Interethnic Proximity and Political Development*

Chun Chee Kok
Monash University

Gedeon Lim
University of Hong Kong

Danial Shariat
UC Berkeley

Abu Siddique
Royal Holloway

Shunsuke Tsuda
University of Essex

February 20, 2025

Abstract

We exploit a population resettlement program of ethnic minorities in Malaysia to identify long-run effects of interethnic proximity on economic and political development. From 1948 to 1951, the colonial government moved 500,000 rural Chinese into hundreds of isolated, mono-ethnic camps. In ethnic majority Malay communities adjacent to these camps, we find greater economic prosperity and lower vote shares for the ethno-nationalist Malay party. Effects are stronger in areas with historical, interethnic economic complementarities. Primary survey data suggests that trust-building and social integration were key channels. Our findings highlight the importance of persistent, localized contact in the co-evolution of economic and political development. (**JEL:** D72, O15, R23, J15.)

*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, 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, 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 participants at UBC, Tufts, Brown, UCSD, UC Irvine (all Spring 2024), LKYSPP (September 2023), Monash, NUS, SMU, DWI Berlin (all 2022), and Boston University DRG (2020). We thank conference participants at ASREC Australasia (Monash, March 2023), PacDev (Stanford, February 2024), Hong Kong Empirical Microeconomics Workshop (HKU-HKUST 2021), Virtual Historical Political Economy Workshop (2021), Korean Economic History Association (Seoul, 2023), Warwick CAGE Summer School 2023, 2023 Australasian Public Choice Conference, 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@monash.edu, gedeonl@hku.hk, danial_shariat@berkeley.edu, abu.siddique@rhul.ac.uk, shunsuke.tsuda@essex.ac.uk.

1 Introduction

Managing threat and social tensions between diverse groups is a pressing issue facing nation-states (Glaeser, 2005; Guriev and Papaioannou, 2022). Throughout history, societies that have built and introduced inclusive national policies and institutions have managed to encourage cooperation, foster social capital, and align incentives between diverse groups, leading to economic growth and political stability (Acemoglu et al., 2001; Alesina and La Ferrara, 2005; North, 1990). Conversely, those that have failed—partially due to historical shocks and colonial legacies—have experienced rising inter-group inequalities, instability, and conflict (Bardhan, 2017; Michalopoulos and Papaioannou, 2013; Nunn, 2009; Putnam, 2007). Yet, how local political economies —– where state institutions interact with societal norms and economic activities to allocate resources, govern, and make collective decisions (Migdal, 1988) —– take shape and evolve in fledging nation-states is less well known. In particular, how is political development shaped by economic integration and interactions between ethnic majority and minority groups?

In this paper, we use a rural resettlement program of ethnic minorities in Malaysia to study the long-term impacts of interethnic proximity (segregation) on contemporary local economic development and political preferences of the ethnic majority for ethno-nationalist politics. From 1948 to 1951, British colonial authorities forcibly resettled 500,000 widely dispersed ethnic minority Chinese (25% of the total ethnic Chinese population and 10% of the total population) into hundreds of fenced, compact villages (Chinese New Villages, *Kampung Baru Cina*). This was done to sever support and supply lines to communist factions during a protracted conflict between British colonial authorities and ethnic Chinese communists. The policy created sharply demarcated ethnic minority Chinese enclaves surrounding pre-existing ethnic majority Malay villages, significantly reshaping the ethnic geography of the Malay Peninsula. In 1960 freedom of movement was reinstated but New Villages have persisted till today due to the award of land titles to the resettled population, which previously lacked formal land rights (Nyce, 1973). In 1963, Malaysia became a first-past-the-post democracy and all Malaysians gained full voting rights.

The policy created granular variation in interethnic proximity. We leverage this variation to study long-term impacts on local economic development and contemporary political preferences of the ethnic Malay majority. We then explore the roles of urbanization, interethnic competition and complementarities, public goods provision, and interethnic attitudes and trust in shaping these outcomes. This is important given that we know relatively little about the granular mechanisms through which local political economies emerge and evolve, especially in developing country contexts where historical legacies and institutional capacities vary widely. To do so, we build a rich dataset combining community-level geospatial, administrative, archival, satellite, and census data with individual-level primary survey data.

In interpreting effect sizes, note that the policy created variation in proximity across ethnic

majority Malay and ethnic minority Chinese villages of roughly equal sizes. Furthermore, Malays and Chinese are culturally and linguistically very dissimilar. This is important to keep in mind given the potentially important role of cultural similarity in determining levels of racial threat (Enos, 2014; Guriev and Papaioannou, 2022).

The Malaysian context offers four key advantages. First, resettlement site criteria were plausibly exogenous to underlying locational fundamentals and unobservable characteristics. Resettlement was carried out rapidly due to military exigencies and sites were largely selected based on defensibility and reinforceability.[1] Second, we geolocate precise locations of Chinese New Villages (CNVs) and combine this with an unusually rich set of initial CNV characteristics from an underutilized historical survey (Malayan Christian Union, 1958). Variation in initial characteristics of CNVs allows us to pin down precise mechanisms of exposure effects on surrounding communities. Third, while many resettlement camps around the world are temporary, uniquely, nearly all CNVs still exist today, allowing us to study long-run effects and interactions with contemporary economic and political factors. Fourth, Malaysian voters rarely change voting addresses (Jomo, 2017). This entails that polling district vote shares can plausibly be interpreted as largely reflecting (i) the political preferences of individuals who were born there; (ii) the effects of interethnic proximity to CNVs during an individual's formative years; and (iii) the effects of interethnic proximity at later-life migration destinations, if any.

For identification, we leverage the plausibly exogenous nature of the resettlement process and site characteristics. We begin by estimating results using ordinary least squares (OLS). To validate these OLS findings and address potential biases, we use a spatial, randomization inference-style approach inspired by Dell and Olken (2020) and Borusyak and Hull (2023). Specifically, we construct 1,000 counterfactual village sites for each CNV using precise military site selection criteria. 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 1,000 spatial configurations. By doing so, we ensure balance across a range of predetermined variables, reducing the likelihood that observed effects are driven by site selection biases or unobservable locational fundamentals.

We analyze the impacts of interethnic proximity by studying the exposures effects of CNVs on ethnic majority Malay communities living in surrounding areas. We do so by applying a "doughnut-hole approach," excluding all grid cells or polling districts containing CNVs from our analyses. This isolates the exposure effects of proximity to CNVs while accounting for potential spillover effects.

Our main measure of political preferences is the vote share for the ethno-nationalist coali-

[1]British military documents explicitly outline four key site selection criteria, providing a clear framework for understanding the resettlement process: (i) distance to main roads; (ii) elevation (defensibility); (iii) exclusion of lands reserved for ethnic Malays (Malay Reservation Lands); (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.

tion, the National Front (*Barisan Nasional*), in the 2013 and 2018 General Elections. The National Front is a coalition of ethnic-based parties that governed Malaysia for nearly six decades. Since independence in 1957, the coalition has consistently advocated for and implemented affirmative race-based policies favoring the ethnic majority Malay population (Jomo, 2017). We measure vote shares at the polling district level—the most disaggregated level possible. We interpret positive vote shares for the National Front as reflecting preferences for ethno-nationalist political parties and policies that prioritize ethnic majority Malays.

We find a long-run negative effect of interethnic proximity on vote shares for the National Front across both elections. Effects are highly localized: polling districts within 0-2 kilometers have about 3 to 6 percentage points lower vote shares for the ethno-nationalistic coalition, in both state and federal elections, implying a shift in political preferences in areas more proximate to Chinese New Villages.[2]

We argue that negative vote shares are likely to reflect the political preferences of ethnic majority Malays. To further quantify this, we would need disaggregated data on voting behavior by ethnicity but such data does not exist systematically for this contemporary period. Instead, we show that our results are robust to restricting our estimation sample to Malay-majority polling districts and conducting a back-of-the-envelope exercise (Becker and Woessmann, 2009; Calderon et al., 2023) to quantify the role of Malay voters by (unrealistically) assuming that all Chinese voters voted against the ethno-nationalist coalition and showing that the range of ethnic Chinese voter turnout rates under such a scenario would have to be implausibly high for Chinese voters alone to explain the entirety of negative ethno-nationalist vote shares.

We then examine potential mechanisms to explain the long-run persistence of differences in political preferences and the key channels that might have led to these differences. A plausible explanation for the persistent impacts is the continued existence of CNVs, which were formalized after 1960 through the awarding of land titles to the resettled population (Nyce, 1973). This policy enabled ethnic Chinese settlers to establish long-term roots in these areas, creating stable communities that remained economically and socially relevant to the surrounding regions.

The establishment of New Villages enabled the ethnic Chinese minority to contribute to local development and public goods provision, likely benefiting neighboring ethnic Malay majority population. These contributions may have arisen through economic complementarities (Alesina and La Ferrara, 2005), where one group's economic activities enhanced the productivity or opportunities of the other, or through improved public goods provision. For example, better access to education could have promoted human capital development and reduced economic disparities, fostering social cohesion and more inclusive political preferences to become supportive of policies that benefit the ethnic Chinese.

Another plausible and related mechanism is urbanization and agglomeration. The trans-

[2]Elections in Malaysia are conducted once every 5 years. We focus on vote shares post-2013 because 2013 is the first election year in which disaggregated, polling-district level voting results exist.

formation of rural areas into economically dynamic and urbanized regions may have shifted traditional political preferences (Dasgupta and Ramirez, 2020; Glaeser and Kahn, 2004). Urbanization, often linked to greater exposure to diverse populations and economic modernization, can weaken identity-based political affiliations, particularly if urbanization fostered ethnic integration or shifted political priorities toward economic issues. Moreover, concentrated economic activity around CNVs likely generated localized productivity spillovers, attracting businesses, enhancing infrastructure, and creating jobs (Duranton and Puga, 2004, 2020). These spillovers may have benefited nearby Malay communities, fostering economic interdependence and reducing incentives for ethnically polarized voting.

Interethnic competition and complementarities could also jointly affect political preferences near CNVs. As the Chinese minority became economically established, their success may have also intensified economic and political competition with the neighboring Malay population, particularly in employment, business, and access to state resources. Where such economic rivalry persisted, perceived competition over resources may have strengthened ethnic voting patterns (Horowitz, 2000). In contrast, proximity could lead to more positive political attitudes if interactions are driven by interethnic complementarities and specialization. This could, in turn, improve productivity, social cohesion, and lead to greater tolerance toward out-groups (Alesina and La Ferrara, 2005; Jha, 2013). Where such economic complementarities dominated, interethnic proximity may have promoted cooperation and political moderation.

Finally, we consider interethnic contact as a plausible mechanism (Allport, 1954; Enos, 2014), as it may have built greater social capital and trust between ethnic groups. For instance, proximity to CNVs might have increased the frequency of casual or formal interactions between ethnic Malays and Chinese, such as through shared markets, community events, or collaborative projects. These interactions may have promoted mutual understanding, improved interactions, and built trust toward the Chinese minority. Increased interethnic contact could also lead to the formation of cross-ethnic friendships or greater familiarity with the out-group, which are key channels through which intergroup biases could be reduced over time.

To test these mechanisms, we use several measures: (i) localized night-time lights satellite data as a proxy for local economic development, capturing the intensity and distribution of economic activity; (ii) localized school, hospital, and road construction data from the Ministry of Education, Ministry of Health, and Open Street Map as a measure of public goods provision, capturing the local investments in infrastructure and human capital; (iii) the Normalized Difference Vegetation Index (NDVI) as a proxy for agricultural productivity (Rao et al., 2022), where a decline in agricultural activity might signal urbanization and a shift in economic structure; and (iv) population data at the grid-cell (1×1 km^2) level from the National Historical Geographic Information System (NHGIS), as population density is highly correlated with economic activity and development. We complement these aggregate measures with individual-level primary survey data to examine the roles of interethnic interactions, social capital, and trust.

First, we find long-run, hyper-local positive effects on economic development and urbanization in areas located immediately around resettlement sites (0-2km). Specifically, in a grid-cell analysis, areas further from resettlement sites exhibit lower nighttime light intensity, lower population density, and higher agricultural crop productivity. In contrast, at a higher level of aggregation (polling districts), effects in further distance bins largely dissipate, except for a decrease in nightlights and an increase in NDVI within the 2-4km bins. The contrast between grid-cell and polling-district level results suggests that highly localized agglomeration benefits with limited spatial spillovers and spatial reallocation are, together, the most likely explanations for the observed spatial patterns in economic outcomes. We also explore scale effects by examining heterogeneity based on the initial population size of CNVs. The results show slightly stronger economic effects near larger CNVs, supporting localized agglomeration benefits, but at the same time, the absence of substantial differences with exposure effects to smaller CNVs, suggests the role of additional mechanisms.

Second, we investigate the role of interethnic competition and complementarities. We do so by analyzing heterogeneity across CNVs with above and below-median initial resettlement employment shares in the agricultural sector; and above and below-median initial resettlement employment shares in the rubber and tin mining sectors. In the former, interethnic competition could have been more salient due to the historically large share of Malays as farmers (Kratoska, 1982a). In the latter, complementarities could have arose due to Chinese businesses in rubber and tin that often employed local Malay workers (Ross, 2014; Siew, 1953).

We find evidence consistent with interethnic complementarities: in the above-median rubber and tin-mining employment sample, ethno-nationalist vote shares are significantly negative in polling districts immediately adjacent to CNVs. We also uncover positive (albeit small) of proximity on contemporary economic outcomes. We interpret this as evidence of economic complementarities which enabled the Malays to benefit from initial economic gains generated by Chinese resettlement and which subsequently shifted political preferences. Positive economic effects, however, could have decreased over time as ethnic Chinese in CNVs transitioned away to smallholder cash crop or commerce occupations which had less scope for interethnic complementarities with Malays. In contrast, we find weak evidence for interethnic competition: effects on vote shares largely disappear in the above-median agricultural sample.

Third, we find significant increases in educational and infrastructural public goods in communities closer to CNVs. Specifically, we find increases in public secondary schools and teachers, suggesting that interethnic proximity might have led to increased investments in educational public goods, benefiting both communities and shifting long-run political preferences. Relatedly, we find significant increases in road density. Roads potentially promoted greater social cohesion and economic mobility, leading to increased support for political parties that prioritized civic-nationalism.

Finally, primary survey results show a large gap between older and younger cohorts in

trust, openness, social interactions, and social capital in areas unaffected by CNVs. However, in areas exposed to CNVs, this gap significantly narrows, suggesting substantial attitudinal and behavioral shifts across age cohorts. These changes in interethnic attitudes likely contributed to the weaker support for ethno-nationalist parties in areas closer to CNVs, as greater trust and social integration possibly reduced the appeal of identity-based political mobilization.

We rule out two key alternative explanations for lower ethno-nationalistic vote shares. First, the observed effects may not be driven by exposure to the ethnic Chinese community per se but by exposure to new people, which could bring new ideas, opportunities, or norms that shape political preferences (Balietti et al., 2021; Chetty et al., 2016). To test this, we replicate our analysis using a small number of Malay New Villages which were established during the Malayan Emergency for similar purposes (Dobby, 1952; Humphrey, 1971). However, we find no comparable political or economic effects near Malay New Villages, indicating that the observed patterns are driven by interethnic proximity to Chinese communities.

Second, it is plausible that higher vote shares for constituent parties in the civic-nationalist (pro-diversity) coalition led to electoral wins and subsequent policies that spurred local economic development and shifts in political preferences. However, we find no evidence of significant electoral wins for these parties. This suggests that while interethnic proximity may have influenced political preferences at the intensive margin, it was insufficient to change electoral outcomes at the extensive margin. We cannot rule out, however, the possibility that improvements may have stemmed from greater grassroots-level interethnic collaboration—a possibility that would be consistent with our survey results—or higher allocations by the ruling ethno-nationalist coalition (Bardhan and Mookherjee, 2010).

Related literature. This paper makes three major contributions. First, it bridges the literature on the effects of forced emigration and resettlement on nation-building and assimilation (Bazzi et al., 2016) with that on spatial growth and economic development (Peters, 2022).[3] Most papers focus on (aggregate) economic or political effects, but the interaction between the two and the distributional effects on natives remain less clear.[4] We use granular data to show that the political effects of resettlement can vary depending on the spatial distribution of economic benefits. Specifically, we document that ethnic-majority individuals living closer (further) to resettled ethnic-minority villages enjoy economic benefits and, as a result, are less (more) likely to vote for ethno-nationalist parties. We hypothesize that one possibility is the reduced demand

[3] 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.

[4] 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 developing country settings.

of natives for political power to protect native economic interests in adjacent communities. Hence, we complement recent studies (e.g. Calderon 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.

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 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 the effects of *inter*-village contact between two large ethnic groups on political preferences at the ballot box and provide novel evidence that positive effects of contact can arise even in the presence of persistently higher inter-village segregation. Separately, our paper closely connects to Jha (2013) by providing evidence that historical interethnic economic complementarities can lead to persistently positive interethnic attitudes even after initial economic benefits have diminished. Using primary survey data, we elucidate micro-level mechanisms: while inter-village segregation limits deep friendship formation, sustained casual contact and economic benefits can still yield long-term political benefits. Our findings, therefore, provide quasi-experimental evidence for the large and generalizable positive effects of *broad* intergroup contact (Lowe, 2024).[5]

Second, we contribute to the broader literature on the political effects of immigration (Alesina and Tabellini, 2024).[6] Importantly, our results contrast with Tabellini (2020) who finds that higher immigration in US cities led to positive economic outcomes but higher political backlash. One possible reason is that, in our largely rural setting, there are greater opportunities for interethnic contact in non-segregated, public spaces. This is consistent with how shared spaces can build social capital through promoting greater contact (Putnam, 2015). Hence, we provide evidence of how the effects of contact might vary with the volume and density of shared spaces (in rural vs urban environs).

Third, 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). 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 demonstrating that negative political preferences towards ethnic minorities can potentially be mitigated through

[5]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).

[6]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. For examples, see Barone et al. (2016); Dustmann et al. (2019); Halla et al. (2017); Otto and Steinhardt (2014); Steinmayr (2021).

positive economic spillovers from interethnic complementarities. Furthermore, we collect novel primary survey data to show that negative, cross-cohort gaps in interethnic attitudes can be narrowed — but only over a sufficiently long period of sustained contact. 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 our data. Section 4 describes our empirical strategy. Section 5 reports empirical results. Section 6 describes underlying mechanisms behind the results. Section 7 rules out alternative mechanisms. Section 8 concludes the paper.

2 Institutional Background

2.1 Ethnic Diversity and Interethnic Relations

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 limited intermarriage between the ethnic Malays and ethnic Chinese. A mere 0.1% of ethnic Malay men married ethnic Chinese women and 0.3% of ethnic Chinese men married ethnic Malay women in 2000 (Nagaraj, 2009). This low rate of intermarriage underscores the distinct boundaries of ethnic identification between Malays and Chinese in contemporary Malaysia.

The earliest waves of Chinese migration to Malaysia can be traced back to the 15th century (Kim, 1998; Wang, 1959) but the resettled ethnic Chinese population were largely those that arrived from the early 20th century (Strauch, 1981). Specifically, these ethnic Chinese comprised of traders, businessman, and laborers driven from China by economic and political upheavals (from the 1860s to the early 1900s), and who were drawn to British Malaya by the demand for vast amounts of labor to fuel the colonial economy. During this time, the ethnic Chinese were often regarded as unassimilable — speaking a distinctive language from the natives and establishing clan-based institutions (Ting, 1976). Later, they were also perceived as politically and culturally oriented towards China, and hence as possible security risks (Suryadinata, 1987).[7]

Interethnic relations were largely shaped by colonial policies and the effects of these policies have been wide-ranging and persistent. The British actively pursued a division of economic roles along ethnic lines. In particular, the native Malay population was often seen as farmers

[7]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).

and tasked with producing rice for workers in rubber estates and tin mines, which yielded substantially lower economic gains (Kratoska, 1982b). Post-independence, these differences in ethnic-based colonial policies resulted in a significant economic gap between the Malays and the Chinese, with the Malays viewing the ethnic Chinese as having achieved greater economic superiority at their expense.

2.2 Electoral Politics

In 1969, the economic gap between Malays and Chinese culminated in widespread racial riots. The violence was partially the result of economic grievances from ethnic Malays who did not fully participate in the rapidly modernizing economy of the newly independent Malaysian nation.

In response, the National Front (*Barisan Nasional*), a ruling coalition of ethnic-based parties led by the dominant and largest ethno-nationalist Malay party, the United Malays National Organization, implemented the New Economic Policy (NEP) in 1970. The twin aims of the NEP was to reduce ethnic inequality and poverty. Through the NEP, the National Front demonstrated its firm commitment to ethnic affirmative action, with ethnic Malays receiving preferential access to education, housing, public-sector jobs, and corporate share ownership (Jomo, 2017).[8] These policies have played a crucial role in maintaining Malaysia's political stability. Until 2018, the National Front had always formed the ruling government and returned to office unopposed.

In 2013, a multi-ethnic coalition called the People's Alliance made substantial in-roads against the ruling National Front government. The People's Alliance (*Pakatan Rakyat*) comprised of three major parties, the People's Justice Party (PKR), Democratic Action Party (DAP), and Parti Islam Malaysia (PAS), and was chiefly led by PKR.[9] All 3 parties ran under a joint banner of anti-corruption, reform, and generally more liberal economic policies than the National Front. In particular, a key policy agenda was the idea that Malaysia is a multicultural nation which should grant equal rights and representation to all citizens, regardless of ethnicity nor religion.[10]

By the end of the 2013 General Elections, the National Front had suffered heavy losses and

[8]Separately, studies have examined the impact of the NEP on inter-ethnic income inequality and redistribution. Ravallion (2020a) shows that after the NEP, all three main ethnic groups (the ethnic Malays, Chinese, and Indians) had rising mean incomes, with Malays and indigeneous groups enjoying the highest rates of income growth. In a subsequent study, Ravallion (2020b) also shows that ethnic redistribution helped reduce poverty, but the impact was large in 1970, the time the NEP was introduced, but declined considerably after.

[9]PKR was and continues to be the party vehicle of then-leader of the opposition, and current Prime Minister Anwar Ibrahim.

[10]Having said that, given the ethnic make-up of the Malaysian electorate, the overall electoral strategy was still ethnic-based in that PKR was to appeal towards urban Malays; DAP was to appeal towards ethnic Chinese and Indian voters; and PAS, being the only outwardly religious party in Malaysia, was tasked with winning over the rural Malay vote bank.

only remained a coalition of ethnic-based parties in principle. In reality, the United Malays National Organization had become the politically most powerful party by virtue of its superior seat count in both federal and state parliaments, spurring a right-ward shift in pro-Malay ethno-nationalist rhetoric.[11] In particular, the National Front began to champion the idea of 'Ketuanan Melayu' (Malay Sovereignty), a political concept emphasizing Malay power and pre-eminence, and that government policy should (continue to) accord Malays special rights given birthright claims on Malay land and soil (Ostwald and Oliver, 2020).

In 2018, the National Front remained intact but the opposition coalition was rechristened the Alliance of Hope (*Pakatan Harapan*) after the departure of the Islamic party, PAS.[12][13] Together with other conservative, right-wing parties, PAS formed a third coalition, *Gagasan Sejahtera*, GS. This culminated in the two main opposition coalitions, the Alliance of Hope and GS running against the National Front. Altogether, the 2018 elections was largely seen as a choice between reform—via progressive politics (Alliance of Hope) or Islamic conservatism (GS)—or a continuation of the old system of ethnic-based policies under the National Front. The result? For the first time in Malaysia's 65-year electoral history, the Alliance of Hope won the popular vote and a majority of seats at both the federal level and in various state assemblies, successfully forming the first non-National Front Federal Government.

2.3 The Electoral System

Malaysia is a first past the post parliamentary democracy. Since 1963, elections have been held at both the federal and state constituency-level in 5-year electoral cycles. At the federal and state-level respectively, voters elect Members to the House of Representatives and State Legislative Assemblies from single-member constituencies. In general, each federal constituency comprises of between two to six state constituencies.

In Malaysia, citizens who are 21 years or older are eligible and automatically registered to vote.[14] Each voter is typically assigned to the polling station located closest to his/her registered voting address. Each polling station is then assigned to the polling district in which the station is located in.[15] In terms of registered voting addresses, Malaysia is unique in that relatively few voters ever change their voting addresses despite out-migration after birth and

[11]The other two constituent parties are the Malaysian Chinese Association, an ethno-Chinese centrist party representing ethnic Chinese, and the Malaysian Indian Congress representing ethnic Indians. By the 2013 General Elections, both parties held just a few seats in the Malaysian Parliament and were largely inconsequential as coalition partners.

[12]This was largely caused by the schism between PKR and PAS. After insurmountable disagreements on topics such as the implementation of sharia law, PAS opted to leave the coalition. This resulted in a number of moderate Islamic politicians leaving PAS to form *Amanah* (The Peace Party), which remained in the Alliance of Hope.

[13]Given that the Alliance of Hope's policy platform remained largely unchanged and continued to be led by both PKR and current Prime Minister Anwar Ibrahim, we continue to refer to it as the People's Alliance in subsequent sections.

[14]Starting from the 2022 General Elections, the minimum voting age has been lowered to 18 years.

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

subsequent changes in actual residential addresses (Jomo, 2017).[16] In Section 3.2, we discuss how this affects our interpretation of results on differences in vote shares.

2.4 The Natural Experiment: Colonial Resettlement of Ethnic Chinese Minorities to New Villages

To study the impact of geographic proximity to ethnic Chinese on electoral support of the National Front, 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 at least 500,000 rural ethnic Chinese (approximately 10% of the entire population and 25% of the entire ethnic Chinese population of Peninsular Malaysia in 1947) to fenced-up New Villages across Peninsular Malaysia. Figure 1 shows the distribution of all New Villages that we successfully geolocated.[17] Following military site selection criteria, nearly all New Villages 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 purposes.

