Interethnic Proximity, Complementarities, and Politics in Malaysia*

Chun Chee Kok
UC Louvain

Gedeon Lim
University of Hong Kong

Danial Shariat
UC Berkeley

Abu Siddique
Royal Holloway

Shunsuke Tsuda
University of Essex

October 10, 2025

Abstract

We exploit a British colonial resettlement policy in Malaysia (1948-1951)—which forcibly relocated over half a million ethnic minority Chinese into segregated "Chinese New Villages" (CNVs)—along with fine-grained spatial variation to examine how persistent interethnic proximity affects the political identity of ethnic majority Malays and long-run local economic development. Malays residing in polling districts closer to CNVs exhibit lower electoral support for the ruling ethno-nationalist coalition, potentially reflecting a moderation of ethno-nationalistic political identity. We also find significant impacts on contemporary local economic development. The political effects are stronger in regions with historical interethnic complementarities, indicating that initial economic interdependence has had lasting consequences. Novel primary survey data further reveal that Malays living near CNVs report greater contact with Chinese minorities, higher trust, weaker zero-sum mindsets, and higher income and wealth. These findings suggest that persistent intergroup contact and economic complementarities can jointly underpin political moderation and local development. ***JEL***: D72, J15, N45, O15, R23.

Keywords: Ethnic Diversity, Interethnic Contact, Political Identity, Economic Complementarities, Forced Resettlement, Colonial Policy, Social Cohesion.

*We thank Sascha O. Becker, Paul Raschky, Sam Bazzi, Dilip Mookherjee, Robert A. Margo, 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, 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, Claudio Labanca, Yu-Hsiang Lei, Bingjing Li, Matt Lowe, Sara Lowes, Pushkar Maitra, Eoin McGuirk, Harrison Mitchell, Chicheng Ma, Salma Mousa, 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 BÆM Development Workshop at Bristol, Diversity and Inclusion Workshop (Milan-Bicocca), Essex, SWET (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), and Early Career Workshop QPE (KCL). 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 Seed Funding (2020-2022). This paper was previously circulated as "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

People identify with political parties in much the same way they identify with racial, ethnic, or religious groups (Campbell et al., 1960). However, unlike many social identities that individuals are born into, political identity can be directly shaped through socialization and economic conditions (Greene, 2004; Jennings and Niemi, 1974; Stigler, 1973). Specifically, who individuals interact with in residential, schooling, and labor markets matters for how they perceive the world, the economic benefits they derive, and ultimately, through these channels, who they vote for. In turn, the aggregation of these individual choices across space and time have direct implications for democratic stability, social cohesion, and economic development (Alesina and La Ferrara, 2005; Glaeser, 2005; Guriev and Papaioannou, 2022; Jha, 2025).

Yet, even as the world grows more diverse, we know little about how routine social and market interactions tied to settlement patterns translate into long-run changes in economic development and political identity, particularly in low- and middle-income countries.[1] Understanding this can offer novel insights into how the physical and economic organization of communities affects the evolution of political identities. It can also push us to think beyond classical models of political behavior (Downs, 1957; Kramer, 1971; Stigler, 1973), which view political preferences primarily as evaluations of party performances and expectations of future utility derived from the parties.

In this paper, we take a historical perspective to answer these questions. Specifically, we study how persistent proximity to an ethnic minority community can influence the co-evolution of political identity and economic prosperity of the ethnic majority group in the long-run. To do so, we leverage a 1950s colonial, military resettlement program in British Malaya (now Malaysia) that forcibly resettled over half a million widely dispersed rural ethnic Chinese—nearly the entire rural Chinese population (about 25% of the total Chinese population and 10% of Malaysia's population)—into hundreds of compact, fenced ethnic Chinese settlements (Chinese New Villages, CNVs). The program was implemented to sever the rural support networks of anti-British communist insurgents, who were predominantly ethnic Chinese.[2] The immediate result was the creation of a dispersed network of sharply demarcated Chinese enclaves surrounding pre-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).[3]

[1] Similar questions about the long-term economic and political effects of diversity (via immigration) have been explored in the context of high-income countries like the US (Sequeira et al., 2020; Tabellini, 2020), showing positive effects on economic outcomes but mixed effects on political outcomes.

[2] For security reasons, small numbers of ethnic majority Malays were also resettled into compact settlements. In Section 7.3, we use these settlements to test if our baseline results are driven by exposure to newcomers in general, or specifically to ethnic Chinese minorities.

[3] In 1963, Malaysia became a first-past-the-post democracy, and all Malaysians gained full voting rights.

We leverage the granular variation in interethnic proximity created by the resettlement program to estimate long-term impacts on Malay political behavior—measured as support for ethnonationalist parties—and on local economic development. We interpret ethnonationalist vote shares as a behavioral proxy for the salience of Malay ethnopolitical identity. We then explore the underlying mechanisms driving these outcomes, including economic competition and complementarities, as well as novel measures of intergroup contact, social capital, trust, empathy, and zero-sum thinking. To do so, we build a rich dataset combining voting data at the polling-district level—the most disaggregated level possible—with geospatial, administrative, archival, satellite, census data (at the finest gradation possible), with individual-level primary survey data.

The Malaysian context offers four advantages. First, we argue that resettlement site criteria were plausibly exogenous to underlying locational fundamentals and unobservable characteristics: the program was implemented rapidly under military exigencies, and locations were chosen primarily for defensibility and ease of reinforcement.[4] Second, we geolocate precise locations of CNVs and link them to a rich set of initial CNV characteristics from an underutilized historical survey by Malayan Christian Union (1958). Variation in initial CNV characteristics allows us to pin down mechanisms behind the exposure effects. Third, because Malaysian voters rarely change their registered voting addresses (Jomo, 2017), polling district vote shares can plausibly be interpreted as reflecting (i) the changing political preferences and identity of Malays who were born 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. Fourth, although Malays and Chinese are culturally and linguistically very different, there have been no large-scale interethnic conflicts in Malaysia since 1969.[5] The Malaysian context hence provides a promising setting to study the potentially positive role of interethnic proximity in preventing conflict and violence.

For identification, we exploit the plausibly exogenous nature of the resettlement process and site characteristics. To address potential biases in ordinary least squares (OLS) estimation, we implement a spatial randomization inference-style approach inspired by Dell and Olken (2020). Specifically, we construct 1,000 counterfactual ("fake") village sites for each *real* CNV, using precise military site selection criteria obtained from archival documents. This method creates a counterfactual distribution of resettlement locations, allowing us to compare the exposure effects of proximity to actual CNVs with those of proximity to the counterfactual sites, averaged across the

[4]British military archival documents explicitly outline key site selection criteria, providing a clear framework for understanding the resettlement process: (i) road accessibility and livelihood continuity; (ii) topographical suitability; (iii) exclusion of lands reserved for ethnic Malays; (iv) drainage/soil suitability. Socio-anthropological studies suggest that (iv) was the least adhered to. We formally embed these criteria into our empirical strategy using a randomization inference-style approach described in Section 4.2.

[5]Malaysia has an ethnolinguistic fractionalization index (ELF) of 0.65, making it the country with the 9th highest ELF in a sample of 79 countries (Desmet et al., 2017). In contrast, Burkina Faso (ELF = 0.66) has had a history of military mutinies, coups, and social unrests (Englebert, 2018).

1,000 spatial configurations. Our procedure ensures balance across a large number of pre-treatment geographical and socio-economic characteristics. The main residual imbalance is the 1947 ethnic Chinese share: polling districts within 0–2 km of actual CNVs have 2.8–4.0 percentage points higher Chinese shares than districts 2–4 km and 4–6 km away (7.3–10.4% relative to a mean of 38.33%). Accordingly, we control for 1947 ethnic Chinese shares in all regressions, so estimates are conditional on the pre-resettlement, spatial distribution of ethnic Chinese communities.

Throughout, we analyze and present results on the impact of interethnic proximity by studying the exposures effect of CNVs on ethnic majority Malays in surrounding areas. In sum, we apply a "doughnut-hole approach," excluding all polling districts containing initial CNVs. This isolates the exposure effects of proximity to CNVs while accounting for potential spillover effects.[6]

Results preview. Our primary political outcome is political behavior: the polling-district level, electoral vote share for the ethnonationalist coalition, the National Front (*Barisan Nasional*, BN) in the 2013 and 2018 general elections. Elections in Malaysia are conducted once every 5 years. Vote shares are observed at the polling district level—the most disaggregated unit available—with a median area of 5.4 km^2, roughly 1,400 voters, and 2-3 villages.[7] Importantly, the use of polling-district level vote shares 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 political candidate and party identity constant throughout all our regressions.

BN is a long-ruling coalition of ethnic-based parties that governed Malaysia for nearly six decades and consistently advocated and implemented affirmative, Malay-first policies (Jomo, 2017). We interpret BN vote share as a behavioral proxy for the salience of Malay ethno-nationalistic identity. Given that identity is arguably broader than any single electoral outcome, we support this interpretation with various supporting mechanisms (social distance, trust, zero-sum beliefs, etc.) and find strong convergent patterns. We detail these below.

We find a persistent, localized 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—consistent with a long-run shift away from ethnonationalism in areas more proximate to CNVs. In 2018, magnitudes are larger and more precisely estimated,

[6] Our treatment estimates potentially reflects the effects of both *increased* exposure to Chinese populations around CNVs and *decreased* exposure in depopulated origin areas. While we are unable to disentangle these effects, the latter is likely minimal, as original Chinese settlements were largely dispersed and located near rivers and jungle peripheries. Hence, our estimates likely capture the *increased* effects of exposure to a higher *concentration* of ethnic Chinese.

[7] Polling district level vote shares are available only from the 2013 general elections.

possibly suggesting the strengthening of effects from longer interethnic exposure.[8][9]

We consider these estimates to be conservative, likely representing a lower bound. Specifically, in our identification strategy, some control polling districts in counterfactual assignments may be located in close proximity to actual CNVs, contaminating our counterfactual estimates and biasing effects toward zero. As a robustness check, we re-estimate the counterfactual regressions excluding polling districts within 0–2 km of *actual* CNVs from the control set. Effect sizes become slightly larger and more significant, suggesting that our main estimates may indeed reflect a lower bound.

We rule out two alternative explanations: differences in ethnic composition and selective migration. First, differences in vote shares may simply be driven by compositional effects. It is possible that a larger number of ethnic Chinese live in closer proximity to CNVs and these individuals simply tend to vote less for the ethno-nationalist coalition.[10] Using full-count voting rolls, we compute and show that there are 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. In addition, a back-of-the-envelope calculation *à la* Becker and Woessmann (2009); Calderon et al. (2023) 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 voting patterns.

Second, using rich 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. In addition, qualitative survey data and the extant socio-political literature suggests that Malaysians rarely change their registered voting address due to high administrative costs and strong social ties to their home villages (Jomo, 2017).[11]

Together with changes in political behavior, we also find meaningful local economic development effects near CNVs: nighttime light intensity is about 19% higher at the polling-district level and about 45% higher at the 1×1km grid-level; population density is about 40% higher at

[8] Anti-ethnonationalist sentiments might also have been boosted by the broader wave of anti-incumbent sentiment that swept Malaysia and resulted in the opposition gaining control of the government for the first time in 60 years.

[9] Beyond average effects, we also uncover a notable non-linear pattern in the data. We observe an inverted-U relationship between proximity to CNVs and ethno-nationalist vote shares, particularly in the 2018 election, often peaking between 2-8 km, implying ethnic-based political preferences are strongest at intermediate distances. We offer three plausible interpretations: (i) racial threat theory (Blalock, 1967) and how perceived competition can be most intense when the outgroup is visible but not familiar, declining with closer contact or greater distance; (ii) Ethnic conflict theory (Caselli and Coleman, 2013), which predicts that conflict peaks at moderate levels of resource competition; (iii) Contextual contact (Stein et al., 2000), wherein intermediate exposure without personal interaction may heighten ethnic awareness, while both close and distant areas experience less such salience.

[10] There is some evidence of in-migration of ethnic Chinese to CNVs over time (Strauch, 1981) that might have led to the expansion of CNVs beyond their original boundaries and into adjacent polling districts.

[11] Relatedly, selective resettlement of Chinese based on pro-Communist leanings, was unlikely as military expediency was key and nearly all rural ethnic Chinese were forcibly relocated into CNVs within just three years, regardless of political leanings.

the polling district and more than double at the grid cell-levels. These patterns indicate highly localized densification and economic activity, consistent with agglomeration close to CNVs. On public goods, we find about 30-40% higher road coverage near CNVs, but little to no change in the number of schools or health facilities, and no effect on school quality (measured by students–to-teacher ratios).[12] We interpret this as connectivity-driven development without parallel growth in service provision or its quality. Furthermore, analysis of electoral victories implies that lower BN vote shares near CNVs did not translate into seat turnover to the civic-nationalist (or pro-diversity) coalition, implying no shift at the extensive margin. Thus, we argue that observed benefits in local economic activities and road connectivity are less likely to be driven by targeted, local electoral competition via coalition-specific policies *within* parliament seats.

Mechanisms. We explore three potential mechanisms of (i) interethnic contact; (ii) historical competition and complementarities; (iii) exposure to newcomers; that might explain the observed changes in the political preferences of Malays, possibly driven by their shifts in political identity.

First, we consider *direct interethnic contact* as a plausible mechanism (Allport, 1954). Proximity to CNVs (and the denser road networks that we observe) may have increased casual and formal interactions between Malays and Chinese and deepened their market relationships. Such contact, together with economic complementarities, can improve economic relationships, build social capital and trust, and encourage more positive attitudes toward the Chinese, potentially changing Malays' political identity and, in turn, their political preferences. To test this hypothesis, we surveyed 1,960 Malays to measure social interactions, trust, attitudes, and economic relationships with the Chinese.[13] Consistent with our hypothesis, we find that Malays living near CNVs report more interactions with Chinese (in villages, schools, and workplaces), significantly higher situational trust, less zero-sum thinking, greater occupational/market overlap with their Chinese neighbors, and higher income and wealth. These patterns support a contact-and-complementarities mechanism through which sustained proximity strengthens trust and economic interdependence, weakens zero-sum beliefs, and persistently moderates Malay people's political identity.

Second, historical interethnic *competition* and *complementarities* could affect political preferences. On one hand, as the Chinese minority became economically established following resettlement, their growing success may have intensified competition with neighboring Malays—particularly over employment, business opportunities, and access to state resources—thereby strengthening ethno-aligned political preferences (Horowitz, 2000). On the other hand, interethnic proximity may have encouraged economic complementarities and specialization, increasing productivity and promoting more inclusive political attitudes through cooperative relationships (Alesina and

[12] In fact, there are slightly more schools 2–4km from CNVs and slightly more health facilities 8–10km away than within 0–2km.

[13] We describe survey sampling, design, and variables in detail in Section 7.1.

La Ferrara, 2005; Jha, 2013). To test these hypotheses, we exploit variation in initial occupational shares from the 1958 Malayan Christian Union data, classifying CNVs by above- or below-median employment in agriculture versus rubber and tin mining. In areas with above-median initial rubber and tin employment—labor-intensive sectors where employers often needed to hire both Malay and Chinese workers (Ross, 2014; Siew, 1953)—we find lower ethnonationalist vote shares and small positive effects on economic outcomes in polling districts adjacent to CNVs, consistent with interethnic complementarities. Conversely, we find little evidence for the role of economic competition. In areas with above-median initial agricultural employment—individualistic sectors where Chinese and Malays likely worked their own land and competed more directly (and which has been long considered as the traditional sector for indigenous Malays (Kratoska, 1982a))—we find few differences.

Finally, observed effects might simply reflect changes in political behavior from exposure to the ideas, opportunities, and norms of newcomers (of any ethnic identity) (Balietti et al., 2021; Chetty et al., 2016) rather than exposure to an ethnic minority group. We test this with a falsification exercise using a contemporaneous, forced resettlement program of ethnic Malays that created Malay New Villages (MNVs). The establishment of MNVs by the British military took place on a much smaller scale to protect dispersed Malay villagers from communist violence (Dobby, 1952; Humphrey, 1971). If generic exposure to newcomers were driving our results, similar political and economic patterns should also appear near MNVs. We find little evidence of this: vote shares near MNVs are statistically indistinguishable from their counterfactuals, suggesting that observed results are likely driven by interethnic proximity to *minority communities*, and not by generic exposure to new settlers regardless of ethnic identity.

Literature and contributions. This paper makes four contributions. First, it bridges the literature on the effects of forced emigration and resettlement on nation-building and assimilation (Bazzi et al., 2016) and that on spatial growth and economic development (Peters, 2022).[14] Most studies focus on aggregate economic or political effects, but the interaction between the two and the distributional effects on natives remain less clear.[15] Using granular data, we show that ethnic-majority individuals living closer to resettled ethnic-minority villages enjoyed economic benefits and are less likely to vote for ethno-nationalist parties. Hence, we extend recent studies (e.g. Calderon

[14] A number of excellent studies have also looked at human capital and labor market outcomes of the forcibly removed in Poland (Becker et al., 2020), South Africa (Carrillo et al., 2023), and the labor market effects of Japanese-Americans who were interned during WWII (Arellano-Bover, 2022). Separately, (Toews and Vézina, 2025) studies long-run exposure effects on the economic development of surrounding communities from Stalin's forced labor camps that targeted intellectuals. Abel (2019) compares economic and social outcomes of resettled vs surrounding communities under South African apartheid. Here, we focus on the interaction between economic and political effects on surrounding communities.

[15] Alix-Garcia et al. (2018) and Tsuda (2022) exceptionally study the distributional effects of forced displacement on host economies, focusing on impacts across different markets in rural developing economies.

et al., 2023; Fouka et al., 2022) finding that intergroup contact leads to coalition-building, positive transmission of political attitudes, and positive downstream political effects. Using fine-grained polling district vote shares and full-count voting rolls, we show that, in the presence of tangible economic benefits, persistently higher exposure to an ethnic minority group can lead to positive effects on the political identity and behavior of the ethnic majority in the long-run—even in the context of extreme ethnic polarization.

Second, this paper connects to work on interethnic complementarities (Diaz-Cayeros and Jha, 2022; Jedwab et al., 2019; Jha, 2013; Montalvo and Reynal-Querol, 2021) by providing evidence that historical economic complementarities are associated with persistently more positive interethnic attitudes. Using novel primary survey data, we elucidate micro-level mechanisms: sustained intergroup contact, more inclusive attitudes and beliefs toward the Chinese, and ongoing market complementarities align with long-run shifts in political behavior. Our findings, therefore, provide quasi-experimental evidence for the large and generalizable positive effects of *broad* intergroup contact (Lowe, 2024).[16] Furthermore, we contribute to studies on the positive impacts of *intra*-community contact (Bazzi et al., 2019; Billings et al., 2021). In particular, Bazzi et al. (2019) studies the effects of contact among multiple ethnic groups on national identity and finds that greater fractionalization (polarization) leads to a stronger (weaker) national identity. In contrast, we study *inter*-village contact between two large ethnic groups and show that interethnic contact, persistent economic interdependence, and shifts in intergroup attitudes and beliefs can jointly explain long-run political moderation even in the context of extreme ethnic polarization.

Third, we contribute to the broader literature on the political effects of immigration (Alesina and Tabellini, 2024).[17] Importantly, our results contrast with Tabellini (2020), who finds that higher immigration in US cities led to positive economic outcomes but stronger political backlash. One possible reason is contextual: in our largely rural, low-income setting, there could have been greater opportunities for interethnic contact in non-segregated, public spaces, leading to positive political attitudes toward the ethnic minority. More importantly, economic complementarities and local development are likely more salient and locally captured in our setting because development arrived on the extensive margin (e.g., new roads and first-time market access) rather than developments on the intensive margin that are more common, but perhaps less immediately salient for racial majority groups in more urban and/or higher-income settings.

[16]This also relates to the growing literature in social sciences that examines how networks affect economic and political outcomes. For instance, studies have shown that social networks influence economic mobility (Chetty et al., 2022; Kling et al., 2007), access to information and opportunities (Jackson, 2008), diffusion of information that can improve economic outcomes of different ethnic groups (Siddique et al., 2024), and political information and voting (Gerber et al., 2009; Kernell and Lamberson, 2023).

[17]Studies have documented mixed evidence for how immigration affects native voting behavior. These studies, however, largely focus on the effects of exposure to refugees or immigrants who do not have voting rights (e.g., Barone et al., 2016; Dustmann et al., 2019; Halla et al., 2017; Otto and Steinhardt, 2014; Steinmayr, 2021).

Finally, we contribute to the literature on the effects of exposure to diversity on social and political preferences (Algan et al., 2016; Bursztyn et al., 2024; Enos, 2014; Lowe, 2021; Rao, 2019; Siddique et al., 2024; Wren-Lewis et al., 2024). Specifically, the closest paper to ours, Billings et al. (2021), studies the end of race-based busing and finds that greater interracial exposure during childhood reduced the likelihood of white voters registering as Republican. We innovate by studying the effects of intergroup exposure over the life-cycle, and show that negative political preferences towards ethnic minorities can be mitigated through positive economic spillovers from interethnic complementarities. Our findings thus provide microfoundations that complement macro-level studies on ethnic segregation and diversity (Alesina and Zhuravskaya, 2011).

Roadmap. Section 2 outlines the institutional background. Section 3 describes the data. Section 4 describes the empirical strategy. Section 5 reports empirical results of the political effects. Section 6 examines local development effects. Section 7 describes underlying mechanisms behind the results. Section 8 concludes.

2 Institutional Background

2.1 Ethnic Diversity and Interethnic Relations in Malaysia

British rule of Peninsular Malaysia began in 1786 and ended in 1957 when Malaysia declared independence. The population of Peninsular Malaysia today comprises 64.5% ethnic Malays and indigeneous people (*Bumiputera*), 25.9% ethnic Chinese, and 8.9% ethnic Indians in 2010. Ethnic Malays differ substantially from ethnic Chinese in terms of language, religious affiliations, appearance, and culture and tradition. There is also limited intermarriage between the ethnic Malays and ethnic Chinese (Nagaraj, 2009).

