535***	18.254***	12.642***	11.639***
	(2.365)	(2.120)	(2.186)	(2.537)	(2.015)	(2.027)
8-10km	12.021***	10.119***	9.301***	15.711***	10.764***	9.965***
	(2.592)	(2.268)	(2.218)	(2.721)	(2.187)	(2.196)
Observations	1004	1004	1004	1011	1011	1011
Adjusted R^2	0.258	0.530	0.555	0.117	0.543	0.560
Mean (Dep. Var.)	43.457	43.457	43.457	43.175	43.175	43.175
SD (Dep. Var.)	16.498	16.498	16.498	16.658	16.658	16.658
Parliamentary Constituency FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: This table reports OLS estimates of Equation 5. In Panel (A), the sample comprises polling districts in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding polling districts that contain Chinese New Village population, polling districts in historically urban areas, and polling districts beyond 10km of a Chinese New Village. In Panel (B), the sample restriction is the same as Panel A. Regressions of state election results additionally include indicators for all possible combinations of party match-ups at the state constituency-level. Standard errors are clustered at the federal parliamentary constituency-level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Effects of Chinese New Villages on Ethno-nationalist Electoral Support (Detailed)

	2013 Election		2018 Election	
	State (DUN)	Federal (Parl)	State (DUN)	Federal (Parl)
	(1)	(2)	(3)	(4)
2–4 km (vs. 0–2 km)				
Real coefficient	3.451	3.868	6.587	6.520
Avg. placebo coefficient	1.177	0.928	0.986	0.734
Centered coefficient	2.274	2.941	5.601	5.786
Empirical p-value	[0.067]	[0.030]	[0.000]	[0.000]
No. of $\lvert\hat{\beta}^{plac}\rvert \geq \lvert\hat{\beta}^{real}\rvert$	67	30	0	0
4–6 km (vs. 0–2 km)				
Real coefficient	3.273	3.439	9.958	9.678
Avg. placebo coefficient	2.484	2.120	2.455	2.266
Centered coefficient	0.789	1.319	7.504	7.411
Empirical p-value	[0.315]	[0.210]	[0.000]	[0.000]
No. of $\lvert\hat{\beta}^{plac}\rvert \geq \lvert\hat{\beta}^{real}\rvert$	315	210	0	0
6–8 km (vs. 0–2 km)				
Real coefficient	5.893	6.305	11.535	11.639
Avg. placebo coefficient	3.691	3.320	4.291	4.186
Centered coefficient	2.203	2.986	7.244	7.454
Empirical p-value	[0.072]	[0.027]	[0.000]	[0.000]
No. of $\lvert\hat{\beta}^{plac}\rvert \geq \lvert\hat{\beta}^{real}\rvert$	72	27	0	0
8–10 km (vs. 0–2 km)				
Real coefficient	3.692	4.386	9.301	9.965
Avg. placebo coefficient	3.122	2.846	4.529	4.763
Centered coefficient	0.571	1.540	4.771	5.202
Empirical p-value	[0.388]	[0.198]	[0.001]	[0.000]
No. of $\lvert\hat{\beta}^{plac}\rvert \geq \lvert\hat{\beta}^{real}\rvert$	388	198	1	0
Controls	Yes	Yes	Yes	Yes
Parliamentary FE	Yes	Yes	Yes	Yes
Mean dep. var.	60.28	60.84	43.46	43.17
Observations	938	938	1,004	1,011

Notes: This table presents the tabular version of Figure 5. The sample comprises polling districts in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, excluding polling districts that contain Chinese New Village population, polling districts in historically urban areas, and polling districts beyond 10 km of a Chinese New Village. The "Real coefficient" is from regression using actual New Village locations. The "Avg. placebo coefficient" is the mean across 1,000 placebo regressions where real New Village locations are replaced with algorithmically generated fake locations on road segments within service areas of real New Villages. The "Centered coefficient" equals Real minus Avg. placebo. Empirical p-values in brackets represent the fraction of 1,000 placebo iterations for which $\lvert\hat{\beta}^{placebo}\rvert \geq \lvert\hat{\beta}^{real}\rvert$. The omitted category is 0–2 km from the nearest New Village.