The colonial resettlement program was a consequence of the political struggle between the British military and ethnic Chinese communists. After the Japanese surrender in 1945, the Malayan Communist Party (M.C.P) gained widespread support, especially among ethnic Chinese communities for waging a war against the Japanese. The M.C.P shifted its focus to an anti-British movement and retreated into the jungle. The communist insurgents targeted rubber estates and tin mines for attacks. A military plan was necessary to maintain political stability.[18]

In response, the British military implemented the resettlement program to weaken the influence of communist insurgents. The resettlement was rapid and often unexpected, with the bulk of resettlement taking place within the first 3 years of the Emergency. Humphrey (1971) emphasized this point by describing how the resettlement essentially "*provided an opportunity to examine a system of communities which have not evolved...in response to normal locational and economic factors, but which have been created in a manner similar to...a laboratory experiment*"

According to historical records, rural ethnic Chinese were resettled to villages located about 3.2 to 9.6 kilometers from widely dispersed ethnic Chinese squatter settlements (Nyce, 1973). The resettled ethnic Chinese were largely rural farmers engaged in food and pig production,

[16] Various sociological explanations have been put forth but potentially the most obvious is that there is a high administrative cost of changing your registered voting address but few upsides given that (i) returning home to vote allows voters to visit their family; (ii) home (rural) polling districts are typically more competitive than urban polling districts and hence a voter's vote might count for more (Jomo, 2017). A case in point: Kuala Lumpur, the capital of Malaysia, does not have a state legislative assembly and has disproportionately fewer seats in the federal assembly despite a much higher population.

[17] We describe the geolocation process in Section 3.

[18] 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-20th century.

tin mining, rubber growing, and other crop production. The resettlement created tightly controlled CNVs. These villages were heavily surveilled, often enclosed by barbed wire and largely isolated from neighbouring communities from 1948-1960.

In the early 1960s, as the British deemed the communist threat to be in retreat, freedom of movement was gradually reinstated. However, there has been a strong persistence of CNVs, many of which continue to exist in the same location. Qualitative evidence suggests that this is due to high moving costs, increased security and amenities, and most importantly, incomplete land titles (Nyce, 1973). Many of the resettled Chinese were originally squatters on government-owned land and had neither legal rights nor titles to the land that their houses and crops were grown on. Hence, the award of land titles in New Villages gave, for the first time, Chinese settlers a stake in their homes and a strong incentive to stay in these areas despite potential out-migration opportunities.

3 Data

3.1 Sampling Frame

Chinese New Villages. We obtain our baseline sample of 236 Chinese New Villages in three steps. First, we obtain a list of 666 New Villages from the 1958 Malayan Christian Council Survey (Malayan Christian Union, 1958). Second, we successfully identify and geolocate 452 New Villages based on official maps published by the Ministry of Housing and Government in 2012 (Lee, 2012). We manually verify that all 452 New Villages still exist today using Google Earth and Google Maps. Figure 1 plots the geographical distribution and Table A.1 shows selected summary statistics.

Third, we impose three sample criteria: (i) we keep all New Villages whose primary medium of language was recorded as Mandarin or a Mandarin dialect in 1958.[19] 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 Chinese New Village. (ii) We exclude New Villages in the two east coast states of Kelantan and Trengganu where resettlement largely involved non ethnic Chinese.[20] (iii) We exclude all New Villages located in 1947 urban census districts.[21] This procedure gives us a total of 236 Chinese New Villages. Figure 2 shows the geographical distribution of our baseline sample of Chinese New Villages.

[19] 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

[20] This is due to the historically low number of ethnic Chinese in these two states.

[21] 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 collect polling-district level electoral results and boundaries from the 2013 and 2018 General Elections (Malaysian Electoral Commission). Polling districts are the smallest geographical unit at which election results are available in Malaysia. In our baseline sample, there are 1,457 polling districts in 2013 and 1,557 in 2018 due to population growth. On average, each polling district has a geographical area of $10.9km^2$ and a population of 1,519 registered voters.[22]

We impose three 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).[23] 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.[24]

3.2 Outcomes and Mechanisms of Interest

Interethnic proximity. We measure interethnic proximity as the fly-by-crow distance from the centroid of each polling district to the nearest CNV. Figure 2 shows the distribution of polling districts with respect to the fly-by-crow distance to the nearest CNV. 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.

Vote shares. We use federal and state election vote shares for the National Front in both the 2013 and 2018 General Elections as our main measure of political preferences for ethno-nationalist policies. In both elections, the National Front ran on an ethno-nationalist, Malay-first platform vis-a-vis the People's Alliance (rechristened the Alliance of Hope in 2018) that ran on a relatively more inclusive, multi-ethnic policy platform. The National Front won 47.38%

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

[23]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.

[24]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.

(33.77%) of the popular vote in 2013 (2018).

We do not distinguish between contests against the opposition Islamic party (PAS) for three reasons. First, there is little distinction between ethnic Malays and Muslims in Malaysia. Nearly all Malays are Muslims and the constitution defines a key part of the ethnic Malay identity as an individual that practices the Muslim faith. Second, the dominant Malay party under the National Front (the United Malays National Organization) has long espoused a right-wing conservative platform based on both ethnic Malay *and* religious Muslim identity. Hence, in 2013, the key distinction between the National Front and PAS was that of PAS running under a multi-ethnic coalition banner.[25] Third, in 2018, PAS had moved firmly to the right of UMNO and the National Front, as part of a third, conservative, right-wing opposition coalition (GS). Having said that, results in 2018 are unlikely to be driven by the presence of PAS since they contested in nearly all parliament seats and won just 16.89% of the popular vote.

Figure 3 shows the polling-district level geographical distribution of vote shares for the National Front 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 mal-apprortionment to favor the National Front 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.[26] In short, analyzing changes in federal election vote shares at the polling-district level allows us to examine the effects of varying proximity to CNVs while keeping both political party and candidate identity fixed. Hence, we uses vote shares in the federal elections as our main outcome variable unless otherwise noted.

We interpret differences in vote shares as capturing a composite effect of differences in interethnic proximity both during an individual's formative years, as determined by proximity of home polling districts to CNVs, and during an individual's latter-life, as determined in migration destinations due to later-life migration choices. As discussed above, Malaysians rarely change voting addresses and this allows us to more accurately interpret vote shares as reflecting (i) the political preferences of individuals who were born there; (ii) differences in interethnic proximity during an individual's formative years, growing up in home polling districts located at varying distances to CNVs. Nonetheless, in polling districts with a greater share of out-migrants who reside outside their home polling districts and return home only to vote, vote

[25]Our results remain largely unchanged when we exclude such contests, possibly attesting to the unity of the multi-ethnic opposition coalition message in 2013.

[26]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. We do, however, have all this data and can conduct simple difference-in-differences analysis to examine the evolution of vote shares at the federal and state seat level across seats with varying numbers of CNVs.

shares could also capture the effects of interethnic proximity at migration destinations. To that end, all regressions control for distance to the nearest 1947, pre-resettlement urban center.

Voter turnout. 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.

Ethnic shares of registered voters. We construct ethnic shares of registered voters across the entire Peninsular Malaysia using voter rolls from the 2013 General Election. Voter rolls in 2013 record the self-identified ethnicity of each voter in Malaysia.[27] This is, to the best of our knowledge, the most granular and accurate measure of ethnic shares available.[28]

Local economic development and urbanization. We measure contemporary economic development using the best available contemporary economic data. Contemporary censuses do not contain data at a disaggregated enough level for our analysis. Since data on income per capita is not available 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).[29] We use the Normalized Difference Vegetation Index (NDVI) in 2010 as a proxy for rurality and the extent of land-use for agricultural production (Rao et al., 2022). The NDVI is derived from NASA's Visible Infrared Imaging Radiometer Suite (VIIRS) and Landsat, and measures vegetation density using satellite sensors: dense vegetation absorbs more red light and reflects more near-infrared light, resulting in higher NDVI values. We also use population data in 2010 from Global Human Settlement Layer (GHSL). These data are aggregated up to obtain the average values at grid cell level of 1km X 1km and polling district level, respectively.

Public goods provision. We measure educational public goods by geo-referencing a complete list of all primary and secondary schools in Malaysia in 2010 from the Ministry of Education.[30] These records contain the point coordinates and number of teachers and students for each school. We aggregate the data up to the polling district level to obtain the number of primary and secondary schools and teachers.

[27] 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.

[28] Existing data on ethnic shares exists only at higher aggregated census and administrative district levels

[29] We use 2014 data as pre-2014 nighttime light intensity measures potentially contain substantial measurement error due to flawed blurring, lack of calibration, and top-coding (Gibson, 2021).

[30] We download the data from https://myschoolchildren.com/list-of-all-primary-schools-in-malaysia/ and https://myschoolchildren.com/list-of-all-secondary-schools-in-malaysia/.

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.[31] We aggregate the 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).[32] 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.

Primay survey data on attitudes and behaviors of Malays. We describe these at length when presenting results in Section 6.4.

4 Empirical Strategy

4.1 Ordinary Least Squares Estimation

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

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

Similarly, we estimate the effects of distances to the nearest CNV on grid-cell level economic outcomes:

$$Y_{g,p} = \alpha + \sum_{k=1}^{4} \beta_k distCNV_g^k + \gamma_d \mathbf{X}_d + \gamma_g \mathbf{X}_g + \theta_c + \epsilon_{g,p} \tag{2}$$

where $Y_{d,p}$ ($Y_{g,p}$) is an outcome of interest in polling district d (grid-cell g) of federal parliamentary constituency p. $distCNV_d^k$ ($distCNV_g^k$) are indicators equal to 1 if the geodesic distance from the centroid of polling district d (grid-cell g) 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, and the natural logarithm of population density in 1947. Whenever possible, we measure and control for these variables at the grid-cell level ($\mathbf{X}_g$). In addition, regressions of polling-district level vote shares in state elections include indicators for every possible combination of party match-ups at the state constituency level. Importantly,

[31] We download the data from https://github.com/MoH-Malaysia/data-resources-public

[32] We download the data from https://data.humdata.org/dataset/hotosm_mys_roads.

we also control for the share of ethnic Chinese in 1947 at the polling-district level. This ensures that all our regressions compare areas with similar levels of pre-treatment diversity. θ_p (θ_c) is a vector of federal parliamentary constituency (nearest Chinese New Village) fixed effects. We cluster standard errors at the federal parliamentary constituency level.

The potentially endogenous location choices of CNVs, however, suggest that the OLS estimates of β_k in Equation 1 could be upward biased. For example, β_k might be picking up the effects of proximity to pre-existing roads given that a key British criterion for site selection was access to roads to facilitate rapid military reinforcements (dhu Renick, 1965). Proximity to roads might, in and of itself, lead to these sites experiencing higher economic growth over time due to higher market access. Similarly, if the British had chosen sites where the existing native population had more positive attitudes towards ethnic Chinese minorities, β_k might be picking up these effects that have little to do with that of interethnic proximity.

In short, the OLS coefficients on economic development and voting patterns in Equation 1 might be picking up effects that have little to do with proximity to Chinese resettlement sites if receiving locations and resettlement sites were endogenously chosen by the British based on their economic potential or pre-existing attitudes of Malays. Hence, we turn to a counterfactual exercise in subsection 4.2 that leverages plausibly exogenous British site criteria to further estimate these coefficients of interest.

4.2 Counterfactual Site Selection Using British Military Criteria

To address the potential endogeneity of CNV locations, we use a counterfactual analysis in the spirit of Dell and Olken (2020), a methodology that has been formalized theoretically in Borusyak and Hull (2023). This approach exploits the fact that there were potentially multiple possible equilibria for CNV resettlement locations, allowing us to reconstruct the universe of possible counterfactuals for CNV resettlement sites. We do this by strictly following a set of British military criteria outlined in previously classified documents. These criteria were plausibly exogenous to pre-existing political and economic conditions, given the exogeneity of fighting an all-out anti-communist war, as evidenced by the rapid resettlement of 500,000 Chinese people into around 500 villages within 3 years (the Briggs' Plan).

To this end, we identify and construct 1,000 sets of feasible, counterfactual CNV sites by imposing the following four requirements based on British military archival documents:[33]

1. *Road accessibility and minimizing disruption to economic livelihoods of squatters:* 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 90th percentile of

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

actual CNVs and contain at least as much NV-suitable land (within 2.5 km radius, which is the average size of a CNV (Nyce, 1973)) as the 10th percentile of actual CNVs.

3. *Exclusion of Malay reservations:* Counterfactual sites cannot be located within Malay Reservations. Malay Reservations refer to lands reserved for the exclusive use of ethnic Malays and, hence, are areas where the British did not resettle any ethnic Chinese to (Kratoska, 1982b, 1983).

4. *Spatial balancing:* To create 1,000 sets of counterfactual CNVs, we shift actual CNVs randomly while ensuring that each set of counterfactual CNV must be approximately balanced North/South and East/West around actual CNV. We use a simulated annealing procedure (Dell and Olken, 2020) to minimize the average distance between counterfactual and actual CNVs.

We measure road accessibility and exclude Malay reservations by digitizing the entire universe of 1947 pre-settlement roads and Malay Reservation polygon boundaries across the entire Peninsular Malaysia using the HIND1035 map series from the Australian National Libraries archives. To our knowledge, the HIND1035 series has one of the highest resolutions (1:63,360 or 1 inch to a mile) available for capturing granular, local-level differences in the immediate pre-resettlement period in Malaysia. Figure A.1 shows an example of a HIND1035 map.

In particular, the fourth condition requires that the distance between each and every other counterfactual site is similar to the distance between each and every other actual CNV site. This is important for two reasons. First, ensuring a similar, adequate distance between sites would have been important given that the British would plausibly have tried to ensure that CNVs were far enough away from each other to prevent coordination among ethnic Chinese villagers towards forming bases for Communist support. Second, a key site criteria (condition one) was to minimize disruptions to economic livelihoods of villagers (dhu Renick, 1965) but systematic, granular data on Chinese pre-resettlement locations does not exist. Instead, we rely on the spatial balancing exercise to partially compensate for this by mimicking the actual distribution of CNVs that would likely have been similarly constrained by pre-resettlement Chinese locations.

Figure 4 illustrates this approach. Panel A shows a real CNV, Buloh Kasap, represented by a purple dot with suitable polygons shaded in green (defined as polygons that fulfill all of the first three conditions), pre-resettlement roads represented by red lines, and polling district boundaries represented as black lines. Panel B shows a single suitable counterfactual CNV site, represented by a blue dot. Panel C shows a single set of 1,000 possible counterfactual village sites for Buloh Kasap after excluding sites that do not fulfill the fourth condition.

The polling district-level 'counterfactual' estimating equation takes the following form:[34]

[34]The grid-cell level 'counterfactual' estimating equation is analogous.

$$Y_{d,p} = \alpha + \sum_{k=1}^{4} \gamma_{k,real(fake)} distCNV_d^{k,real(fake)} + \gamma_2 X_d + \theta_p + \epsilon_{d,p} \quad (3)$$

The key difference with Equation 1 is that, for every outcome, Equation 3 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 explicitly purge the effects of any unobserved factors correlated with the combination of British military site suitability criteria.

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 3 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.

4.3 Counterfactual Site Selection Balance Checks

We begin by examining the balance in geographic, soil, and pre-resettlement characteristics. If our procedure yields counterfactual locations that were equally suitable for resettlement, we should expect differences to be similar between actual and counterfactual CNVs.

We illustrate the patterns in the data by plotting the difference between $distCNV_d^{k,real}$ coefficients for the proximity to the actual CNVs and the mean of each of the $distCNV_d^{k,fake}$ coefficients for proximity to the counterfactual CNVs. We denote the significance of each of the $distCNV_d^{k,real}$ coefficients relative to the counterfactual $distCNV_d^{k,fake}$ distributions. Crosses indicate coefficients that are above 95^{th} percentile of the counterfactual distributions, solid dots denote coefficients above the 90^{th} percentile, and hollow dots indicate coefficients below the

90^{th} percentile.

Figure A.2 shows the balance checks. The geographic and soil characteristics we consider are slope, elevation, percentage of east-facing grids, percentage of topsoil organic carbon, percentage of topsoil sodicity, an indicator for drainage is very poor, an indicator for soil is coarse, an indicator for soil is medium. We also examine important pre-resettlement variables: distance to the nearest urban center, distance to the nearest coast, the share of ethnic Chinese in 1947, and the natural logarithm of population density in 1947. We do not find any statistically significant differences on a wide range of characteristics and distances to the nearest CNVs, except for elevation and slope. Moving forward, we control for differences in elevation and slope throughout all our analyses.

Importantly, throughout all our analyses, we focus on exposure effects on surrounding Malay communities by conducting a "doughnut hole" analysis, whereby we exclude all polling districts that contains CNVs. We next turn to our main results.

5 Interethnic Proximity and Vote Shares for the National Front

5.1 Main results

OLS analysis. We begin by examining OLS results on the relationship between ethnic proximity and vote shares in Table 1. Panel A reports the results for the 2013 election and Panel B for the 2018 election, both measured at the polling district level. This table examines two types of elections—state (columns 1–3) and federal (columns 4–6). The dependent variable is the vote share for the ethno-nationalist coalition, the National Front, and the independent variables are distance bins from the CNVs. The omitted category among the independent variables is the 0–2 kilometers distance bin, which is directly adjacent to CNVs, thereby excluding the CNVs themselves. Consequently, the estimated coefficients for the 2–4, 4–6, 6–8, and 8–10 km distance bins reflect the vote share for the ethno-nationalistic coalition relative to the 0–2 km bin. Positive coefficients indicate higher support for the ethno-nationalistic coalition compared to the 0–2 km bin, while negative coefficients indicate the opposite.

In the 2013 election (Panel A), the results show that in both state and federal elections, the vote share for the ethno-nationalist coalition in polling districts located in the 2–10 km bins is approximately 6–9 percentage points (pp) higher than in polling districts within the 0–2 km bin (columns 3 and 6, with the full set of controls). This corresponds to a 11–16% higher vote share relative to the mean (56%) for polling districts farther from CNVs. In other words, these results indicate significantly reduced support for the ethno-nationalist coalition in areas close to CNVs (0–2 km), suggesting weaker ethno-nationalist sentiments in these areas, possibly due to greater interethnic contact with the Chinese, local economic development, or shared public goods near CNVs.

In the 2018 election (Panel B), the mean support for the ethno-nationalist coalition fell sharply from 56% to 39%, reflecting a 17 percentage point or 30% decline in average vote share. Despite this decline, the vote share for the ethno-nationalist coalition in polling districts within the 2–10 km bins remained approximately 7–11pp higher than in polling districts within the 0–2 km bin (columns 3 and 6, with the full set of controls). Given the lower average vote share in 2018, the relative difference in vote share between the 2–10 km bins and the 0–2 km bin increased to approximately 18–28%. This suggests that, while the national trend moved against the ethno-nationalist coalition, the spatial patterns of support persisted, with significantly lower ethno-nationalist sentiments in areas close to CNVs compared to farther away.

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 1, 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, we observe that polling districts within the 2–4 km distance bin show approximately 5pp higher support for the ethno-nationalist coalition relative to those in the 0–2 km bin, with effects tapering off at greater distances. Effects become statistically insignificant in the 8–10 km bin for both state and federal elections. A similar pattern can be observed for the 2018 election, with effect sizes ranging from 2-6pp across the 2–10 km bins and all effects significant at the 5% level.

These findings reveal the localized nature of political preference shifts induced by proximity to CNVs. The attenuation of effects at greater distances suggests that the mechanisms driving changes in political preferences—whether economic, social, or a combination of both —are strongest in areas immediately adjacent to CNVs. That is, nearly six decades after resettlement, ethnic majority Malays living near segregated, mono-ethnic Chinese villages continue to show ethno-nationalistic voting preferences, emphasizing the lasting and positive impact of interethnic proximity on contemporary political behavior.

5.2 Alternative Explanations and Robustness Checks: Voter turnout and Malay political preferences

Our results thus far suggest that Malays living closer to CNVs are less likely to vote for the ethno-nationalist coalition and we interpret this as possibly reflecting a shift in the political preferences of ethnic majority Malays. Here, we rule out two key alternative explanations. First, that voters living in greater proximity to CNVs might have turned out at a higher rate due to greater ethno-nationalist sentiments and racial threat (Enos, 2014). In this case, differences in ethno-nationalist vote shares would reflect differences in voter turnout rather than shifts in underlying political preferences of the broader electorate.

Second, that differences in vote shares might reflect changes in both Malay *and* Chinese political preferences. This would be the case if the ethnic composition of voters at the polling-district level varied with proximity to CNVs. In a nutshell, even though we conduct a "doughnut-hole" exercise and exclude all polling districts containing CNVs from our estimation sample, it is still possible that differences in vote shares are a result of more ethnic Chinese living in polling districts adjacent to CNVs. Qualitative fieldwork suggests that, upon marriage, sons of ethnic CNV villagers often move out of their parents' homes to setup their matrimonial homes in adjacent areas. Given that voting data is only available at the polling-district level, it is possible that differences in vote shares might be a result of higher ethnic Chinese shares in areas adjacent to CNVs that fall within the 0-2km distance bin.

To quantify the role of Malay vis-a-vis Chinese voters, one would need disaggregated data on voting behavior by ethnicity, which is not systematically available for this contemporary period. We proceed instead in three steps. First, we test for differences in ethnic composition across distance bins. Second, we further restrict our estimation sample to Malay-majority areas. Specifically, we restrict our estimation sample to, respectively, only polling districts with below-median Chinese voters and polling districts where the number of Malay voters is greater than Chinese voters. Third, following Becker and Woessmann (2009); Calderon et al. (2023), we quantify the role of Malay voters by assuming that all Chinese voters voted *against* the ethno-nationalist coalition and showing 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.

Throughout, 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. We focus on 2013 electoral results as ethnic shares of registered voters are only available in 2013.

Turnout. We examine the effects of distances to CNVs on turnout in the 2013 elections. In Figure A.3, effects are quantitatively small and statistically insignificant. Differences in vote shares do not appear to be driven by differential turnout rates.

Ethnic composition of voters. Panel A of Figure A.4 shows that areas located 0-2km from CNVs have a roughly 8.3pp (4.2pp; 6.3pp; 6.2pp) higher share of registered ethnic Chinese voters compared to areas that are located 2-4km (4-6km; 6-8km; 8-10km) from CNVs. Similarly, Panel B shows that there is a slightly lower share of around 4.0pp (0.9pp; 2pp) of registered ethnic Malay voters in areas located 2-4km of CNVs compared to those located 0-2km (4-6km; 6-8km; 8-10km) away. There is, however, little difference in the share of registered ethnic Malay voters between areas located 4-6km from CNVs.

Taken together, these results suggest that there are substantial differences in ethnic shares across polling districts within 0-2km of a CNV vis-a-vis those further away. Specifically, these results suggest that the 5pp increase in ethno-nationalist vote share in the 2-4km bin (Figure 6), could potentially be more than accounted for by the 8.3pp decrease in registered ethnic Chinese voters in the same distance bin (Figure A.4).

Addressing differences in ethnic composition: keeping only Malay-dominant polling districts. To that end, Figure A.5 shows that our results are robust to further restricting our sample to Malay-dominant polling districts only. Specifically, Panel A restricts our sample to polling districts containing a below-median number of Chinese voters and Panel B restricts our sample to polling districts where the number of registered Malay voters is larger than Chinese voters. Reassuringly, results remain positive and statistically significant across both specifications. In a further robustness check, Figure A.6 restricts our estimation sample to polling districts with $\leq$ 10%, 20%, 30%, 40%, and 50% ethnic Chinese voters and our estimates continue to be similarly positive and statistically significant across most distance bins.

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*).

As mentioned, results in Figure 6 and Figure A.4 suggests that polling districts within 0-2 kilometers of CNVs, experienced a decrease in vote share for the National Front that was greater than the increase in the share of registered ethnic Chinese voters. However, it is 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.

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 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.7 plots coefficient estimates where we vary ethnic Chinese voter turnout rates from 0.1 (10%) to 1 (100%). The first subfigure (on the top left), shows a clear decrease in $VotebyNonBN - ChiVoters_d$ in bins beyond 2 kilometers of CNVs when the Chinese turnout rate is assumed to be 0. Subsequent plots demonstrate that negative results on National Front vote shares persist up until the point where ethnic Chinese voter turnout rates exceed 0.5-0.6. Given typical ethnic Chinese turnout rates that rarely exceed 0.5 (Malay Mail, 2024), this exercise suggests that differences in the share of registered ethnic Chinese voters, are insufficient for explaining the entirety of the estimated decrease in the vote share for the National Front (or, conversely, the increase in the observed vote share of opposition coalition). Taken together, these results bolster our argument that a substantial number of ethnic majority Malay voters, that lived in close proximity to CNVs, voted against the ethno-nationalistic coalition.

5.3 Discussion of persistent impacts

A plausible explanation for the persistent impacts on vote shares is the close proximity of ethnic Malay communities to CNVs that have continued to exist even after the end of the Malayan Emergency in 1960. The persistence of CNVs can be largely attributed to their formalization after 1960 through the award of land titles to the resettled ethnic Chinese population (Nyce, 1973). The establishment of land ownership during and after the Malayan Emergency was driven by a combination of military, strategic, economic, and social considerations. As noted by Strauch (1981), the primary goal of resettlement was to disrupt the communist insurgents' support network. However, the provision of land titles was part of a broader effort to secure the cooperation of the relocated Chinese by encouraging them to view the villages as permanent homes rather than temporary camps. This policy provided the settlers with economic stability and a stake in the country's development, fostering loyalty to the state and reducing the appeal of communist ideologies.

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). Their persistence was further reinforced by better economic opportunities and the provision of public goods within these villages (Strauch, 1981).