The earliest Chinese migration to Malaysia dates back to the 15^{th} century, but most resettled Chinese arrived in the early 20^{th} century, driven by upheaval in China and drawn to British Malaya's labor demands (Kim, 1998; Strauch, 1981; Wang, 1959). These migrants, often traders and laborers, remained culturally distinct, maintained clan-based institutions (Ting, 1976), and were viewed as aligned with China, raising security concerns (Suryadinata, 1987).[18]

British colonial policy further entrenched ethnic divisions by assigning Malays to subsistence farming and Chinese to lucrative sectors like mining and trade (Kratoska, 1982b). This institutionalized economic disparity laid the groundwork for persistent post-independence tensions, with many Malays perceiving the Chinese as disproportionately advantaged.

[18]Malaysia hosts the second largest ethnic Chinese population after Thailand across Southeast Asia. In 1981, there were as many as 4 million ethnic Chinese living in Malaysia (Suryadinata, 1987).

2.2 Electoral Politics in Malaysia

In 1969, the economic gap between Malays and Chinese culminated in widespread racial riots, which were driven by economic grievances from ethnic Malays who did not fully participate in the rapidly modernizing economy of the newly independent Malaysia.

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 poverty and ethnic inequality. The NEP granted ethnic Malays preferential access to education, public-sector jobs, housing, and corporate equity (Jomo, 2017).[19] These policies helped maintain political dominance: until 2018, the National Front governed without interruption.

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 anti-corruption and inclusive, multi-ethnic governance.[20] 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).[21]

In 2018, PAS exited the opposition over ideological splits—mainly on implementing Islamic sharia law—and formed the conservative *Gagasan Sejahtera* (GS) bloc.[22] The opposition rebranded itself as the Alliance of Hope (*Pakatan Harapan*), still led by PKR and Anwar Ibrahim.[23] The election became a three-way race: reformist (Alliance of Hope), conservative Islamic (GS), and ethno-nationalist (National Front). In a historic result, the Alliance of Hope won the popular vote and formed Malaysia's first non–National Front federal government.

2.3 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.

[19] 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.

[20] PKR, led by Anwar Ibrahim, appealed to urban Malays; DAP to Chinese and Indian voters; and PAS to rural Malays.

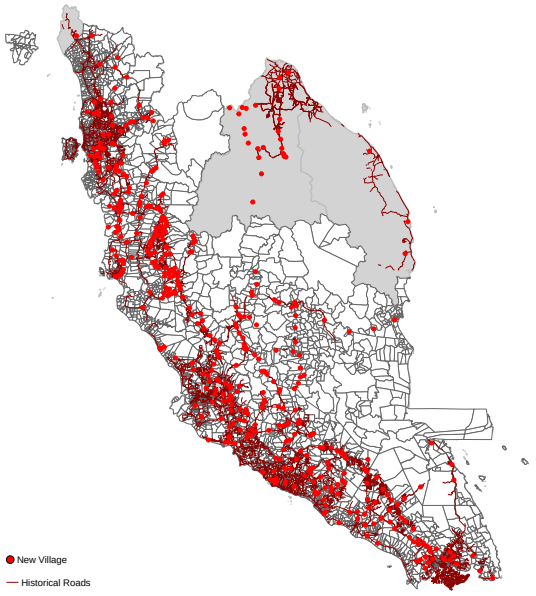
[21] The Malaysian Chinese Association and Malaysian Indian Congress—UMNO's coalition partners—had become politically marginal by 2013.

[22] This schism led to the formation of *Amanah*, a moderate Islamic party that remained in the opposition coalition.

[23] We refer to the Alliance of Hope as the People's Alliance throughout for continuity.

Citizens aged 21 and above are automatically registered to vote.[24] Voters are generally assigned to polling stations near their registered voting address, which almost always maps directly to a single polling district.[25] Notably, Malaysians rarely update their registered voting addresses despite later-life migration, often preferring to vote in their hometowns.[26] In Section 3.2, we explain how this informs our interpretation of vote share differences.

Figure 1
Location of New Villages and Historical Roads



Notes: The location of 452 New Villages are represented by red dots and the pre-resettlement road network is represented by dark red lines. The white polygons indicate polling district boundaries in 2013. Polygons shaded in grey are the states of Kelantan and Trengganu. We exclude these states from our analyses as they largely contain New Village resettlements of non ethnic Chinese villagers. Source: (Lee, 2012; Lim and Song, 2002), authors' own geo-referencing and HIND1035-series maps from 1947 (Australian National Library Archives).

2.4 The Colonial Resettlement of Ethnic Chinese to New Villages

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

[24]The voting age was lowered to 18 starting from the 2022 General Elections.

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

[26]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.

(nearly the entire rural Chinese 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 CNVs that we successfully geolocated, as described 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 rising communist insurgency, led by the Malayan Communist Party (MCP), which had gained legitimacy among ethnic Chinese after resisting Japanese occupation. As the MCP turned against British rule, 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.[27]

The program rapidly displaced rural Chinese—mostly engaged in farming, pig-rearing, and smallholder mining—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. Described by Humphrey (1971) as a "laboratory experiment", the resettlement imposed strict movement restrictions as Chinese New Villages were heavily surveilled, often enclosed by barbed wire and largely isolated from neighboring communities from 1948-1960.

Freedom of movement returned after the Emergency, 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).

3 Data

3.1 Sampling Frame

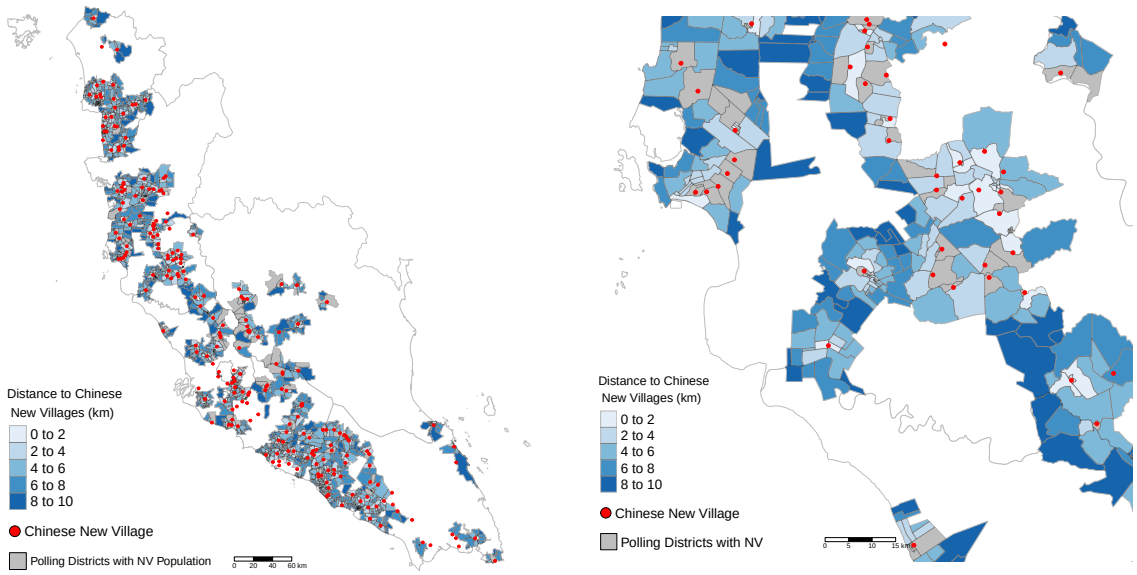
Chinese New Villages. We construct a baseline sample of 236 CNVs in three steps. First, we obtain a list of 666 CNVs from the Malayan Christian Union (1958). Second, using maps published by the Ministry of Housing and Government (Lee, 2012), we geolocate 452 CNVs and manually verify their existence today using Google Earth and Google Maps. Figure 1 shows their distribution across Peninsular Malaysia.

Third, we impose three sample criteria: (i) we keep all CNVs whose primary medium of lan-

[27] 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.

guage was recorded as Mandarin or a Mandarin dialect in 1958.[28] 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. (ii) We exclude CNVs in the two east coast states of Kelantan and Trengganu, where the ethnic Chinese population has historically been low and resettlement largely involved non-ethnic Chinese. Hence, we focus on seven states. (iii) We exclude all CNVs located in 1947 urban census districts.[29] This procedure gives us a total of 208 CNVs in our regression sample. Table A.1 reports the summary statistics of these 208 CNVs and compares them to 244 out-of-sample CNVs. Figure 2 shows the geographical distribution of our baseline sample of CNVs.

Figure 2
Interethnic Proximity: Distances to the nearest Chinese New Village at the Polling District-Level



Panel A **Panel B**

Notes: Panel A shows the location of Chinese New Villages in red dots and the distances to the nearest Chinese New Villages in the estimating sample. Panel B is an example of an enlarged area for illustrative purposes. Polygons depict polling district boundaries. Polygons shaded in darker colors indicate greater distances to the nearest Chinese New Village.

[28] 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 7.

[29] 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.

Polling districts. We use polling-district level electoral data from the 2013 and 2018 Malaysian General Elections. Polling districts, the smallest electoral unit in Malaysia, had 1,457 units in 2013 and 1,557 in 2018, each averaging 10.9 km^2 and 1,519 registered voters (Longuet-Marx, 2024).[30]

We impose four sampling restrictions. First, to ensure that we are comparing outcomes only across polling districts 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 that we describe in Section 4.2 where Chinese were resettled between 3.2 to 9.6 kilometers from their original locations (dhu Renick, 1965).[31] 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.[32] Last, given our focus on studying the effects of inter-ethnic proximity between ethnic Chinese and Malays—the two largest ethnic groups in Malaysia—we exclude all polling districts with an ethnic Indian-majority.[33]

3.2 Treatment and Primary Outcomes

Interethnic proximity. We define interethnic proximity—our treatment variable—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 as a substantial number of pre-existing, non-resettlement villages in the pre-resettlement period were (i) not accessible by main roads and (ii) it is possible that inter-village movement and contact was more likely to occur as a result of villagers traversing unrecorded dirt paths rather than paved roads.

[30]The average population of a polling district is roughly equivalent to an electoral precinct in the US (Longuet-Marx, 2024).

[31]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.

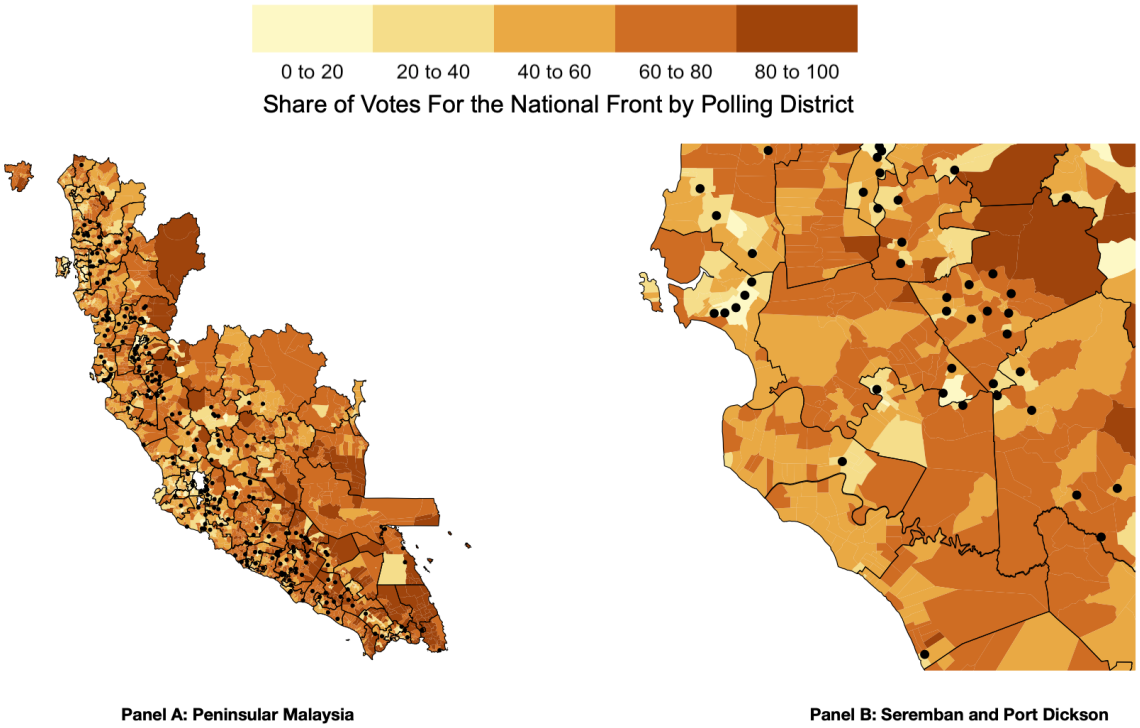
[32]Specifically, we digitize census district polygons from the 1947 Malayan Population Census and classify a census district as urban if it contains a major town. A major town is defined by the Census as any town containing a population of more than 10,000 inhabitants. We then overlay contemporary polling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

[33]Specifically, given 3 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 a mere 8.9% of Malaysia's population. Furthermore, the political economy of these polling districts are extremely different — these polling districts would have comprised of plantation lines, where rubber 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, 1982b).

Vote shares (a behavioral proxy for Malay political identity). 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 ethno-nationalist 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 opposition Islamic party (PAS) for three reasons. First, in Malaysia, ethnicity and religion are closely linked—nearly all Malays are Muslims by constitutional definition. Second, the 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 as part of a multi-ethnic coalition, and our results are robust to excluding such contests. Third, in 2018, PAS shifted further right as part of a separate conservative coalition (GS), but it won only 16.89% of the popular vote, making its impact limited.[34]

Figure 3
Vote Shares For the National Front in 2013, Polling District-Level



Notes: This figure displays vote shares for the National Front at the polling district-level in 2013. Federal parliament constituency boundaries are outlined in black and polling districts are shaded in yellow. Chinese New Villages are represented as black dots. Each parliament constituency contains an average of 18 polling districts.

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

Figure 3 shows the polling-district level geographical distribution of vote shares for BN in the 2013 federal elections. Using polling district-level data offers three advantages. First, polling districts are much smaller and more demographically homogenous. An average federal constituency contains 18 polling districts. Second, unlike federal and state constituencies where cases of malapprortionment to favor BN coalition have been well-documented (Ostwald, 2017), polling districts largely contain a roughly equal number of voters and have been less subject to political efforts at re-drawing boundaries. Third, the much finer granularity allows us to leverage granular distances to CNVs and control for parliament seat fixed effects.[35] Hence, we use vote shares in the federal elections as our main outcome variable unless otherwise noted.

We interpret vote share differences as reflecting both formative and later-life exposure to CNVs. Since Malaysians rarely change their voting addresses, vote shares primarily capture the political *preferences* shaped by growing up near CNVs. However, in areas with more out-migrants who return to vote, preferences may also reflect interethnic exposure at migration destinations. To account for this, all estimations control for proximity to the nearest 1947 urban center.

Voter turnout and registered voters. We calculate turnout rates at the polling-district level. During our sample period, federal and state elections were held simultaneously. Datasets were obtained from the Malaysian Electoral Commission. We construct ethnic shares of registered voters across the entire Peninsular Malaysia using full-count voter rolls from the 2013 General Election. Voter rolls in 2013 record the self-identified ethnicity of each voter in Malaysia. This is, to the best of our knowledge, the most granular and accurate measure of ethnic shares available, as existing data on ethnic shares exists only at higher aggregated census and administrative district levels. We are unable to do so using voter rolls from the 2018 General Election as the self-reported ethnicity variable is missing for a substantial number of voters.

Local economic development and urbanization. Contemporary censuses do not provide information at a sufficiently disaggregated level for our analysis. As income per capita data is unavailable at the very local level in this context, 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).[36] We also use population data in 2010 from the Global Human Settlement Layer (GHSL). These data are aggregated to obtain the average values at the grid cell level of 1km$\times$1km and the polling district level, respectively.

We measure educational public goods by geo-referencing a complete list of all primary and

[35] Pre-2013 electoral results are only available at the parliament or state constituency seat level which is too aggregate for exploiting the fine-grained variation of within-parliament seat distance to New Villages.

[36] We use 2014 data because earlier nighttime light intensity measures may contain substantial measurement error due to issues such as flawed blurring, lack of calibration, and top-coding (Gibson, 2021).

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. 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 total road length by the area (in square kilometers) of each polling district.

4 Empirical Strategy

4.1 Ordinary Least Squares Estimation

We estimate the effects of distances to the nearest CNV on polling-district level political and economic outcomes via OLS:

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

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. $\mathbf{X}_d$ is a vector of polling-district level controls that includes slope, elevation, percentage of east-facing grids, percentage of topsoil organic carbon, percentage of topsoil sodicity, an indicator for drainage being very poor, an indicator for soil being coarse, an indicator for soil being medium, distance to the nearest coast, distance to the nearest urban center in 1947, the natural logarithm of population density in 1947, and ethnic Chinese shares in 1947. Regressions using state legislative seat election results further control for indicators for every possible combination of party match-ups at the state constituency level.

Importantly, controlling for ethnic Chinese shares and population density in 1947 ensures that we compare areas with similar levels of pre-treatment diversity and economic prosperity. Furthermore θ_p is a vector of federal parliamentary constituency fixed effects. This ensures that our analyses compares voting outcomes across polling districts in which voters are voting for both the same party and candidate. We cluster standard errors at the federal parliamentary constituency level.

However, the potentially endogenous placement of CNVs raises concerns that OLS estimates of β_k in Equation 1 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 hence better long-run market access, confounding the estimated effects. Similarly, if CNVs were placed in areas where locals were already more accepting of Chinese settlers, observed voting patterns may reflect pre-existing attitudes rather than the effects of interethnic proximity. To address this concern, we complement the OLS strategy with a counterfactual analysis in subsection 4.2, which leverages plausibly exogenous British site selection criteria to better isolate the causal effects of proximity to Chinese resettlement sites.

4.2 Counterfactual Site Selection Using British Military Criteria

To address concerns about the potential endogeneity of CNV locations, we implement a counterfactual analysis inspired by Dell and Olken (2020). The approach leverages the fact that CNV sites were selected rapidly under military duress during the Malayan Emergency—when about 500,000 ethnic Chinese were resettled into nearly 500 villages in just three years under the Briggs Plan. This urgency meant that site selection followed a set of practical military criteria rather than pre-existing economic or political conditions.

Using declassified British military planning documents, we reconstruct the set of feasible counterfactual CNV locations that could have been chosen. We generate 1,000 sets of counterfactual CNVs, each satisfying the following plausibly exogenous site selection rules:[37]

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 NV-suitable land (within 2.5km radius, which is the average size of a CNV (Nyce, 1973)) as the 10^{th} percentile of actual CNVs.

In addition to the explicit military planning criteria documented in archives, we introduce two additional constraints to improve the realism and comparability of our counterfactual analysis:

3. *Exclusion of Malay reservations:* Because the British military could not resettle Chinese communities on lands reserved for ethnic Malays, counterfactual sites must lie outside designated Malay reservations (Kratoska, 1982b, 1983).

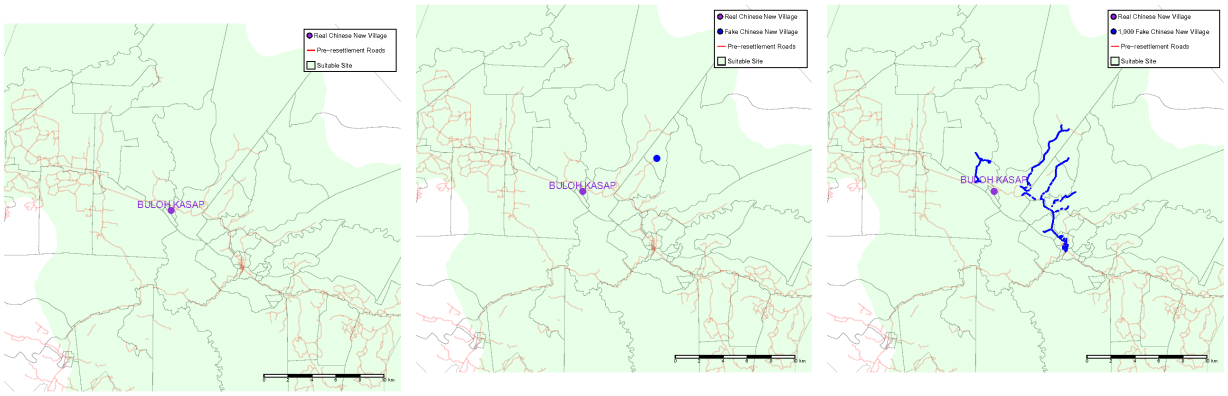
4. *Spatial balancing:* To ensure geographic comparability, each set of counterfactual CNVs is constructed by randomly shifting actual CNV locations while maintaining balance in the north–south and east–west directions. We implement a simulated annealing procedure to minimize the average distance between actual and counterfactual CNVs following Dell and Olken (2020).

[37] See Appendix A: Site Selection Criteria for more details.

Data. To measure road accessibility and exclude sites within Malay Reservations, we digitize the full universe of 1947 pre-resettlement roads and Malay Reservation boundaries across Peninsular Malaysia using the high-resolution HIND1035 map series (scale: 1:63,360 or 1 inch to a mile) from the Australian National Library archives. To our knowledge, this is the most detailed mapping source available for capturing local-level variation in the pre-resettlement period. Figure A.1 presents an example of these historical maps.

In implementing the fourth site selection condition, *spatial balancing*, we ensure 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 unavailable, spatial balancing helps replicate the distribution of actual CNVs, which were likely sited to minimize disruption to villagers' economic activities (dhu Renick, 1965).

Figure 4
Counterfactual Chinese New Villages Example



Panel A: A Real New Village | **Panel B: A Counterfactual New Village** | **Panel C: 1,000 Counterfactual New Villages**

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

Figure 4 illustrates this procedure. Panel A shows an actual CNV (Buloh Kasap) in 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. The polling district-level 'counterfactual' estimating equation takes the following form, controlling for the same set of covariates and parliament fixed effects, θ_p, as before

in OLS Equation (1):[38]

$$Y_{d,p} = \alpha + \sum_{k=1}^{4} \gamma_{k,real(fake)} distCNV_d^{k,real(fake)} + \gamma_2 X_d + \theta_p + \epsilon_{d,p} \quad (2)$$

The key difference with Equation 1 is that, for every outcome, Equation 2 is estimated twice. Once for the real CNVs and once for the 1,000 sets of counterfactual/fake CNVs. Hence, the point estimate of the effect of being distance k away from a CNV is given by the difference between the coefficient of $distCNV_d^{k,real}$ and the average of the coefficients of 1,000 $distCNV_d^{k,fake}$ from 1,000 sets of counterfactual regressions. The key advantage over the simple OLS in Equation 1 is that here, we purge the effects of any unobserved factors that might be correlated with British military site suitability criteria. We also continue to control for $\mathbf{X}_d$, a vector of pre-determined controls, that include 1947, pre-treatment population density and ethnic Chinese shares. As before, θ_p is a vector of federal parliamentary constituency fixed effects—this ensures that we continue to compare voting outcomes across polling districts in which voters are voting for both the same party *and* candidate.

