CIRJE-F-1237

Infrastructure and Structural Transformation: Evidence from Satellite, Administrative, and Multi-Generation Household Data in a Developing Country

Soyoung Kim
The University of Tokyo

Yuki Higuchi
Sophia University

Kei Kajisa
Kyoto University

Yasuyuki Sawada
The University of Tokyo

November 2024

Infrastructure and Structural Transformation: Evidence from Satellite, Administrative, and Multi-Generation Household Data in a Developing Country

SOYOUNG KIM, YUKI HIGUCHI, KEI KAJISA, AND YASUYUKI SAWADA*

November 7, 2024

Despite emerging academic interest in place-based policies, their impact on long-term structural transformation remains under-investigated, especially in developing countries. This study explores the combined effects of infrastructure development (highway, industrial park, and school establishments) in transforming agrarian communities in the Philippines using 40 years of family dynasty data, combined with satellite imagery and public administrative data. The results suggest that infrastructure development has led to structural transformations by increasing the probability of male employment in modern sectors and facilitating female human capital investments. Additionally, both the demand and supply sides of labor are key to successful modernization through place-based policies.

JEL Codes: I2, J2, O1

Keywords: Place-based Policy; Human Capital; Occupational Choice; Structural Transformation; The Philippines

---

* Kim: University of Tokyo, soyoung.kim@e.u-tokyo.ac.jp; Higuchi: Sophia University, higuchi@sophia.ac.jp; Kajisa: Kyoto University, kei.kajisa@gmail.com; Sawada: University of Tokyo, sawada@e.u-tokyo.ac.jp. We would like to thank the late Professor Nobuhiko Fuwa for his tremendous contribution in initiating this research project. The Japan Society for the Promotion of Science provided funding (JSPS, Grant Nos. 26220502, 17H01653, and 22K21341). We also thank the residents of the villages in the study and Agnes Chupungco and Elvie Dumayas of the College of Public Affairs and Development, University of the Philippines Los Banos, for their cooperation in the tracking surveys. Further, we appreciate the support from the late Professor Robert Evenson, who allowed us to use his datasets, and Professor Barry M. Popkin, who provided us with detailed information and data on the initial surveys. We also thank Tomohiro Hara, Takafumi Kawakubo, Saori Nishimura, and Haruki Suto for their excellent research assistance, and Marianne Bertrand, Patricia Cortes, Simon Fuchs, Robin Gong, Kohei Kawaguchi, Bryant Kim, Hisaki Kono, Takashi Kurosaki, Costas Meghir, Mushfiq Mobarak, Kentaro Nakajima, Sangyoon Park, Yasuharu Shimamura, Chikako Yamauchi, and other participants at the Asian Development Bank Economists' Forum, the Japanese Economic Association Conference, the Asian Meeting of the Econometric Society, the Western Economic Association International Annual Conference, the Asian and Australasian Society of Labor Economics Conference, the Kansai Labor Conference, Nihon University, and the Hong Kong University of Science and Technology for their constructive comments. We are responsible for any errors in this article; the opinions expressed in this article are our own and do not reflect the views of affiliated organizations.

1. Introduction

The theoretical effects of place-based policies are mixed, and empirical evidence is inconclusive (Glaeser and Gottlieb, 2008; Kline and Moretti, 2014a; Neumark and Simpson, 2015). One strand of literature focuses on short-term effects, and another strand investigates the issue of long-term structural transformation from agriculture to non-agriculture as an indispensable process of economic development (Lewis, 1955; Ranis and Fei, 1961; Matsuyama, 1992; Hayashi and Prescott, 2008; Duarte and Restuccia, 2010; Bustos et al., 2016). While these studies conventionally employ aggregated data built on a dynamic general equilibrium framework, more recent studies have addressed key heterogeneities, particularly in terms of informality, family, gender, and intergenerational dynamics, using microdata (Doepke and Tertilt, 2016; Gollin and Kaboski, 2023).[1]

Within these two strands of literature, the question arises as to whether and how place-based policies facilitate long-term structural transformation. It is imperative to adopt and analyze microdata to fully understand the influence of these policies on long-term individual- and household-level heterogeneous outcomes, which are critical for assessing their effectiveness in driving structural changes. However, the existing studies on the nexus between place-based policies and structural transformation utilize aggregated data (Fan and Zou, 2021; Heblich et al., 2022). Moreover, to the best of our knowledge, no study has examined its individual-level heterogeneities because “long-T” panel datasets have been rarely available until recently (Fernández-Val and Weidner, 2018).