To that end, Malay communities who lived in close proximity to these villages likely experienced persistently higher exposure to Chinese communities after the Malayan Emergency and reinstatement of free movement in 1960. In addition, a higher level of local amenities and development around these communities could also have led to greater economic benefits and contributed to a sustained shift in political preferences. We explore these possibilities in the next section.

6 Potential Mechanisms

In this section, we explore potential mechanisms that may have influenced the political preferences of the Malay population near resettled Chinese communities. The extant literature suggests that economic and social factors are important in shaping political behavior and attitudes (Algan et al., 2023; Autor et al., 2020; Panunzi et al., 2024). To that end, we examine the potential role of (i) contemporary economic development and urbanization, (ii) interethnic competition and economic complementarities, (iii) public goods provision, and (iv) interethnic contact and attitudes in shaping political preferences. Throughout, when presenting results on vote shares, we focus on polling-district level vote shares in the federal elections.[35]

Importantly, to examine the role of scale effects in urbanization, and interethnic competition and complementarities, we leverage a rich and under-utilized dataset of initial CNV characteristics collected by the Malayan Christian Council (Malayan Christian Union, 1958). This dataset contains CNV-level data on the initial population and occupational employment shares of ethnic Chinese residing in CNVs, and was collected 6-8 years after initial population resettlement (but where, throughout, there continued to be strict movement restrictions). We describe these variables and the analyses we conduct in the following sections.

In addition, we 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.

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, creating shared economic interests across ethnic groups. OLS results in Table A.2 present significant associations with nightlight luminosity, population density, and NDVI, in all distance bins and in both polling district-level and grid cell-level analyses. We next present results from the counterfactual exercise.

Grid-cell level. Figure 7 presents the results of our grid-cell level (1×1 km^2) analysis of economic activity neighboring CNVs. Panel A shows the results using 2014 nightlight luminosity data. We find that communities located within 2–4 km of CNVs, compared to those within 0–2 km, show significantly lower economic activity. Specifically, nightlight luminosity in the 2–4 km

[35] Results using vote shares in the state elections are largely similar and available upon request.

range is reduced by approximately 2.5 units ($p < 0.05$). These effects, of both similar and larger magnitudes, persist across the 4–6 km, 6–8 km, and 8–10 km ranges (all $p < 0.05$), indicating that contemporary economic activities are more local and concentrated around CNVs.

Panel B examines effects on NDVI. NDVI measures vegetation density using satellite sensors: dense vegetation absorbs more red light and reflects more near-infrared light, resulting in higher NDVI values. We find that areas further from CNVs (e.g., beyond 2 km) have higher NDVI values, approximately 0.03 units greater than those within the 0–2 km radius ($p < 0.05$). In other words, areas beyond 2 km rely 23% more on agriculture than areas in 0-2 km radius of CNVs, which could also explain lower levels of economic activity reported in Panel A. These effects, of similar and larger magnitude, persist across the 4–6 km, 6–8 km, and 8–10 km ranges (all $p < 0.05$), supporting the interpretation that areas further away remain more rural and less urbanized.

Panel C reports the effects on population density, a key driver of economic activity and urbanization. We find that areas 2–10 km away from the 0–2 km radius have significantly lower population density, with approximately 250 fewer people on average, all statistically significant at the $p < 0.05$ level. This reduction in population density further emphasizes the concentration of economic activity and urbanization in areas closer to CNVs.

In all, these patterns suggest that areas distant from CNVs remained predominantly agricultural and less urbanized, while areas within the 0–2 km radius of CNVs appear to have experienced very localized structural transformation. The observed spatial distribution of economic activity aligns with the presence of agglomeration benefits around CNVs (Duranton and Puga, 2004, 2020). As economic activity increased near CNVs, resulting improvements in living standards and opportunities may have encouraged Malays to support political parties that prioritize ethnic unity and collaboration over ethno-nationalist political agendas. To dig deeper into this mechanism, we next examine economic outcomes at the polling district level, which provides both a higher level of aggregation and aligns more closely with our unit of analysis for voting outcomes.

Polling-district level. Polling district-level analyses are useful for two reasons. First, it allows us to directly compare economic and political outcomes, as voting data is only available at the polling district-level. Second, it allow us to test whether the observed grid-cell level economic benefits near CNVs extend to Malay populations throughout the entire polling district or are confined to these localized pockets.

Figure 8 reports results. Panel A and Panel C show that effects on nightlights and population density largely dissipate when aggregated to the polling district level, underscoring the potentially highly localized economic benefits of CNVs. In particular, we observe a significantly large dip in nighttime light intensity in the 2-4 km bin. This is consistent with spatial reallocation of economic activity toward the core leading to limited spillovers or even declines

in economic activity in the periphery.

Panel B, however, shows that there continues to be a significant reduction in NDVI near CNVs at the polling district level (relative to 2-4km and 6-8km bins). This is consistent with greater urbanization around CNVs and the presence of agglomeration forces even in the presence of limited spatial spillovers.

Taken together, the contrast between grid-cell and polling-district level results suggest that highly localized agglomeration benefits and spatial reallocation are the most likely explanations for observed spatial patterns in the exposure effects of CNVs on economic outcomes. In particular, the relative absence of economic effects at the polling-district level does not rule out economic development as a mechanism for changes in political preferences. Rather, it suggests a nuanced relationship between interethnic proximity and economic development that depends crucially on how economic activity and development are distributed and experienced by Malays *within* polling districts.

To that end, we next turn to a heterogeneity analysis which provides supporting evidence that localized benefits in core regions (immediately proximate to CNVs) might partially explain observed differences in vote shares.[36]

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 in 1958. The intuition is that a larger population size of CNV villagers could bring about stronger agglomeration benefits to surrounding Malay populations and possibly attenuate political grievances. Conversely, a larger CNV population size could trigger the saliency (of interethnic competition) and lead to greater grievances against the ethnic Chinese that are expressed at the ballot box (Glaeser, 2005).[37]

We present OLS results, followed by estimates from our counterfactual analysis. Table A.3 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.

Figures 9 and Figures 10 present results from our counterfactual analysis. At the polling district-level, we observe weak effects on surrounding areas exposed to CNVs with a below-median initial population (see Figure 9). Panel A shows that the effects on vote shares remain relatively flat across most distance bins, except for the 2-4 km bin where the vote share for the ethno-nationalist coalition in this bin is approximately 3.8 units (6.3%) higher ($p < 0.05$). Similarly, Panel B shows that the effect on nightlight luminosity is moderately significant only in the 2-4 km bin ($p < 0.1$). Panel C shows that the effects on NDVI are mostly insignificant (except for the 6-8 km bin), and Panel D shows that the effects on population density remain

[36] Ideally, we would test this using grid-cell level data on vote shares but such data does not exist.

[37] Ideally, we would like to construct measures of historical ethnic polarization and segregation but we do not have data on the historical population of Malay villages.

relatively flat across most distance bins. These patterns are broadly consistent with our grid-cell level results (see Figures A.8).

Turning to our above-median sample, the effects on vote shares are broadly similar to our pooled sample but statistically more significant than those with the below-median sample (Panel A of Figure 10). That is, political effects in our baseline analysis appear to be concentrated in areas that were exposed to CNVs with a larger initial Chinese population size. Panel B shows that nightlight luminosity in the 2–4 km and 4-6 km ranges are reduced by about 2.3 units (11.6% decrease compared to the 0–2 km range, $p < 0.05$). Again, these effects are slightly stronger than those in the below-median sample. Panels C and D show few differences in NDVI and contemporary population density. These patterns are broadly consistent with grid-cell level results (see Figure A.9)

Taken together, the statistically significant and similar effects of proximity to CNVs in both the below- and above-median samples indicate that the presence of CNVs consistently benefits local Malay populations, irregardless of initial population size. Moreover, these effects are slightly stronger in the above-median sample, providing further support for the presence of localized agglomeration benefits and suggesting that racial threat is unlikely to be a key mechanism in our setting.

The lack of substantial differences in the *magnitude* of estimated effects, however, suggests that additional mechanisms—beyond localized agglomeration effects—could also be at play in driving observed economic benefits for local Malays. Motivated by this, the next subsection explores the role of interethnic competition and complementarities by exploiting initial heterogeneity in cross-sectoral, CNV occupational employment shares.

6.2 Interethnic Competition and Complementarities

On the one hand, interethnic proximity could lead to more negative political attitudes if economic interactions are largely driven by competition in production and labor markets (Horowitz, 2000). On the other, proximity could lead to more positive political attitudes if interactions are driven by interethnic complementarities and specialization. This could, in turn, improve productivity, social cohesion, and lead to greater tolerance toward out-groups (Alesina and La Ferrara, 2005; Jha, 2013). Building on these arguments, we investigate the extent to which the economic impact of CNVs on local Malays was (i) shaped by interethnic competition vis-a-vis economic complementarities and (ii) how these forces might have influenced contemporary voting behavior.

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 initial CNV-level occupational shares in 1958 (Malayan Christian Union, 1958). Specifically, we test for heterogenous 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.[38]

OLS analysis. We present OLS estimates followed by counterfactual analyses. Table A.4 presents OLS results including an interaction term for whether a CNV had an above-median share of workers in agriculture. These results suggest a clear role for interethnic competition and complementarities.

Interethnic competition in agriculture—Malay communities within 0-2km of agricultural CNVs (lower tin and rubber employment) exhibit higher ethno-nationalist vote shares, while economic indicators remain largely unchanged. This suggests that competition between Malay farmers and Chinese agricultural workers may have reinforced ethnic-based political preferences.

Economic complementarities in rubber and tin—Malay communities within 0-2km of non-agricultural CNVs (higher tin and rubber employment) exhibit lower ethno-nationalist vote shares, and both nightlight luminosity and population density are higher in areas immediately surrounding these CNVs. This suggests that economic benefits from these Chinese-dominant industries might have led to more moderate political attitudes among Malays.

Overall, these patterns indicate that interethnic competition in agriculture may have fueled ethnic polarization, while economic complementarities in non-agricultural sectors promoted better economic outcomes and more inclusive political preferences.

Counterfactual analysis. To further validate OLS results on interethnic complementarities, we conduct a counterfactual analysis using CNVs with below- (Figure 11) and above-median initial employment (Figure 12) in rubber and tin.

We find few differences in vote shares in the below-median sample (Panel A of Figure 11) but a negative and statistically significant impact on the vote shares of polling districts immediately

[38]Grid-cell level results are largely consistent with polling district level findings and would be made available upon request.

adjacent to CNVs in the above median sample ($p < 0.05$) (Panel A of Figure 12). Specifically, we observe a 8p.p. 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 (5.2p.p. in Panel B of Figure 6) and the above-median initial CNV population size sample (5.8p.p. in Panel A of Figure 10), suggesting a substantial role for interethnic complementarities in driving observed results on voting behavior.

In addition, while Panel B of both figures show similar impacts on nightlight luminosity, Panel D of Figure 12 (Figure 11) indicates a greater (lower) concentration of population around CNVs that had initially higher (lower) rubber and tin employment, further underscoring how economic benefits from interethnic complementarities between local Malay and Chinese populations might be a key driver behind lower vote shares for the ethno-nationalist coalition.

Figures A.10 and A.11 present 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 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.10), 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.11). 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.

Lastly, we turn back to results based on below and above-median rubber and tin employment shares to discuss the persistence of differences in vote shares in the presence of quantitatively small differences in contemporary economic development (Figures 11 and 12). Specifically, Figure 12 shows large differences in vote shares alongside modest differences in contemporary economic development: impacts on nightlight luminosity are only slightly larger in the 2-4km bin in the above-median sample (Panel B) and there are few differences in NDVI across both samples (Panel C). The only exception is in Panel D where we see statistically significant decreases in population density in areas further away from CNVs (4-6 km ($p < 0.05$) and 8-10 km bins ($p < 0.1$)) in the above-median sample, suggesting a concentration of economic opportunities immediately around CNVs.

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 provide suggestive evidence for this in Section 6.4. 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.

6.3 Public Goods Provision

We explore the impacts of CNVs on public goods provision, focusing on the contemporary density of schools, teachers, roads, and hospitals at the polling district level, offering a comparable analysis that aligns with observed changes in polling-district level vote shares. We find positive effects on the provision of schools and roads.

Public goods provision may have an important role in improving access to essential services, increasing neighborhood quality, and uplifting the overall standard of living. In turn, political preferences could depend on the availability of public goods where voters reside (Calabrese et al., 2006). In particular, by promoting goodwill and equitable access, these investments could shift political preferences, encouraging support for parties that prioritize interethnic collaboration and inclusion.

Schools and teachers. Table A.5 reports OLS results. Columns 1 and 2 examine the number of contemporary primary and secondary schools in polling districts within varying distances from CNVs. The results indicate mixed effects. For primary schools (column 1), the coefficient is not statistically significantly different from zero ($p > 0.10$), suggesting no measurable differences in primary school accessibility based on proximity to CNVs. However, for secondary schools (column 2), the coefficients are generally negative and statistically significant, particularly in the 4–6 km, 6–8 km, and 8–10 km bins. This implies that secondary schools are relatively fewer in these further distance bins.

Columns 3 and 4 explore the number of teachers in primary and secondary schools, revealing similar patterns. The results show large (albeit statistically insignificant) decreases in the number of secondary school teachers in polling districts farther from CNVs (4–6 km, 6–8 km, and 8–10 km ranges). This is consistent with the findings for secondary school density and further supports the hypothesis that public goods provision related to education was concentrated closer to CNVs. Columns 5 and 6 further examine student-teacher ratio as a proxy measure for school quality. Here, we see that school quality in polling stations further from CNVs is poorer, but only in primary schools (column 1). While the OLS results for the number of schools and teachers hold in the counterfactual analysis (Figure 13), results on school quality are not robust.

The observed patterns in school and teacher density suggest that proximity to CNVs facilitated greater access to secondary education for local Malay populations, particularly in the 0–4 km range. These findings have two key implications for political preferences. First, greater access to schools may have directly improved educational opportunities for Malays, promoting goodwill toward parties that support local development and collaboration. Second, improved human capital through better access to secondary education could have reshaped political attitudes of Malays, leading to more inclusive political behavior. This aligns with existing research linking education to more progressive and inclusive political preferences (Glaeser et al., 2007).

Roads. Column 1 of Table A.6 presents the impacts on road density (length of roads per area in km^2) near CNVs. The results show consistently negative and significant coefficients across all distance bins relative to the 0–2 km range (all $p < 0.01$), indicating higher road density in polling districts closer to CNVs. This pattern persists across all specifications, with the magnitude of the effect declining as the distance from CNVs increases. We also see a similar pattern in Panel A of Figure 14 where, in particular, the difference in the 2-4km distance bin is substantively lower than that in the 0-2km bin. These results suggest that areas near CNVs benefited from greater road infrastructure development compared to more distant areas.

Higher road density could have impacted political preferences in a number of ways. First, improved road networks could have facilitated localized economic activity by reducing transportation costs and increasing market access for both Chinese and Malay communities, encouraging integration into the local economy. Such changes could shift vote shares by exposing communities to new economic opportunities and reducing reliance on ethno-nationalist policies and would be consistent with highly localized economic benefits or interethnic complementarities uncovered in Section 6.1. Second, better road infrastructure might have directly facilitated greater interethnic interactions between Malay and Chinese communities by increasing mobility and communication. Such social connectivity could have reduced de facto segregation, increased social cohesion, and built trust between the two groups, potentially shifting Malay political preferences toward more inclusive and collaborative policies. We find strong evidence for this interpretation in Section 6.4.

Health facilities. Finally, we examine the OLS impacts of CNVs on the public provision of health facilities in 2022, the only year in which we have the most comprehensive data for health public goods. We focus on the number of clinics and hospitals (Table A.6). The analysis shows no significant impacts on the number of clinics (column 2) and hospitals (column 3). Results from counterfactual analysis are similar (Panels B and C of Figure 14).

Although not significant, a simple visual inspection of the combined effect on health facilities (i.e., aggregating clinics and hospitals) in our counterfactual analysis, reveals positive effects in the 0-2km relative to all 2–10 km distance bins. One plausible explanation would be

the strategic placement of health facilities to serve larger populations across multiple communities. That is, clinics and hospitals in these areas may have been sited to maximize accessibility for both the resettled Chinese populations and the surrounding Malay-majority communities, ensuring broader geographic coverage.

The evidence here suggests that the availability of health facilities is unlikely to be a mechanism for shaping political preferences. However, one crucial limitation is that our data on health facilities is from 2022, rather than from pre-2013 or pre-2018.

6.4 Attitudes and Behaviors Toward Chinese

Having observed some aggregated economic benefits, we now examine micro-level mechanisms related to Malays' attitudes and behaviors toward the Chinese. We use novel primary survey data to assess changes in attitudes and behaviors resulting from interethnic interactions facilitated by Chinese resettlements. This is important given that economic and political effects could have been shaped by varying degrees of interethnic exposure across locations. For example, ethnic Malays living closer to CNVs may have experienced more frequent and sustained interethnic contact with the ethnic Chinese minority. This proximity could have facilitated more interactions, affected social capital, and influenced their views.

Primary survey. To better understand the micro-level mechanisms, we conducted an in-person survey in partnership with one of Malaysia's top survey firms, *Ilham*. We conducted a primary survey because there are no existing datasets that capture micro-level measures of interethnic attitudes and behaviors in this context. We collected detailed individual-level data from December 2024 to January 2025 in both 'treated' and 'control' villages across Johor—one of the two Malaysian states with the largest number of CNVs–in two districts (Batu Pahat and Kluang).[39] Our survey targeted about 350 randomly selected Malay males aged 18 and above from 19 Malay Villages. We do not have female respondents because the pilot revealed major logistical challenges in recruiting female respondents given the predominantly Muslim survey sample. The sample was stratified by age to ensure representation of both older Malays (60+), who directly experienced colonial resettlement, and younger Malays (under 60), who were born after it. The survey targeted four key dimensions:

1. Trust—general trust in Chinese, and willingness to entrust a Chinese person with childcare, which measures situational trust.

2. Openness—willingness to have Chinese neighbors, and openness to intermarriage for themselves or their children/grandchildren.

[39]The data presented here is part of an ongoing data collection effort (about 3,000 respondents across two states) that will be completed in June 2025. The survey is funded by the Hong Kong Research Grants Council.

3. Interactions—frequency of interactions with Chinese people in primary and secondary schools, and workplace.

4. Friendship—whether they have close Chinese friends, and the number of Chinese contacts saved in their mobile phone.

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 variables of elevation, slope, market access to pre-existing villages, and distance to the nearest pre-existing Malay village, to select a single counterfactual (control) resettlement site. These four variables are our best available proxies for military site selection suitability following military archival documents.[40] Importantly, just as in our counterfactual exercise, we restrict the set of possible counterfactual sites only to grid-cells that are located along main roads. We perform matching at the 1 km $\times$ 1 km grid-cell level to ensure comparability. Second, we select a maximum of two treatment and two control Malay villages based on their proximity (within 0-2km) to each real and counterfactual resettlement site. Third, after doing the matching, we collected our individual-level data using in-person, door-to-door surveys in both treatment and control villages as described above.

To estimate the effects of proximity on the four key dimensions of attitudes and behaviors, we estimate the following equation:

$$Y_{iv} = \alpha + \beta_1 T_{iv} + \beta_2 Age \geq 60_{iv} + \beta_3 (T \times Age \geq 60)_{iv} + \theta_{enum} + \theta_c + \epsilon_{iv} \tag{4}$$

Let i denote an individual and v a village. Y_i is the outcome of interest. T_{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. $Age \geq 60$ is an indicator for respondents that are 60 years or older. θ_{enum} takes the value of 1 (0) if a enumerator was from Peninsular Malaysia (East Malaysia). θ_c denotes nearest CNV fixed effects. We cluster standard errors two-way: by whether an individual is aged greater than 60 and the village in which an individual resides.[41]

Results. Table 2 reports our main survey results on trust (columns 1-2), openness (columns 3-4), interethnic interactions (columns 5-7), and social capital (columns 8-9) of ethnic majority Malays towards ethnic minority Chinese. Since the older cohort (60+) was alive during the

[40] 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: Site Selection Criteria for details.

[41] Results are largely similar when we control for pre-determined Malay village-level controls. We do not control for individual-level covariates such as education and income as these are potentially endogenous outcomes of interethnic proximity to ethnic Chinese.

colonial resettlement period, while the younger cohort (18–59) was not born yet, we expect these two groups to behave differently. Moreover, the older cohort in treatment villages (i.e., those near real CNVs) would have had greater direct exposure to CNVs and interactions with the Chinese than their younger counterparts. On the other hand, both younger and older cohorts in control villages (i.e., near counterfactual resettlement sites) likely had minimal interethnic exposure, allowing us to better isolate the effects of proximity to CNVs.

We begin by analyzing results for the control group (Table 2, coefficients on $Age \geq 60$), which reflects attitudes and behaviors toward the Chinese in the absence of exposure to CNVs. The results suggest that older Malays in control villages hold lower situational trust toward the Chinese compared to the younger cohort (column 2), though there is no significant difference in general trust (column 1). Additionally, older Malays are significantly less open to interethnic relationships, both in terms of willingness to have Chinese neighbors (column 3) and acceptance of interethnic marriage (column 4). Interethnic interactions also appear to differ between the older and younger cohorts, primarily interactions in the secondary school (column 6), suggesting that opportunities for direct contact with the Chinese were also different across generations in control villages. Finally, regarding social capital, older Malays report fewer connections to Chinese people (column 8), yet the number of close friendships with Chinese is similar across age cohorts (column 9). Overall, in the absence of direct exposure to CNVs, older Malays tend to show more conservative attitudes and behaviors toward the Chinese compared to younger Malays. This pattern provides an important baseline for understanding how proximity to CNVs may have shaped interethnic attitudes in treatment villages, which we explore next.

We next examine the same patterns between older and younger cohorts but focus only on the treatment villages. In comparison to the control villages, we find that the differences in trust, openness, and social capital between older and younger Malays either diminish (column 3) or dissipate (columns 1, 4-9) in treatment villages (see 'Estimate (age$\geq$ 60+Int.)' and 'P-value (age$\geq$ 60+Int.)' rows at the end of the table, which come from adding the interaction term with the $age \geq 60$ dummy). This implies that sustained exposure to Chinese communities improved attitudes (e.g., openness) and behaviors (e.g., trust and friendship) of the older cohort, who would otherwise hold conservative opinions. More importantly, the difference between the two differences across cohorts (within treatment and control), which is captured by the interaction term, shows that the cohort gaps we observe in treatment villages are statistically different from that in control villages, especially in terms of situational trust (column 2), openness to having Chinese neighbors (column 3), interactions in secondary schools (column 6), and social connections (column 8).

These results align with the localized patterns of reduced ethno-nationalist vote shares near CNVs, suggesting that interethnic exposure encourages political moderation through two key channels. First, sustained interethnic contact appears to foster trust-building and social integration, particularly among older cohorts who would otherwise hold more conservative views.

This highlights the role of prolonged exposure in breaking down stereotypes and promoting openness to outgroups. Second, relating to our discussion in Section 6.2, the persistently positive attitudes we observe here may also be rooted in historical interethnic complementarities. These complementarities could have established a foundation of economic interdependence and goodwill, leading to a lasting impact on political preferences even as initial economic benefits diminished. Together, these channels underscore the complex and enduring impacts of interethnic proximity on social cohesion and political moderation.

7 Alternative Mechanisms

In this section, we rule out two alternative mechanisms: (i) That effects are driven by exposure to new people and not ethnic Chinese per se. (ii) That economic effects and effects on public good provision are driven by policy changes made by the civic-nationalist coalition.

7.1 Exposure to New People

The observed effects discussed in Sections 5 and 6 may not be driven by exposure to the ethnic Chinese community per se, but by exposure to new people, which could bring new ideas, opportunities, or norms that shape political preferences (Balietti et al., 2021; Chetty et al., 2016). To test this, we replicate our analysis using the resettlement of ethnic Malays into Malay New Villages (MNVs). During the Malayan Emergency, the majority of British resettlement involved ethnic Chinese. However, in rural areas where Communist activity was especially high, a small number of ethnic Malays were also resettled into MNVs (Dobby, 1952; Humphrey, 1971). MNVs were largely provided with the same amenities as CNVs.

If the effects arose from exposure to new people rather than ethnic Chinese communities, we would expect effects of MNVs to be similar to that of CNVs. The results in Table A.7 show that the estimates of the impacts on the ethno-nationalistic coalition vote share and ethnic shares (of Chinese and Malays) are statistically insignificant throughout ($p > 0.10$). Furthermore, Table A.8 indicates that the effects on nightlight luminosity (column 1) and population density (column 3) are also statistically insignificant (all $p > 0.10$). Although there are some positive effects on NDVI (column 2) implying greater rurality in areas distant from MNVs.

These analyses suggest that exposure effects to CNVs are unlikely to be driven by exposure to new populations. Instead, they are likely to be a consequence of proximity to Chinese communities. The absence of significant changes in population density at the polling district level (Section 6.1) strengthens this interpretation — effects on vote shares are unlikely to be driven by exposure to more people, whether due to exogenous shocks or broader demographic shifts within these regions.

7.2 Policy Changes Made by the Civic-Nationalist Coalition

Another alternative mechanism that might have influenced observed differences in vote shares is the role of local politicians and political campaigns (Manacorda et al., 2011). Local politicians from the civic-nationalist coalition could have shifted political preferences by targeting funds and projects towards Malays living in close proximity to CNVs. Such interventions in market conditions or facilitation of economic activities could have increased the appeal of the civic-nationalist coalition and generated trust and cooperation. This would be consistent with observed positive economic impacts and a decline in support for the ethno-nationalist party but could happen only if the civic-nationalist coalition won a state or federal constituency seat.