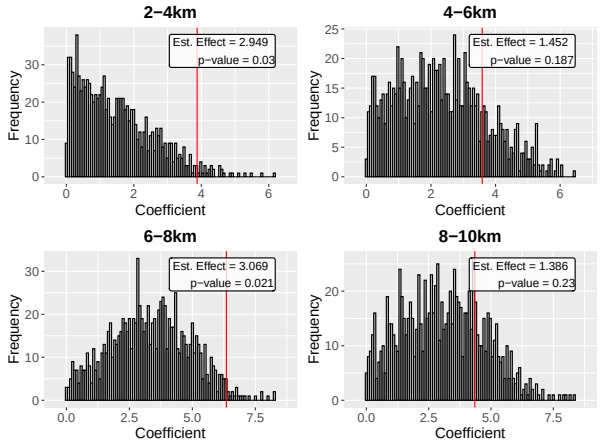
To illustrate this, consider a stylized example with only one criterion: proximity to roads. Given this criterion and that nearly all CNVs are located along main roads (Figure 1), the coefficients in Equation 1 implicitly compare outcomes of places that are progressively located further away from main roads to that of CNVs sites located right next to a main road. In contrast, the estimation of the coefficients in Equation 2 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 5 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 1 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 5 illustrates this procedure.

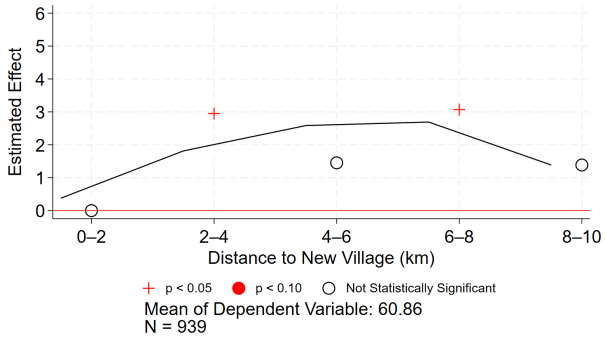
Balance Checks We check whether real and counterfactual CNV sites are balanced on key geographic, soil, and pre-treatment characteristics. If our counterfactual approach is valid, character-

[38] In addition, we further estimate economic outcomes at both the polling district and grid-cell level. The grid-cell level 'counterfactual' estimating equation is analogous.

Figure 5
Illustration of Methodology: Ethno-nationalistic Vote Share at the Federal Election in 2013



Panel A: Independent Shifts: Counterfactuals



Panel B: Independent Shifts: Plotted Coefficients

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. 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.

istics should be similar across both sets of locations.

We visualize this in Figure A.2, which plots the difference between real CNV distance coefficients and the average of 1,000 counterfactual coefficients across distance bins. Significance is denoted by crosses (above 95^{th} percentile), solid dots (above 90^{th}), and hollow dots (below 90^{th}). Variables include slope, elevation, soil quality indicators, and pre-treatment characteristics such as distance to urban centers, ethnic Chinese shares, and population density—all measured in 1947.

We do not find any statistically significant differences across a wide range of characteristics, except for elevation, slope, and 1947 ethnic Chinese shares. In any case, we continue to control for these three and all other pre-treatmant covariates in our regressions. Throughout all our analyses, we also continue to focus on exposure effects on surrounding Malay communities by conducting a "doughnut hole" analysis, whereby we exclude all polling districts that contain CNVs.

5 Interethnic Proximity and Vote Shares for the National Front

5.1 Main results

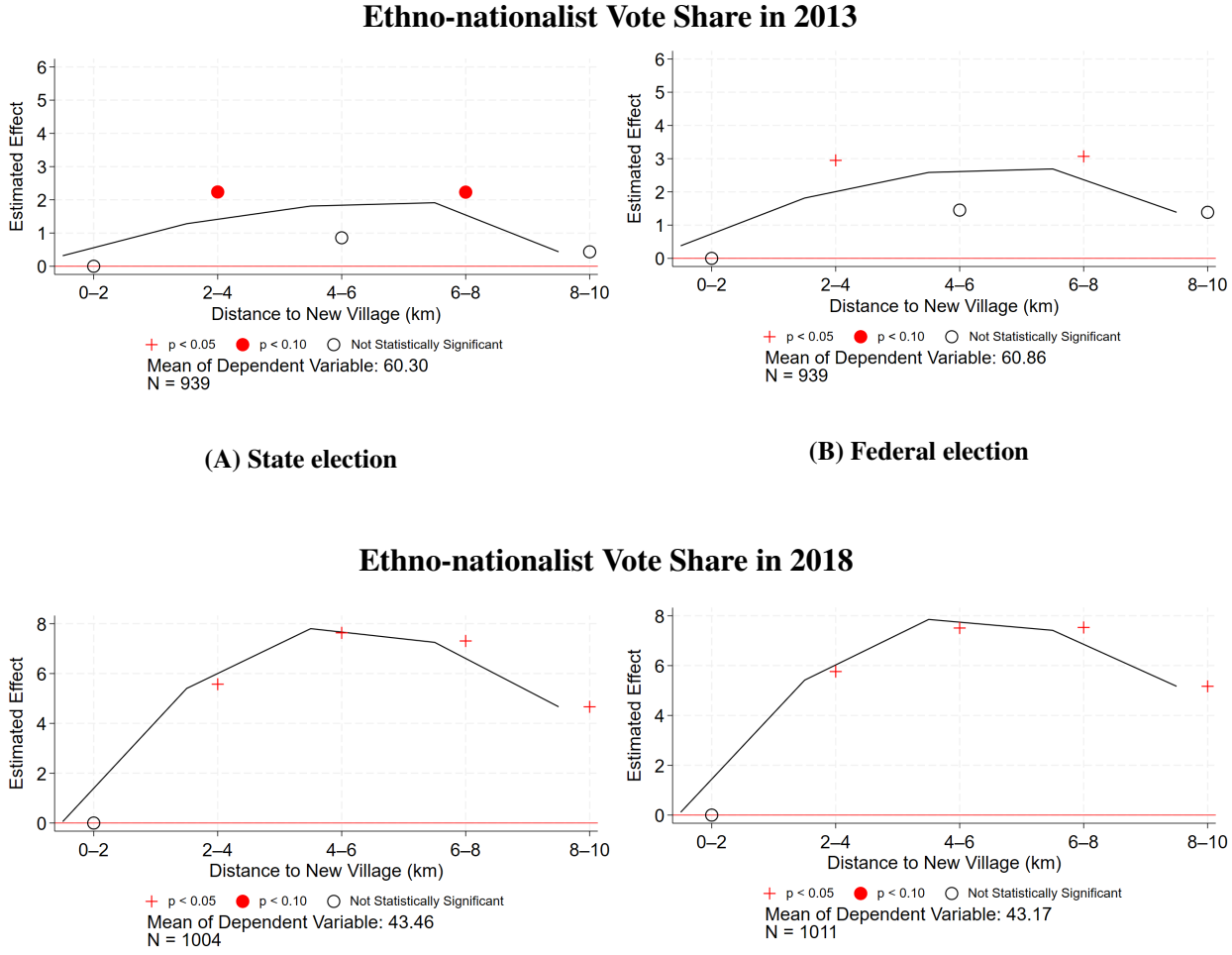
OLS analysis. We begin by examining OLS estimates in Table A.2, which reports the relationship between proximity to CNVs and vote shares for the ethno-nationalist coalition (National Front or BN) in the 2013 and 2018 elections. The regressions are run at the polling district level for both state and federal elections, using distance bins from CNVs as independent variables. The 0–2 km bin serves as the omitted category, allowing comparisons of vote shares at greater distances to those closest to CNVs. We find that in the 2013 election (Panel A, columns 3 and 6 with full set of controls), polling districts 2–10 km from CNVs had 3–6pp higher support for the ethnonationalist coalition than those within 0–2km, representing an 5–10% increase relative to the 60% mean. This implies significantly lower ethnonationalist sentiment in areas closest to Chinese resettlement sites. The 2018 election (Panel B) shows similar but much stronger spatial patterns, with a national drop in support (from 60% to 43%): vote shares in 2–10 km bins were still 7–12pp higher, or 16–28% relative to the mean.

Counterfactual analysis. We now turn to the counterfactual analysis explained in Section 4.2. This analysis is an important step to validate and refine the OLS estimates reported above, by accounting for potential biases arising from the possibly endogenous placement of CNVs. As outlined in Section 4.2, the counterfactual approach generates 1,000 sets of hypothetical resettlement sites following British military site-selection criteria. In Figure 6, we present the results for vote shares after applying the counterfactual adjustment. This figure plots the estimated coefficients on each distance bin, which are then fit with a linear spline.

Panel A shows results for the 2013 election, and Panel B shows results for the 2018 election. The results broadly align with the OLS findings in Table A.2, confirming that close proximity to CNVs is associated with a reduction in support for the ethno-nationalist coalition. However, the magnitude of the effects is consistently smaller than the OLS estimates, which is expected since the OLS results likely capture upward bias from unobservables, such as better market access, economic fundamentals, or pre-existing political attitudes in areas that are located further away from CNVs.

Specifically, in the 2013 election, polling districts 2–4km from CNVs show about 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. Notably, however, effects are smaller and statistically insignificant in the 4-6km bins and is plausibly a logical result of higher pre-resettlement exposure of Malay communities in the 4-6km bins to ethnic Chinese squatter settlements given resettlement

Figure 6
Effects of Chinese New Villages 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.

distances of 3.2 to 9.6km (Section 2.4).[39]

In 2018, magnitudes are stronger: relative to 0–2km, support is 5–8pp higher across the 2–10km bins, and all estimates are significant at the 5% level.[40] These findings reveal the localized

[39]We are in the process of analyzing primary survey data that we collected on the precise locations of pre-resettlement Chinese communities which were resettled into Chinese New Villages.

[40]There may also be potential bias based on counterfactual CNV locations as some fake CNV locations may be close enough in proximity to the corresponding actual CNVs. In particular, if some "control" polling districts in

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, social, or a combination of both—are strongest in areas immediately adjacent to CNVs. The fact that effects persist across two election cycles and strengthen by 2018 is consistent with slow-moving changes in Malay ethnopolitical identity that are accumulating through continued contact and cohort replacement rather than election-specific shocks or short-run campaign factors.

Discussion of persistent impacts. The long-term political effects are likely explained by the continued presence of CNVs after the Malayan Emergency. Their persistence stems from the post-1960 formalization of CNVs through the grant of land titles to resettled Chinese populations (Nyce, 1973). While the initial purpose of resettlement was to weaken communist support networks, land ownership was also used strategically to encourage settlers to establish permanent roots and loyalty to the state (Strauch, 1981). Over time, the Chinese population remained in these villages due to the strong social and economic foundations they built, as CNVs evolved into self-sufficient communities (Strauch, 1981).

The "inverted-U" shaped relationship. We observe an inverted-U relationship between proximity to CNVs and vote shares for the ethno-nationalist coalition in both the 2013 and 2018 elections (Figure 6), with the maximum being at the 2–4 km distance bins in 2013. Besides the possibility mentioned above that Malay communities in 4-6 km distance bins might have experienced higher pre-resettlement interethnic exposure, this "inverted-U" possibly suggests that ethnic-based political preferences are strongest at intermediate distances from CNVs.[41] We present three interpretations.

First, we argue that this could be due to "racial threat" being more salient in intermediate distances (Blalock, 1967), as such perceptions could be weaker closest to CNVs (0–2 km) (Allport, 1954) and in distant areas (8–10 km), where the outgroup's presence is minimal. A second interpretation draws from theories of ethnic conflict (Caselli and Coleman, 2013), which predict that conflict is most likely at intermediate levels of resource competition where the perceived benefits of exclusion outweigh the costs. If proximity to CNVs is associated with perceived competition over economic or social resources, then those living at intermediate distances may feel the

the fake regressions (i.e., polling districts within 2-10 km from a fake village) are located near the real CNVs, the estimated effects in the main specification above may be underestimating the true treatment effects. Therefore, while the reported estimates above are conservative as these might represent the lower bound, we also check that the results are robustness to running the counterfactual regressions in which we exclude polling districts that are 0-2 km from real CNVs for the fake regressions. These results are available upon request.

[41]The effects of pre-resettlement exposure in the 4-6km distance bins could have been attenuated during the 2018 elections given the broader wave of anti-incumbent sentiment that swept Malaysia and resulted in the opposition gaining control of the government for the first time in 60 years.

strongest need to support ethnic-based politics, because they see the Chinese as close enough to be a threat. The third interpretation relates to the distinction between interpersonal and contextual contact. At close distances, direct intergroup contact can promote positive attitudes through everyday interactions (Allport, 1954). However, at intermediate distances, Malays may be indirectly exposed to the Chinese community—through visible institutions, shared spaces, or reputational knowledge—without actual contact. This form of contextual contact (Stein et al., 2000) can raise ethnic awareness or perceptions of competition, especially in the absence of personal relationships. At further distances, this contextual exposure likely fades, reducing the relevance of ethnic identity in political choices. However, we are unable to empirically test the other interpretations due to data limitations.

Distaste for coercive policies or interethnic proximity? We interpret our voting result as a shift in political identity among the ethnic Malay majority driven by proximity to CNVs. A plausible alternative explanation could be that lower vote shares for BN may instead reflect a greater distaste for coercive policies, given that Malay communities living closer to CNVs might have had greater indirect exposure to policies targeted at the resettled. This explanation is unlikely for two reasons: (i) resettlement was carried out by the British government under colonial rule, while the ethnonationalist coalition was formed only after Malaysian independence. Since the coalition did not implement these policies, it is unlikely that vote shares reflect punishment for them. Second, in Section 7.3, we present results from a placebo test exploiting a nearly identical forced resettlement program targeting Malays (instead of Chinese) and creating Malay New Villages (MNVs), which shows no impact of MNVs on exposed Malays' political behavior. This further implies that our results are not driven by a generalized backlash against coercive resettlement policies.

5.2 Alternative Explanations and Robustness Checks

Our results suggest that Malays living closer to CNVs are less likely to vote for the ethno-nationalist coalition. We interpret this result as a shift in political preferences among the ethnic Malay majority driven by proximity to CNVs. In this subsection, we address four alternative explanations based on differential turnout; ethnic composition; selective migration; and contamination bias from any potential overlap between counterfactual and treated polling districts. Specifically, lower vote shares near CNVs could reflect higher turnout driven by racial threat or ethno-nationalist mobilization, rather than genuine changes in political attitudes (Enos, 2014). In addition, differences in vote shares might stem from variation in ethnic composition if more ethnic Chinese voters reside closer to CNVs. While our doughnut-hole design excludes polling districts containing initial CNVs, qualitative fieldwork indicates that Chinese residents often settle in adjacent areas after marriage. As

such, observed lower vote shares near CNVs could simply reflect higher ethnic Chinese shares rather than actual shifts in Malay preferences.

Turnout We examine the effects of distances to CNVs on voter turnout in the 2013 elections. In Panel A of Figure A.3, effects are statistically significant but quantitatively insignificant. Specifically, differences in turnout range from 0.5-0.8p.p. Given we observe differences in vote shares of 3p.p., our results are unlikely to be driven solely by differential turnout rates.

Addressing differences in ethnic composition. We construct the ethnic Chinese and ethnic Malay shares of registered voters across the entire Peninsular Malaysia using full-count voter rolls from the 2013 General Election. Voter rolls in 2013 record the self-identified ethnicity of every registered voter in Malaysia. Reassuringly, Panels B and C of Figure A.3 tests and shows that there are little differences in the ethnic composition of registered voters across all distance bins. This suggests that any differences in ethnic composition are unlikely to fully explain our observed results and lower ethnonationalist vote shares in less proximate distance bins are almost surely a reflection of changes in Malay political behavior.

Quantifying Malay voting behavior. To further quantify the role of ethnic Malay voters in driving the decreases in ethno-nationalist 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 (Becker and Woessmann, 2009; Calderon et al., 2023) to estimate the share of ethnic Chinese voters that would have had to turnout to vote, to explain away the observed lower vote shares for the National Front (*BN*). In so doing, we (unrealistically) assume that all Chinese voters voted *against* the ethno-nationalist coalition and show that the range of ethnic Chinese voter turnout rates would have to be implausibly high for Chinese voters alone to explain the entirety of negative ethno-nationalist vote shares.[42][43] In Section 7, we further present primary survey data on individual-level Malay political and ethnic attitudes to further support our argument.

To that end, we estimate Equation (1) via OLS using the outcome variable, $VotebyNonBN - ChiVoters_{d,p}$, which indicates 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

[42] 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.

[43] 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, 2017; Ostwald and Oliver, 2020).

under (i) the extreme assumption that all ethnic Chinese voted against the National Front and (ii) by varying ethnic Chinese voter turnout rates from 0.1 to 1.

Figure A.5 shows that, even under optimistic assumptions (i.e., all ethnic Chinese voters that voted, voted against *BN*), the observed vote share differences persist unless turnout exceeds 50–60%. Given that typical Chinese turnout rates rarely surpass 50% (Malay Mail, 2024), this suggests that changes in ethnic Chinese turnout nor composition are sufficient to explain the full decline in *BN* support. These findings imply that a substantial share of ethnic Malay voters living near CNVs likely voted against the ethno-nationalist coalition.

Selective migration. Next, we rule out selective migration as an alternative mechanism. Using primary survey data on the migration histories of Malays residing in villages near CNVs, collected through our primary survey on about 288 Malay village leaders and 1990 non-leaders, we examine whether there was selective in- or out-migration of Malays in response to the resettlement program.[44] Understanding selection and attrition is important, as systematic differences in who moved into or out of areas near CNVs following the resettlement program could alternatively explain observed differences in political behavior.

Figure A.12 shows that only around 1% of people have ever moved into, and very few have permanently left, Malay villages since 1960, across Malay villages located both closer to (treated) and further from (control) CNVs (Panels A and B).[45] Moreover, about 80% of respondents were themselves born in their village of residence (Panel C). These figures are similar across both treatment and control Malay villages (all $p > 0.10$). These results suggest minimal in- or out-migration of Malays in response to ethnic Chinese resettlement.[46]

6 Interethnic Proximity and Contemporary Local Development

In this section, we examine the effects of CNVs on contemporary local development, recognizing that economic and political development may be interlinked. In comparison to significant effects on political preferences, we find only modest impacts on local economic development and public goods provision.

[44] See Section 7.1 for more details about the primary survey we conducted.

[45] In our survey, we used 1960 as a cutoff date as pilot surveys suggested that respondents were more likely to recall events from the 1960 onwards as opposed to using 1949/1950s as a benchmark.

[46] Moreoever, given our primary interest in studying exposure effects on Malays, we are less concerned about selective in- and out-migration of Chinese villagers across time. Nonetheless, selection on the Chinese side is also likely minimal: the total ethnic Chinese population in Malaya in the 1940s was about 1.8 million, of which roughly one-third lived in rural areas (Sandhu, 1964). Since about 573,000 rural Chinese were forcibly resettled into CNVs within a short span of three years, it is most likely that nearly all rural ethnic Chinese were relocated, leaving little scope for selective inclusion.

6.1 Contemporary Economic Development and Urbanization

Contemporary economic development and urbanization around CNVs could have shifted political preferences by creating new economic opportunities. Agglomeration benefits resulting from increased population density and localized economic activity could increase productivity and generate economic spillovers (Duranton and Puga, 2004, 2020), creating shared economic interests across ethnic groups.[47] Here, we only discuss the results of the counterfactual exercise (Tables of all OLS estimates can be found in the Online Appendix).[48]

Polling-district level. The polling district-level analysis is useful because it allows us to directly compare economic and political outcomes, as the voting data is only available at the polling district level. Graphs A and B of Figure 7 show the results. Graph A shows the results using 2014 nightlight luminosity data. Communities located within 2-4 km of CNVs, compared to those within 0-2 km, show significantly lower economic activity: nightlight luminosity in the 2–4 km range is reduced by approximately 1.5 units ($p < 0.05$). These effects also persist across the 4–6km and 6–8km bins, which is similar to the political effect patterns we observed.

Graph B reports the effects on population density, a key driver of economic activity and urbanization. As is similar to the nightlight luminosity result, we find a significant dip in population density in 2-4km and further distant bins (all $p < 0.05$). These results suggest sizable impacts of CNVs on local economic development and urbanization.

Grid-cell level. As these data on economic outcomes are available at higher spatial resolution, Figure 7 also presents the results of the grid-cell level (1×1 km^2) analysis. Graph C shows that, compared to those within 0–2km, the nightlight luminosity communities located within 2–4 km of CNVs is lower by approximately 1 unit ($p < 0.05$), which is around 40% lower. These effects, of both similar and larger magnitudes, persist across 4–10km ranges (all $p < 0.05$). Graph D reports that areas 2–10km away from the 0–2 km radius have significantly lower population density, with approximately 200 fewer people on average, all statistically significant at the $p < 0.05$ level.

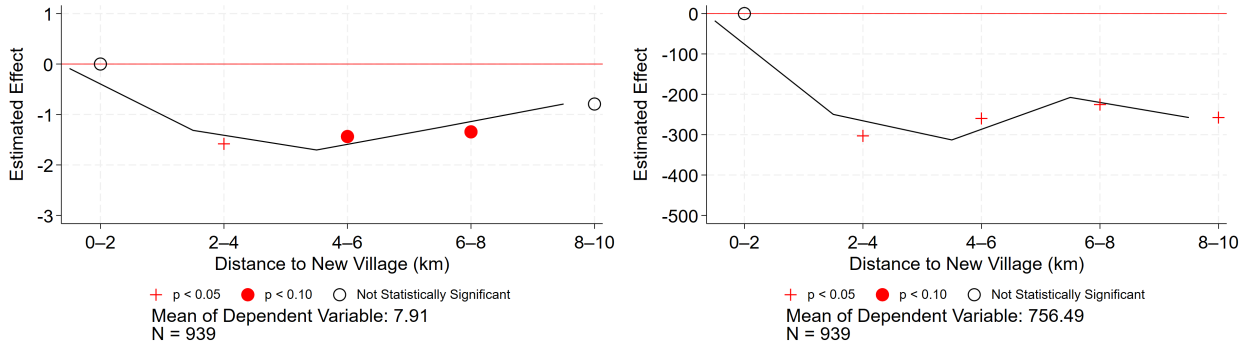
Scale effects. To further investigate the role and degree of the agglomeration externality, we examine heterogenous effects by below vs. above-median initial population size of CNV villagers

[47] While agglomeration has often been discussed in urban settings, our focus is the rural resettlement program. Agglomeration in rural areas is also plausible, even in the agricultural sector, as empirically illustrated by Tsuda et al. (2023). Hsu (2025) focuses on the economic effects in the same setting, finding agglomeration benefits for Chinese with limited positive spillovers to Malays.