To fill this critical research gap, this study constructs and analyzes unique individual-level family dynasty data covering the four decades from

[1] Banerjee et al. (2020) and Faber (2014) investigate the effects of large infrastructure using county-level data in China. Ghani et al. (2016) use establishment-level data such as entrants and incumbents. Shenoy (2018) exploits nighttime light data from 1992 to 2012 in India. Chaurey (2017) uses aggregated and firm-level panel data as well as data from household surveys to ascertain the impact of location-based tax incentive schemes in India.

1977 to 2017 in the regions surrounding Lake Laguna in the Philippines using a quasi-experimental setting. Since the late 1970s, development policy implemented through highway construction and the consequent establishment of industrial parks have targeted the west side of Lake Laguna; the east side was not included in the policy. While both sides of the lake were the agrarian periphery of the Philippines' capital city until the 1970s, the western side experienced rapid socioeconomic development from the 1980s. By exploiting such a quasi-experimental setting, we investigate the effects of place-based policy on individuals' choices of education and occupation.

We apply a counterfactual framework to a hybrid of three data sets, the primary one being our original long-term tracking data. Based on a household survey conducted in 1977 (Evenson, 1980), we track all of the original households, their descendants, and the household members of the descendants over a period of more than four decades. The second dataset comprises long-term satellite imagery from the Landsat dating back to the 1970s, and the third consists of administrative data on school openings from the Department of Education of the Philippines (DepEd). Using these combined data sets and following the lead of Jensen (2012) and Heath and Mobarak (2015), we adopt the difference-in-differences (DID) approach to exploit the cohort- and geography-specific variations and to identify impacts.

Our study yielded three key findings. First, we found that place-based policy facilitated rapid structural transformation and occupational changes across generations, with the construction of highways and industrial parks significantly increasing the probability of employment in the modern sector. Second, although the policy-induced creation of modern jobs increased both the expected return on education and the opportunity cost of educational investment, the latter seemed to be dominant among males. Although the number of years of schooling increased over time with improved school access, the availability of modern jobs discouraged males from obtaining tertiary education. These findings suggest that the successful modernization of developing countries hinges on the interrelated effects of labor demand-side (i.e., job creation through land transportation and industrial parks) and supply-side (i.e., school establishment) policies. Third, the occupational dynamics outside the treated area, which involved transitioning into a service-led economy prior to industrialization, appear to be consistent with Rodrik's (2016) hypothesis of premature deindustrialization.

Our study contributes to three lines of study. First, by employing a quasi-experimental approach to evaluate large-scale place-based policies, our study offers, at the very least, a partial response to the existing critiques regarding the "narrowness" of the experimental approach in recent development economics (Rosenzweig, 2012; Deaton, 2020; Ravallion, 2020). The novelty of our research lies in the fact that randomized controlled trials of large-scale place-based policies cannot be designed and implemented to identify factors driving long-term economic growth and structural transformation. We also consider our approach to be a mechanistic experiment, as defined by Ludwig et al. (2011), designed to test the structural transformation hypothesis.

Second, given that transport and industrial infrastructure are crucial components of place-based policies, our study expands the related literature (Faber et al., 2014; Ghani et al., 2016; Baum-snow et al., 2017; Donaldson 2018; Adukia et al., 2020; Asher and Novosad, 2020; Banerjee et al., 2020; Brooks and Donovan, 2020; Akresh et al., 2023). Our findings suggest that infrastructure development significantly affects regional development by lowering transport and travel costs, enhancing access to labor and other markets, and improving public services such as education.

Finally, our research contributes to the literature on the intergenerational transmission of education, occupation, and earnings. Although such dynamics have been extensively studied (e.g., Solon, 1999; Black et al., 2005; Black and Devereux, 2011), most existing studies have focused on two generations (from parents to children) or at most three generations (from grandparents to grandchildren). Our long-term tracking data allowed us to conduct a deeper investigation into intergenerational dynamics as we tracked all of the descendants of the original households and gathered and analyzed information from the first to the fifth generation.