Table A.9 provides evidence against this mechanism. While proximity to CNVs appears to influence voting behavior, this influence does not translate into major electoral victories for parties in the civic-nationalist (pro-diversity) coalition in either the 2013 (columns 1-2) or 2018 (columns 3-4) elections. We observe a slight reduction in the probability of a civic-nationalist coalition electoral victory but effects are largely statistically insignificant. Table A.10 further shows that there were few differences in political alignment — the civic-nationalist coalition was not more likely to achieve an electoral victory at both the federal and state-level. This is an important result given that political alignment across both levels are crucial for politicians to implement projects given funding allocation rules that favor the government of the day (Ostwald, 2017).

These results suggest that observed increases in public goods provision in Section 6.3, cannot be attributed to policies enacted by civic-nationalist parties nor political cycles and fluctuations between policies implemented by opposing parties (Cheremukhin et al., 2024). Instead, these improvements may reflect greater grassroots-level collaboration between Malay and Chinese communities (consistent with Section 6.4) or strategic allocations in public investment by the ethno-nationalist coalition. Lower winning margins might have incentivized the ruling ethno-nationalist coalition to allocate resources to these areas to secure re-election (Bardhan and Mookherjee, 2010).

Taken together, interethnic proximity appears to have shifted political preferences at the intensive margin (i.e., reducing support for ethno-nationalist parties, as observed in Section 5), but these shifts were not large enough to engender coalition victories at the extensive margin. Rather, improvements in economic conditions appear to have fostered political moderation through sustained interethnic interactions and social cohesion.

8 Conclusion

This paper uses disaggregated grid-cell data and novel, polling district-level data to study the long-run economic and political effects of a mass forced resettlement program. Between 1949

and 1952, about 500,000 ethnic Chinese were forcibly resettled to New Villages, resulting in sharp differences in interethnic proximity across Peninsular Malaysia. Nearly all of these villages still exist today, allowing us to leverage persistent spatial variation in out-group proximity to study contemporary political and economic effects on receiving areas.

We find both positive political and economic effects. Nearly half a century later, areas in greater proximity to Chinese resettlement sites exhibit lower vote shares for the ethno-nationalistic coalition. These effects are unlikely to be driven by differences in turnout or the ethnic composition of voters. Examining economic mechanisms, we find positive, hyper-local economic impacts in areas immediately surrounding resettlement sites (at the grid-cell level). We also find evidence of urbanization, interethnic complementarities, and improved public goods provision. Beyond economic effects, proximity promoted greater trust and social capital between older and younger cohorts, reducing biases among older individuals who, in the absence of such exposure, tend to hold more conservative views toward ethnic minority Chinese. These transformations in economic and social conditions likely contributed to the marked shift in vote shares and political preferences away from the ethno-nationalist coalition.

While our findings suggest that interethnic proximity can have positive long-run effects, they may not generalize to all settings. In particular, urban ethnic enclaves often remain economically disadvantaged due to systemic barriers such as residential segregation, discrimination in labor markets, and limited access to high-quality education and public services. In contrast, our largely rural setting may have facilitated greater interethnic interactions in shared public spaces, thereby promoting trust and integration. This aligns with prior evidence on how social capital is built through everyday intergroup contact (Putnam, 2015).

Therefore, our findings contribute to broader debates on the long-run effects of resettlement and intergroup proximity on economic and political outcomes. We show that even when relocation occurs across distinct, segregated communities, it can generate positive effects on political outcomes under two key conditions: (1) economic benefits for the 'native' population and (2) sustained, casual interactions in shared public spaces that facilitate social integration.

These results have important implications for resettlement policies, particularly in light of ongoing global displacement from conflict and climate change. The persistence of ethnic minority resettlement sites provides a rare opportunity to study the long-term consequences of forced migration and intergroup proximity. By leveraging high-resolution spatial and electoral data, we offer, to the best of our knowledge, the first causal estimates of how resettlement influences the co-evolution of economic and political development. Our findings suggest that under the right conditions, resettlement policies can promote peaceful interethnic relations and economic benefits, ultimately shaping long-run political preferences in ways that reduce ethnic divisions.

References

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

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

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., Dalvit, N., Do, Q.-A., Le Chapelain, A., and Zenou, Y. (2023). Friendship networks and political opinions: A natural experiment among future french politicians. *CESifo Working Paper*.

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.

Autor, D., Dorn, D., Hanson, G., and Majlesi, K. (2020). Importing political polarization? the electoral consequences of rising trade exposure. *American Economic Review*, 110(10):3139–3183.

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. (2017). Corruption and development: a review of issues. *Political Corruption*, pages 321–338.

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 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.

Borusyak, K. and Hull, P. (2023). Nonrandom exposure to exogenous shocks. *Econometrica*, 91(6):2155–2185.

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.

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.

Cheremukhin, A., Golosov, M., Guriev, S., and Tsyvinski, A. (2024). The political development cycle: The right and the left in people's republic of china from 1953. *American Economic Review*, 114(4):1107–1139.

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.

Dasgupta, A. and Ramirez, E. (2020). Explaining rural conservatism: Political consequences of technological change in the great plains. *American Political Science Review*, pages 1–23.

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

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

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

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.

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.

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

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. and Kahn, M. E. (2004). Sprawl and urban growth. In *Handbook of Regional and Urban Economics*, volume 4, pages 2481–2527. Elsevier.

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

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.

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

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

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

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.

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.

Lim, G. J. (2022). Local elites, land rents and incentives for development: Evidence from village chiefs in indonesia.

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.

Migdal, J. S. (1988). *Strong societies and weak states: state-society relations and state capabilities in the Third World*. Princeton University Press.

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

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

Nunn, N. (2009). The importance of history for economic development. *Annual Review of Economics*, 1(1):65–92.

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.

Panunzi, F., Pavoni, N., and Tabellini, G. (2024). Economic shocks and populism. *The Economic Journal*, page ueae042.

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.

Putnam, R. D. (2007). E pluribus unum: Diversity and community in the twenty-first century the 2006 johan skytte prize lecture. *Scandinavian Political Studies*, 30(2):137–174.

Putnam, R. D. (2015). Bowling alone: America's declining social capital. In *The city reader*, pages 188–196. Routledge.

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

Rao, M., Eberhard, J., and Bharadwaj, P. (2022). Towns and rural land inequality in india. Technical report, Working Paper.

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*.

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.

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

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.

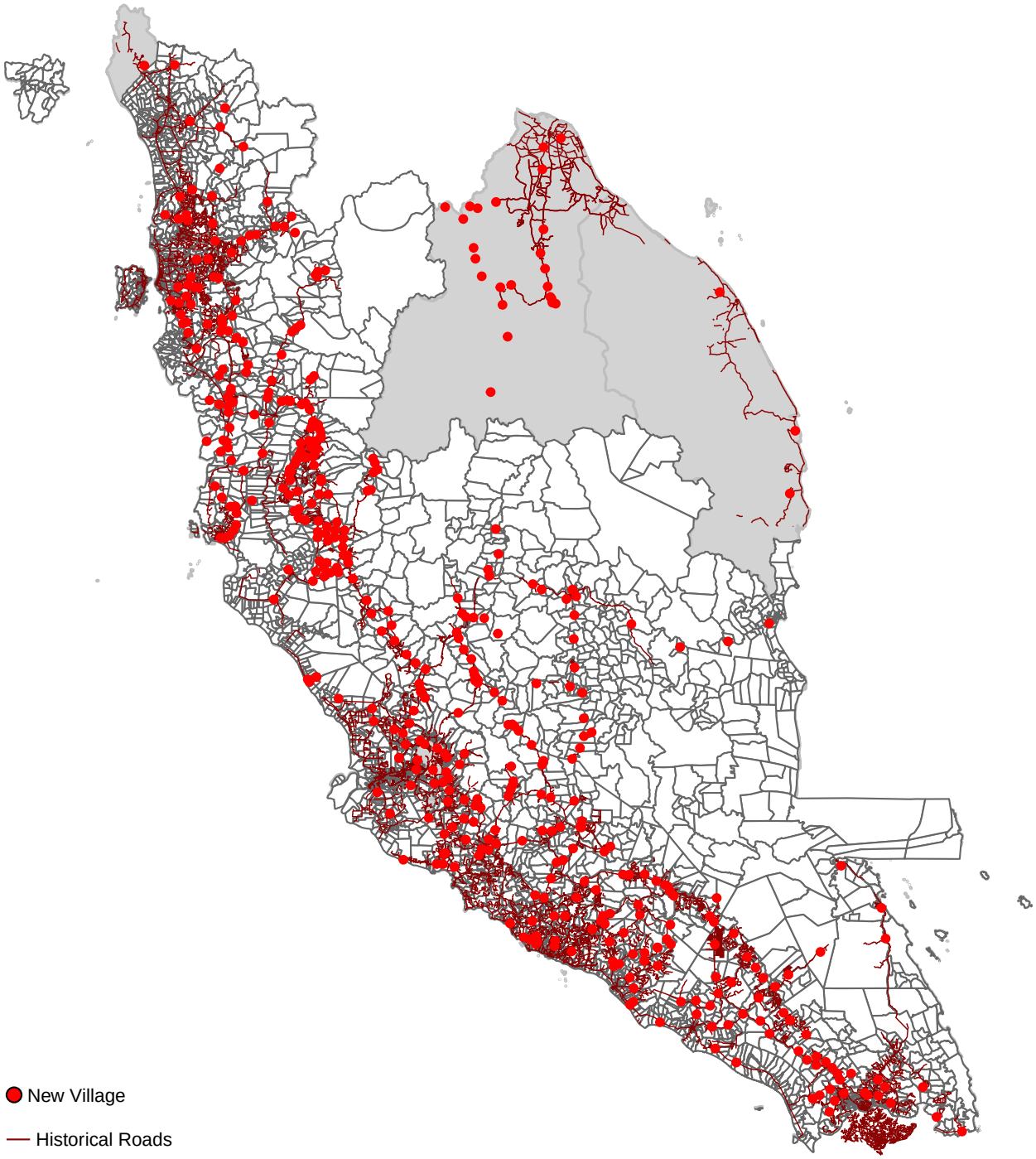
Ting, C. P. (1976). *The Chinese in Peninsular Malaysia: A study of race relations in a plural society*. PhD thesis, University of Warwick.

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.

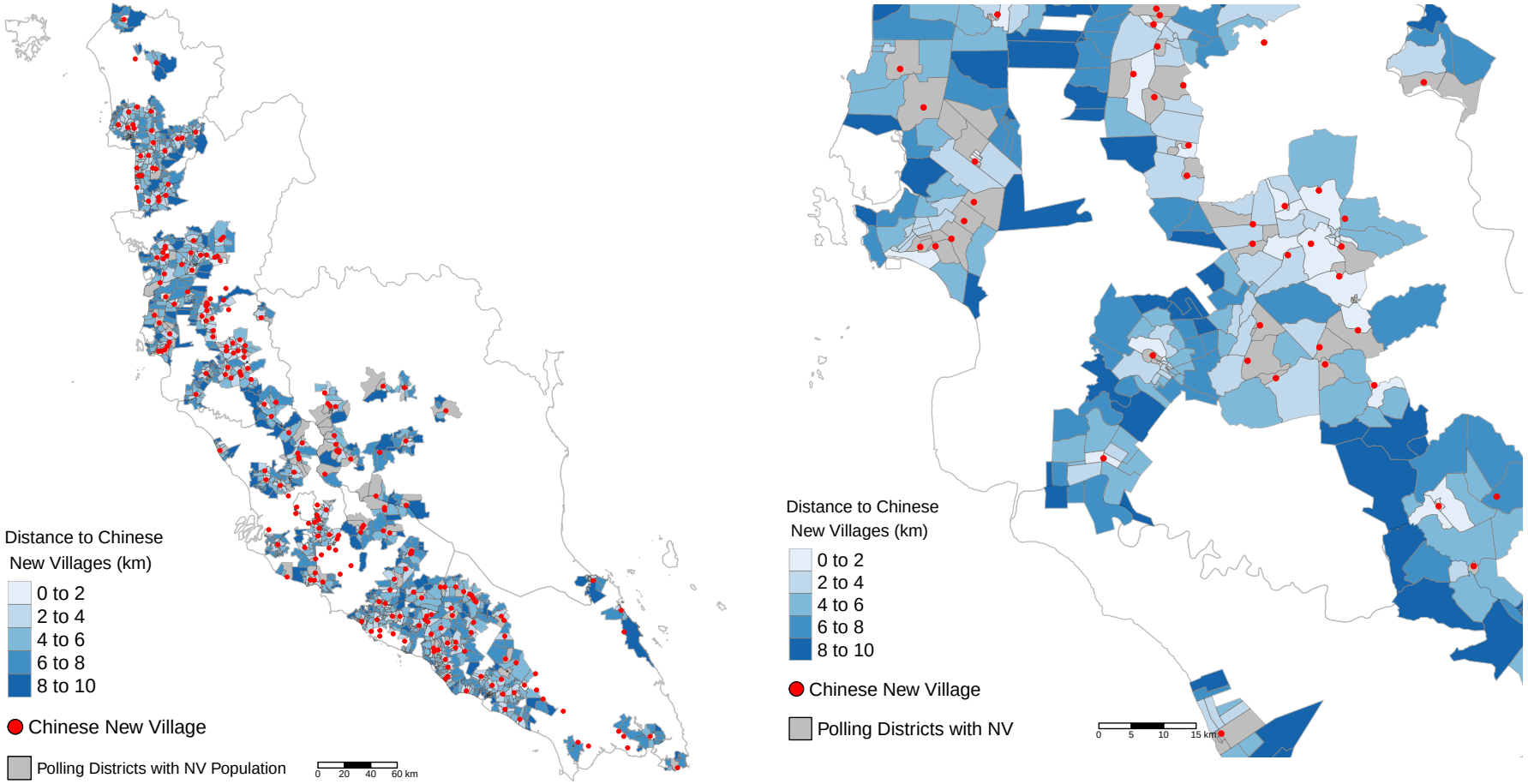
Wang, G. (1959). *A short history of the Nanyang Chinese*. Number 13. D. Moore.

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.

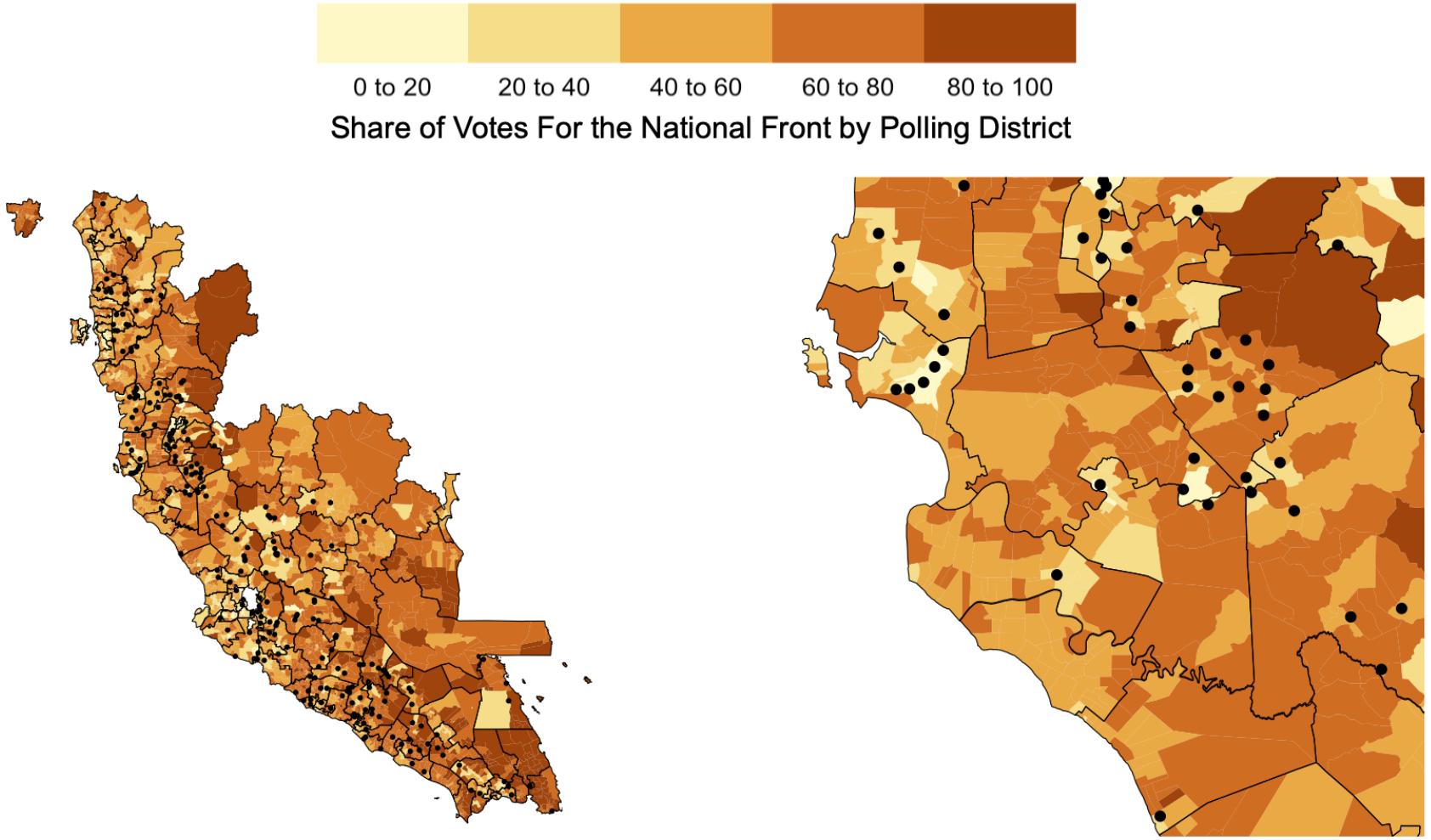
Figure 2
Interethnic Proximity: Distances to the nearest Chinese New Village at the Polling District-Level



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.

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

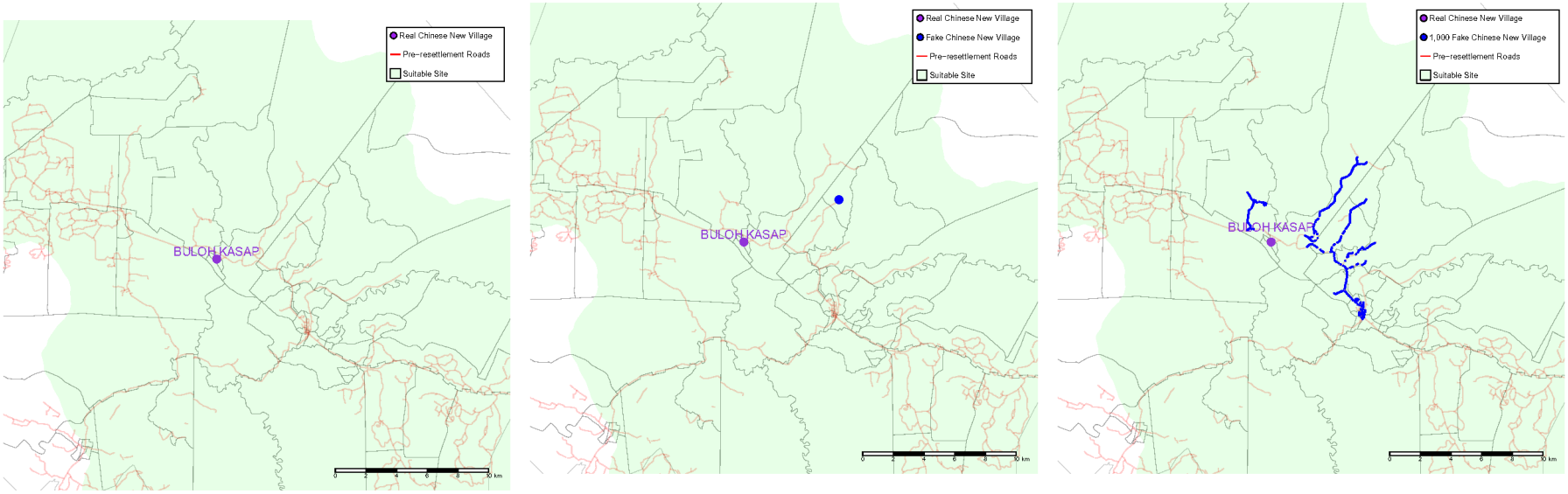


Panel A: Peninsular Malaysia

Panel B: Seremban and Port Dickson

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.

Figure 4
Counterfactual Chinese New Villages Example



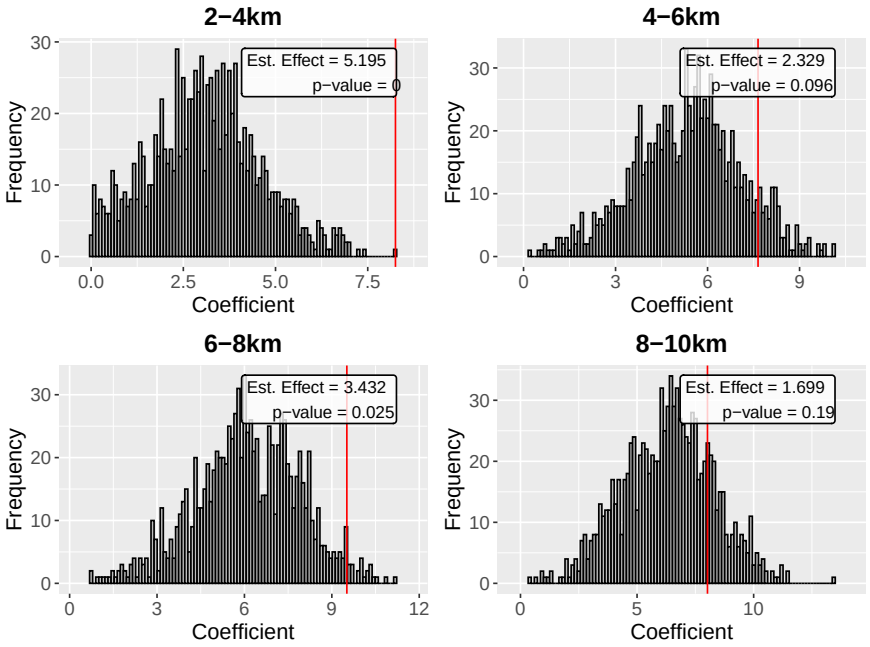
Panel A: A Real New Village

Panel B: A Counterfactual New Village

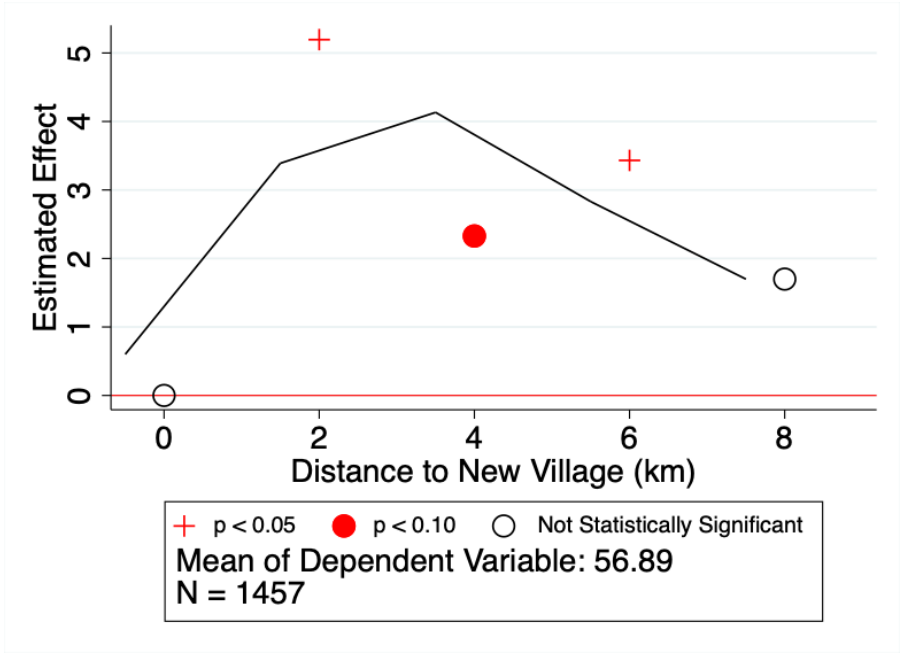
Panel C: 1,000 Counterfactual New Villages

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

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



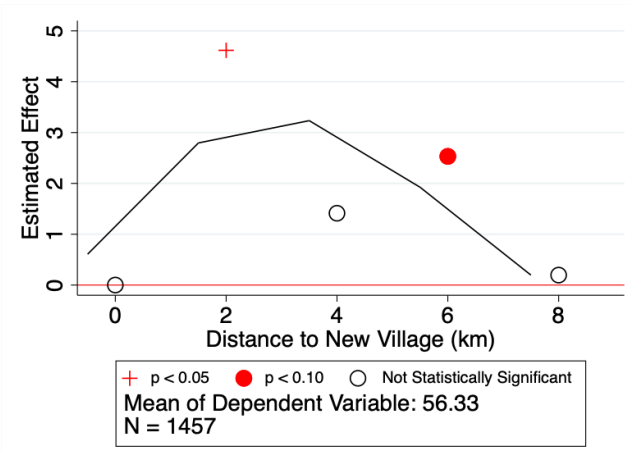
Panel A: Independent Shifts: Counterfactuals



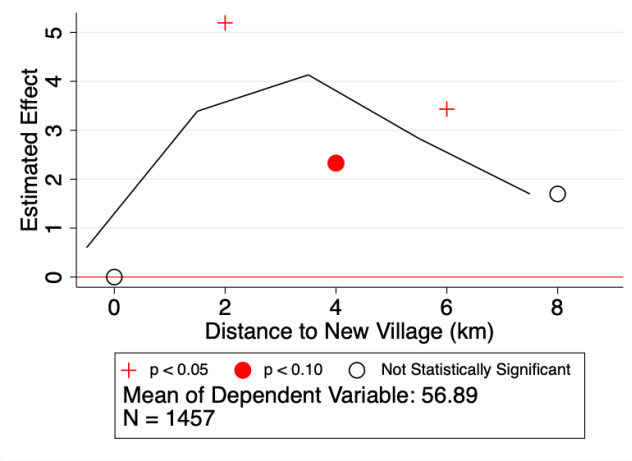
Panel B: Independent Shifts: Plotted Coefficients

Notes: The outcome variable is the vote share for ethno-nationalistic coalition in 2013 at the federal parliamentary constituency level. Panel (A) plots histogram of absolute coefficients from a regression of the outcome variable on bins in distance to counterfactual Chinese New Villages, controlling for federal parliamentary constituency fixed effects, geographic and pre-treatment controls. For each New Village, a counterfactual was selected at random from the region of the road network that was suitable and within 2.5 to 10 kilometers via historical road from the real New Village. This procedure was repeated to construct a 1,000 sets of counterfactual New Villages. The coefficients for distance to the real New Villages are shown as vertical lines. Panel (B) plots differences between real coefficients for each bin and mean counterfactual coefficients, with the symbols indicating the real coefficients' position in the distribution of counterfactual coefficients shown in Panel (A).