[48] OLS results in Table A.3 present significant associations with nightlight luminosity, and population density in all distance bins and in both polling district-level and grid cell-level analyses. In the following counterfactual exercise, we focus on nightlight luminosity and population density. Results for NDVI are also consistent with the main findings and available upon request.

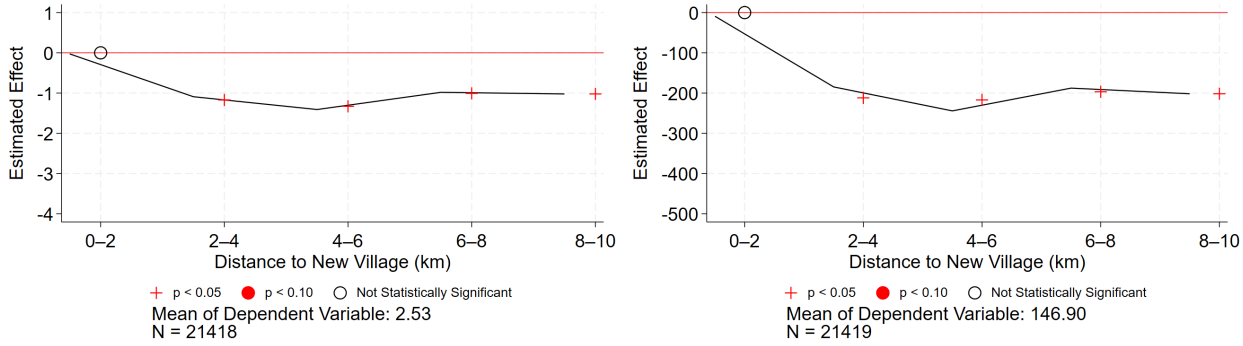
Figure 7
Economic Outcomes

Polling district-level



(A) Nightlight Luminosity **(B) Population Density**

Grid cell-level



(C) Nightlight Luminosity **(D) Population Density**

Notes: These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for nearest Chinese New Village 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.

in 1958, using data from Malayan Christian Union (1958). We present the results of the counterfactual analysis. At the polling-district level, Figure A.6 shows that the impacts on vote share and nighttime luminosity are slightly stronger and more significant in the above median sample, although these effects are quantitatively large but albeit, largely statistically insignificant in the below-median sample.[49]

[49]Similarly, in the grid-cell level analysis (Figure A.7), the effects on nighttime luminosity and population density appear slightly stronger in the above median sample, while remaining statistically significant in the below-median

Main takeaways. There are two takeaways from these results. First, the contrast between polling-district and grid-cell level results suggests that modest, highly localized agglomeration benefits, coupled with spatial reallocation, best explain the observed spatial patterns in the CNV exposure effects on economic outcomes. More broadly, positive economic effects at the polling-district level suggest that economic prosperity might be one possible reason for changes in political behavior. We further interpret these results as pointing to a more nuanced relationship between interethnic proximity and economic development that depends on how such development effects are distributed and experienced by Malays *within* polling districts.[50] Second, similar effects in both below- and above-median initial population samples suggest that the presence of CNVs consistently benefits nearby Malay populations, regardless of initial population size. The similar magnitude of the estimated effects further implies that mechanisms beyond localized agglomeration—such as interethnic complementarities—may also be driving the observed political effects. We explore this potential channel in Section 7.2.

6.2 Public Goods Provision

As another dimension of local development and plausible explanation for differences in political behavior, we explore the impacts of CNVs on contemporary public goods provision, focusing on schools and teachers, roads, and health facilities at the polling district level, offering a comparable analysis that aligns with the vote share changes discussed in Section 5. Public goods play an important role in improving access to essential services, neighborhood quality, and overall standard of living. In turn, political preferences could depend on the availability of public goods where voters reside (Calabrese et al., 2006). Hence, expanded access to public goods could promote support for parties favoring interethnic collaboration and inclusion. Again, we present the results of the counterfactual exercise below.[51]

Schools and teachers. Figure A.9 shows that the only marginal difference in the number of schools appears in the 2-4km bin (Graph A), while coefficients in all the other distance bins are insignificant. In terms of school quality, we find few differences as proxied by average student-

sample as well. Table A.4 also presents OLS results where we include an interaction term for whether a CNV was of above-median initial population. Across all distance bins and outcomes, we find little evidence of any heterogeneity by initial population size.

[50] Ideally, we would test this using grid-cell level data on vote shares, but grid-cell level data on vote shares is unavailable.

[51] We also present OLS results in Appendix. Table A.6 shows that proximity to CNVs is significantly associated with a greater number of secondary schools, a higher number of school teachers (though not statistically significant), and a slightly higher student-teacher ratio. Table A.7 indicates that road density decreases consistently with distance from CNVs, with all effects statistically significant, while there is no significant relationship between CNV proximity and the number of clinics or hospitals.

teacher ratio (Graph B). In theory, improved human capital through better access to education could have reshaped the political attitudes of Malays, encouraging more inclusive political behavior (Glaeser et al., 2007). However, these results suggest that increases in educational public goods are unlikely to explain differences in political behavior.[52]

Contemporary roads. Graph C of Figure A.9 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). This increased connectivity through more road infrastructure near CNVs may have influenced political preferences by lowering transport costs and improving market access. In this way, roads could have facilitated localized economic growth and integration, consistent with findings in Section 6.1.

Health facilities. Graphs D of Figure A.9 indicate no significant impacts on the presence of health clinics, except for a marginally statistically significant increase in the 8-10km bin. While not statistically significant—more broadly, health facilities appear more concentrated near CNVs, potentially reflecting a strategic placement/re-allocation to serve both Chinese and nearby Malay communities. However, as this data is from 2022, well after the elections studied, we cannot definitively assess the role of health infrastructure in shaping political preferences.

Taken together, we interpret few changes in public goods, save roads, as possibly reflecting connectivity-driven development without parallel growth in public good service provision or improvements in quality.

6.3 Political Targeting by the Civic-Nationalist Coalition

We conclude this section by showing that the observed local development effects in this section are unlikely to have been driven by policy changes enacted by the civic-nationalist coalition (i.e., the pro-diversity party). A plausible alternative mechanism is targeted engagement by civic-nationalist politicians who, if elected, might have steered public projects toward Malays near CNVs to build political support (Manacorda et al., 2011). However, Figure A.10 (and Tables A.9–A.10 for OLS) show no systematic gains in electoral victories or state-federal alignment for the civic-nationalist coalition near CNVs in either the 2013 or 2018 elections. Given that fiscal allocations in Malaysia are typically biased toward the ruling coalition (Ostwald, 2017), this pattern suggests that the observed modest economic gains (Sections 6.1 and 6.2) are less likely to be politically driven. Rather, it is plausible that it arose from proximity-induced economic complementarities, which then shaped political preferences. However, due to data limitations, we cannot rule out that lower

[52] Section 7 further provides survey evidence of few individual-level differences in the education of Malays residing close to and further from CNVs.

winning margins might have incentivized the ruling ethno-nationalist coalition to allocate resources to these areas to secure re-election (Bardhan and Mookherjee, 2010).

7 Potential Mechanisms for Political Outcomes

Our results show that Malays living in close proximity to Chinese are less likely to support ethnonationalist parties. To further understand the underlying mechanisms through which interethnic proximity continues to influence contemporary political behavior, we investigate three potential channels. First, we explore whether proximity led to meaningful intergroup contact, attitudinal changes, and economic interdependence between ethnic groups, using a novel, individual-level primary survey data focusing on attitudes and behaviors of Malays. Second, we examine the role of interethnic competition versus productive complementarities, using a historical survey dataset. Finally, we disentangle exposure effects to ethnic minorities from generic resettlement effects by leveraging a concomitant resettlement program that only resettled ethnic Malays.

7.1 Interethnic Contact—Attitudes, Behaviors, and Economic Outcomes

We begin with micro-level mechanisms, drawing on 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. In areas closer to CNVs, Malays were likely to experience more frequent and sustained social interactions with the Chinese minority, shaping their perceptions, trust, economic opportunities, and ultimately political identity.

Primary survey on micro-level contact and attitudes: Sampling & Variable Description. We conducted an in-person survey in partnership with one of Malaysia's leading survey firms, *Ilham*. We do so as there are no existing datasets that capture micro-level measures of interethnic contact or attitudes in this context. We collected novel, individual-level data from December 2024 to June 2025 in both 'treated' and 'control' villages (we define these in our empirical strategy below) across the states of Johor and Perak—the two states with the largest number of resettled ethnic Chinese.[53] Our survey targeted about 1,960 randomly selected Malay males, aged 18 and above from 75 Malay villages.[54] The sample was stratified by age to ensure representation of both older

[53] Specifically, we survey individuals from 3 Johor districts of Batu Pahat; Kluang; and Muar, and 3 Perak districts of Kuala Kangsar; Kinta; and Batang Padang. The data presented here is part of an ongoing data collection effort (about 3,000 respondents across two states) that will be completed in late 2025. The survey is funded by the Hong Kong Research Grants Council.

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

(60+) and younger (under 60) Malays. To minimize concerns around self-selection (discussed in Section 5.2), our analysis here focuses only on respondents who were born in the surveyed villages (about 1,369 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) *Direct contact/interactions*—frequency of contact with Chinese peers in villages, schools, and workplaces. (B) *Social capital*—presence of close Chinese friends, number of Chinese contacts in one's phone, 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*—self-reported monthly income and educational attainment, and enumerators' visual assessment of respondents' wealth. (F) *Zero-sum thinking*—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 situation*—whether the presence of Chinese affects Malays' wage and business profits, and whether both ethnic groups work in the same occupation.

Empirical strategy. There are three key steps in our empirical strategy. 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.[55] These four variables are our best available proxies for CNV site selection following the extant literature and archival documents that described the key role of the British military in carrying out the mass resettlement program (Section 2.4).[56] 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 0-2km of an actual (counterfactual) CNV resettlement site. Third, we collect individual-level data using in-person, door-to-door surveys in both treatment and control villages.

To estimate the effects of proximity on the six key dimensions listed above, we estimate the

[55] We perform matching at the 1 km $\times$ 1 km grid-cell level, the lowest-level of disaggregation for which we can measure these variables.

[56] 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.

following equation using OLS:

$$Y_{iv} = \alpha + \beta_1 Treat_{iv} + \theta_{enum} + \theta_c + \epsilon_{iv} \tag{3}$$

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 (0) if an enumerator was from Peninsular Malaysia (East Malaysia); θ_c denotes nearest CNV fixed effects. We cluster standard errors in two ways: by whether an individual is aged greater than 60 and the village in which an individual resides.[57]

Table 1
Attitudes and Behaviors toward Chinese

	Interactions				Social capital			Trust		Emotions	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	= 1 if ever visited CNV	= 1 if ≥ some Chinese pri school	= 1 if ≥ some Chinese sec school	= 1 if ≥ some Chinese colleagues	= 1 if ≥ 10% Chinese phone contacts	= 1 if have good Chinese friend	= 1 if ≥ somewhat willing Chinese marriage	= 1 if ≥ somewhat trust	= 1 if ≥ somewhat entrust kid	= 1 if sad if Chinese treated unfairly	= 1 if feel bad if Chinese have bigger houses
Treat	0.080***	0.011	-0.027	0.048***	0.017	-0.023	0.020	-0.022	0.030***	0.019	0.022**
	(0.026)	(0.007)	(0.018)	(0.017)	(0.028)	(0.017)	(0.018)	(0.021)	(0.010)	(0.022)	(0.010)
R^2	0.052	0.020	0.100	0.036	0.085	0.134	0.070	0.097	0.107	0.068	0.140
Mean Dep. Var. (Control)	0.161	0.007	0.135	0.097	0.214	0.158	0.699	0.317	0.134	0.739	0.056
Std. Dev. Var. (Control)	0.368	0.084	0.342	0.296	0.411	0.365	0.459	0.466	0.341	0.440	0.229
Observations	1335	1316	1084	1045	1354	1354	1357	1315	1357	1351	1352
Cluster	Cohort × 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 sample comprises two administrative districts in Johor. The dependent variables in columns 1 through 11 are indicator variables that take the value of 1 when a Malay: has ever visited the nearest Chinese New Village (col 1); had at least some (>= 10% to more than half) Chinese primary school classmates (col 2); had at least some (>= 10% to more than half) Chinese secondary school classmates (col 3); has at least some (>= 10% to more than half) Chinese colleagues (col 4); has at least 10% Chinese cell phone contacts (col 5); has at least one good friend who is Chinese (col 6); is somewhat willing or very willing to have a family member marry a Chinese (col 7); somewhat trusts or trusts a Chinese a lot (col 8); somewhat trusts or trusts a lot, a Chinese neighbor to take care of their child (col 9); feels bad when Chinese are treated unfairly (col 10); and feels bad if Chinese have bigger houses (col 11). All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results on interethnic contact. Table 1 reports our main survey results on Malays' interactions with, and attitudes and behaviors toward the Chinese. We find that Malays living near CNVs report significantly more interethnic interactions compared to those in control areas: in terms of visiting CNVs (column 1, $p < 0.01$) and in workplaces (column 4, $p < 0.01$). Interactions in primary schools are also higher but not statistically significant (column 2, $p > 0.10$). These results suggest that early-life and professional contact with the Chinese was more frequent among Malays in treatment areas.

Results on social capital and trust. Increased interactions observed do not, however, appear to translate into deeper social integration, as we find no significant differences in social capital measures (columns 5-7, $p > 0.10$). Similarly, while general trust toward the Chinese remains unchanged (column 8, $p > 0.10$), we observe a statistically significant improvement in situational

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

trust (column 9, $p < 0.01$). While both trust measures are attitudinal, situational trust possibly captures behavioral intention more directly than general trust, giving better insight into how Malays might behave toward Chinese in real-life situations. This pattern aligns with Allport (1954), suggesting that repeated contact can reduce bias and improve intergroup attitudes and behaviors.

Results on emotions. We also examine empathy-related outcomes. Malays living near CNVs are slightly more likely to feel sad when Chinese are treated unfairly, but this effect is statistically insignificant (column 10, $p > 0.10$). Interestingly, envy emerges as a salient emotional response: Malays are significantly more likely to report feeling jealous when Chinese people have larger homes (column 11, $p < 0.05$). Given the positive impacts on several attitudinal and behavioral dimensions, this result may reflect benign envy, which is a motivational response to upward comparison, rather than a negative reaction. Benign envy can encourage effort and economic aspiration when the higher status appears attainable through one's own means (Van de Ven et al., 2009). Alternatively, this might partially reflect persistence in the anecdotal perception that CNVs were awarded a larger number of development projects from their inception by the British Colonial government.[58]

Table 2
Individual-level Economic Outcomes

	Economic status			Zero-sum thinking			Business & Labor market		
	(1) = 1 if high income (≥ RM3,500)	(2) Wealth class (1=lower, 2=lower middle, 3=upper middle 4=top)	(3) Years of education	(4) Chinese take away Malay jobs	(5) Chinese take away business/trade opportunities of Malays	(6) If Chinese richer, other ethnic groups poorer	(7) = 1 if lower business profits without Chinese	(8) = 1 if Chinese in the same occupation	(9) = 1 if lower wage/income without Chinese
Treat	0.045**	0.053**	0.077	-0.060**	-0.059*	-0.018	0.049*	0.065**	-0.056***
	(0.018)	(0.025)	(0.155)	(0.025)	(0.033)	(0.038)	(0.028)	(0.026)	(0.018)
R^2	0.038	0.076	0.053	0.165	0.106	0.113	0.267	0.044	0.082
Mean Dep. Var. (Control)	0.134	1.608	10.066	2.766	2.833	2.760	0.104	0.206	0.140
Std. Dev. Var. (Control)	0.341	0.641	2.610	0.780	0.792	0.854	0.307	0.405	0.347
Observations	1237	1342	1293	1338	1321	1321	350	1036	1030
Cluster	Cohort × 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 sample comprises two administrative districts in Johor. The dependend variables are defined as follows: monthly income higher than or equal to RM3,001 - RM3,500 (col 1); the categorical wealth class based on enumerators' assessment (col 2); years of education (col 3); in this country, the Chinese frequently take away Malay jobs (Strongly disagree=1 to Strongly agree=4) (col 4); in this country, the Chinese frequently take away business/trade opportunities of Malays (Strongly disagree=1 to Strongly agree=4) (col 5); in this country, if one ethnic group becomes richer, other ethnic groups typically become poorer (col 6); =1 if answering to the question: if there were no Chinese villages nearby, how much do you think your profits would have changed? (col 7); if answering more than to the question in your area, excluding your colleagues, how many Chinese are typically in the same occupation as you? (col 8); =1 if answering to the question: if there were no Chinese working in the same occupation as you, in the region where your job is located, how much would your wage or income be changed? (col 9). All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results on economic status. Next, we report micro-level economic impacts in Table 2. Malays living near CNVs report significantly higher monthly income compared to those living farther away (column 1, $p < 0.05$).[59] This finding is corroborated by enumerator assessments of respondents'

[58]Primary survey data analysis, however, suggests that there are few, if any, differences in these perceptions across our treated and control villages. Results available upon request.

[59]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.

visible wealth, which is based on house size, quality, and plot size (column 2, $p < 0.05$). This also indicates that Malays closer to CNVs appear to be economically better off. On the other hand, we find no statistically significant difference in years of education (column 3), suggesting that these economic gains are unlikely to be driven by differences in human capital accumulation.

Results on zero-sum thinking. We also examine zero-sum thinking among Malays—the belief that one group's gains come at the expense of another (Foster, 1965). We find that Malays living in close proximity to Chinese communities are significantly less likely to agree with statements suggesting that Chinese people take away jobs (column 4, $p < 0.05$) or business and trade opportunities (column 5, $p < 0.10$) from Malays. Notably, however, we find no statistically significant difference in respondents' agreement with the statement that: "Chinese becoming rich means Malays becoming poor" (column 6), though the coefficient is also negative.[60] 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., 2025).

Results on contemporary economic complementarity. Finally, we explore whether Malays perceive their economic outcomes to be affected by the presence of Chinese, as a proxy for assessing perceived economic complementarity. Specifically, we ask whether Malays believe the absence of Chinese would affect their business profits or personal income, and whether they tend to work in the same occupations. We find that Malays living near CNVs are more likely to agree that business profits would decline without the presence of Chinese (column 7, $p < 0.10$), indicating perceived complementarity at the extensive margin, potentially through supply chains, partnerships, or market access. On the other hand, however, treated respondents are more likely to report being in the same occupations as the Chinese (column 8, $p < 0.05$), and to believe that their own wages or incomes would increase in the absence of Chinese workers (column 9, $p < 0.01$). This latter finding suggests that proximity increases economic overlap and there is perceived labor market competition at the intensive margin (i.e., within shared occupations). These results are not contradictory because Malays may see the Chinese as economically beneficial in broader business contexts, while also feeling competitive pressure in day-to-day wage work. Such perceptions may still shift political preferences away from ethnonationalist appeals, as the economic benefits of coexistence become more tangible and the costs of exclusion more salient (Bursztyn et al., 2024).

Furthermore, Panels (A) and (B) of Table 3 and Table 4 report the heterogeneous effects on attitudes and economic outcomes among individuals engaging in agricultural jobs and those engaging

[60] Based on our discussions with enumerators, one possibility is the possibly more abstract nature of the question "If one ethnic group (Chinese) becomes richer, other ethnic groups typically become poorer) that rural Malays with lower education might have been less likely to comprehend fully.