The remainder of this paper is organized as follows. Section 2 explains the research background, Section 3 describes our data and identification strategy, and Section 4 presents a descriptive analysis. Section 5 develops an empirical framework, and Section 6 presents the estimation results. Finally, Section 7 concludes the paper.

2. Background

Our study focuses on the Laguna Province in the Philippines, located south of Manila, the capital city. As one of the largest provinces in the country, it has a population exceeding 3 million. Laguna Province is indicated in green in Figure 1. The province is home to Lake Laguna, the largest lake in the country, which is located centrally within its boundaries. Before the 1970s, the province's economy was driven primarily by the agriculture and fisheries that developed along the lake. The National Economic and Development Authority (2011) stated in its regional development plan that the province remained predominantly rural until the 1970s (p.7). Furthermore, as presented in Appendix A, the built-up areas on both sides of the lake were equally limited in the 1970s.

[Insert Figure 1 Here]

In 1965, during national address, President Macapagal announced a plan to establish the Pan-Philippine Highway System as a top-priority project to connect the northern and southern ends of the country.[2] The Laguna Province

[2] https://www.officialgazette.gov.ph/1965/01/25/diosdado-macapagal-fourth-state-of-the-nation-address-january-25-1965/ (retrieved on 23 January, 2024)

is situated along the proposed route, and there are two potential routes for the passage of the highway: the westbound or eastbound route of Lake Laguna. At that time, both sides of the lake had roads, which were constructed under American colonial rule in 1898. Starting from Metro Manila, Manila South Road traversed the western side of the lake, while the Manila East Road was connected to the eastern side of Laguna Province via Rizal Province.

Several factors influenced the decision to pursue a westbound plan for the Pan-Philippine Highway System. First, the topography of the westbound route was slightly more favorable for construction (NEDA, 2011; JICA and NEDA, 2014). In addition, political considerations most likely affected the choice of the westbound route (Manapat, 1991; Dalisay, 2010). It is also worth noting, as mentioned in the 2021 Asian Development Bank (ADB) report, the decision-making process, even for major infrastructure projects in the Philippines, is often conducted without an exhaustive feasibility study. Thus, it appears that that the westbound route was not selected because of its superior developmental potential.

A segment of the Pan-Philippine Highway System, extending from Metro Manila (Point A in Figure 1) to Alabang (Point B), was established in 1969. Owing to the opening of this segment, the highway network was extended from Point B to Calamba City (Point c). In 1978, operations commenced to construct (SLEx) to facilitate connectivity between Points B and C, running alongside the western side of the lake, (Hayami and Kikuchi, 2000). The Southern Luzon Expressway runs parallel to Manila South Road, which has historically experienced increasing transportation demand, thus significantly improving accessibility to the west side of Laguna Province from Manila.

Another factor contributing to the modernization of local infrastructure is the establishment of industrial parks along highways. The first industrial park in this area opened in 1980, three years after the SLEx. From that point until the 2000s, several industrial parks were constructed in the area. A prime example of this is the Laguna Technopark in Calamba, which was established in 1989.

This industrial park spans 460 ha and, according to the 2014 annual report from the Philippine Economic Zone Authority, has generated over 100,000 jobs. A comprehensive list of all the industrial parks constructed in our study area is provided in Appendix Table B1.

In sum, the westbound route was selected for the SLEx, and highway construction led to the opening various industrial parks and industrial development on the west side of the Laguna Province from 1978. While both the western and eastern sides of the lake had an existing road at the time of highway construction, only the western side was equipped with the highway and subsequent industrial parks. Therefore, we considered the east side of Laguna Province as a counterfactual scenario of the west side, which would not have industrial development.

3. Data and Identification Strategy

In this section, we describe the datasets employed and present our identification strategy for evaluating the impacts of place-based policies.