Figure 6
Effects of Chinese New Villages on Ethno-nationalist Electoral Support
Panel A: Ethno-nationalist Vote Share in 2013

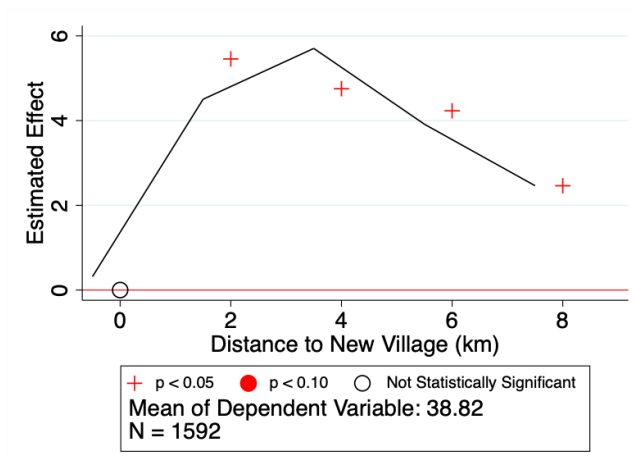


(A) State election

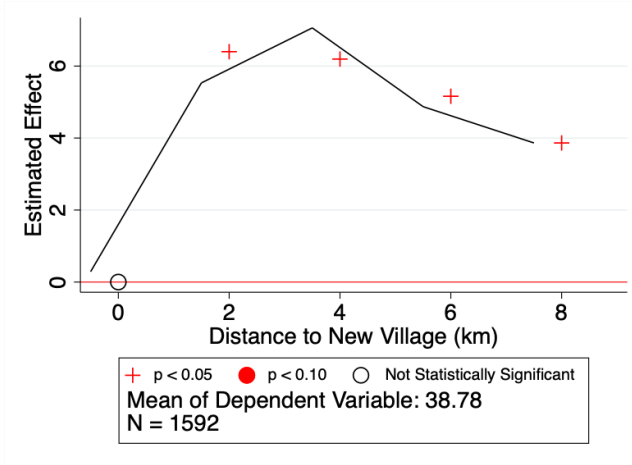


(B) Federal election

Panel B: Ethno-nationalist Vote Share in 2018



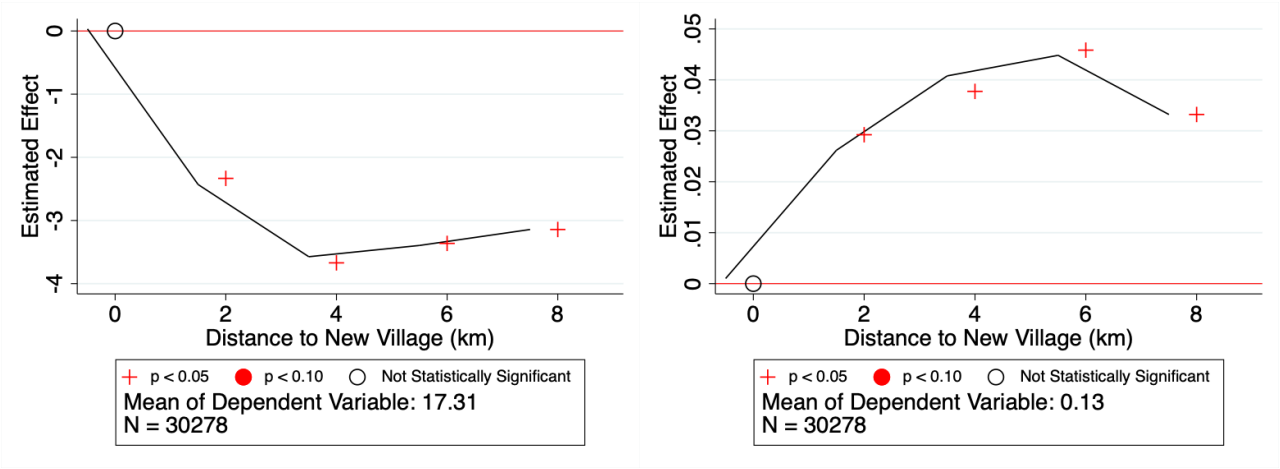
(C) State election



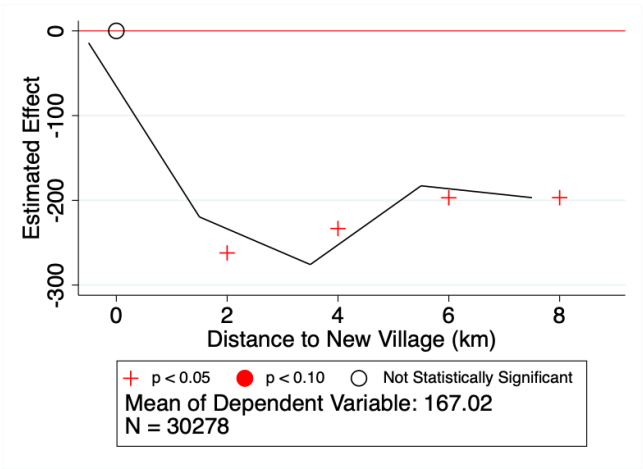
(D) Federal election

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

Figure 7
Economic Outcomes: Grid Cell-level



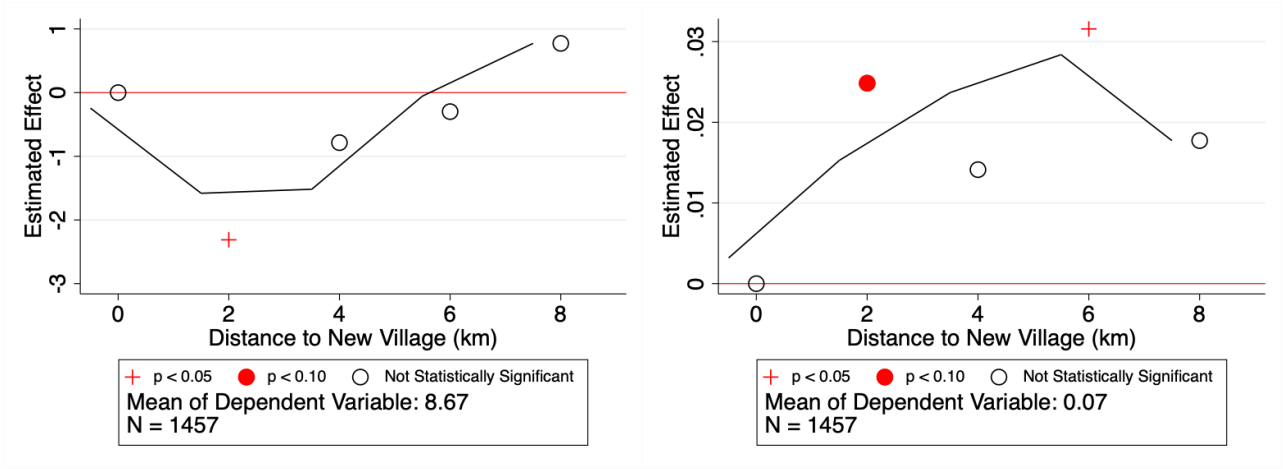
(A) Nightlight Luminosity **(B) NDVI**



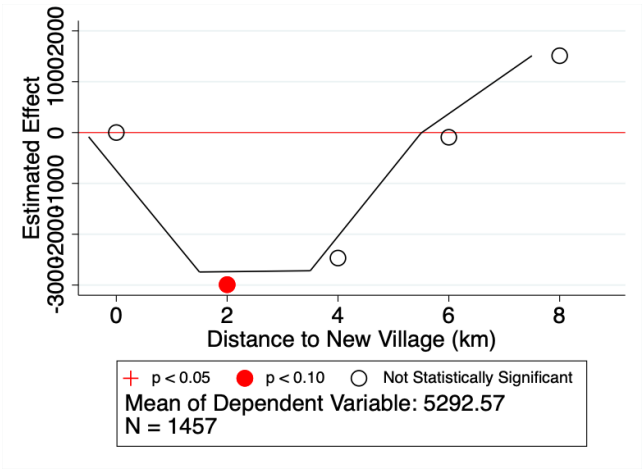
(C) Population Density

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 8
Economic Outcomes: Polling District-Level



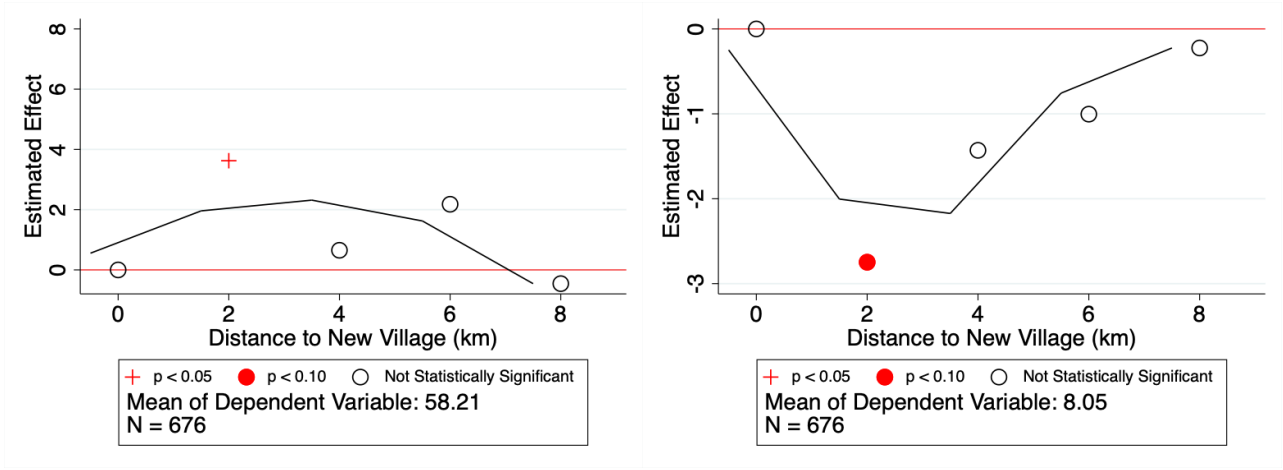
(A) Nightlight Luminosity **(B) NDVI**



(C) Population Density

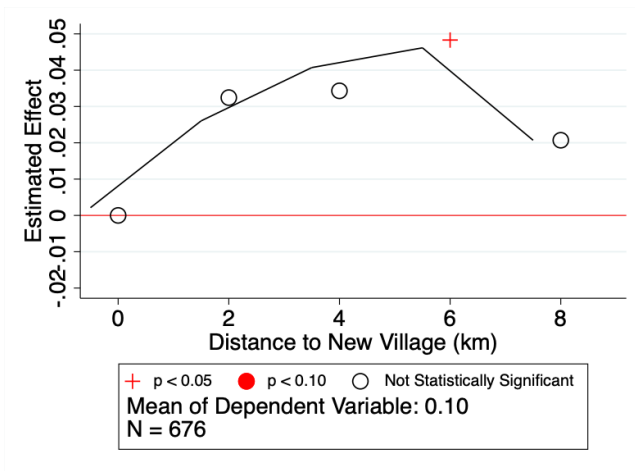
Notes: These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects and geographical and pre-treatment controls. The means of analogous estimates computed from 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 9
Polling district-level: Below Median Initial Population Size of Chinese New Villages

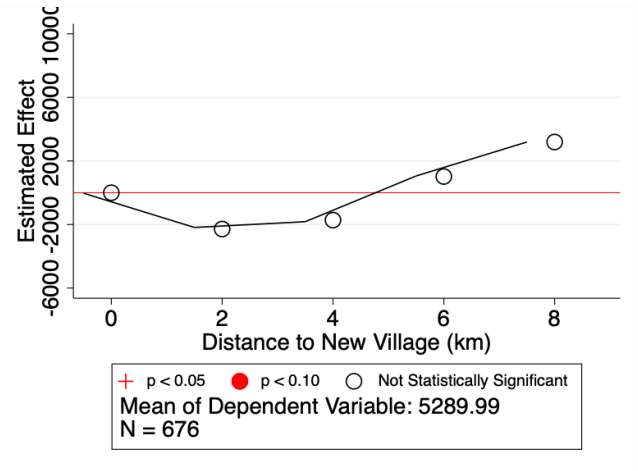


(A) Vote Share

(B) Mean Luminosity



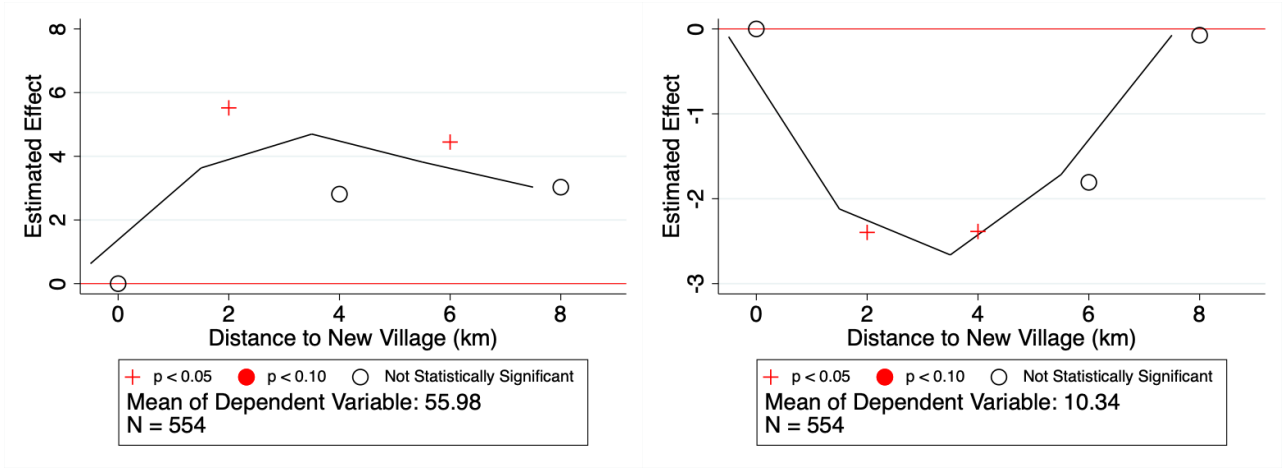
(C) NDVI



(D) Population Density

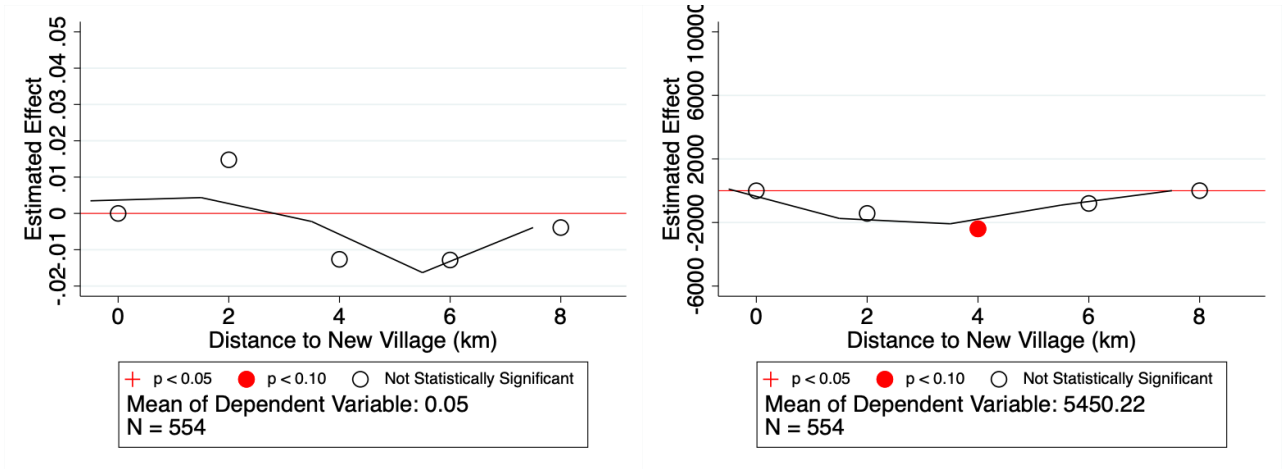
Notes: The outcome variable for (A) is the vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal 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 10
Polling district-level: Above Median Initial Population Size of Chinese New Villages



(A) Vote Share

(B) Mean Luminosity

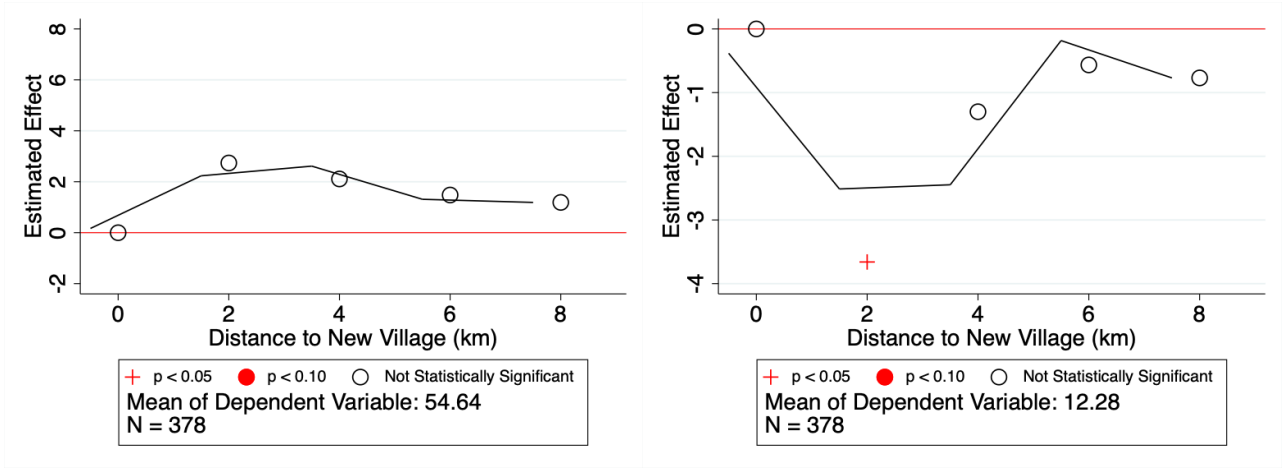


(C) NDVI

(D) Population Density

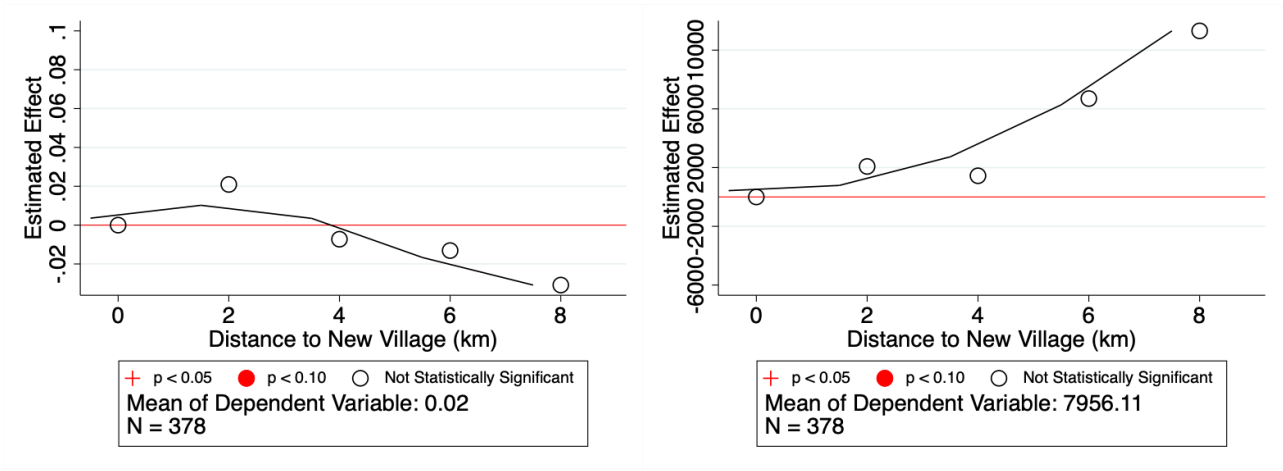
Notes: The outcome variable for (A) is the vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal 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 11
Polling district-level: Below Median Rubber and Tin Employment of Chinese in Chinese New Villages



(A) Vote Share

(B) Mean Luminosity

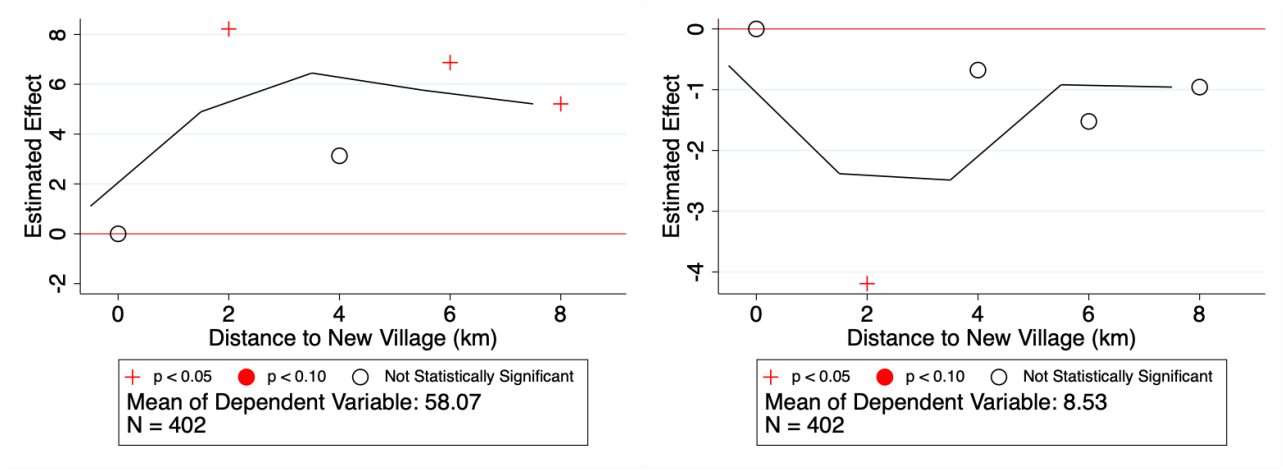


(C) NDVI

(D) Population Density

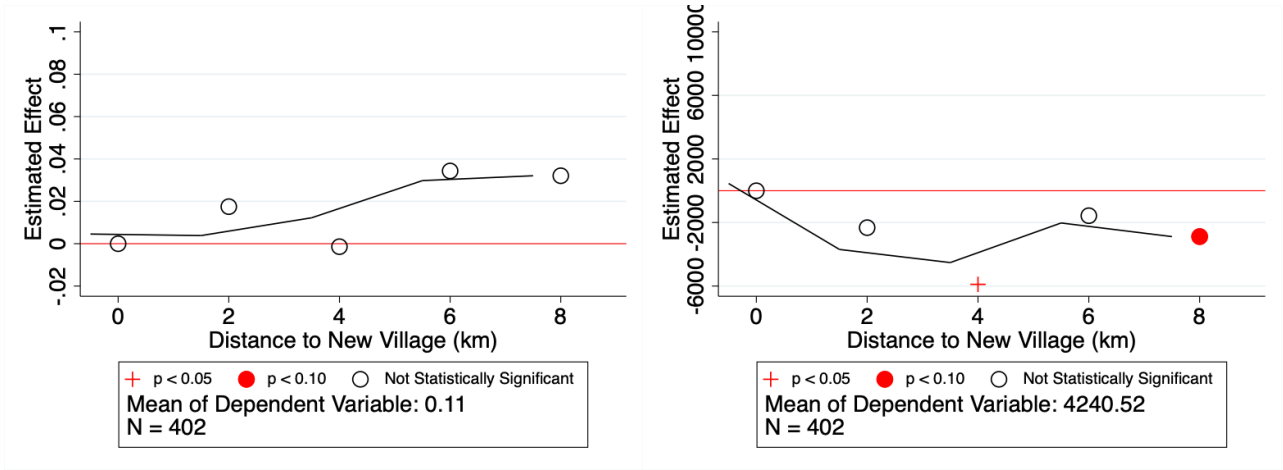
Notes: The outcome variable for (A) is the vote share for ethno-nationalistic coalition at the federal parliament constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary 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 12
Polling district-level: Above Median Rubber and Tin Employment of Chinese in Chinese New Villages



(A) Vote Share

(B) Mean Luminosity

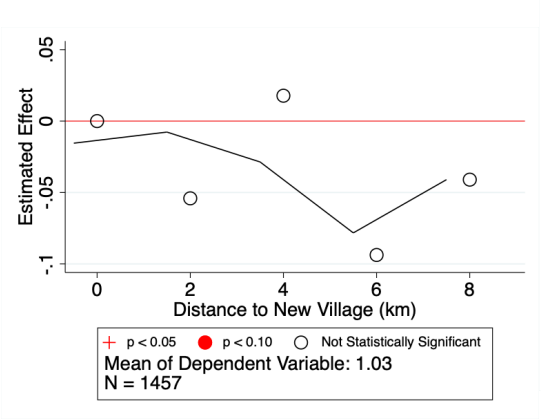


(C) NDVI

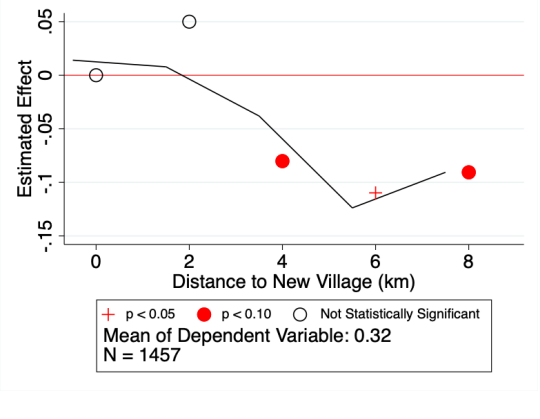
(D) Population Density

Notes: The outcome variable for (A) is the vote share for ethno-nationalistic coalition at the federal parliament 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.