Table C.4: Balance Tests of the Survey Data

(A) Individual-level variables	Control (C)		Treated (T)		*t*-test (T=C)
	Mean	N	Mean	N	*p*-value
Respondent's age	49.01	731	49.45	1027	0.63
Respondent is married	0.68	732	0.69	1028	0.76
Grandfather interacted with Chinese	0.50	395	0.48	549	0.52
Grandfather trusted Chinese	0.52	387	0.47	534	0.16
Grandfather would have accepted marriage with Chinese	0.46	420	0.38	594	0.01**
Grandfather would have been happy to be near Chinese	0.45	409	0.42	582	0.40
Grandfather mentioned conflicts with Chinese	0.36	526	0.33	735	0.32
Family was rich ($> 50\%$) before Chinese resettlement	0.03	526	0.03	732	0.90
(B) Village-level variables	Control (C)		Treated (T)		*t*-test (T=C)
	Mean	N	Mean	N	*p*-value
Elevation	37.55	43	36.48	58	0.84
Slope	4.25	43	3.69	58	0.15
Market Access to pre-existing villages	12.20	43	12.24	58	0.97
Distance to the nearest Chinese town in 1947 (km)	4.89	43	4.61	58	0.67
Year of village establishment	1918.05	32	1918.00	44	0.99
Village population (post-1945, pre-1947)	104.02	32	131.59	44	0.32
Number of houses (post-1945, pre-1947)	39.26	32	43.11	44	0.65

Notes: This table reports balance test results from the survey data. Panel (A) presents balance tests using individual-level information from the survey, including retrospective data on the older generation. Panel (B) presents balance tests using villege-level information for the surveyed villages. The first four rows use geographical characteristics measdured at the grid cell in which each surveyed village is located, while the last four rows use village characteristics reported by village leaders (aggregating responses from multiple village leaders into the village level). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

D Empirical Results Appendix

D.1 Political Effects

Back-of-the-envelope calculation. To further quantify the potential role of ethnic Malay voters in driving the decreases in ethnonationalist vote share, we would need disaggregated data on vote shares by ethnicity. To the best of our knowledge, such data does not exist. Instead, we perform a back-of-the-envelope calculation *à la* Becker and Woessmann (2009); Calderon et al. (2023) to estimate the minimum levels of ethnic Chinese voter turnout needed to explain away the observed lower vote shares for the National Front. To do so, we assume, unrealistically, that all Chinese voters voted *against* the ethnonationalist coalition and show that the range of Chinese voter turnout rates would have to be implausibly high for Chinese voters alone to explain the entirety of negative ethnonationalist vote shares.[88]

To that end, we estimate Equation (5) via OLS using ethnonationalistic vote shares constructed from the division of $VotebyNonBN - ChiVoters_{d,p}$ by total votes cast in polling district d. The numerator measures the difference between the number of votes received by the opposition coalition (non-*BN*) and the estimated number of ethnic Chinese who would have voted for the opposition coalition in polling district d of parliamentary constituency p. The latter is computed under the extreme assumption that all ethnic Chinese voted against the National Front.

Figure C.7 presents results showing that, even under the extreme (assumption) that all ethnic Chinese voters voted *against BN*, observed vote share differences persist unless Chinese turnout exceeds 50–60%. Given that typical Chinese turnout rates rarely surpass 50% (Malay Mail, 2024), this suggests that neither changes in ethnic Chinese turnout nor ethnic composition are sufficient to explain the full decline in *BN* support.

Appendix References

Becker, S. O. and Woessmann, L. (2009). Was weber wrong? a human capital theory of protestant economic history. *The Quarterly Journal of Economics*, 124(2):531–596.

Hodler, R. and Raschky, P. A. (2014). Regional favoritism. *The Quarterly Journal of Economics*, 129(2):995–1033.

Kratoska, P. H. (1982). Rice cultivation and the ethnic division of labor in British Malaya. *Comparative Studies in Society and History*, 24(2):280–314.

[88] We focus on polling-district level vote shares for 2013 federal constituency seats. Results using vote shares for 2013 state constituency seats are largely similar. Note that it is both extremely unlikely that (i) *all* registered ethnic Chinese voters turned out to vote in the 2013 elections and (ii) that *all* ethnic Chinese voters voted against the National Front (Jomo, 2017; Ostwald and Oliver, 2020).

Kratoska, P. H. (1983). "Ends that we cannot foresee": Malay Reservations in British Malaya. *Journal of Southeast Asian Studies*, 14(1):149–168.

Markandan, P. (1954). *The Problem of the New Villages in Malaya*. Number 1. D. Moore.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.

Phee, T. T. (2012). Oral history and people's memory of the malayan emergency (1948-60): the case of pulai. *Sojourn: Journal of Social Issues in Southeast Asia*, 27(1):84–119.

Sidhu, M. S. (1976). Chinese dominance of west malaysian towns, 1921-1970. *Geography*, pages 17–23.