Table 3
Attitudes and Behaviors toward Chinese by Contemporary and Historical Occupations

(A) Among individuals engaging in agricultural jobs

	Interactions				Social capital			Trust		Emotions	
	(1) = 1 if ever visited CNV	(2) = 1 if ≥ some Chinese pri school	(3) = 1 if ≥ some Chinese sec school	(4) = 1 if ≥ some Chinese colleagues	(5) = 1 if ≥ 10% Chinese phone contacts	(6) = 1 if have good Chinese friend	(7) = 1 if ≥ somewhat willing Chinese marriage	(8) = 1 if ≥ somewhat trust	(9) = 1 if ≥ somewhat entrust kid	(10) = 1 if sad if Chinese treated unfairly	(11) = 1 if feel bad if Chinese have bigger houses
Treat	0.030	-0.003	-0.069	0.095**	0.016	-0.030	-0.002	-0.106**	-0.037	-0.107**	0.031
	(0.047)	(0.015)	(0.044)	(0.037)	(0.029)	(0.043)	(0.053)	(0.049)	(0.035)	(0.050)	(0.028)
R^2	0.131	0.072	0.211	0.128	0.213	0.093	0.140	0.145	0.125	0.126	0.250
Mean Dep. Var. (Control)	0.148	0.013	0.112	0.033	0.161	0.099	0.679	0.318	0.118	0.765	0.043
Std. Dev. Var. (Control)	0.356	0.114	0.317	0.180	0.369	0.299	0.468	0.467	0.324	0.425	0.204
Observations	352	330	221	262	353	354	355	348	354	353	353

(B) Among individuals engaging in non-agricultural jobs

	Interactions				Social capital			Trust		Emotions	
	(1) = 1 if ever visited CNV	(2) = 1 if ≥ some Chinese pri school	(3) = 1 if ≥ some Chinese sec school	(4) = 1 if ≥ some Chinese colleagues	(5) = 1 if ≥ 10% Chinese phone contacts	(6) = 1 if have good Chinese friend	(7) = 1 if ≥ somewhat willing Chinese marriage	(8) = 1 if ≥ somewhat trust	(9) = 1 if ≥ somewhat entrust kid	(10) = 1 if sad if Chinese treated unfairly	(11) = 1 if feel bad if Chinese have bigger houses
Treat	0.099***	0.019**	-0.029	0.028	0.039	-0.019	0.045*	0.023	0.046**	0.051*	0.026*
	(0.036)	(0.009)	(0.023)	(0.022)	(0.047)	(0.023)	(0.026)	(0.031)	(0.019)	(0.026)	(0.014)
R^2	0.054	0.041	0.133	0.056	0.075	0.156	0.097	0.094	0.119	0.118	0.128
Mean Dep. Var. (Control)	0.169	0.006	0.154	0.121	0.220	0.199	0.716	0.311	0.140	0.736	0.059
Std. Dev. Var. (Control)	0.375	0.077	0.362	0.327	0.415	0.400	0.452	0.464	0.348	0.441	0.235
Observations	824	828	706	778	840	839	841	816	842	839	839

(C) In villages where rubber & tin employment did not have the highest share in 1960

	Interactions				Social capital			Trust		Emotions	
	(1) = 1 if ever visited CNV	(2) = 1 if ≥ some Chinese pri school	(3) = 1 if ≥ some Chinese sec school	(4) = 1 if ≥ some Chinese colleagues	(5) = 1 if ≥ 10% Chinese phone contacts	(6) = 1 if have good Chinese friend	(7) = 1 if ≥ somewhat willing Chinese marriage	(8) = 1 if ≥ somewhat trust	(9) = 1 if ≥ somewhat entrust kid	(10) = 1 if sad if Chinese treated unfairly	(11) = 1 if feel bad if Chinese have bigger houses
Treat	0.025	0.000	0.022	0.056	-0.006	0.042	-0.103***	-0.147	0.081	0.221	-0.018
	(0.050)	(.)	(0.023)	(0.088)	(0.061)	(0.057)	(0.023)	(0.149)	(0.080)	(0.222)	(0.029)
R^2	0.125	0.082	0.187	0.172	0.117	0.162	0.155	0.134	0.153	0.161	0.348
Mean Dep. Var. (Control)	0.266	0.016	0.122	0.065	0.159	0.109	0.769	0.306	0.154	0.708	0.015
Std. Dev. Var. (Control)	0.445	0.126	0.331	0.250	0.368	0.315	0.425	0.465	0.364	0.458	0.124
Observations	167	163	118	117	166	167	168	161	168	167	167

(D) In villages where rubber & tin employment had the highest share in 1960

	Interactions				Social capital			Trust		Emotions	
	(1) = 1 if ever visited CNV	(2) = 1 if ≥ some Chinese pri school	(3) = 1 if ≥ some Chinese sec school	(4) = 1 if ≥ some Chinese colleagues	(5) = 1 if ≥ 10% Chinese phone contacts	(6) = 1 if have good Chinese friend	(7) = 1 if ≥ somewhat willing Chinese marriage	(8) = 1 if ≥ somewhat trust	(9) = 1 if ≥ somewhat entrust kid	(10) = 1 if sad if Chinese treated unfairly	(11) = 1 if feel bad if Chinese have bigger houses
Treat	0.100***	0.012	-0.047**	0.042	0.029	-0.037	0.027	-0.012	0.029**	-0.015	0.029**
	(0.029)	(0.009)	(0.018)	(0.026)	(0.045)	(0.029)	(0.024)	(0.024)	(0.014)	(0.029)	(0.013)
R^2	0.056	0.036	0.105	0.037	0.083	0.142	0.083	0.107	0.115	0.063	0.129
Mean Dep. Var. (Control)	0.143	0.007	0.149	0.099	0.226	0.159	0.702	0.296	0.117	0.739	0.067
Std. Dev. Var. (Control)	0.350	0.083	0.356	0.298	0.419	0.366	0.458	0.457	0.322	0.439	0.251
Observations	1001	976	814	795	1011	1010	1013	978	1012	1008	1009

Cluster	Cohort × 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 sample comprises two administrative districts in Johor. The dependent variables in columns 1 through 11 are indicator variables that take the value of 1 when a Malay: has ever visited the nearest Chinese New Village (col 1); had at least some (>= 10% to more than half) Chinese primary school classmates (col 2); had at least some (>= 10% to more than half) Chinese secondary school classmates (col 3); has at least some (>= 10% to more than half) Chinese colleagues (col 4); has at least 10% Chinese cell phone contacts (col 5); has at least one good friend who is Chinese (col 6); is somewhat willing or very willing to have a family member marry a Chinese (col 7); somewhat trusts or trusts a Chinese a lot (col 8); somewhat trusts or trusts a lot, a Chinese neighbor to take care of their child (col 9); feels bad when Chinese are treated unfairly (col 10); and feels bad if Chinese have bigger houses (col 11). All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4
Individual-level Economic Outcomes by Contemporary and Historical Occupations

(A) Among individuals engaging in agricultural jobs

	Economic status			Zero-sum thinking			Business & Labor market		
	(1) = 1 if high income (≥ RM3,500)	(2) Wealth class (1=lower, 2=lower middle, 3=upper middle 4=top)	(3) Years of education	(4) Chinese take away Malay jobs	(5) Chinese take away business/trade opportunities of Malays	(6) If Chinese richer, other ethnic groups poorer	(7) = 1 if lower business profits without Chinese	(8) = 1 if Chinese in the same occupation	(9) = 1 if lower wage/income without Chinese
Treat	0.003	-0.019	0.315	0.037	0.137*	0.065	0.080*	0.061	-0.045
	(0.031)	(0.068)	(0.293)	(0.066)	(0.080)	(0.103)	(0.042)	(0.054)	(0.033)
R^2	0.087	0.132	0.244	0.304	0.217	0.154	0.553	0.159	0.276
Mean Dep. Var. (Control)	0.081	1.553	8.486	2.658	2.675	2.753	0.136	0.168	0.120
Std. Dev. Var. (Control)	0.273	0.580	2.877	0.780	0.783	0.843	0.345	0.376	0.326
Observations	354	354	322	351	345	345	193	259	256

(B) Among individuals engaging in non-agricultural jobs

	Economic status			Zero-sum thinking			Business & Labor market		
	(1) = 1 if high income (≥ RM3,500)	(2) Wealth class (1=lower, 2=lower middle, 3=upper middle 4=top)	(3) Years of education	(4) Chinese take away Malay jobs	(5) Chinese take away business/trade opportunities of Malays	(6) If Chinese richer, other ethnic groups poorer	(7) = 1 if lower business profits without Chinese	(8) = 1 if Chinese in the same occupation	(9) = 1 if lower wage/income without Chinese
Treat	0.048*	0.059	-0.143	-0.104**	-0.186***	-0.057	0.046	0.070**	-0.063***
	(0.024)	(0.042)	(0.151)	(0.039)	(0.044)	(0.050)	(0.042)	(0.026)	(0.023)
R^2	0.047	0.090	0.064	0.153	0.099	0.150	0.297	0.061	0.066
Mean Dep. Var. (Control)	0.165	1.642	10.396	2.788	2.889	2.766	0.067	0.221	0.147
Std. Dev. Var. (Control)	0.372	0.661	2.254	0.758	0.769	0.849	0.252	0.416	0.355
Observations	839	840	823	827	821	821	148	771	768

(C) In villages where rubber & tin employment did not have the highest share in 1960

	Economic status			Zero-sum thinking			Business & Labor market		
	(1) = 1 if high income (≥ RM3,500)	(2) Wealth class (1=lower, 2=lower middle, 3=upper middle 4=top)	(3) Years of education	(4) Chinese take away Malay jobs	(5) Chinese take away business/trade opportunities of Malays	(6) If Chinese richer, other ethnic groups poorer	(7) = 1 if lower business profits without Chinese	(8) = 1 if Chinese in the same occupation	(9) = 1 if lower wage/income without Chinese
Treat	0.091	0.117***	1.016	-0.041	-0.262**	-0.062	-0.000	0.310*	0.112
	(0.081)	(0.038)	(0.715)	(0.093)	(0.121)	(0.162)	(0.000)	(0.177)	(0.128)
R^2	0.106	0.133	0.109	0.097	0.112	0.118	0.329	0.108	0.153
Mean Dep. Var. (Control)	0.220	1.662	9.797	2.769	2.813	2.891	0.000	0.239	0.196
Std. Dev. Var. (Control)	0.418	0.668	3.277	0.724	0.814	0.928	0.000	0.431	0.401
Observations	151	167	161	166	162	162	36	117	117

(D) In villages where rubber & tin employment had the highest share in 1960

	Economic status			Zero-sum thinking			Business & Labor market		
	(1) = 1 if high income (≥ RM3,500)	(2) Wealth class (1=lower, 2=lower middle, 3=upper middle 4=top)	(3) Years of education	(4) Chinese take away Malay jobs	(5) Chinese take away business/trade opportunities of Malays	(6) If Chinese richer, other ethnic groups poorer	(7) = 1 if lower business profits without Chinese	(8) = 1 if Chinese in the same occupation	(9) = 1 if lower wage/income without Chinese
Treat	0.031	0.066*	0.154	-0.105***	-0.116**	-0.100*	0.034	0.060*	-0.039**
	(0.023)	(0.036)	(0.196)	(0.034)	(0.047)	(0.054)	(0.048)	(0.035)	(0.017)
R^2	0.051	0.089	0.054	0.203	0.128	0.118	0.320	0.052	0.083
Mean Dep. Var. (Control)	0.129	1.585	10.146	2.775	2.842	2.774	0.129	0.202	0.127
Std. Dev. Var. (Control)	0.336	0.638	2.514	0.786	0.783	0.829	0.337	0.402	0.333
Observations	919	998	956	995	983	983	264	786	781

Cluster	Cohort × 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 sample comprises two administrative districts in Johor. The dependend variables are defined as follows: monthly income higher than or equal to RM3,001 - RM3,500 (col 1); the categorical wealth class based on enumerators' assessment (col 2); years of education (col 3); in this country, the Chinese frequently take away Malay jobs (Strongly disagree=1 to Strongly agree=4) (col 4); in this country, the Chinese frequently take away business/trade opportunities of Malays (Strongly disagree=1 to Strongly agree=4) (col 5); in this country, if one ethnic group becomes richer, other ethnic groups typically become poorer (col 6); =1 if answering to the question: if there were no Chinese villages nearby, how much do you think your profits would have changed? (col 7); if answering more than to the question in your area, excluding your colleagues, how many Chinese are typically in the same occupation as you? (col 8); =1 if answering to the question: if there were no Chinese working in the same occupation as you, in the region where your job is located, how much would your wage or income be changed? (col 9). All regressions include nearest Chinese New Village fixed effects and two-way clustered standard errors at the cohort and nearest new village-level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

in non-agricultural jobs, respectively. The motivation behind this exercise is that we presume that Malays engaging in non-agricultural jobs, due to its relatively more collectivistic nature of their profession, perceive economic complementarities with the Chinese than those engaging in agricultural jobs, due to its invidualistic nature.[61] Among Malays in the agricultural sector, the results are indeed more consistent with economic competition. Proximity in this context translates into significantly lower trust and empathy (columns 8 and 10 in Table 3, Panel A) and is linked to a heightened perception that Chinese businesses pose a competitive threat (columns 5 and 7 in Table 4, Panel A). Conversely, among Malays in non-agricultural jobs, the evidence points towards strong contemporary complementarities. Proximity leads to higher situational trust and empathy (columns 9 and 10 in Table 3, Panel B). Economically, these individuals report higher incomes and are significantly less likely to view the economic landscape through a zero-sum lens, perceiving less competition for jobs and business opportunities (columns 4 and 5 in Table 4, Panel B). These contrasting findings underscore that interethnic contact alone is insufficient; the presence of economic complementarities, as seen in the non-agricultural sector, is a crucial catalyst for translating proximity into improved social cohesion and shared economic sentiment.

7.2 Historical Economic Competition vs Complementarities

Interethnic proximity can promote either negative political attitudes through economic competition (Becker and Pascali, 2019; Horowitz, 2000) or more positive ones through complementarities and specialization, which could, in turn, improve 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 competition over scarce agricultural resources, such as land and water, led to greater economic strain for local Malays, we would expect higher support for ethno-nationalist parties in areas immediately surrounding agricultural CNVs relative to more distant areas. Conversely, non-agricultural CNVs may have generated economic complementarities given that these labor-intensive industries were more likely to employ both Malays and Chinese (Ross, 2014; Siew, 1953), potentially fostering positive, downstream economic interactions. In this case, we would expect lower ethno-nationalist vote shares near CNVs associated with these sectors, as economic benefits or intergroup contact improved interethnic relations and political attitudes.

To test these hypotheses, we use two different datasets. First, we leverage a rich but underutilized dataset from the Malayan Christian Union (1958), which records CNV-level population

[61] Relatedly, Leibbrandt et al. (2013) show that fishermen from individualistic societies are more competitive than those from collectivistic societies. We also check that indivuals' likelihood of engaging in agricultural or non-agricultural jobs is fairly balanced, and its result is available upon request. Note also that, for individuals who are not currently working, we are using the information on their latest jobs engaged.

and occupational data collected 6–8 years after resettlement, during a period of continued movement restrictions.[62] We test for heterogeneous exposure effects with respect to CNVs with below- and above-median employment shares of ethnic Chinese in (i) non-cash-crop, agricultural employment and (ii) rubber and tin employment. We focus on employment in rubber estates and tin mines as these were two of the largest non-farm economic sectors both for CNV villagers. Throughout, we focus on polling-district level economic outcomes to ensure direct comparability with voting results.[63] Second, we again draw on our primary survey data, in which village leaders were asked about the occupation with the largest employment share in 1960. Based on this information, we classify villages where rubber or tin accounted for the largest share in 1960 separately from all other villages.

Results. We present OLS results in Table A.5 to explore how interethnic competition versus complementarities shaped political preferences. Malay communities near agricultural CNVs show higher ethno-nationalist vote shares—consistent with competition—while those near CNVs with higher rubber and tin employment show lower vote shares, suggesting economic complementarities could have promoted political moderation. To further validate OLS results on interethnic complementarities, we conduct a counterfactual analysis using CNVs with below- and above-median initial employment in rubber and tin (Figure 8).

We find few differences in vote shares in the below-median sample (Panel A of Figure 8) but a negative and statistically significant impact on the vote shares of polling districts immediately adjacent to CNVs in the above median sample ($p < 0.05$) (Graph A of Figure 8). Specifically, we observe a 8pp higher vote share for the ethno-nationalist coalition in polling districts within the 2–4 km distance bin relative to those in the 0–2 km bin. Notably, the size of this effect is larger than both that of the pooled sample (5.2pp in Graph B of Figure 6) and the above-median initial CNV population size sample (5.8p.p. in Panel A of Figure A.6), suggesting a substantial role for *historical* interethnic complementarities in driving observed results on voting behavior.[64]

Figure A.8 presents results for the above- and below-median agricultural employment samples. Panel A of both figures provide three key takeaways. First, in contrast to the simple OLS

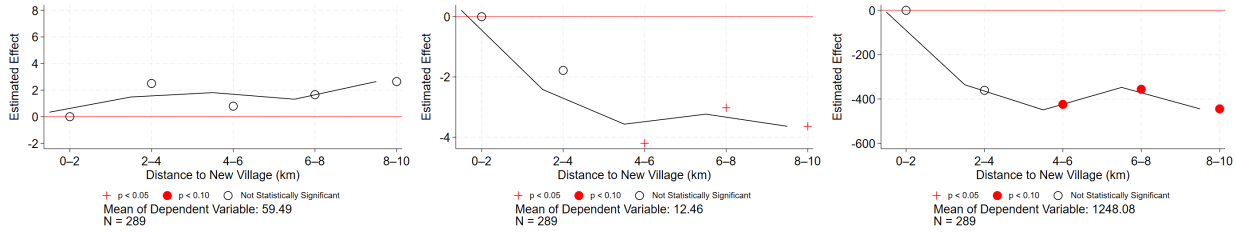
[62]Note that due to historical data limitations, we have about 43% missing observations and hence these results should be interpreted with caution. To that end, we are in the process of analyzing primary survey data based on the heterogeneity of (historical) occupational shares. These results, preliminarily, appear to point in the same direction.

[63]Grid-cell level results are largely consistent with polling district level findings and available upon request. We also note that the estimation sample in some of these analyses are smaller than our baseline sample, as a number of sparsely populated, rural CNVs were excluded from this historical survey. This, however, is unlikely to be a threat to our analysis given that our estimates look at the effect of CNVs on receiving areas that would have had to be populated in the first place to observe any effect.

[64]Specifically, in contrast, the economic development effects, measured by nightlight luminosity and population density, are, instead, more pronounced in the below-median initial rubber and tin employment sample (Panels B and C of Figure 8), while the contemporary development effects are insignificant in the above-median sample (Panels E and F of Figure 8).

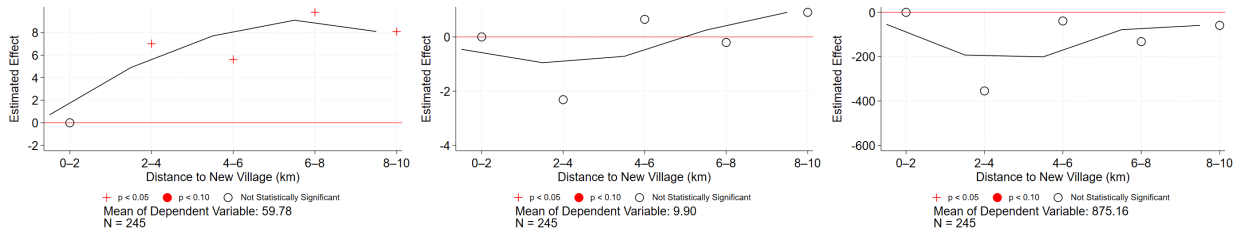
Figure 8
Polling district-level: Rubber and Tin Employment of Chinese in Chinese New Villages

Below Median Sample



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

Above Median Sample



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

Notes: There are 405 (43.1%) missing observations due to incomplete data on employment of Chinese in Chinese New Villages. 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.

results, effects on vote shares largely disappear in the above-median sample except for the 6-8km bin ($p < 0.10$) (Panel A of Figure A.8), suggesting that interethnic competition is unlikely to be the key mechanism driving earlier results on vote shares. Second, the contrast between OLS and counterfactual analyses results suggests that OLS is potentially upward biased due to greater selection bias for resettlement site locations of initially agricultural CNVs. This is plausible given the politically sensitive nature of resettling agricultural ethnic Chinese communities near agricultural Malay communities.

Third, and more importantly, we continue to observe large and statistically significant effects on vote shares in the immediate environs of CNVs with a below-median initial agricultural employment share (2-4 km and 6-8 km bins, $p < 0.05$) (Panel A of Figure A.8). As villages with lower agricultural employment shares typically have higher employment shares in the rubber, tin,

and other non-agricultural sectors, the concentration of significant effects in the below-median sample are consistent with the presence of significant interethnic complementarities observed in our above-median rubber and tin employment analyses.

Next, Panels (C) and (D) of Table 3 and Table 4 show that the analysis using our primary survey data also yields consistent results. In villages where historical rubber and tin employment share was the highest, Malays today exhibit more positive intergroup attitudes. Proximity to CNVs in these regions is linked to more frequent visits, higher situational trust, and more envy (columns 1, 9, and 11 in Table 3, Panel D). Moreover, it also leads to a robust and statistically significant rejection of zero-sum thinking: Malays are far less likely to believe that Chinese take away Malay jobs; Chinese take away business/trade opportunities of Malays; and if Chinese become richer, other ethnic groups become poorer (columns 4-6 in Table 4, Panel D). In contrast, in villages where historical rubber and tin employment was not high, proximity has failed to foster the same level of trust and is even associated with greater reluctance toward intermarriage (column 7 in Table 3, Panel C). At the same time, proximity to CNVs in these villages leads to higher wealth status of individuals than in villages where historical rubber and tin employment share was the highest. This contrast also aligns with the pattern that we observed in the above polling district-level analysis.

Taken together, we interpret these patterns as evidence that Malays living immediately near CNVs might have experienced initial economic benefits from employment in Chinese 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. We provided suggestive evidence for this in Section 7.1. Economic benefits, however, might have decreased over time given that ethnic Chinese in CNVs have largely transitioned to smallholder cash crop or commerce (Jomo, 2017) which employ fewer Malays and afford relatively fewer opportunities for interethnic interactions. Together, these results imply that sustained, localized interethnic exposure can yield lasting benefits for social cohesion through initial economic interdependence, even without persistent economic gains. This also suggests that contemporary economic and political development do not necessarily co-evolve in the same area.

7.3 Ethnic vs Resettlement Effects

The observed effects in Section 5 may not be driven by exposure to the ethnic Chinese community per se, but by exposure to new people through resettlement (i.e., generic resettlement effects), which could bring new ideas, opportunities, or norms that shape political preferences (Balietti et al., 2021; Chetty et al., 2016). To test this, we conduct a placebo-style analysis using Malay New Villages (MNVs). Like CNVs, MNVs were created during the Malayan Emergency and involved the resettlement of rural Malay populations into fortified villages (Dobby, 1952; Humphrey,

1971). The British colonial administration implemented this policy, also part of the Briggs Plan, in regions considered to be high-risk for Communist insurgency activities, where ethnic Malays were believed to be vulnerable to influence or violence from the Malayan Communist Party. Because the resettlement of Malays was at a much smaller scale, our sample includes 51 MNVs only. Note that MNVs were largely provided with the same amenities as CNVs.

Importantly, since MNVs were created under similar conditions and in comparable rural Malay-majority settings, they serve as a useful control/placebo group for isolating the ethnic-specific effects of CNVs. If the effects arose from exposure to new people rather than exposure to Chinese communities, we would expect effects of MNVs to be similar to those of CNVs.

In our counterfactual analysis (Figure A.11), differences in ethno-nationalist vote shares remain largely flat and statistically insignificant. Simply eyeballing and comparing coefficients across the same distance bins in CNVs (Figure 6) vs. MNVs (Figure A.11) shows that exposure effects to ethnic-minority Chinese communities were substantively and statistically different from generic resettlement effects. Exposure to individuals from the same, ethnic majority group appears to have had little, long-run effects on ethnic majority Malay political behavior.

8 Conclusion

Our study leverages a historical forced resettlement program to examine how persistent interethnic proximity shapes long-run political and economic development in Malaysia. Exploiting the sharp and durable variation created by the colonial relocation of nearly half a million rural ethnic Chinese into newly created villages, we examine how the spatial reconfiguration of ethnic geography continues to affect political preferences of the Malay majority, local economic development, and intergroup relations across Malaysia today.