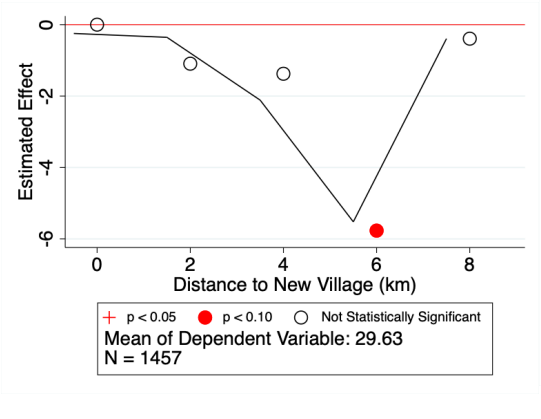
Figure 13
Effects of Chinese New Villages on Public School Provision



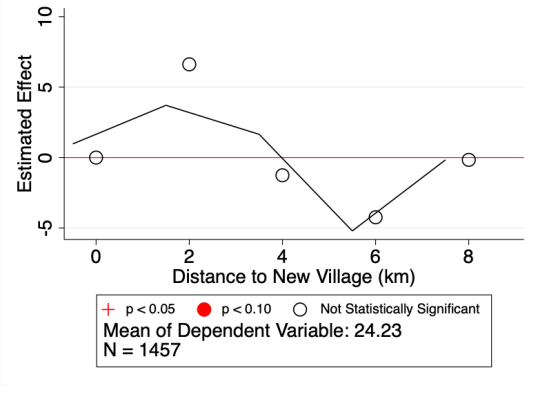
(A) Number of Primary Schools



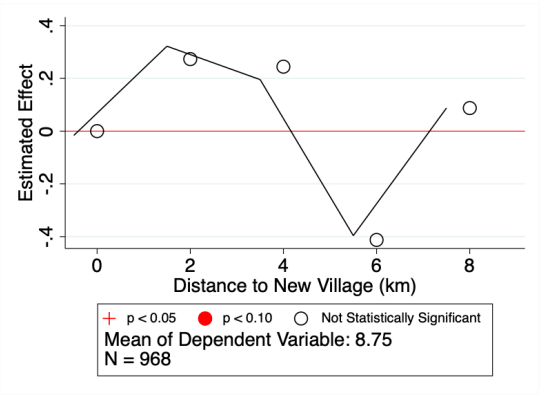
(B) Number of Secondary Schools



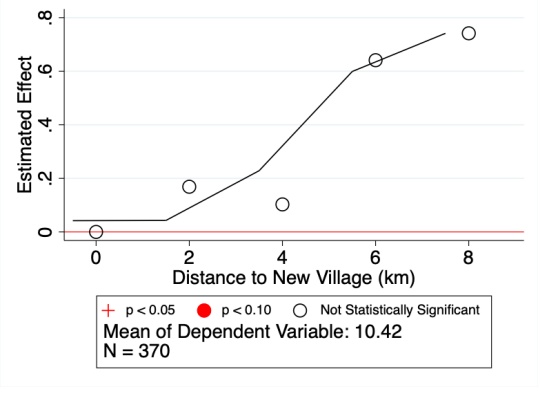
(C) Number of Teachers in Primary Schools



(D) Number of Teachers in Secondary Schools



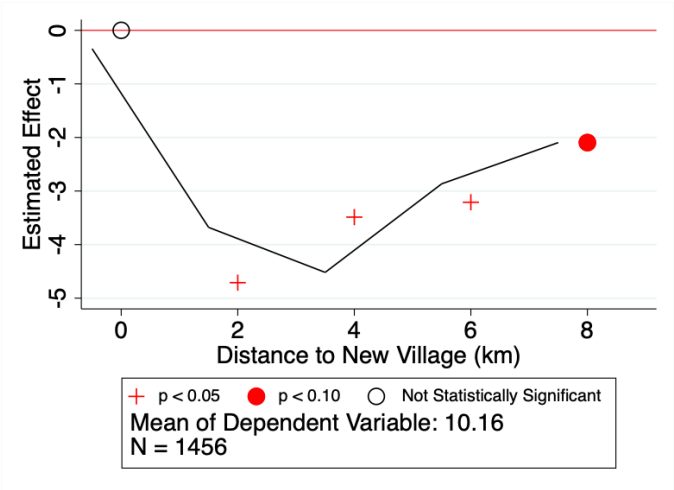
(E) Average Student-Teacher Ratio in Primary Schools



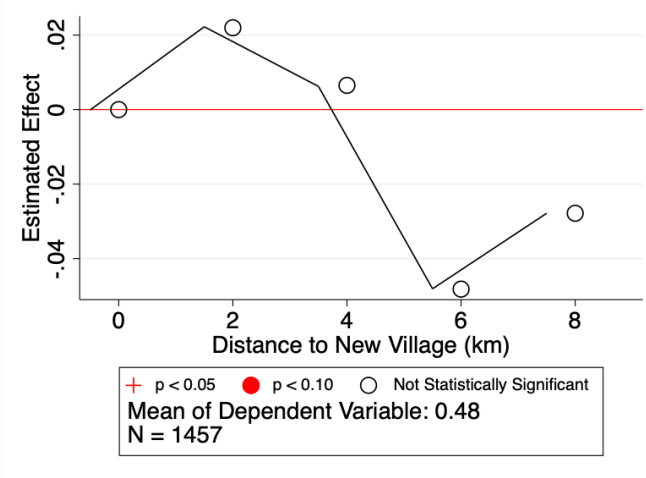
(F) Average Student-Teacher Ratio in Secondary Schools

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.

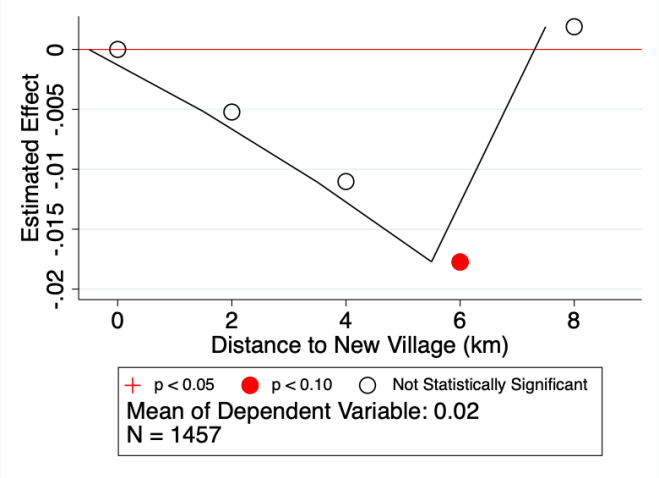
Figure 14
Effects of Chinese New Villages on Infrastructure and Public Health Facilities



(A) Road Density



(B) Number of Clinics



(C) Number of Hospitals

Notes: Points plot coefficients estimated from regression of 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.

Table 1
OLS: Chinese New Villages and Ethno-Nationalist Support

	Panel A: Ethno-nationalistic Coalition Vote Share, 2013					
	State Election			Federal Election		
	(1)	(2)	(3)	(4)	(5)	(6)
2-4km	9.113***	8.604***	7.278***	10.123***	9.865***	8.252***
	(1.414)	(1.816)	(1.767)	(1.762)	(2.078)	(1.866)
4-6km	10.455***	7.850***	6.116***	12.722***	9.807***	7.646***
	(2.069)	(2.057)	(1.999)	(2.259)	(2.347)	(2.107)
6-8km	13.147***	10.240***	7.917***	14.911***	12.276***	9.515***
	(2.456)	(2.360)	(2.355)	(2.825)	(2.710)	(2.486)
8-10km	10.010***	7.702***	5.809**	13.143***	10.350***	8.020***
	(2.608)	(2.856)	(2.778)	(2.862)	(3.120)	(2.834)
Observations	1,457	1,457	1,457	1,457	1,457	1,457
Dep. Var. Mean	56.331	56.331	56.331	56.891	56.891	56.891
Dep. Var. SD	17.393	17.393	17.393	17.404	17.404	17.404
Number of Clusters	83	83	83	83	83	83
Parliamentary Constituency FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes
	Panel B: Ethno-nationalistic Coalition Vote Share, 2018					
	State Election			Federal Election		
	(1)	(2)	(3)	(4)	(5)	(6)
2-4km	8.945***	9.868***	7.476***	11.656***	11.595***	8.786***
	(1.257)	(1.560)	(1.441)	(1.238)	(1.460)	(1.402)
4-6km	12.647***	12.295***	8.447***	16.847***	15.015***	10.485***
	(1.763)	(1.862)	(1.763)	(1.840)	(1.800)	(1.731)
6-8km	14.076***	13.943***	9.025***	18.354***	16.495***	10.717***
	(2.078)	(2.072)	(1.963)	(2.588)	(2.045)	(1.981)
8-10km	11.861***	12.030***	6.686***	17.837***	15.168***	8.977***
	(2.048)	(2.289)	(2.337)	(2.678)	(2.385)	(2.418)
Observations	1,592	1,592	1,592	1,592	1,592	1,592
Dep. Var. Mean	38.82	38.82	38.82	38.78	38.78	38.78
Dep. Var. SD	17.00	17.00	17.00	17.06	17.06	17.06
Number of Clusters	85	85	85	85	85	85
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 dependent variables are ethno-nationalist coalition vote share in the state elections of 2013 (columns 1-3) and ethno-nationalist coalition vote share in the federal elections of 2013 (columns 4-6). In Panel B, the dependent variables are ethno-nationalist coalition vote share in the state elections of 2018 (columns 1-3) and ethno-nationalist coalition vote share in the federal elections of 2018 (columns 4-6). 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. In columns 2 and 5, federal parliamentary constituency fixed effects are included. In columns 3 and 6, controls are additionally included. Controls are mentioned in the main text. Standard errors are clustered at the federal parliamentary constituency level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 2
Attitudes and Behaviors Toward Chinese

	Trust		Openness		Interactions			Social Capital	
	=1 if ≥ somewhat trust	=1 if ≥ somewhat entrust kid	=1 if ≥ somewhat willing Chinese neighbor	=1 if ≥ somewhat willing Chinese marriage	=1 if ≥ some Chinese pri school	=1 if ≥ some Chinese sec school	=1 if ≥ some Chinese colleagues	=1 if ≥ 10% Chi phone contacts	=1 if have good Chinese friend
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.025*	-0.062***	-0.064**	0.011	0.030	-0.036	0.026	-0.119*	-0.103*
	(0.012)	(0.017)	(0.023)	(0.020)	(0.024)	(0.041)	(0.045)	(0.055)	(0.047)
Treated X age≥ 60	-0.001	0.099**	0.058*	0.084	0.005	0.307***	0.036	0.150**	0.014
	(0.054)	(0.036)	(0.031)	(0.069)	(0.028)	(0.074)	(0.061)	(0.063)	(0.056)
age≥ 60	0.009	-0.121***	-0.117***	-0.126*	-0.027	-0.378***	-0.052	-0.225***	-0.029
	(0.044)	(0.029)	(0.030)	(0.060)	(0.026)	(0.082)	(0.056)	(0.034)	(0.043)
cluster	Cohort X nearest New Village								
FE	nearest New Village								
R2	0.175	0.141	0.173	0.102	0.033	0.157	0.027	0.157	0.110
Mean Dep. Var.	0.370	0.237	0.803	0.636	0.056	0.312	0.174	0.366	0.293
Std. Dev. Var.	0.483	0.426	0.398	0.482	0.230	0.464	0.380	0.482	0.456
Observations	397	397	396	396	377	333	316	393	392
Estimate (age≥ 60+Int.)	0.008	-0.022	-0.058	-0.042	-0.022	-0.071	-0.016	-0.075	-0.016
P-Value (age≥ 60+Int.)	0.802	0.453	0.071	0.276	0.424	0.138	0.708	0.192	0.654
Estimate (Treated+Int.)	-0.026	0.037	-0.006	0.095	0.035	0.270	0.062	0.030	-0.089
P-Value (Treated+Int.)	0.633	0.257	0.801	0.177	0.063	0.000	0.174	0.319	0.011

Notes: Source: Primary survey data. This table reports OLS estimates of Equation 4. The dependent variables are indicator variables that take the value of 1 when a Malay: somewhat trusts or trusts a Chinese a lot (col 1); somewhat trusts or trusts a lot, a Chinese neighbor to take care of their child (col 2); is somewhat willing or very willing to have a Chinese neighbor (col 3); is somewhat willing or very willing to have a family member marry a Chinese (col 4); had at least some (>= 10% to more than half) Chinese primary school classmates (col 5); had at least some (>= 10% to more than half) Chinese secondary school classmates (col 6); has at least some (>= 10% to more than half) Chinese colleagues (col 7); has at least 10% Chinese cell phone contacts (col 8); has at least one good friend who is Chinese (col 9). The sample comprises two administrative districts in Johor. 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.

APPENDIX

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 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.[42] Second, Chinese New Villages were on high ground to improve defensibility from the communists. The resettled often lived under the surveillance regime during the Malayan Emergency.[43] 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).

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

[43] 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 Chinese New Villages. 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 produced and published by the Ministry of Housing and Local Government, Malaysia in 2012 (Lee, 2012). In this way, we successfully identify and geolocate a total of 452 New Villages.

B.2 Voting Variables

We use data from the Malaysian General Election in 2013. We employ the following variables at the polling district level in our analysis:

1. Vote count of *Barisan Nasional*, vote count of *Pakatan Rakyat*, and vote count of other coalition.

2. Number of registered voters by ethnic groups.

For our analysis, we construct 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. 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. We construct voter turnout, which is the total number of votes cast over the total number of registered voters in each polling district.

B.3 Spatial, Topographical, and Agroclimatic Variables

We include geographical characteristics as controls in our regressions. These include measures of (i) topography (elevation, slope, and aspect), (ii) soil quality (% of topsoil carbon, % of topsoil sodicity, type of soil, class of drainage), and (iii) climatic variables (temperature and precipitation).

B.3.1 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.

B.3.2 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. For each polling district, we created the following measures of soil types: percentage of land covered by coarse, medium, and fine soils respectively. We use the HWSD to compute indicators for each polling district that equal one if majority of grids of a polling district is covered by soils with very poor, imperfect, moderately good, or excessive drainage.

B.4 Pre-treatment Demographic Variables in 1947

We use the population census in 1947 to construct pre-treatment demographic variables. 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. 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.

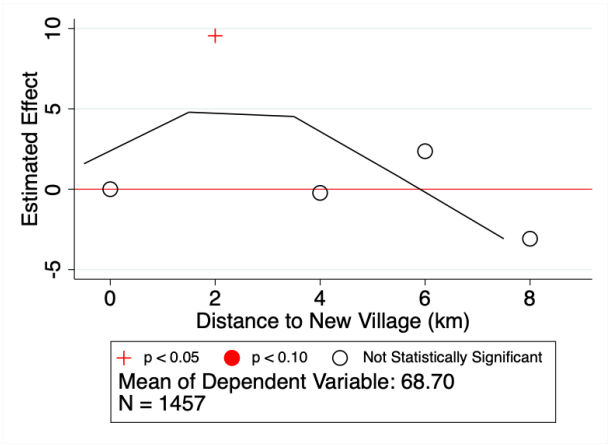
Appendix Figures and Tables

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

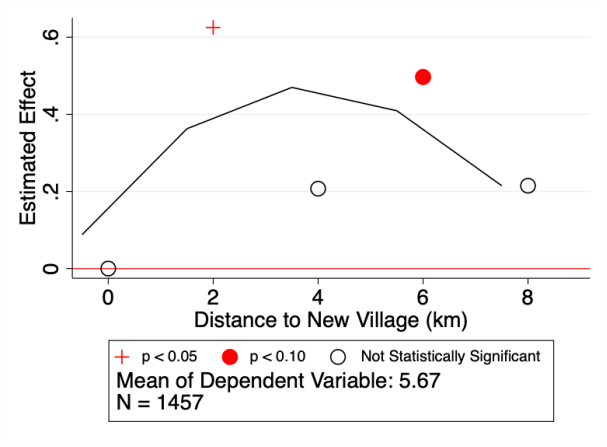


Notes: The orange lines indicate main roads and the red box indicates an example of a Malay Reservation Area.
Source: HIND 1035, Sheet 2N/14

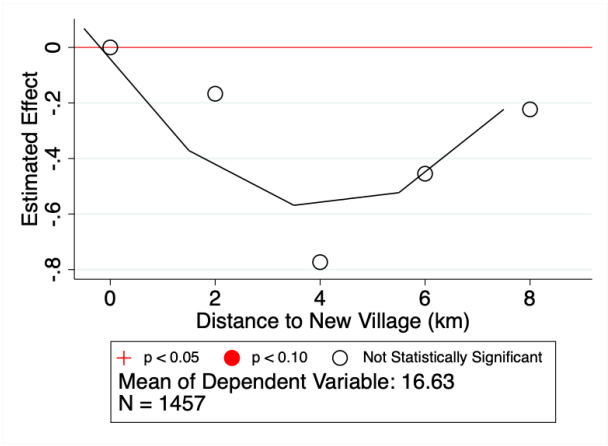
Figure A.2
Geographic, Soil, and Pre-Resettlement Balance



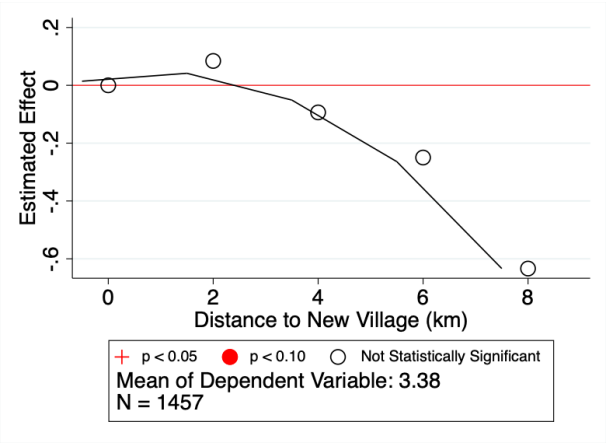
(A) Elevation



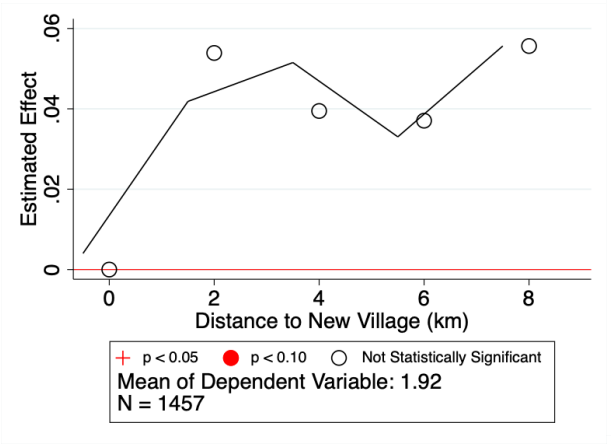
(B) Slope



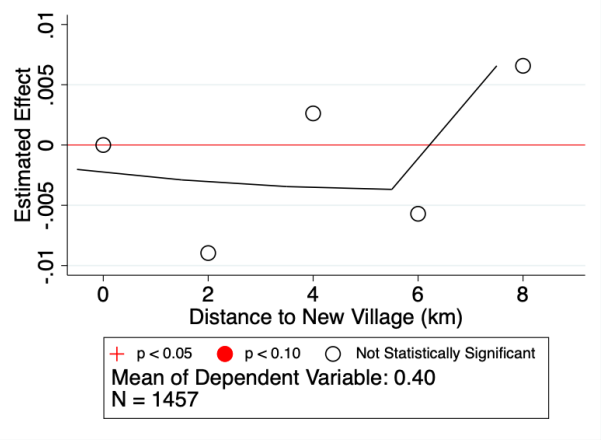
(C) % of East-facing Grids



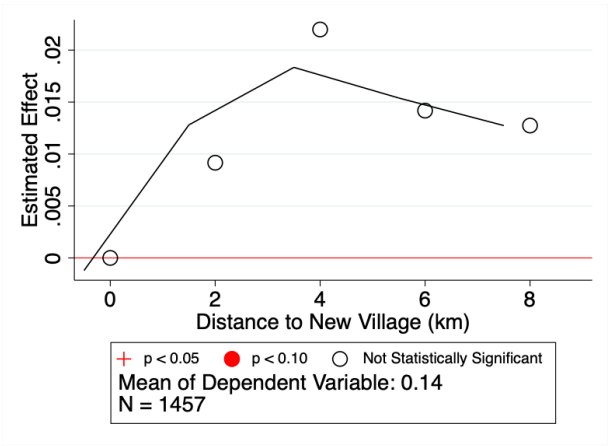
(D) % of Topsoil Organic Carbon



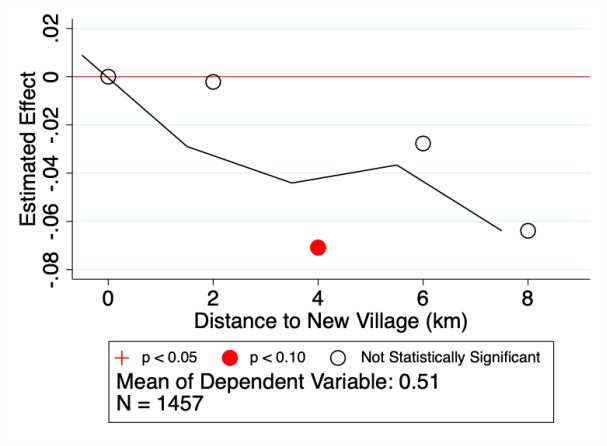
(E) % of Topsoil Sodicity



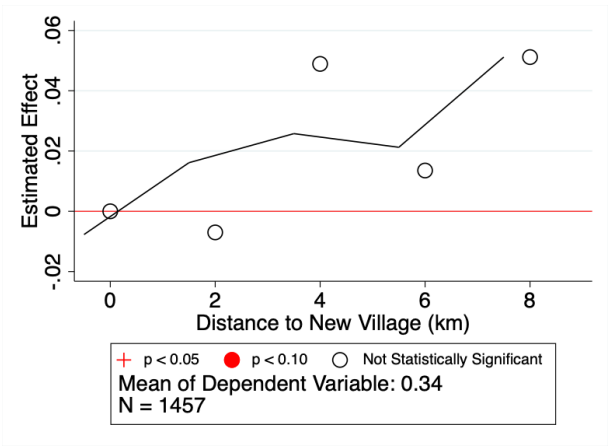
(F) Drainage is very poor (=1)



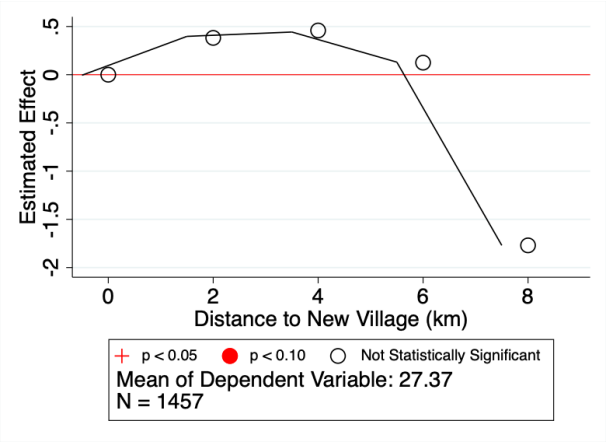
(G) Soil is Coarse (=1)



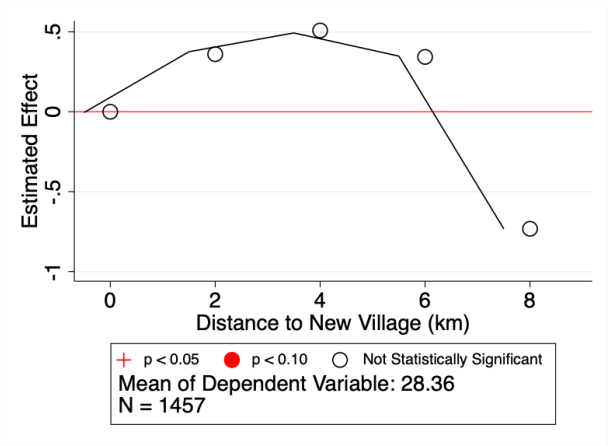
(H) Soil is Medium (=1)