3.1. Tracking Survey

The primary data set used in our analysis was derived from our original tracking survey, which was based on the Laguna Multipurpose Household Survey (LMHS) (Evenson, 1980). The initial survey was conducted in 1975, targeting 34 villages in Laguna Province and interviewing 576 households. However, as the original data files and respondent lists from the 1975 survey were unavailable, we conducted a tracking survey based on the second wave of the LMHS, conducted in 1977. This survey, which targeted a subset of households sampled in the first wave, was conducted just before the completion of the SLEx. We maintain that the households and villages sampled in the 1977 survey adequately represent the socioeconomic conditions of Laguna Province. Data from 322 households, referred to hereafter as the "original households," from the 23 villages surveyed in 1977 were used as baseline information for our tracking survey. Figure 1 shows the locations of the 23 sample villages (in yellow). The survey was later conducted in 1979 (Rosenzweig and Wolpin, 1986) as well as in 1982, 1985, 1990, 1992, and 1998 (Ejrnæs and Pörtner, 2004). Further information on the LMHS is provided in Appendix A.

We conducted a tracking survey to follow all members of the original household, including those who were born in or joined the household after 1977, and their descendants. These descendants included the children, grandchildren, great-grandchildren, and great-great-grandchildren of the original household head. Additionally, we gathered data on the birth year, educational achievements, and lifetime occupations of those who had died by the time of our tracking survey. This unique dataset enables us to analyze the educational and occupational choices of individuals during the initial phase of the SLEx and industrial parks. It should be noted that we could accurately distinguish and track both emigrant and non-emigrant households because we had access to the original list of respondents from the 1977 survey. The strength of our research design of 40 years of tracking is that we can accurately capture dynamic structural changes from the viewpoint of representative families in the 1970s despite a lack of information about immigrants in the study area.

We endeavored to identify and interview as many descendants as possible from the original 322 households and used proxy reports to collect information from all available members and descendants.[3] After interviewing respondents, we asked them to introduce other potential respondents who could provide the missing information. We repeated this data collection process until we had gathered all of the information about the members in the family tree. Appendix A provides a detailed description of the tracking procedure.

[3] Rosenzweig (2003) demonstrates that the mean difference in years of education between self-reports and proxy reports is not statistically different. Furthermore, the distributions of primary occupations, as determined by two interview methods, show no significant difference. This suggests that proxy reports are a reliable method for gathering basic information.

3.2. Tracking Data

Out of the 322 households initially surveyed in 1977, we were able to identify 318, indicating that we managed to reach 98.7% of the original households, a percentage that surpassed that of other existing surveys. Of these 318 households, 4,992 were established at the time of the 2017 survey. This number includes current households that inherited their original households, with the majority being spin-offs. The structure of the tracking survey is illustrated in Figure 2. We gathered information on 23,650 individuals from 4,992 households comprising 318 family trees. Almost half of the current households had remained in the same village as their original counterparts, and nearly two-thirds were located in the same municipality as the original village.

[Insert Figure 2 Here]

The survey employed a nested structure encompassing four generations descending from the original household. We initiated the study with the original household (first generation) and surveyed the children (second generation), grandchildren (third generation), great-grandchildren (fourth generation), and great-great-grandchildren (fifth generation). This allowed us to construct original and unique long-term panel data.

3.3. Modernization Treatment

In our study, we used the counterfactual framework, defining the treatment as the cohort- and geography-specific "exposure to modernization" at the village level. The treatment was determined based on the proximity of each village to the nearest entrance/exit of the SLEx. Figure 3 demonstrates how proximity was measured and provides the proximity distribution among the sample villages. The average distance from the village centroid to the nearest entrance/exit point was 0.17 (in decimal degrees), with a standard deviation of 0.12. All eight villages situated on the western side of the lake fell below the mean (inside the red box), and were categorized as our "treatment" zone. The average distances were 0.28 and 2,60 decimal degrees in the treatment and control zones, respectively. We used this binary indicator and the distance variable as treatment variables in our estimation.