More than seven decades later, Malays living closer to Chinese resettlement sites show lower support for the ethno-nationalist coalition that prioritizes the interests of the Malay majority over those of ethnic minorities such as the Chinese. We also find that areas closer to resettlement sites experienced stronger economic complementarities and deeper interethnic trust, while local economic gains were limited. Importantly, the shift in political preferences we observe did not rely on formal integration, but rather emerged from routine exposure, shared institutions, and possibly hyper-local economic interdependence.

Of course, not all contexts lend themselves to these interethnic relationships. The conditions for successful proximity, like localized economic complementarities, institutional neutrality, and the opportunity for repeated low-stakes interaction, may be absent in highly segregated urban ethnic enclaves.

As forced displacement continues to rise globally, from climate shocks, civil conflict, and

geopolitical instability, these findings carry renewed relevance. Understanding how spatial integration, even under adverse beginnings, can recalibrate identity and political behavior is important for policies aimed at promoting social cohesion. Clearly, more research is needed in this space to unpack more complex social, psychological, and institutional mechanisms at play. More comparative work across different historical and contemporary resettlement contexts can help clarify when, how, and for whom such proximity promotes integration versus exacerbates division. Investigating the role of local governance, civic institutions, and exposure across different stages of an individual's life cycle may also shed light on the durability and replicability of these findings.

References

Abel, M. (2019). Long-run effects of forced resettlement: evidence from apartheid south africa. *The Journal of Economic History*, 79(4):915–953.

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

Alesina, A. and Tabellini, M. (2024). The political effects of immigration: Culture or economics? *Journal of Economic Literature*, 62(1):5–46.

Alesina, A. and Zhuravskaya, E. (2011). Segregation and the quality of government in a cross section of countries. *American Economic Review*, 101(5):1872–1911.

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.

Alix-Garcia, J., Walker, S., Bartlett, A., Onder, H., and Sanghi, A. (2018). Do refugee camps help or hurt hosts? the case of kakuma, kenya. *Journal of Development Economics*, 130:66–83.

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

Arellano-Bover, J. (2022). Displacement, diversity, and mobility: Career impacts of japanese american internment. *The Journal of Economic History*, 82(1):126–174.

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.

Bardhan, P. and Mookherjee, D. (2010). Determinants of redistributive politics: An empirical analysis of land reforms in west bengal, india. *American Economic Review*, 100(4):1572–1600.

Barone, G., D'Ignazio, A., De Blasio, G., and Naticchioni, P. (2016). Mr. rossi, mr. hu and politics. the role of immigration in shaping natives' voting behavior. *Journal of Public Economics*, 136:1–13.

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., Grosfeld, I., Grosjean, P., Voigtländer, N., and Zhuravskaya, E. (2020). Forced migration and human capital: Evidence from post-wwii population transfers. *American Economic Review*, 110(5):1430–1463.

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.

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.

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.

Blalock, H. M. (1967). *Toward a theory of minority-group relations*. Wiley.

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.

Campbell, A., Converse, P. E., Miller, W. E., and Stokes, D. E. (1960). *The American Voter*. University of Chicago Press.

Carrillo, B., Charris, C., and Iglesias, W. (2023). Moved to poverty? a legacy of the apartheid experiment in south africa. *American Economic Journal: Economic Policy*, 15(4):183–221.

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.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022). Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921):108–121.

Chinoy, S., Nunn, N., Sequeira, S., and Stantcheva, S. (2025). Zero-sum thinking and the roots of us political differences. *American Economic Review (forthcoming)*.

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.

Desmet, K., Ortuño-Ortín, I., and Wacziarg, R. (2017). Culture, ethnicity, and diversity. *American Economic Review*, 107(9):2479–2513.

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

Diaz-Cayeros, A. and Jha, S. (2022). Conquered but not vanquished: Complementarities and indigenous entrepreneurs in the shadow of violence. *Manuscript. Stanford University. At https://web. stanford. edu/˜ saumitra/pa pers/DiazJhaCochineal. pdf, accessed June*, 12:2021.

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.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2):135–150.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of Regional and Urban Economics*, volume 4, pages 2063–2117. Elsevier.

Duranton, G. and Puga, D. (2020). The economics of urban density. *Journal of economic perspectives*, 34(3):3–26.

Dustmann, C., Vasiljeva, K., and Piil Damm, A. (2019). Refugee migration and electoral outcomes. *The Review of Economic Studies*, 86(5):2035–2091.

Englebert, P. (2018). *Burkina faso: Unsteady statehood in west africa*. Routledge.

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.

Gerber, A. S., Karlan, D., and Bergan, D. (2009). Does the media matter? a field experiment measuring the effect of newspapers on voting behavior and political opinions. *American Economic Journal: Applied Economics*, 1(2):35–52.

Gibson, J. (2021). Better night lights data, for longer. *Oxford Bulletin of Economics and Statistics*, 83(3):770–791.

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.

Greene, S. (2004). Social identity theory and party identification. *Social science quarterly*, 85(1):136–153.

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

Halla, M., Wagner, A. F., and Zweimüller, J. (2017). Immigration and voting for the far right. *Journal of the European Economic Association*, 15(6):1341–1385.

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

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

Hsu, S. H.-M. (2025). Coercive growth: Forced resettlement and ethnicity-based agglomeration. *Available at SSRN 5243409*.

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

Jackson, M. O. (2008). *Social and economic networks*, volume 3. Princeton university press Princeton.

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.

Jennings, M. K. and Niemi, R. G. (1974). *Political character of adolescence: The influence of families and schools*. Princeton University Press.

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

Jha, S. (2025). Breaking through the ethnic growth trap. *Economic Policy (forthcoming)*.

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

Kernell, G. and Lamberson, P. (2023). Social networks and voter turnout. *Royal Society Open Science*, 10(10):230704.

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.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kramer, G. H. (1971). Short-term fluctuations in us voting behavior, 1896–1964. *American Political Science Review*, 65(1):131–143.

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

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

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

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

Leibbrandt, A., Gneezy, U., and List, J. A. (2013). Rise and fall of competitiveness in individualistic and collectivistic societies. *Proceedings of the National Academy of Sciences*, 110(23):9305–9308.

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

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)*.

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.

Manacorda, M., Miguel, E., and Vigorito, A. (2011). Government transfers and political support. *American Economic Journal: Applied Economics*, 3(3):1–28.

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

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.

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.

Otto, A. H. and Steinhardt, M. F. (2014). Immigration and election outcomes—evidence from city districts in hamburg. *Regional Science and Urban Economics*, 45:67–79.

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

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.

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 1: Inequality. *World Development*, 134:105040.

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

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

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

Siddique, A., Vlassopoulos, M., and Zenou, Y. (2024). Leveraging edutainment and social networks to foster interethnic harmony. *CEPR Working Paper*.

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

Stein, R. M., Post, S. S., and Rinden, A. L. (2000). Reconciling context and contact effects on racial attitudes. *Political Research Quarterly*, 53(2):285–303.

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

Stigler, G. J. (1973). General economic conditions and national elections. *American Economic Review*, 63(2):160–167.

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

Suryadinata, L. (1987). Ethnic chinese in southeast asia: Problems and prospects. *Journal of International Affairs*, pages 135–151.

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.

Ting, C. P. (1976). *The Chinese in Peninsular Malaysia: A study of race relations in a plural society*. PhD thesis, University of Warwick.

Toews, G. and Vézina, P.-L. (2025). Enemies of the people. *American Economic Journal: Macroeconomics*, 17(1):310–342.

Tsuda, S. (2022). Refugee inflows, surplus farm labor, and crop marketization in rural africa. *Journal of Development Economics*, 155:102805.

Tsuda, S., Takasaki, Y., and Tanaka, M. (2023). Human and nature: economies of density and conservation in the amazon rainforest. *Available at SSRN 4652170*.

Van de Ven, N., Zeelenberg, M., and Pieters, R. (2009). Leveling up and down: the experiences of benign and malicious envy. *Emotion*, 9(3):419.

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

Wren-Lewis, L., Steinhardt, M. F., Merlino, L. P., and Rosenbaum, P. (2024). The impact of childhood inter-ethnic contact on managers' hiring decisions. *Available at SSRN*.

ONLINE APPENDIX

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

A 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.[65] Second, CNVs were on high ground to improve defensibility from the communists. The resettled often lived under the surveillance regime during the Malayan Emergency.[66] According to the site selection criteria, a New Village should possess basic amenities and sufficient agricultural land, however, the rapid strategic demands meant that criteria that did not concern the military objectives were not complied with in practice (Phee, 2012).

[65] 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.

[66] 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.

B Data Appendix

In this section, we provide further description of key variables, data sources and the detailed steps we took to construct them.

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 identify and geolocate a total of 452 New Villages.

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

The 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 ethnic share by dividing the number of registered Chinese or Malay voters over 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 construct measures of local economic activity using 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 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.

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.

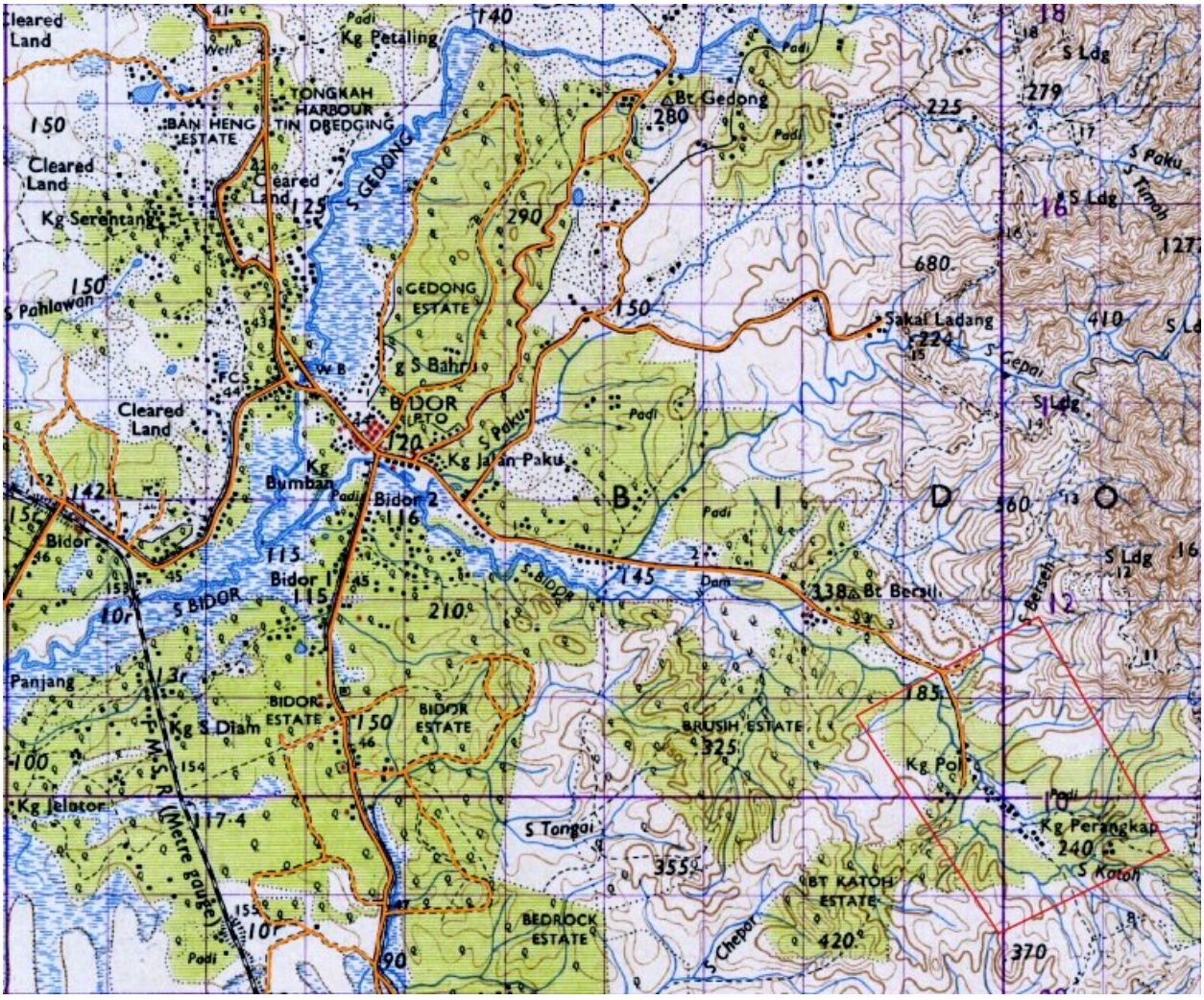
Elevation, slope, aspect. Topographical variables were created using raster data from the *Harmonized World Soil Database* (HWSD). The raster files are compiled from high-resolution source data and aggregated to 30-arc-second grids. We compute elevation for each polling district as the average elevation 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 sub-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.

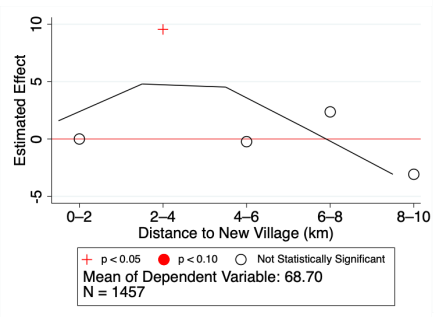
Appendix Figures and Tables

Figure A.1
Digitizing Historical Roads and Malay Reservation Polygons:
Extract from HIND1035 1947 Map (State of Perak)

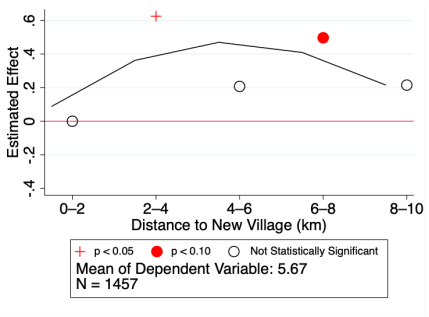


Notes: The orange lines indicate main roads and the red box indicates an example of a Malay Reservation Area.
Source: HIND 1035, Sheet 2N/14

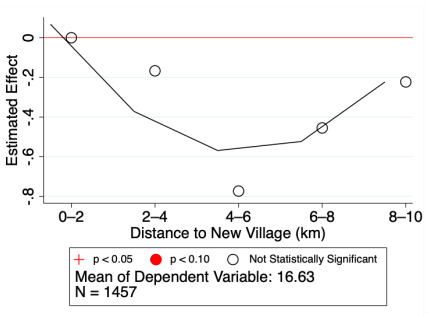
Figure A.2
Geographic, Soil, and Pre-Resettlement Balance



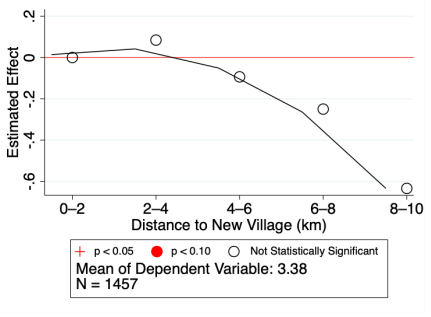
(A) Elevation



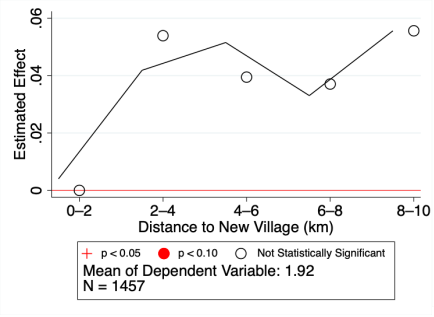
(B) Slope



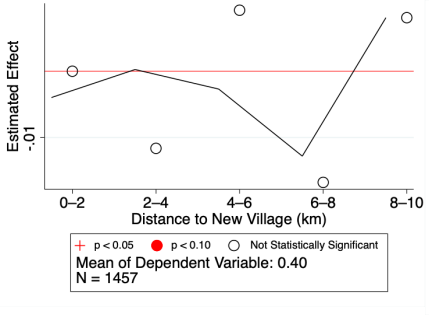
(C) % of East-facing Grids



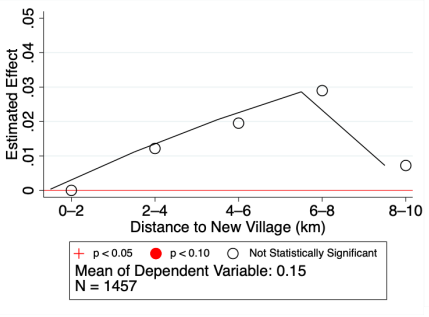
(D) % of Topsoil Organic Carbon



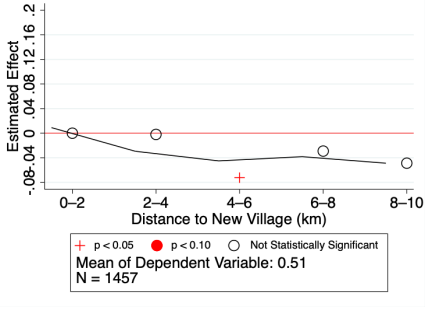
(E) % of Topsoil Sodicity



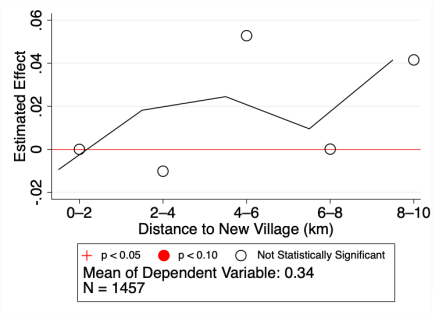
(F) Drainage is very poor (=1)



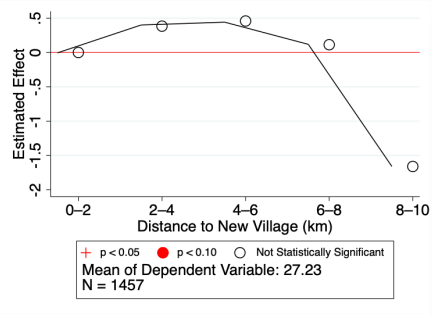
(G) Soil is Coarse (=1)



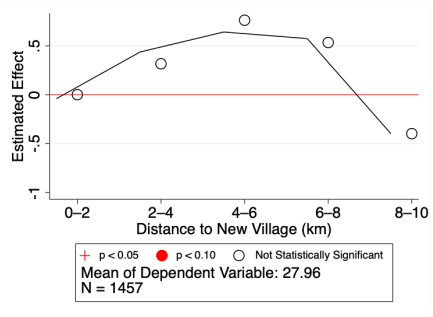
(H) Soil is Medium (=1)



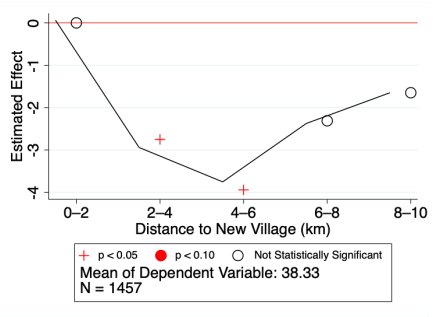
(I) Soil is Fine (=1)



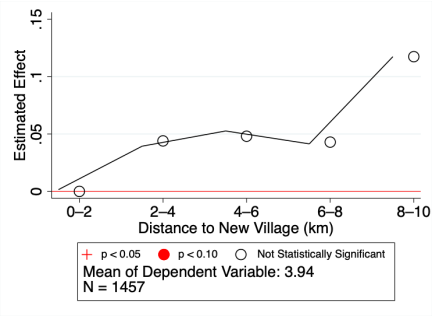
(J) Distance to urban centre in 1947 (km)



(K) Distance to coast (km)



(L) Share of Chinese in 1947

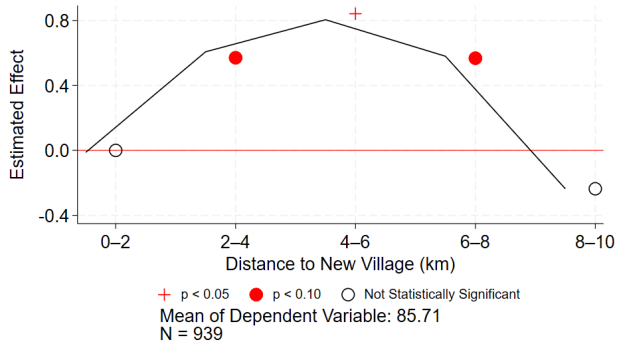


(M) Log Population Density in 1947

Notes: 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.