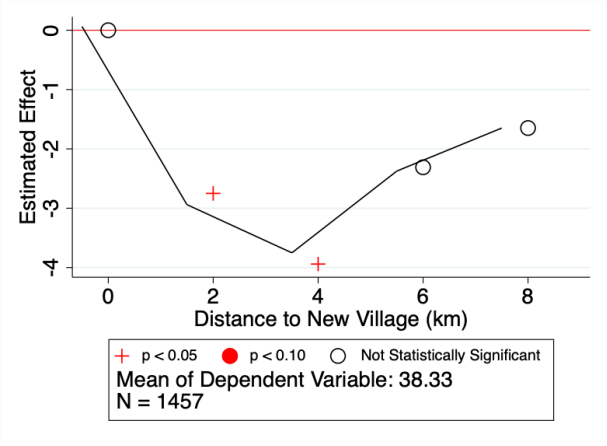
(I) Soil is Fine (=1)



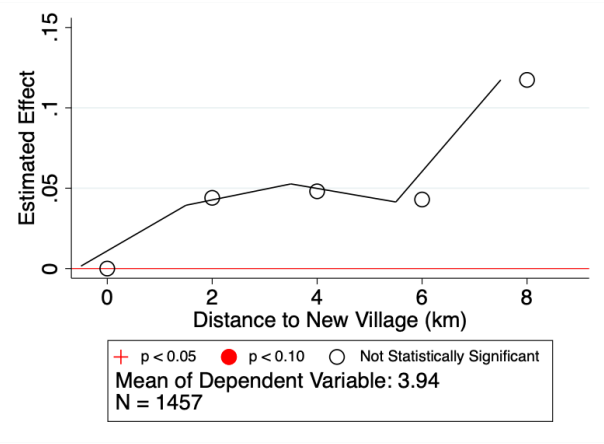
(J) Distance to urban centre in 1947 (km)



(K) Distance to coast (km)



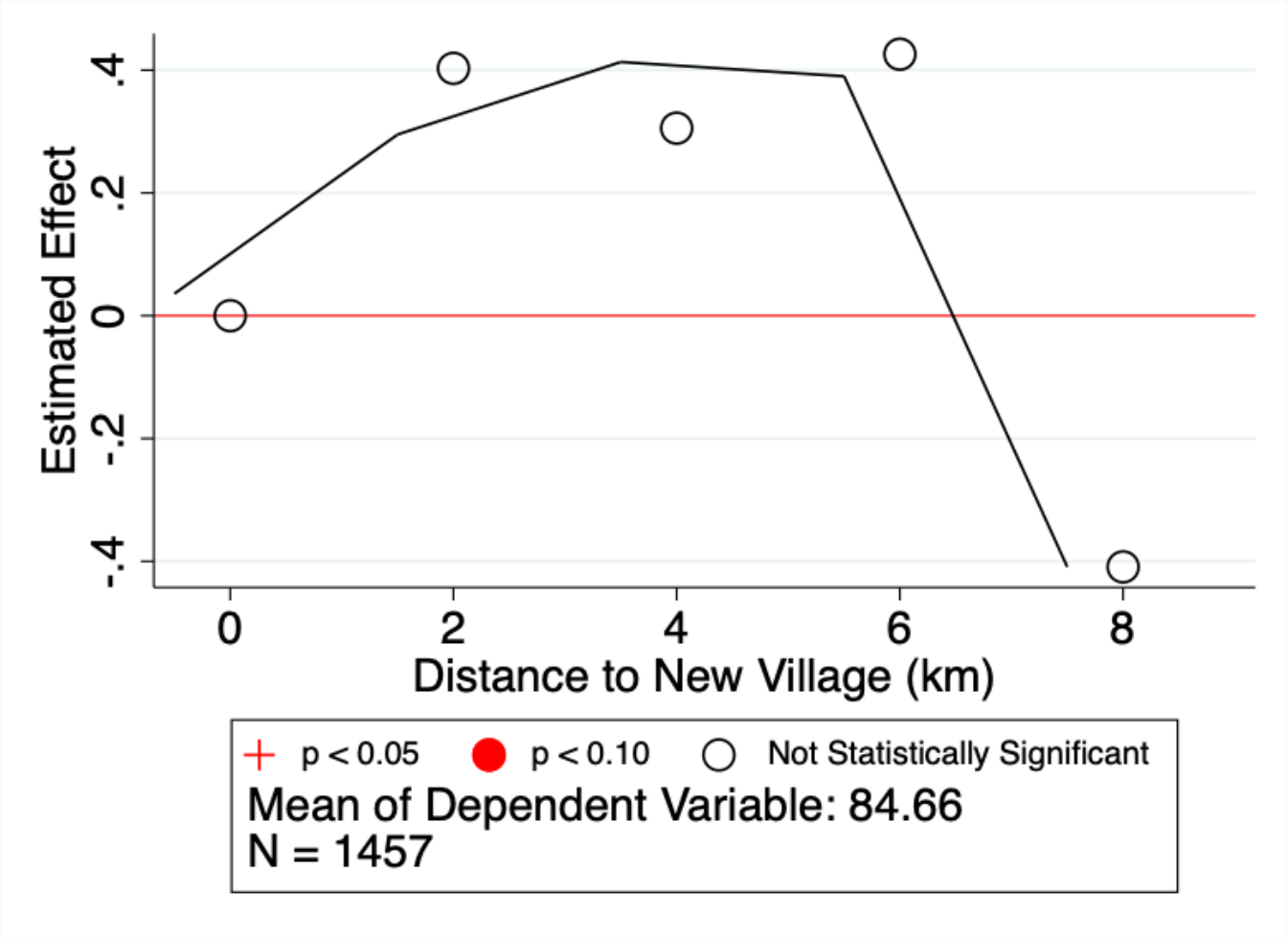
(L) Share of Chinese in 1947



(M) Log Population Density in 1947

Notes: Points plot coefficients estimated from regression 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.

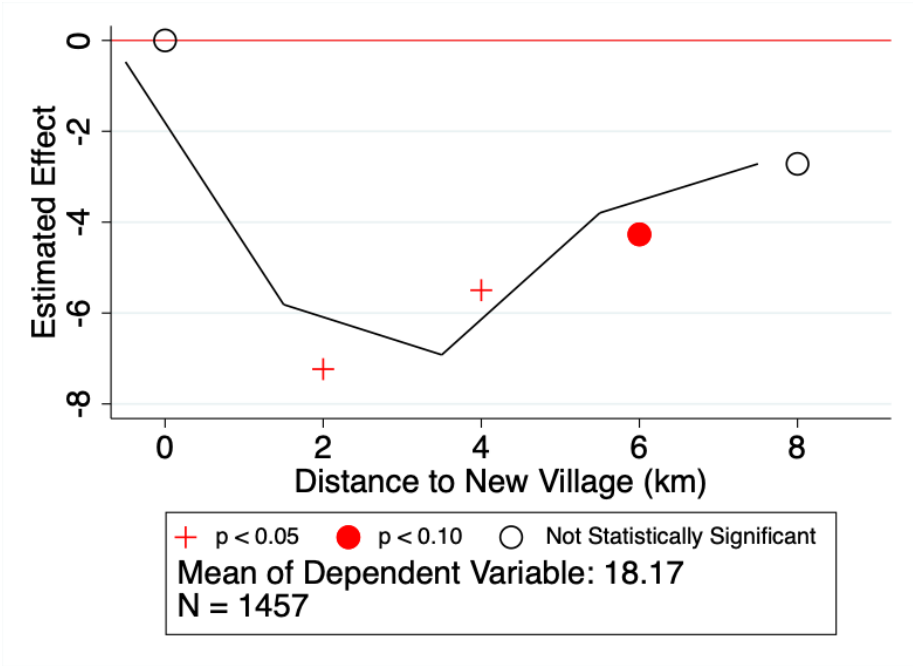
Figure A.3
Effects of Proximity to Chinese New Villages on Voter Turnout, 2013



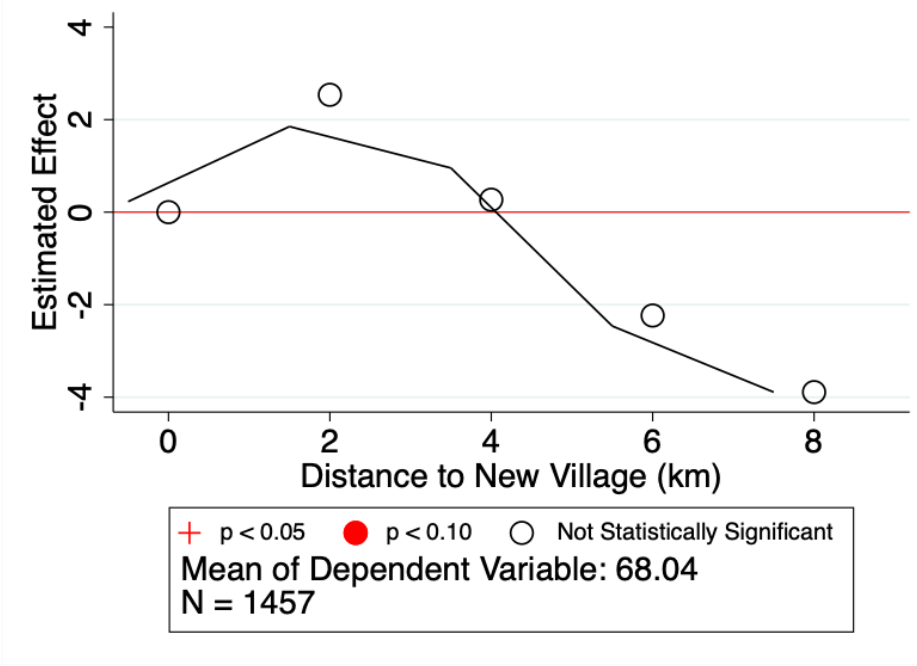
Turnout, 2013

Notes: This figure plots 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.4
Effects of Chinese New Villages on Ethnic Composition



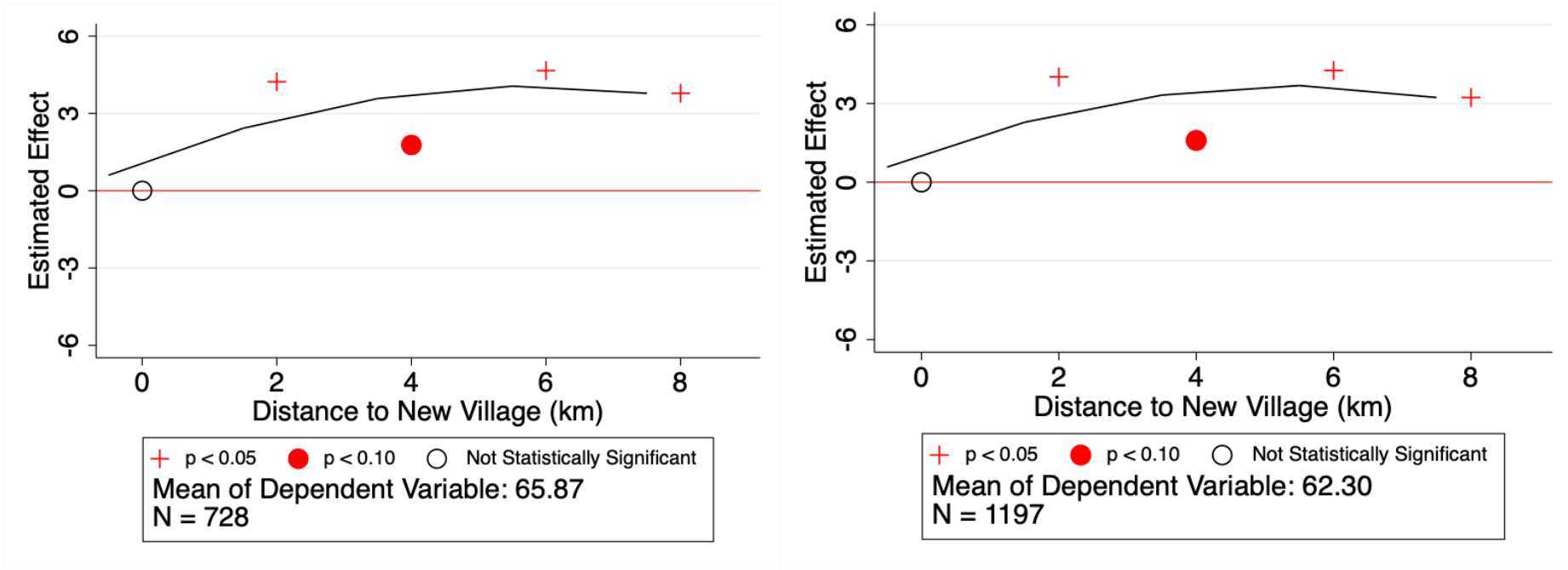
(A) Share of Registered Ethnic Chinese Voters, 2013



(B) Share of Registered Ethnic Malay Voters, 2013

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 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.5
Effects of Chinese New Villages on Ethno-nationalist Electoral Support: Malay dominant samples

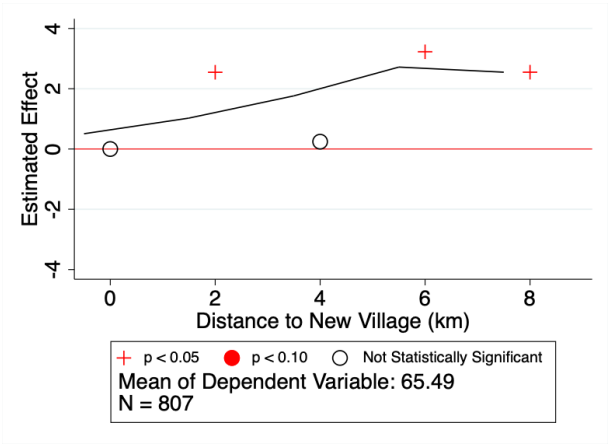


(A) Effects in Polling Districts where Number of Chinese Voters is Below Median

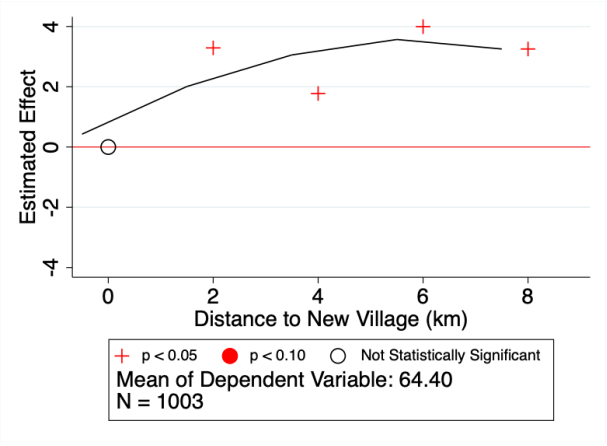
(B) Effects in Polling Districts Where Number of Registered Malay Voters > Chinese Voters

Notes: 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. 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.

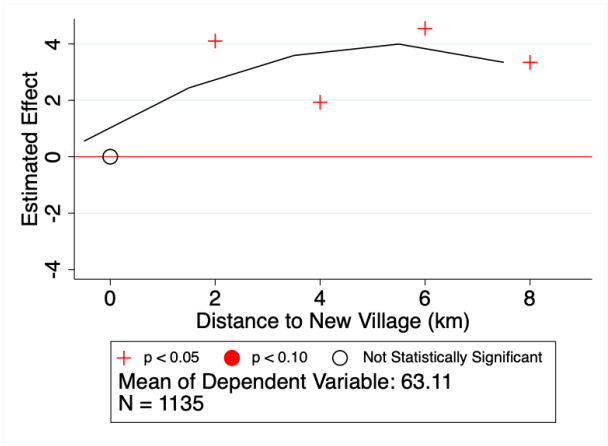
Figure A.6
Effects of Chinese New Villages on Ethno-nationalist Electoral Support: Samples with varying Chinese shares



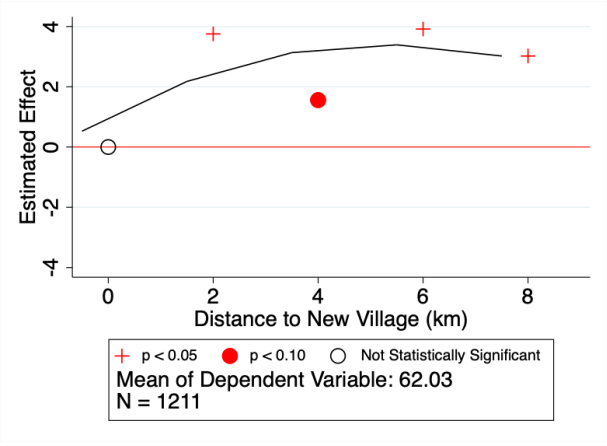
(A) Sample: Chinese Share <= 10%



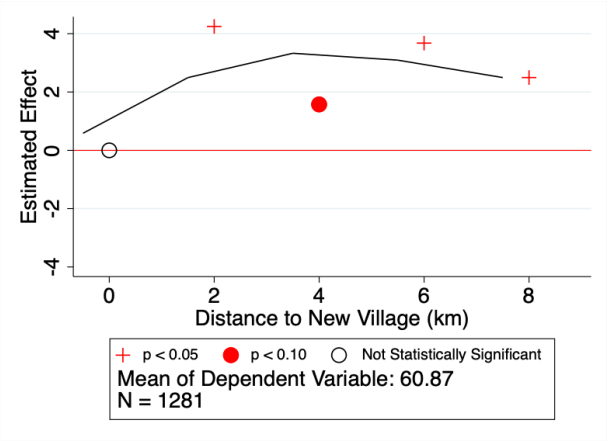
(B) Sample: Chinese Share <= 20%



(C) Sample: Chinese Share <= 30%



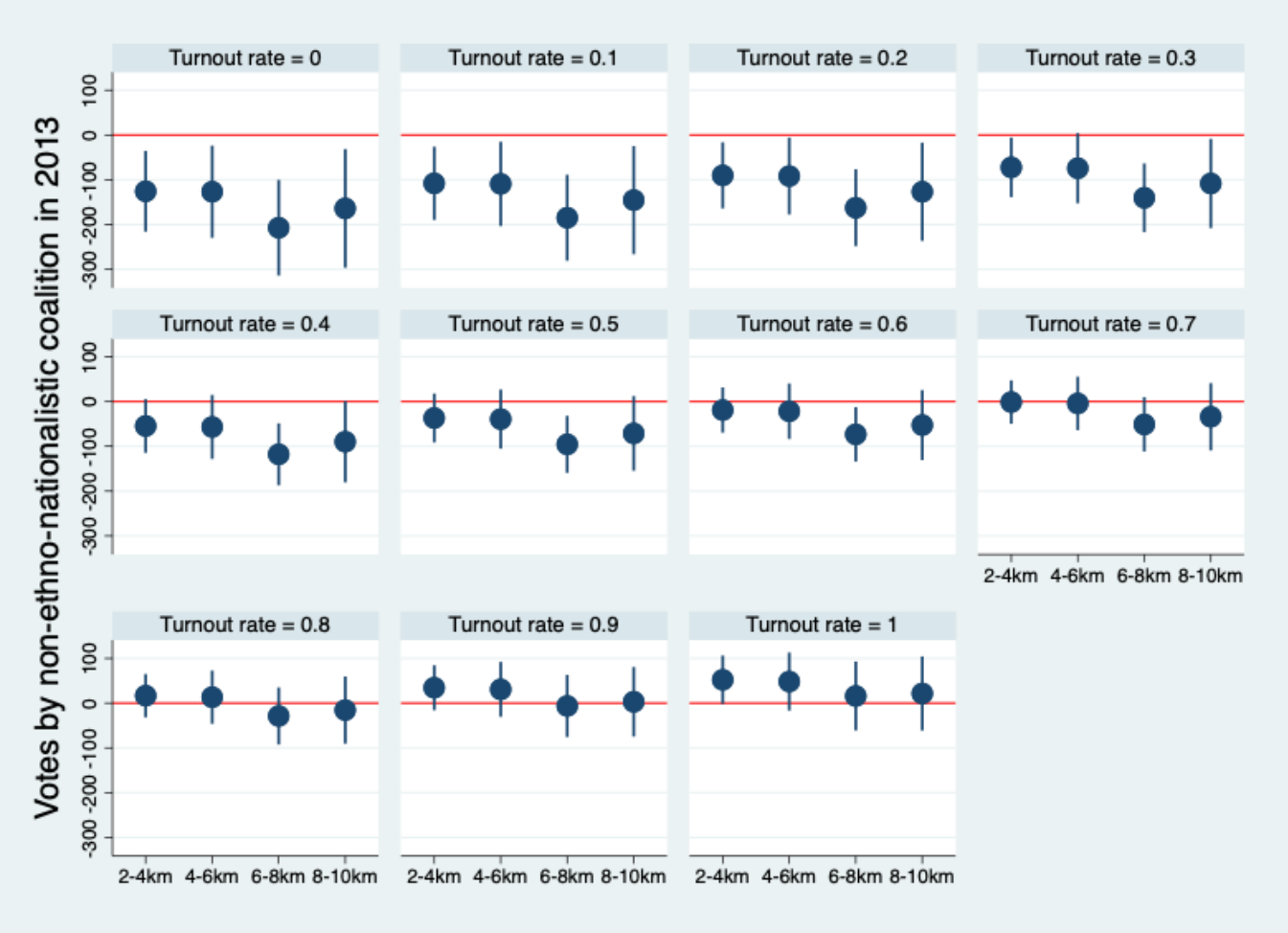
(D) Sample: Chinese Share <= 40%



(E) Sample: Chinese Share <= 50%

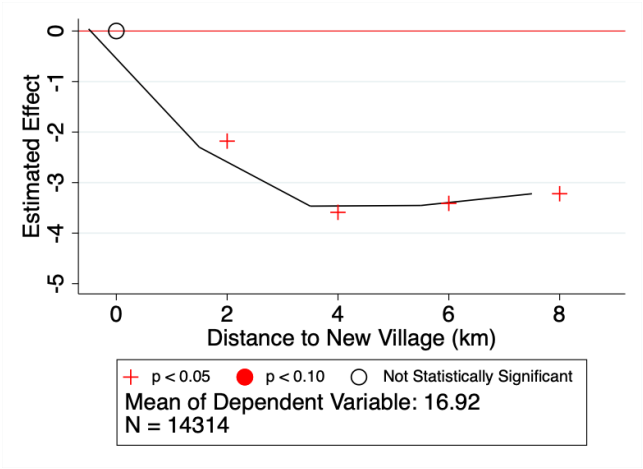
Notes: The outcome variable is vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects and geographical and pre-treatment controls. 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
Quantifying Malay Voting Behavior: The Effects of Varying Ethnic Chinese Voter Turnout Rates on Ethno-Nationalist Vote Shares: Polling District Level

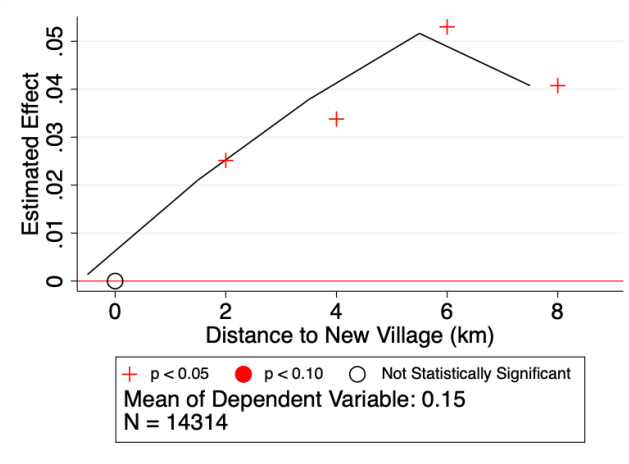


Notes: This figure plots OLS estimates of Equation 1. The outcome variable is vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New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

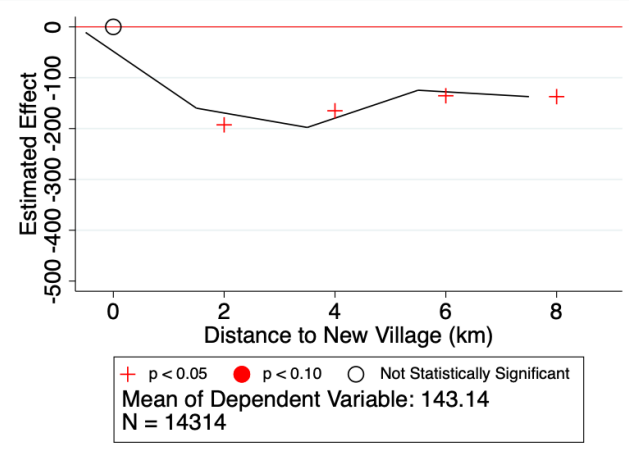
Figure A.8
Grid cell-level: Below median Initial Population Size of Chinese New Villages



(A) Mean Luminosity



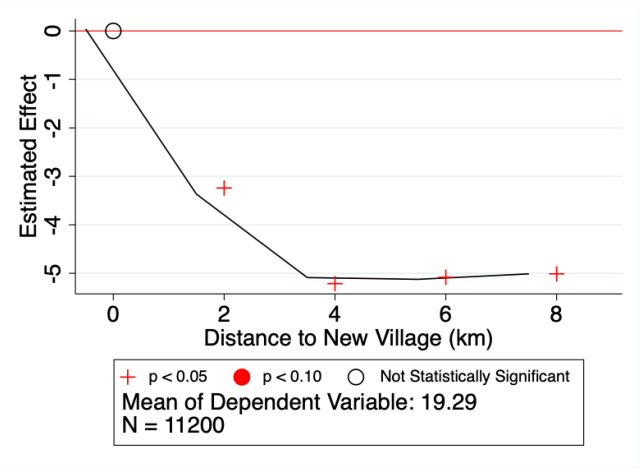
(B) NDVI



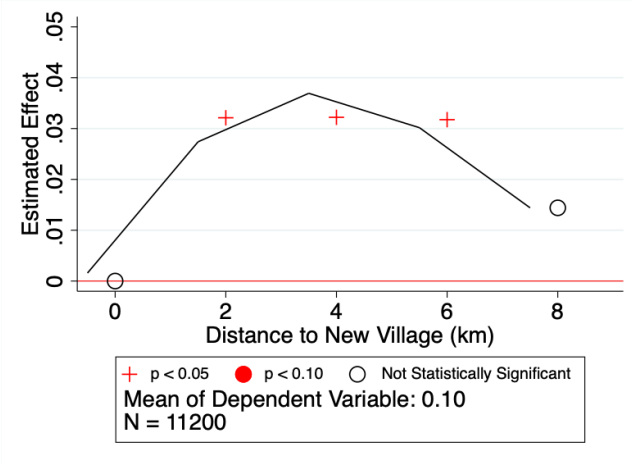
(C) Population Density

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.