[Insert Figure 3 Here]

To focus our analysis on the relevant working-age population and exclude those still educated, we established an age restriction for individuals aged 30−79 years in 2017. Because we sought to examine lifetime occupations, we set the minimum age at 30 years, above which individuals are less likely to make major occupational changes in the Philippines.[4] Due to the small number of observations over the age of 80 years, we restricted the age cohort to < 80 years. We considered individuals who were 20 years old or younger in 1978 (or, equivalently, under 60 years old in 2017) as the cohort likely to be affected by the treatment, which would thus influence their decisions regarding education and career. This age range accounted for 69% of the total sample in our tracking data. In other words, we considered individuals aged 30–59 years as the treated cohort and those aged 60–79 years as the control cohort. Additionally, we excluded spouses who had joined the family tree only through marriage or cohabitation because they were likely to have spent their childhood elsewhere. Given the availability of reliable occupational and family background data, our sample in the statistical analysis included 4,086 individuals from 3,296 households across 294 family trees.

[4] According to the International Labor Organization (2024), the labor force participation rates for individuals aged 15−24 and 25−54 were 44.7% and 74.1%, respectively, in 1991. The corresponding numbers were 37.1% and 73.1% in 2017. Based on these figures, we decided to set 30 years old as the minimum age threshold for occupational decisions.

3.4 Parallel Trend Test Using Satellite Imageries and Census Data

It is crucial to conduct a formal test to determine whether the western (treated) and eastern (control) sides of the lake differ solely in terms of modernization. Because the data from the first LMHS survey in 1975 were unavailable, we alternatively used Landsat satellite imagery and Census data to examine parallel trends. In addition, we conducted a placebo test using a subsample of emigrated individuals who may not have been exposed to the treatment, and the results supported the assumption of a parallel trend (see Section 6.3).

Regarding satellite imagery, cloud-free Landsat images were used to classify the land cover. Each image was classified into four categories (i.e., water body, vegetation, bare land, and built-up land) using a random forest algorithm. We postulated the following event study model to examine parallel trends:

$$(1) \qquad Y_{vt} = \gamma_v + \lambda_t + \sum\nolimits_{\mathrm{t}} \beta_t d_v \lambda_t + \varepsilon_{vt},$$

where Y_{vt} represents the proportion of built-up areas within the boundary of village v in year t (t = 1972, 1976, 1989, 1992, 1996, 1999, 2008, 2011 and 2016) based on the availability of Landsat images. The village and year fixed effects are denoted by γ_v and λ_t, respectively. The treatment variable (either the distance from the nearest entrance or the treatment indicator) is denoted by d_v. The final term ε_{vt} is a well-behaved error term. In this specification, the estimated coefficient for each year β_t represents the treatment effect specific to each year.

Using 1976 as the reference year, Panel A in Figure 4 presents the event study plots to examine the pre-trends with 95% and 90% confidence intervals derived from Equation (1). The left and right figures are based on the continuous distance variable and binary treatment indicator, respectively. Both figures depict a gray area representing the period of Western modernization, which commenced in 1977 and continued until the establishment of the largest industrial park in 1989. In both figures, we have insignificant coefficients prior to the modernization period, suggesting that parallel pre-trend assumptions among villages hold. This finding reinforces the internal validity of the identification strategy.

[Insert Figure 4 Here]

Furthermore, we observed consistent treatment effects after the modernization period with reasonable statistical significance. Based on the binary treatment variable, modernization increased the built-up proportion by 29 percentage points in 2016. This increase is economically significant given that the average built-up proportion in the control villages was only 10% in 2016.

Next, using Census data from 1960 to 2015, we checked the parallel trend of the population as an alternative dependent variable, Y_{vt}, in Equation (1), which is estimated for years t = 1960, 1970, 1975, 1980, 1990, 1995, 2000, 2007, 2010, and 2015 based on the years the Census was conducted. We used 1975 as the reference year. Because Census data are available at the municipality (not village) level, v stands for the municipality in the population analysis. Our 23 sample villages were located in ten municipalities, three of which were in the treated area and seven in the control area. Panel B presents the coefficient of the event study estimation. The results illustrate that the population growth in the two groups was concurrent prior to the onset of modernization but began to diverge after the treatment. Appendix A provides additional checks for parallel trends using educational data from the Census.