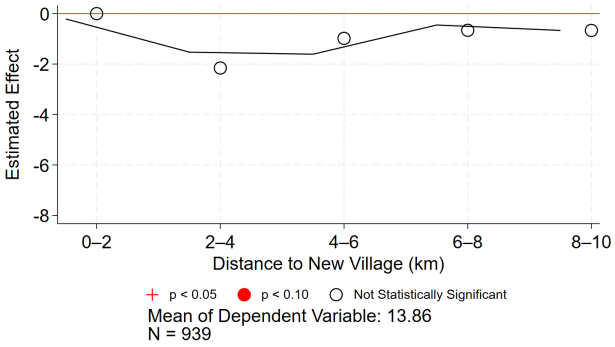
Figure A.3
Effects of Proximity to Chinese New Villages on Voter Turnout and Ethnic Composition

Voter Turnout in 2013

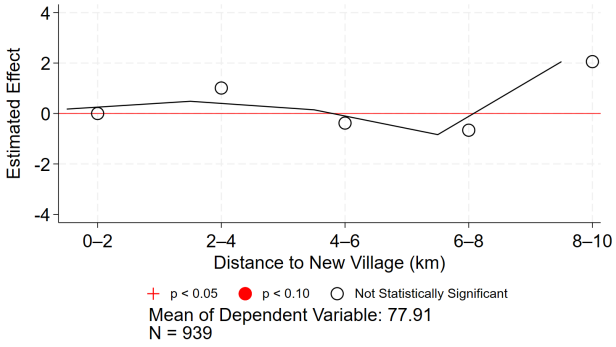


(A) Turnout in 2013

Ethnic Composition in 2013



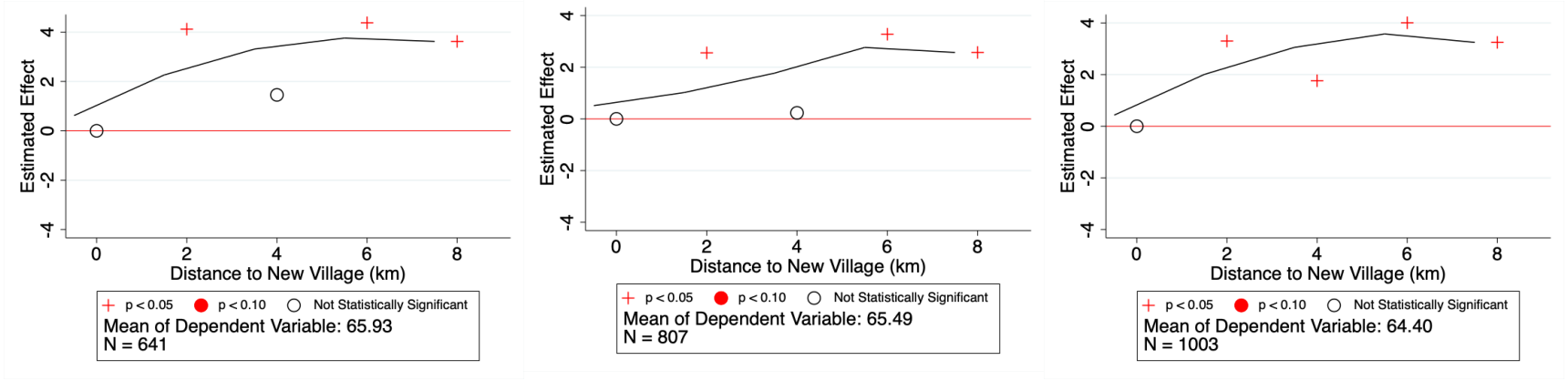
(B) Share of Registered Ethnic Chinese Voters in 2013



(C) Share of Registered Ethnic Malay Voters in 2013

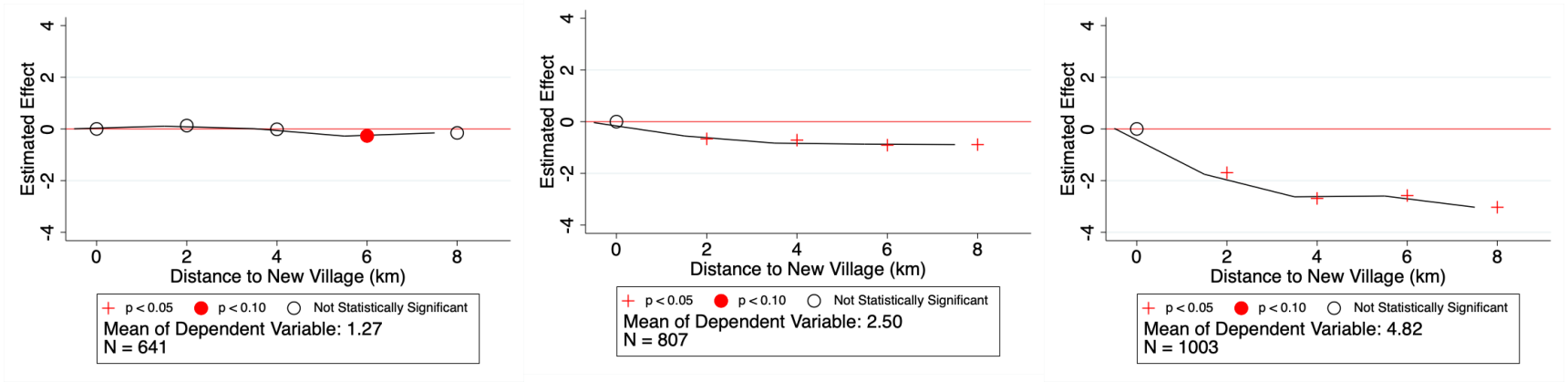
Notes: Figures plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects, geographical and pre-treatment controls. The means of analogous estimates computed from 1,000 counterfactual 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 A.4
Figure: Effects of Chinese New Villages: Samples with varying Chinese shares
Dep. Var.: Ethno-nationalist Vote Share at the Federal Election in 2013



(A) Sample: Chinese Share <= 5% **(B) Sample: Chinese Share <= 10%** **(C) Sample: Chinese Share <= 20%**

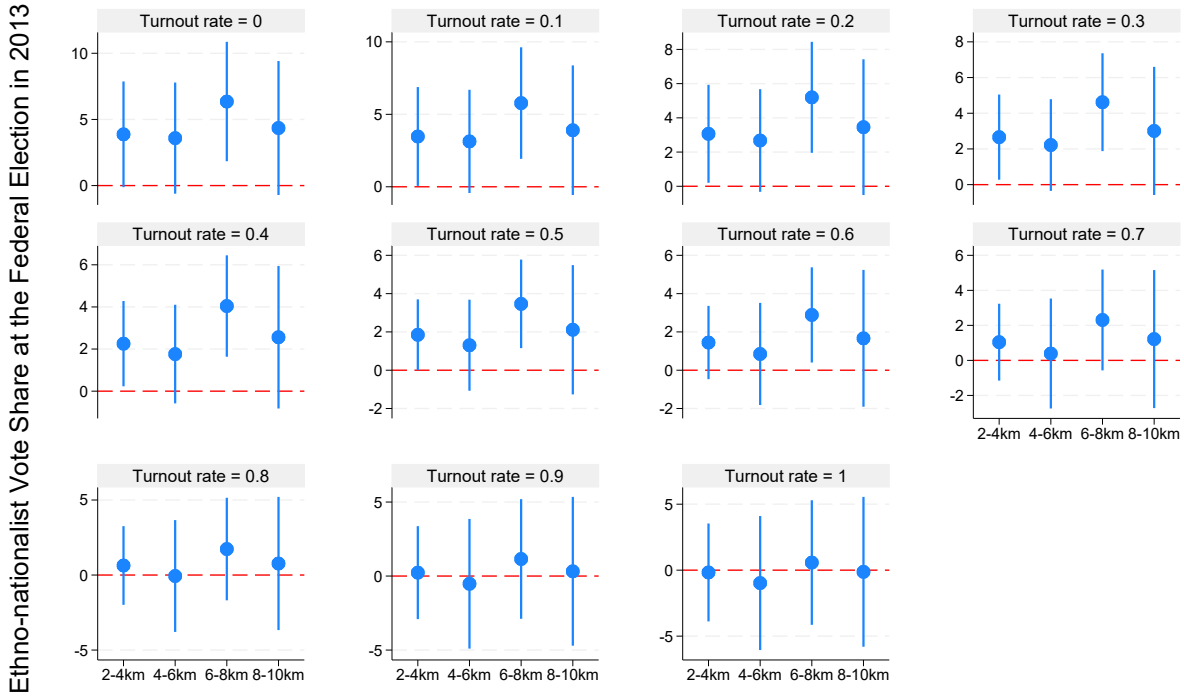
Dep. Var.: Share of Registered Ethnic Chinese Voters in 2013



(D) Sample: Chinese Share <= 5% **(E) Sample: Chinese Share <= 10%** **(F) Sample: Chinese Share <= 20%**

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. 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.

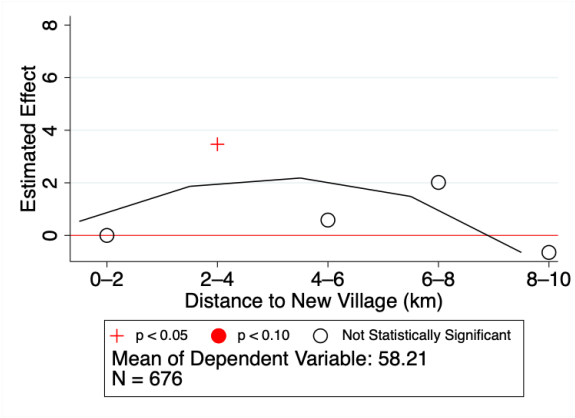
Figure A.5
Quantifying Malay Voting Behavior: The Effects of Varying Ethnic Chinese Voter Turnout Rates on Ethno-Nationalist Vote Shares: Polling District Level



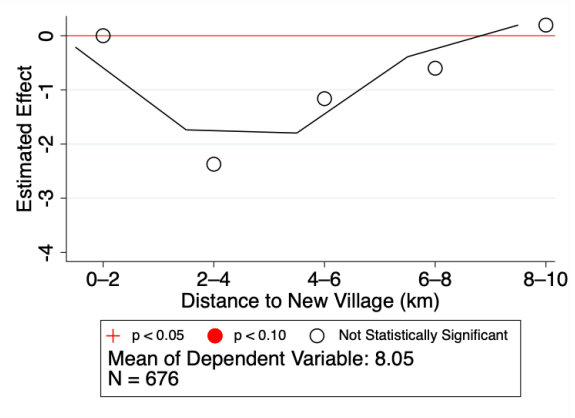
Notes: This figure plots OLS estimates of Equation 1. The outcome variable is vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New 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 A.6
Polling district-level: Initial Population Size of Chinese New Villages

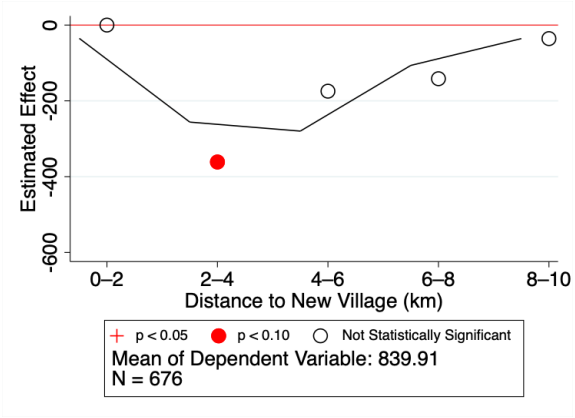
Below Median Sample



(A) Vote Share

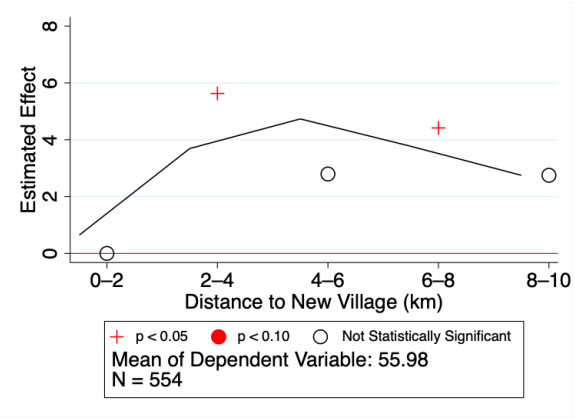


(B) Mean Luminosity

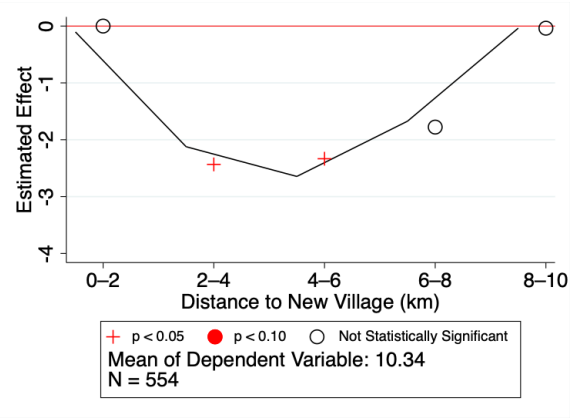


(C) Population Density

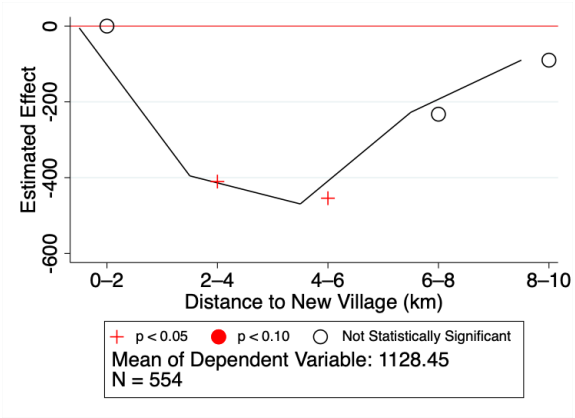
Above Median Sample



(D) Vote Share



(E) Mean Luminosity

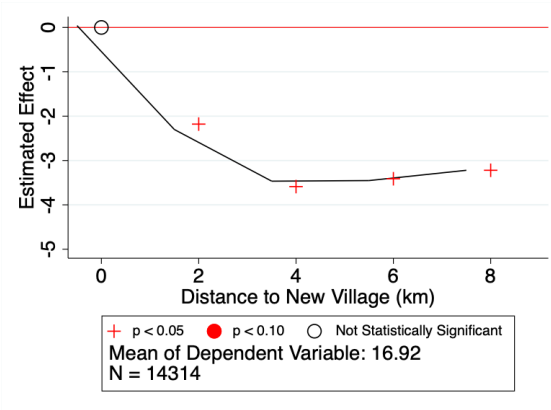


(F) Population Density

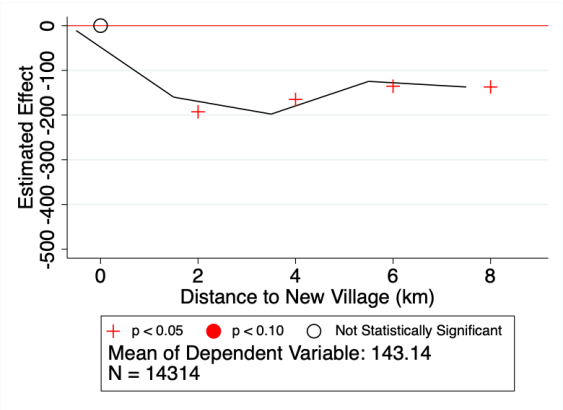
Notes: The outcome variables for (A) and (D) are the vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for 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.

Figure A.7
Grid cell-level: Initial Population Size of Chinese New Villages

Below Median Sample

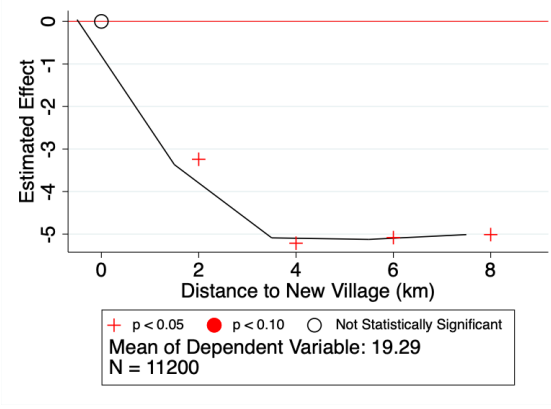


(A) Mean Luminosity

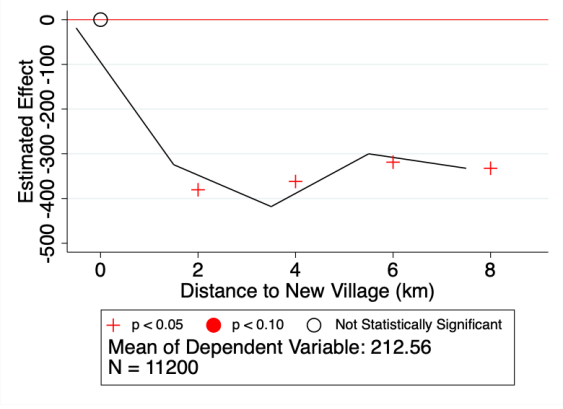


(B) Population Density

Above Median Sample



(C) Mean Luminosity

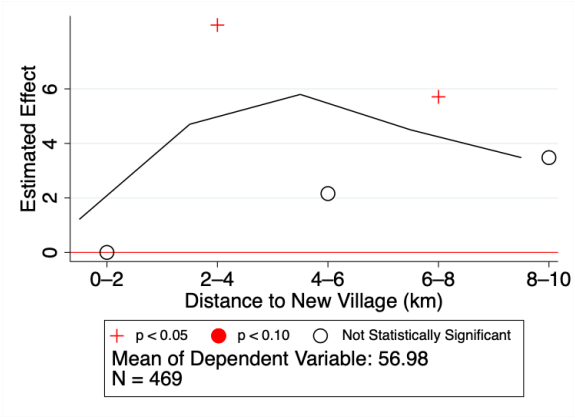


(D) Population Density

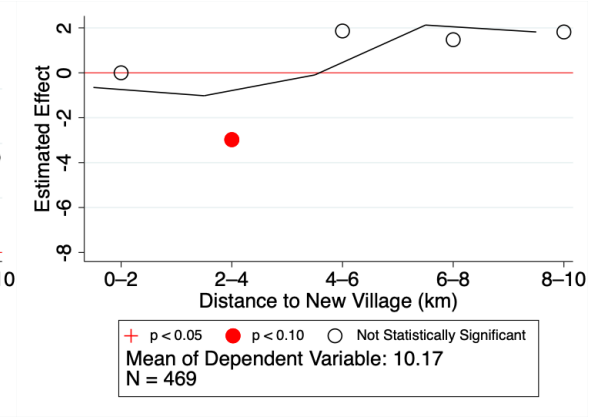
Notes: These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for nearest Chinese New Village fixed effects and geographical and pre-treatment controls. The means of analogous estimates computed from 1,000 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.

Figure A.8
Polling district-level: Agricultural Employment of Chinese in Chinese New Villages

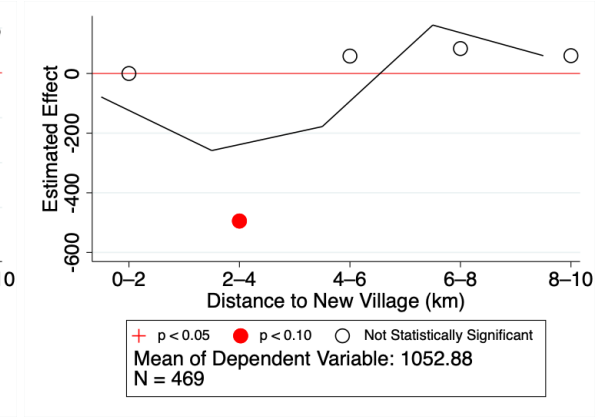
Below Median Sample



(A) Vote Share

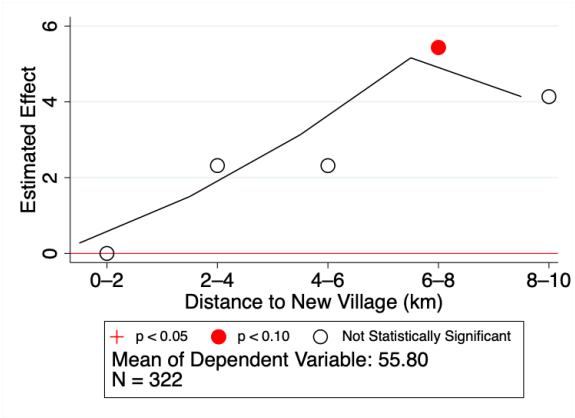


(B) Mean Luminosity

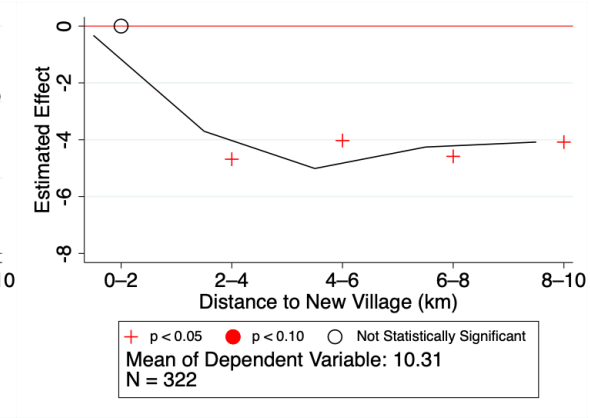


(C) Population Density

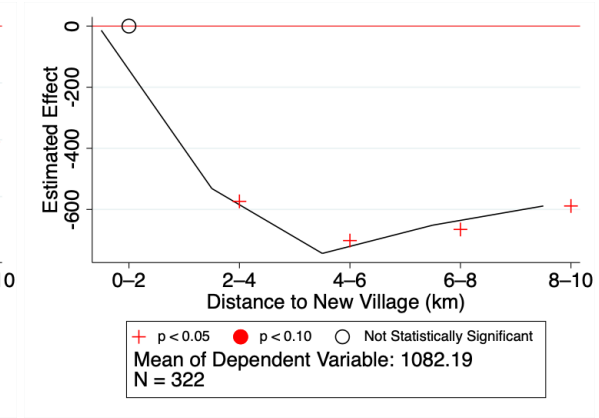
Above Median Sample



(D) Vote Share



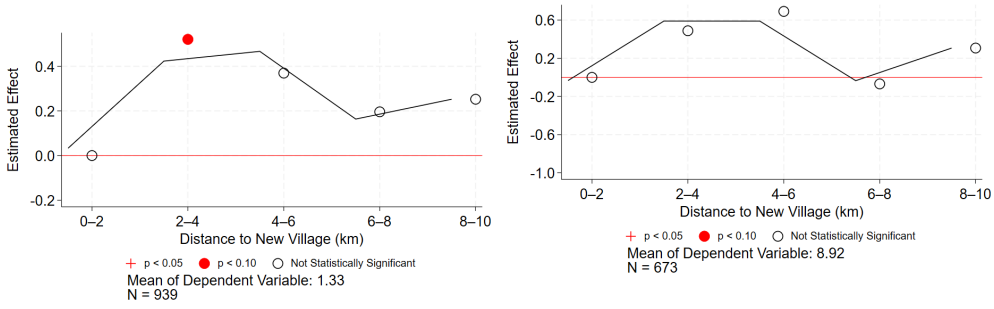
(E) Mean Luminosity



(F) Population Density

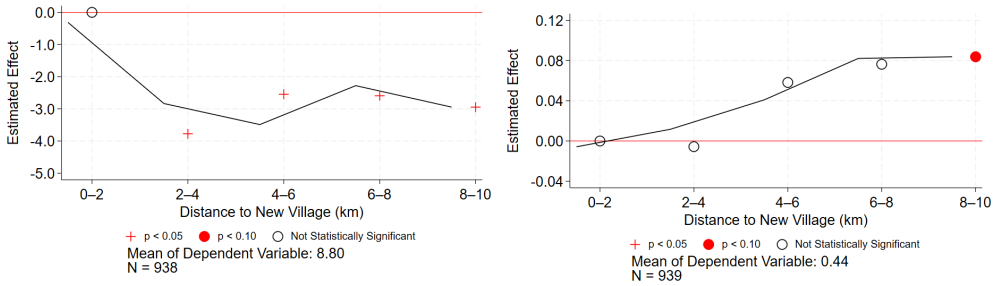
Notes: The outcome variables for (A) and (D) are the vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. There are 405 (43.1%) missing observations due to incomplete data on employment of Chinese in Chinese New Villages. These figures plot coefficients estimated 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.