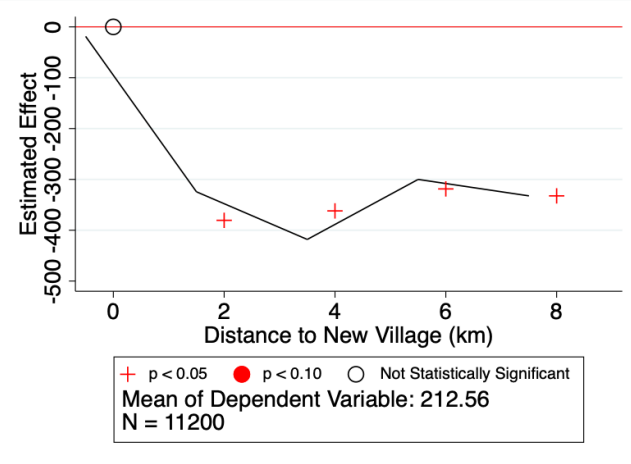
Figure A.9
Grid cell-level: Above median Initial Population Size of Chinese New Villages



(A) Mean Luminosity



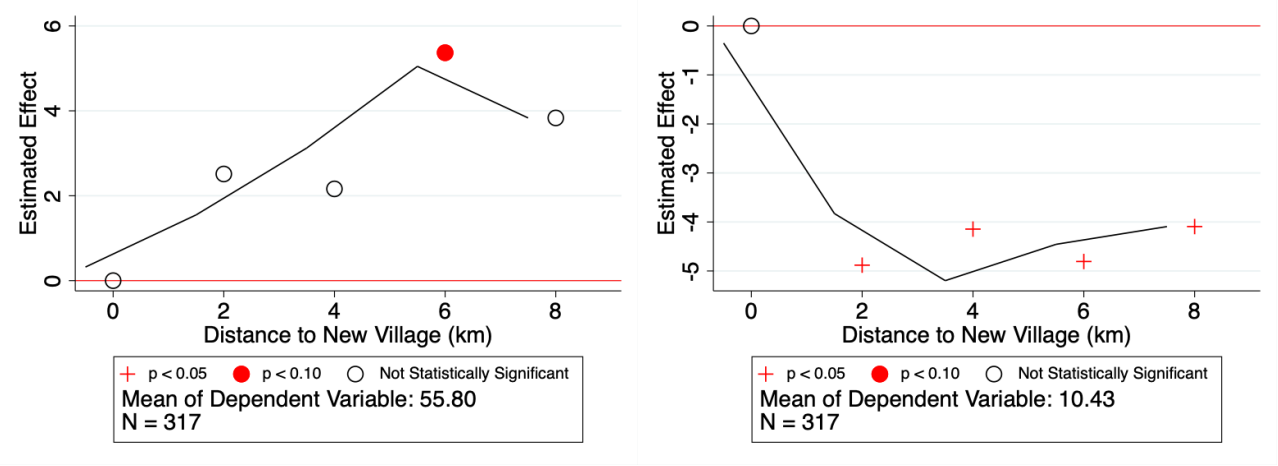
(B) NDVI



(C) Population Density

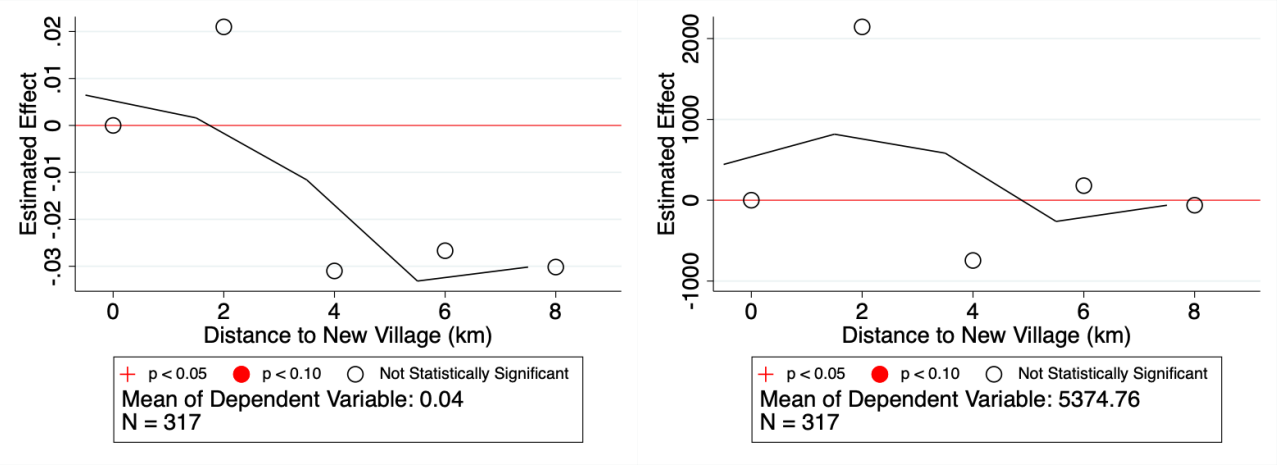
Notes: These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for nearest Chinese New Village fixed effects and geographical and pre-treatment controls. The means of analogous estimates computed from 1,000 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.10
Polling district-level: Above Median Agricultural Employment of Chinese in Chinese New Villages



(A) Vote Share

(B) Mean Luminosity

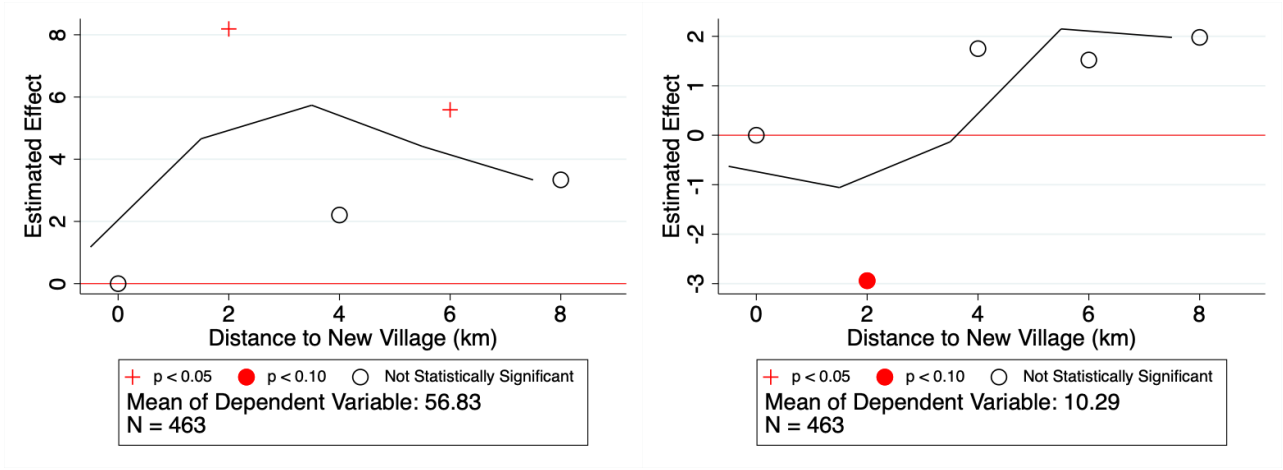


(C) NDVI

(D) Population Density

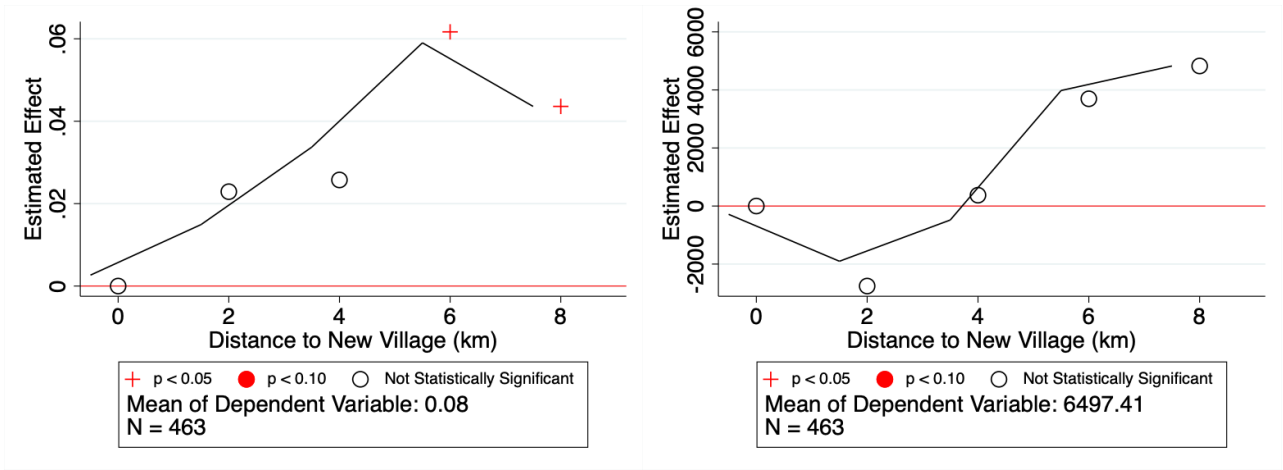
Notes: The outcome variable for (A) is vote share for ethno-nationalistic coalition at the federal parliamentary constituency level. These figures plot coefficients estimated from regressing the outcome variable on 2-km bins of distance to the nearest Chinese New Village, controlling for federal parliamentary constituency fixed effects and geographical and pre-treatment controls. 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
Polling district-level: Below Median Agricultural Employment of Chinese in Chinese New Villages



(A) Vote Share

(B) Mean Luminosity



(C) NDVI

(D) Population Density

Notes: The outcome variable for (A) 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. 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.

Table A.1
Summary Statistics of 452 New Villages

	Mean	S.D	Min	Max	N
Distance to Urban Center					
Distance within 5miles (=1)	0.12	0.32	0.00	1.00	452
Distance within 10miles (=1)	0.15	0.35	0.00	1.00	452
Distance within 15miles (=1)	0.13	0.33	0.00	1.00	452
Distance within 30miles (=1)	0.24	0.43	0.00	1.00	452
Distance more than 30miles (=1)	0.18	0.39	0.00	1.00	452
Missing data (=1)	0.19	0.39	0.00	1.00	452
Transportation Access					
Main Highway (=1)	0.76	0.43	0.00	1.00	452
Railroad (=1)	0.02	0.15	0.00	1.00	452
River (=1)	0.03	0.17	0.00	1.00	452
Missing data (=1)	0.19	0.39	0.00	1.00	452
Type of Government					
City Council (=1)	0.01	0.08	0.00	1.00	452
District Council (=1)	0.45	0.50	0.00	1.00	452
Municipal Council (=1)	0.47	0.50	0.00	1.00	452
Self Governance (=1)	0.05	0.21	0.00	1.00	452
Missing data (=1)	0.03	0.17	0.00	1.00	452
Medical Facilities					
None (=1)	0.08	0.28	0.00	1.00	452
Provided by adjacent towns (=1)	0.02	0.12	0.00	1.00	452
Static or mobile clinic (=1)	0.71	0.45	0.00	1.00	452
Missing data (=1)	0.19	0.39	0.00	1.00	452
Settlement Type					
Assimilated to existing settlement (=1)	0.12	0.33	0.00	1.00	452
Attached to existing settlement (=1)	0.29	0.45	0.00	1.00	452
New settlement (=1)	0.40	0.49	0.00	1.00	452
Nucleation (=1)	0.01	0.08	0.00	1.00	452
Missing data (=1)	0.18	0.39	0.00	1.00	452
Population Statistics					
Population 1954	1460.20	1540.82	132.00	13000.00	305
Population 1970	2290.91	2610.49	190.00	27436.00	347
Population 1980	2789.67	3072.14	184.00	29191.00	254
Population 1995	3724.00	5505.79	100.00	45000.00	443
Population Statistics 1958					
Population 1958	1863.54	2027.71	109.00	16700.00	342
Chinese population 1958	1439.77	1879.61	0.00	13912.00	128
Indian population 1958	0.16	1.98	0.00	24.00	147
Malay population 1958	45.07	215.27	0.00	2260.00	165
Other population 1958	1.42	20.65	0.00	300.00	211
Languages spoken					
Chinese dialects only (=1)	0.70	0.46	0.00	1.00	452
Malay and Javanese (=1)	0.07	0.25	0.00	1.00	452
Tamil and Sikh (=1)	0.00	0.05	0.00	1.00	452
Thai (=1)	0.00	0.07	0.00	1.00	452
Malay, Javanese, and Chinese dialects (=1)	0.01	0.09	0.00	1.00	452
Tamil, Sikh, and Chinese dialects (=1)	0.02	0.15	0.00	1.00	452
Missing (=1)	0.19	0.39	0.00	1.00	452

Notes: The table shows descriptive statistics of 452 New Villages. Source: Lim (2022); Malayan Christian Union (1958)

Table A.2
OLS: Chinese New Villages and Economic Effects

	Luminosity	Population Density	NDVI
	Panel A: Grid-cell level		
	(1)	(2)	(3)
2-4km	-4.136***	-357.966***	0.041***
	(0.421)	(40.690)	(0.009)
4-6km	-7.584***	-372.582***	0.058***
	(0.750)	(47.403)	(0.012)
6-8km	-8.450***	-337.643***	0.077***
	(0.925)	(49.915)	(0.014)
8-10km	-8.578***	-320.680***	0.072***
	(1.017)	(50.708)	(0.015)
Observations	30,278	30,278	30,278
Dep. Var. Mean	17.312	167.016	0.132
Dep. Var. SD	15.02	599.20	0.28
Number of Clusters	271	271	271
Nearest Chinese Village FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
	Panel B: Polling District level		
	(1)	(2)	(3)
2-4km	-4.904***	-5102.882***	0.036**
	(1.331)	(1633.056)	(0.018)
4-6km	-4.551***	-4892.240**	0.041**
	(1.526)	(1884.340)	(0.019)
6-8km	-4.294**	-2898.789	0.065***
	(1.678)	(2175.099)	(0.021)
8-10km	-3.528*	-1383.039	0.061***
	(1.814)	(3131.660)	(0.019)
Observations	1,457	1,457	1,457
Dep. Var. Mean	8.671	5,292.568	0.075
Dep. Var. SD	12.348	16,785.598	0.183
Number of Clusters	83	83	83
Parliamentary Constituency FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table displays OLS estimations of Equation 2 in Panel A and Equation 1 in Panel B. The sample comprises grid cells in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding grid cells that contain Chinese New Village population, grid cells in historically urban areas and grid cells beyond 10km of a Chinese New Village. In Panel A, all regressions include nearest Chinese village fixed effects and controls. In Panel B, all regressions include federal parliamentary constituency fixed effects and controls. The controls are mentioned in the main text. Standard errors are clustered at the nearest Chinese Village-level in Panel A and at the federal parliamentary constituency level in Panel B. * p<0.1, ** p<0.05, *** p<0.01.

Table A.3
Scale Effects: Above vs Below-Median Initial Population Size of Chinese New Villages

	Vote Share	Luminosity	Population Density	NDVI
	(1)	(2)	(3)	(4)
2-4km	7.314**	-5.672**	-4647.400*	0.033
	(2.845)	(2.293)	(2752.475)	(0.028)
4-6km	5.561*	-4.149*	-3813.620	0.048*
	(3.070)	(2.321)	(2851.312)	(0.028)
6-8km	7.267**	-2.938	-601.215	0.064**
	(3.548)	(2.491)	(3322.610)	(0.028)
8-10km	5.508	-2.633	1424.485	0.043
	(4.480)	(2.713)	(5195.032)	(0.031)
2-4km * Initial Pop. Above Median	0.058	0.858	1237.276	0.001
	(3.087)	(2.161)	(2886.405)	(0.034)
4-6km * Initial Pop. Above Median	2.201	-1.385	176.395	-0.031
	(3.241)	(2.450)	(3428.754)	(0.035)
6-8km * Initial Pop. Above Median	2.944	-3.055	-2074.184	-0.036
	(3.801)	(2.498)	(3536.083)	(0.035)
8-10km * Initial Pop. Above Median	3.749	-1.159	-2160.331	0.008
	(4.673)	(2.764)	(4839.784)	(0.037)
Observations	1,230	1,230	1,230	1,230
Dep. Var. Mean	57.207	9.079	5,362.159	0.079
Dep. Var. SD	17.616	12.905	17,629.579	0.184
Number of Clusters	78	78	78	78
Parliamentary Constituency FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table displays OLS estimations of Equation 1. The dependent variables are vote share for ethno-nationalistic coalition at the parliamentary constituency level (column 1), luminosity (column 2), population density (column 3), and Normalized Difference Vegetation Index (NDVI) (column 4). 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. All regressions include parliamentary constituency fixed effects and controls. The controls are mentioned in the main text. Standard errors are clustered at the federal parliamentary constituency level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A.4
Interethnic Competition: Above vs Below-Median Agricultural Employment of Chinese in Chinese New Villages

	Vote Share	Luminosity	Population Density	NDVI
	(1)	(2)	(3)	(4)
2-4km	13.667***	-7.798***	-6089.585**	0.029
	(3.929)	(2.052)	(2549.565)	(0.023)
4-6km	11.497***	-6.162**	-4935.137*	0.038
	(3.806)	(2.340)	(2828.348)	(0.026)
6-8km	15.963***	-6.512**	-2281.639	0.062**
	(4.124)	(2.482)	(3045.953)	(0.029)
8-10km	15.266***	-6.244**	680.570	0.075**
	(4.100)	(2.940)	(5437.506)	(0.029)
2-4km * Agriculture Above Median	-10.298**	1.652	6355.779**	0.021
	(4.629)	(2.725)	(2405.707)	(0.036)
4-6km * Agriculture Above Median	-6.235	0.212	4045.308*	-0.038
	(4.191)	(2.296)	(2269.147)	(0.041)
6-8km * Agriculture Above Median	-10.592***	2.148	5111.546*	-0.074*
	(3.883)	(2.125)	(2708.339)	(0.043)
8-10km * Agriculture Above Median	-10.713***	2.602	2801.922	-0.099*
	(3.933)	(2.410)	(3373.156)	(0.054)
Observations	780	780	780	780
Dep. Var. Mean	56.412	10.347	6,041.155	0.066
Dep. Var. SD	19.049	14.041	20,249.740	0.186
Number of Clusters	65	65	65	65
Parliamentary Constituency FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table displays OLS estimations of Equation 1. The dependent variables are vote share for ethno-nationalistic coalition at the federal parliamentary constituency level (column 1), luminosity (column 2), population density (column 3), and Normalized Difference Vegetation Index (NDVI) (column 4). 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. All regressions include federal parliamentary constituency fixed effects and controls. The controls are mentioned in the main text. Standard errors are clustered at the federal parliamentary constituency level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A.5
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	0.015	0.035	0.376	5.895	-0.052	0.352
	(0.092)	(0.066)	(4.581)	(5.303)	(0.423)	(0.472)
4-6km	0.064	-0.115*	-0.890	-4.886	-0.148	0.219
	(0.095)	(0.061)	(3.326)	(4.903)	(0.460)	(0.536)
6-8km	-0.059	-0.140**	-6.059	-8.519	-0.868*	0.492
	(0.087)	(0.070)	(3.831)	(5.350)	(0.491)	(0.551)
8-10km	-0.012	-0.170**	-1.710	-8.094	-0.489	0.354
	(0.116)	(0.068)	(4.479)	(5.423)	(0.534)	(0.629)
Observations	1,457	1,457	1,457	1,457	968	370
Dep. Var. Mean	1.030	0.321	29.627	24.228	8.753	10.423
Dep. Var. SD	1.005	0.614	40.218	50.070	3.720	2.753
Number of Clusters	83	83	83	83	81	76
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 dependent variables are the number of schools in primary schools (column 1) and in secondary schools (column 2), number of teachers in primary schools (column 3) and in secondary schools (column 4), and average student to teacher ratio in primary schools (column 5) and in secondary schools (column 6). 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 all columns, federal parliamentary constituency fixed effects and controls are included. Controls are mentioned in the main text. 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 Infrastructure and Public Health Facilities Provision

	Road Density	Number of ...	
		Clinics	Hospitals
	(1)	(2)	(3)
2-4km	-7.014***	0.055	-0.010
	(1.263)	(0.061)	(0.015)
4-6km	-7.003***	0.086	-0.014
	(1.522)	(0.057)	(0.015)
6-8km	-7.181***	0.054	-0.019
	(1.632)	(0.064)	(0.015)
8-10km	-6.506***	0.066	-0.008
	(1.757)	(0.081)	(0.018)
Observations	1,456	1,457	1,457
Dep. Var. Mean	10.156	0.483	0.019
Dep. Var. SD	11.223	0.663	0.137
Number of Clusters	83	83	83
Parliamentary Constituency FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. The dependent variable is road density (column 1), number of clinics (column 2), and number of hospitals (column 3). 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 all columns, federal parliamentary constituency fixed effects and controls are included. Controls are mentioned in the main text. Standard errors are clustered at the federal parliamentary constituency level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A.7
OLS: Malay New Villages, Ethno-Nationalist Support and Ethnic Shares

	Vote Share		Chinese Share	Malay Share
	State	Federal		
	(1)	(2)	(3)	(4)
2-4km	0.895	1.303	-3.630	6.311
	(3.416)	(3.670)	(5.394)	(5.996)
4-6km	0.046	-0.096	-0.257	0.283
	(2.680)	(2.816)	(4.265)	(4.801)
6-8km	2.040	2.045	-6.136	5.866
	(3.469)	(3.348)	(4.897)	(5.265)
8-10km	2.705	1.221	-3.488	4.013
	(3.431)	(3.402)	(4.660)	(5.548)
Observations	722	722	722	722
Dep. Var. Mean	53.363	53.257	26.391	61.503
Dep. Var. SD	18.723	18.728	30.503	34.140
Number of Clusters	51	51	51	51
Parliamentary Constituency FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. The dependent variables are ethno-nationalist coalition vote share in the state elections of 2013 (column 1) and in the federal elections of 2013 (column 2), ethnic Chinese share of registered voters in 2013 (column 3) and ethnic Malay share of registered voters in 2013 (column 4). The sample comprises polling districts in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding polling districts in historically urban areas and polling districts beyond 10km of a Malay New Village. In all columns, federal parliamentary constituency fixed effects and controls are included. Controls are mentioned in the main text. 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 Economic Outcomes

	Luminosity	NDVI	Population Density
	(1)	(2)	(3)
2-4km	-2.430	0.050*	-3574
	(2.084)	(0.027)	(2728)
4-6km	-2.284	0.041	-3955
	(2.079)	(0.030)	(2990)
6-8km	-2.379	0.063**	-853.9
	(2.332)	(0.031)	(3780)
8-10km	-1.139	0.057	-3305
	(2.609)	(0.036)	(3439)
Observations	722	722	722
Dep. Var. Mean	8.223	0.086	7,056
Dep. Var. SD	11.829	0.196	21,718
Number of Clusters	51	51	51
Parliamentary Constituency FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. The dependent variables are luminosity (column 1), Normalized Difference Vegetation Index (NDVI) (column 2), and population density (column 3). The sample comprises polling districts in Johor, Kedah, Melaka, Negeri Sembilan, Perak, Selangor, and Pahang, after excluding polling districts in historically urban areas and polling districts beyond 10km of a Malay New Village. In all columns, federal parliamentary fixed effects and controls are included. Controls are mentioned in the main text. 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

	Winning Coalition is PR, 2013		Winning Coalition is PH, 2018	
	State	Federal	State	Federal
	(1)	(2)	(3)	(4)
2-4km	-0.044	0.023	-0.046	-0.022
	(0.038)	(0.045)	(0.043)	(0.036)
4-6km	-0.107*	-0.071	-0.021	-0.022
	(0.061)	(0.062)	(0.059)	(0.051)
6-8km	-0.087	-0.035	-0.004	-0.041
	(0.074)	(0.076)	(0.074)	(0.071)
8-10km	-0.154*	-0.087	0.021	-0.095
	(0.082)	(0.086)	(0.082)	(0.086)
Observations	1,457	1,457	1,592	1,592
Dep. Var. Mean	0.353	0.340	0.542	0.640
Dep. Var. SD	0.478	0.474	0.498	0.480
Number of Clusters	83	83	85	85
Controls	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. The dependent variable equals 1 if a party in the People's Alliance coalition won a: state constituency seat in 2013 (column 1); federal constituency seat in 2013 (column 2); state constituency seat in 2018 (column 3); federal constituency seat in 2018 (column 4). 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 all columns, controls are included. Controls are mentioned in the main text. 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	-0.004	-0.016
	(0.040)	(0.035)
4-6km	-0.063	0.028
	(0.061)	(0.045)
6-8km	-0.077	0.022
	(0.073)	(0.061)
8-10km	-0.154*	0.011
	(0.087)	(0.073)
Observations	1,457	1,593
Dep. Var. Mean	0.466	0.714
Dep. Var. SD	0.499	0.452
Number of Clusters	83	85
Controls	Yes	Yes

Notes: This table reports OLS estimates of Equation 1. The dependent variable equals 1 if the civic-nationalist coalition, won the state and federal elections in 2013 (column 1) and in 2018 (column 2). 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 all columns, controls are included. Controls are mentioned in the main text. Standard errors are clustered at the federal parliamentary constituency level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.