Figure A.9
Effects of Chinese New Villages on Public Goods Provision



(A) Number of Schools

(B) Average Student-Teacher Ratio in Schools



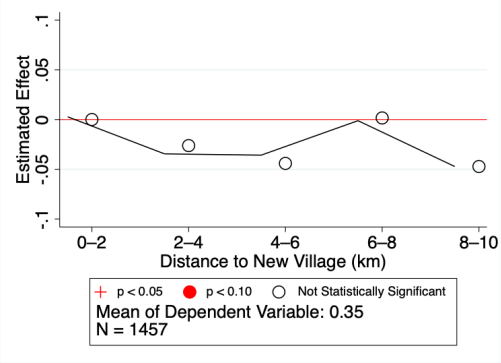
(D) Road Density

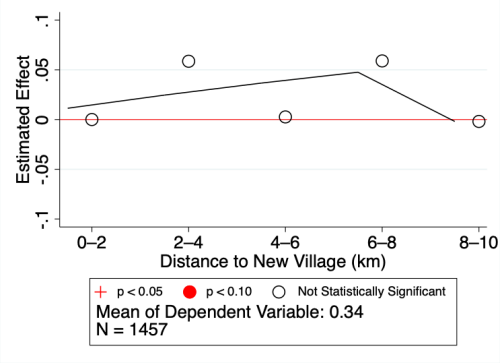
(D) Clinics (=1)

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 A.10
Effects of Chinese New Villages on Winning Coalition is Civic-nationalist

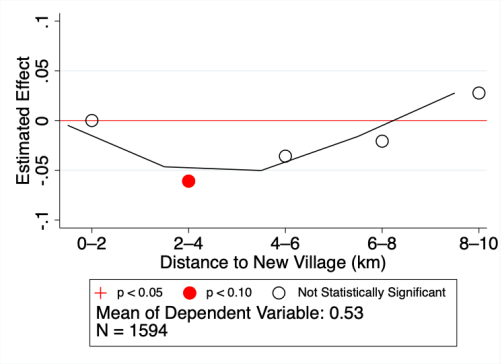
Winning Coalition is PR in 2013

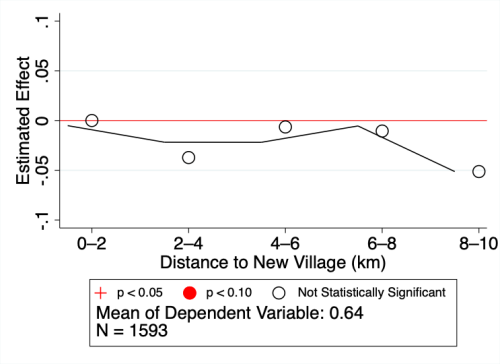




(A) State election **(B) Federal election**

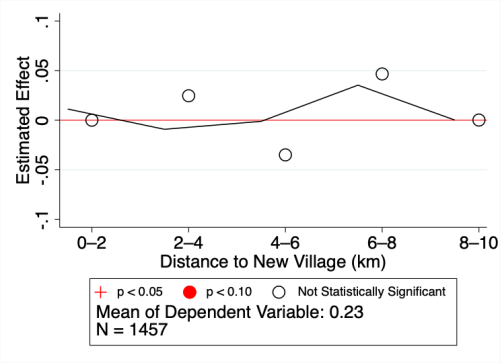
Winning Coalition is PH in 2018

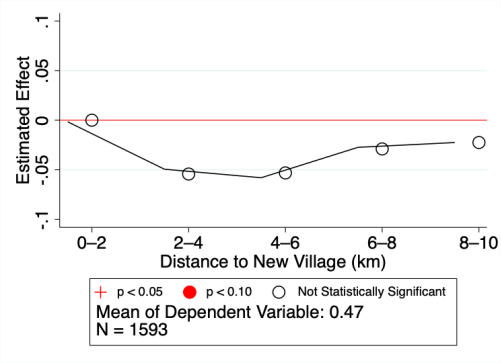




(C) State election **(D) Federal election**

Civic-nationalist Winning Coalition at both Federal and State Election

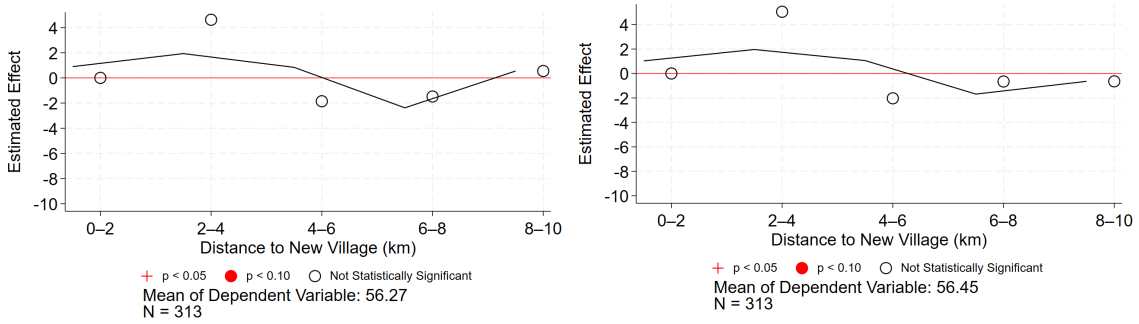




(E) Election in 2013 **(F) Election in 2018**

Notes: Figures plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for 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.

Figure A.11
Effects of Malay New Villages on Ethno-nationalist Electoral Support
Ethno-nationalist Coalition Vote Share in 2013

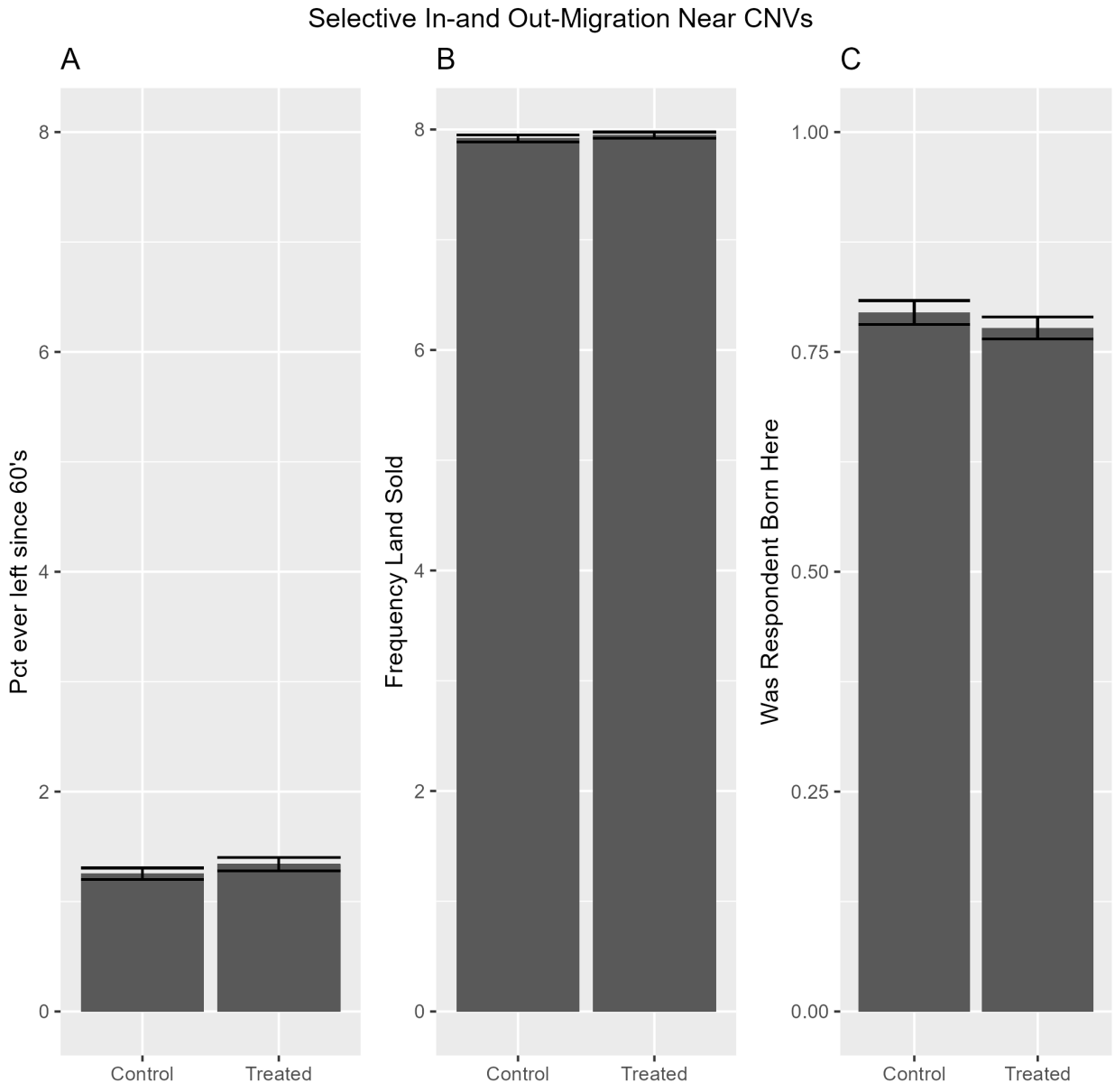


(A) State election

(B) Federal election

Notes: Figures plot coefficients from regressing the outcome variable on 2-km bins of distance to the nearest Malay 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 A.12



Notes: A *Treated* village is defined as a village being 0-2km from a real CNV. Panel A plots categorical responses to the question of "What percentage of residents since the 1960s have sold their house and never returned." The responses are: less than 1% = 1, between 1-5% = 2, between 5-10% = 3, 11-20% = 4, 21-30% = 5, 31-40% = 6, 41-50% = 7, 51-100% = 8. Panel B plots the frequency with which land in the village is sold to someone not from the village itself, with everyday = 1, weekly = 2, every two weeks = 3, every month = 4, every 3 months = 5, every 6 months = 6, every year = 7, not even yearly = 8. Panel C plots responses to the question of whether a respondent was born in the village, with 0 being no and 1 being yes. Panels A and B are responses from village leader surveys, while Panel C are from non-leader surveys. P values of a difference of means t test are 0.272, 0.492, and 0.342 for Graphs A, B, and C, respectively. The sample size for the village leader survey (Graph A and B) is 288 and the sample size for the non leader survey (Graph C) is 1990.

Table A.1
Summary Statistics of New Villages

Variable	(1) Sampled villages	(2) Non-sampled villages	(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: Sampled villages refer to the 208 CNVs included in our analysis out of the 452 CNVs that we were able to geolocate and verify. Non-sampled 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 A.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**	3.429*	4.706**	4.379**	3.876*
	(1.880)	(1.914)	(1.947)	(2.098)	(2.066)	(2.008)
4-6km	6.866***	3.944**	3.394*	6.800***	4.234*	3.587*
	(1.975)	(1.957)	(1.883)	(2.457)	(2.263)	(2.111)
6-8km	11.725***	7.135***	5.950***	12.118***	7.747***	6.352***
	(1.867)	(2.127)	(2.155)	(2.305)	(2.355)	(2.268)
8-10km	7.356***	4.637*	3.739	8.176***	5.467**	4.351*
	(2.175)	(2.502)	(2.500)	(2.556)	(2.660)	(2.545)
Observations	939	939	939	939	939	939
Adjusted R^2	0.088	0.462	0.479	0.046	0.461	0.481
Mean (Dep. Var.)	60.301	60.301	60.301	60.863	60.863	60.863
SD (Dep. Var.)	16.338	16.338	16.338	16.358	16.358	16.358
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***	6.527***	8.080***	7.128***	6.499***
	(1.776)	(1.904)	(1.846)	(2.093)	(1.861)	(1.794)
4-6km	13.558***	10.746***	9.874***	13.380***	10.399***	9.648***
	(2.258)	(1.796)	(1.798)	(2.801)	(1.720)	(1.694)
6-8km	16.486***	12.559***	11.338***	18.254***	12.642***	11.523***
	(2.365)	(2.120)	(2.115)	(2.537)	(2.015)	(1.947)
8-10km	12.021***	10.119***	8.915***	15.711***	10.764***	9.683***
	(2.592)	(2.268)	(2.205)	(2.721)	(2.187)	(2.149)
Observations	1004	1004	1004	1011	1011	1011
Adjusted R^2	0.258	0.530	0.554	0.117	0.543	0.559
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 1. 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 A.3
OLS: Chinese New Villages and Economic Effects

Panel (A)	Luminosity	Population Density
	Grid cell-level	
	(1)	(2)
2-4km	-2.264***	-296.663***
	(0.290)	(35.039)
4-6km	-3.014***	-348.714***
	(0.402)	(43.021)
6-8km	-2.962***	-346.812***
	(0.400)	(43.154)
8-10km	-3.109***	-355.305***
	(0.407)	(42.803)
Observations	21418	21419
Adjusted R^2	0.529	0.468
Mean (Dep. Var.)	2.526	146.896
SD (Dep. Var.)	5.756	464.318
Parliamentary Constituency FE	Yes	Yes
Controls	Yes	Yes

Panel (B)	Luminosity	Population Density
	Polling district-level	
	(1)	(2)
2-4km	-2.814***	-379.508***
	(0.903)	(119.934)
4-6km	-3.488***	-455.581***
	(1.166)	(123.909)
6-8km	-3.455***	-441.296***
	(1.054)	(128.169)
8-10km	-3.267***	-480.885***
	(1.174)	(114.486)
Observations	939	939
Adjusted R^2	0.851	0.795
Mean (Dep. Var.)	7.909	756.491
SD (Dep. Var.)	13.740	1517.974
Parliamentary Constituency FE	Yes	Yes
Controls	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. In Panel (A), the sample comprises grid cells in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding grid cells in 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, but at the polling district-level. 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 A.4
Scale Effects: Above vs Below-Median Initial Population Size of Chinese New Villages

	Vote Share	Luminosity	Population Density
	(1)	(2)	(3)
2-4km	1.780	-1.975***	-280.528**
	(2.253)	(0.574)	(118.288)
4-6km	1.415	-1.039	-169.726
	(2.209)	(0.733)	(128.109)
6-8km	5.349**	-0.930	-148.144
	(2.367)	(0.764)	(123.936)
8-10km	0.243	0.040	-108.762
	(3.462)	(1.292)	(144.042)
Initial Pop. Above Median	-6.084*	5.992**	707.043**
	(3.363)	(2.272)	(308.807)
2-4km * Initial Pop. Above Median	1.956	-0.777	-90.609
	(3.387)	(1.535)	(224.865)
4-6km * Initial Pop. Above Median	2.106	-3.786**	-399.627*
	(3.646)	(1.766)	(232.788)
6-8km * Initial Pop. Above Median	0.646	-3.794*	-415.789
	(4.097)	(1.940)	(282.658)
8-10km * Initial Pop. Above Median	8.981*	-5.170**	-590.804**
	(4.587)	(2.424)	(275.550)
Observations	799	799	799
Adjusted R^2	0.522	0.865	0.812
Mean (Dep. Var.)	60.762	8.649	818.639
SD (Dep. Var.)	16.887	14.638	1602.378
Parliamentary Constituency FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. In Panel (A), the sample comprises grid cells in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding grid cells in 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, but at the polling district-level. 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 A.5
Interethnic Competition: Above vs Below-Median Agricultural Employment of Chinese in Chinese New Villages

	Vote Share	Luminosity	Population Density
	(1)	(2)	(3)
2-4km	7.981**	-2.846*	-433.464**
	(3.306)	(1.462)	(214.283)
4-6km	8.142**	-2.124	-377.139*
	(3.125)	(1.347)	(212.416)
6-8km	13.830***	-2.744	-505.778*
	(3.296)	(1.992)	(284.109)
8-10km	9.956***	-1.556	-467.495*
	(3.222)	(2.056)	(250.195)
Agriculture Above Median	8.738**	2.731	104.847
	(4.230)	(2.812)	(381.812)
2-4km * Agriculture Above Median	-9.505**	-0.537	87.159
	(4.323)	(2.344)	(349.414)
4-6km * Agriculture Above Median	-11.315**	-2.982	-101.084
	(4.565)	(2.604)	(327.931)
6-8km * Agriculture Above Median	-14.833***	-1.428	175.526
	(4.755)	(2.948)	(385.420)
8-10km * Agriculture Above Median	-10.202*	-5.095*	-109.968
	(5.280)	(3.018)	(360.668)
Observations	543	543	543
Adjusted R^2	0.568	0.872	0.811
Mean (Dep. Var.)	59.548	11.202	1064.712
SD (Dep. Var.)	18.508	16.878	1855.241
Parliamentary Constituency FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. In Panel (A), the sample comprises grid cells in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding grid cells in 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, but at the polling district-level. 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 A.6
OLS: Chinese New Villages and Public School Provision

	Number of ...		Number of Teachers in ...		Average Student-Teacher Ratio	
	Primary	Secondary	Primary	Secondary	Primary	Secondary
	(1)	(2)	(3)	(4)	(5)	(6)
2-4km	0.308**	0.081	12.794***	8.911	0.672	-0.243
	(0.120)	(0.086)	(4.302)	(6.504)	(0.547)	(0.565)
4-6km	0.285**	-0.035	10.326***	-0.136	0.843*	-0.022
	(0.110)	(0.078)	(3.538)	(5.899)	(0.480)	(0.760)
6-8km	0.215*	-0.144	7.143*	-7.581	-0.076	0.509
	(0.111)	(0.088)	(3.740)	(6.766)	(0.617)	(0.958)
8-10km	0.207	-0.090	6.668	-4.339	0.023	-0.312
	(0.133)	(0.101)	(4.468)	(7.768)	(0.677)	(0.774)
Observations	939	939	939	939	635	223
Adjusted R^2	0.028	-0.008	0.087	0.007	0.416	0.369
Mean (Dep. Var.)	1.030	0.301	28.675	21.652	8.759	10.540
SD (Dep. Var.)	0.993	0.602	36.723	46.841	3.509	2.798
Parliamentary Constituency FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. 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. Standard errors are clustered at the federal parliamentary constituency-level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7
OLS: Chinese New Villages and Infrastructure and Public Health Facilities Provision

	Road Density	Number of ...	
		Clinics	Hospitals
	(1)	(2)	(3)
2-4km	-4.307***	0.071	-0.007
	(1.071)	(0.065)	(0.013)
4-6km	-4.090***	0.127*	-0.016
	(1.140)	(0.067)	(0.016)
6-8km	-4.652***	0.148*	-0.011
	(1.130)	(0.079)	(0.015)
8-10km	-5.094***	0.151	0.002
	(1.142)	(0.092)	(0.024)
Observations	938	939	939
Adjusted R^2	0.689	0.069	-0.022
Mean (Dep. Var.)	8.795	0.542	0.012
SD (Dep. Var.)	10.603	0.690	0.108
Parliamentary Constituency FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. 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. Standard errors are clustered at the federal parliamentary constituency-level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8
OLS: Malay New Villages and Ethno-nationalist Support

	Ethno-nationalist Vote Share in 2013	
	State	Federal
	(1)	(2)
2-4km	4.873**	4.935*
	(2.136)	(2.507)
4-6km	0.572	-0.032
	(3.040)	(3.526)
6-8km	1.558	2.063
	(3.369)	(3.867)
8-10km	4.769	3.587
	(3.140)	(3.202)
Observations	313	313
Adjusted R^2	0.370	0.338
Mean (Dep. Var.)	56.274	56.454
SD (Dep. Var.)	18.085	18.239
Parliamentary Constituency FE	Yes	Yes
Controls	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. The sample comprises polling districts in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding polling districts that contain Malay New Village population, polling districts in historically urban areas, and polling districts beyond 10km of a Malay New Village. In addition, regression of state election results include indicators for all possible combination 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 A.9
OLS: Chinese New Villages and Winning Coalition is Civic-nationalist

Panel (A)	Winning Coalition is PR, 2013	
	State	Federal
	(1)	(2)
2-4km	-0.048	0.022
	(0.038)	(0.045)
4-6km	-0.116*	-0.076
	(0.061)	(0.062)
6-8km	-0.095	-0.040
	(0.075)	(0.076)
8-10km	-0.156*	-0.091
	(0.081)	(0.085)
Observations	1457	1457
Adjusted R^2	0.185	0.203
Mean (Dep. Var.)	0.353	0.340
SD (Dep. Var.)	0.478	0.474
Controls	Yes	Yes

Panel (B)	Winning Coalition is PH, 2018	
	State	Federal
	(1)	(2)
2-4km	-0.077*	-0.035
	(0.040)	(0.032)
4-6km	-0.068	-0.028
	(0.055)	(0.047)
6-8km	-0.071	-0.039
	(0.065)	(0.067)
8-10km	-0.054	-0.086
	(0.070)	(0.079)
Observations	1593	1593
Adjusted R^2	0.233	0.224
Mean (Dep. Var.)	0.527	0.640
SD (Dep. Var.)	0.499	0.480
Controls	Yes	Yes

Notes: 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. Standard errors are clustered at the federal parliamentary constituency-level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10
OLS: Chinese New Villages and Civic-nationalist Winning Coalition at both Federal and State Elections

	2013	2018
	(1)	(2)
2-4km	-0.018	-0.071*
	(0.035)	(0.040)
4-6km	-0.120**	-0.090
	(0.058)	(0.057)
6-8km	-0.051	-0.085
	(0.076)	(0.070)
8-10km	-0.089	-0.101
	(0.077)	(0.078)
Observations	1457	1593
Adjusted R^2	0.183	0.234
Mean (Dep. Var.)	0.227	0.468
SD (Dep. Var.)	0.419	0.499
Controls	Yes	Yes

Notes: 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. Standard errors are clustered at the federal parliamentary constituency-level.