3.5 School Availability Treatment

Duflo (2001) highlighted that physical accessibility to schools in each community can significantly affect the accumulation of human capital in developing countries. In our empirical model, we characterized school availability as the presence of public primary schools within each residential village when each individual was 6 years old. This variable was formulated using administrative data from the Department of Education. Figure 5 illustrates the locations of public primary schools before and after the modernization period. A comparison of Panels A (1979) and B (2019) reveals a substantial increase in the number of primary schools. A rigorous examination of this supply side factor in education is crucial to accurately estimate the treatment effect of modernization on human capital accumulation.

[Insert Figure 5 Here]

4. Descriptive Analysis

In this section, we present a descriptive analysis based on our tracking data.

4.1 Main Outcomes

Occupation data were collected based on the 24 industrial classifications specified in the 2009 Philippine Standard Industrial Classification. While we collected information on both the current and lifetime main occupations, we focused on the lifetime main occupation. This included determining those who passed away or retired at the time of the survey in 2017 and conducting an analysis of non-temporal occupational decisions over their lifespan.

To simplify the empirical analysis, we aggregated these classifications into six occupational categories: agriculture, manufacturing, industry, traditional services, and modern services. A comprehensive list of all 24 classifications and their corresponding categories used in our analysis is provided in Appendix Table B2. As the manufacturing and modern services sectors are important industries developed owing to the province's modernization, we combined these two categories into a single category, which we labeled the "modern sector," to investigate the structural transformation. We measured educational attainment by the number of years of schooling using information on the highest grade completed.

4.2 Socioeconomic Characteristics

Table 1 presents the sample's descriptive statistics. Our study sample, comprising individuals aged between 30 and 79 years in 2017, had an average age of 44.7 years, with females constituting 49% of the group. For the lifetime occupation as of 2017, 42% of the sample were employed in the modern sector for lifetime occupations. Of these, 20% worked in manufacturing and 22% worked in modern services. The average educational attainment was 9.96 years of schooling, a duration that in the educational system of the Philippines indicates the discontinuation of upper secondary school. To investigate the intergenerational transmission of human capital, we used data on the fathers' completed years of education, which averaged 5.98 and was slightly higher than that of the mothers.

[Insert Table 1 Here]

Table B3 in the Appendix displays the descriptive statistics for the non-migrant subsample, which comprises individuals who stayed in their original municipalities. As depicted by Figure 2, this group includes those residing in the same village (48%) and other villages in the same municipality (17%). The values for this subsample are comparable to those for the entire sample, although non-migrant individuals are slightly older and have fewer years of education.[5]

4.3 Educational Attainment and Occupational Choice by Treatment

Panel A in Figure 6 shows the proportion of completed grades against age in 2017, demonstrating an overall progression in education across generations in both the treatment and control zones. The shift from primary to secondary

5 The regression results obtained from the subsample of non-migrants are presented in Appendix C. The results are qualitatively similar to those obtained from the entire sample as presented in the main text.

graduate "dominance" to secondary graduate "dominance" was slightly more pronounced in the treatment zone compared to the control zone. We also noted the simultaneous presence of a relatively high percentage of primary and tertiary graduates in the control zone. Since tertiary-level education is required in many modern service sectors, this feature may be associated with the possibility of "premature deindustrialization" toward the servicification of the economy.

[Insert Figure 6 Here]

Panel B illustrates the age-specific distributions of the main lifetime occupations by treatment. The treatment status was based on the location of the original village regardless of the individual's current residential location (see Section 5.1). In both the treated and control zones, we noted a significant decrease in the employment share of the agricultural sector across generations and a shift in sector-specific employment away from agriculture. Furthermore, the figure shows a more pronounced presence of manufacturing in the treated zone. In the treated zone, the manufacturing sector emerged as the dominant sector among those under 40 years, whereas the agricultural sector remained prevalent among cohorts over 60 years. The traditional service sector was common among the middle-aged group; that is, those in their 40s and 50s.

Conversely, in the control zone, the manufacturing sector did not dominate in any age group. Instead, traditional services held the highest share among those under the age of 50, followed by modern services. The solid and dashed vertical lines represent the transition points from agriculture to manufacturing and modern services, respectively. A notable contrast is evident in the earlier manufacturing-based industrialization in the treatment zone (solid line) compared with the earlier transition to modern services in the control zone (dashed line). This contrast suggests premature deindustrialization; that is, leapfrogging from agriculture directly to services without adequate manufacturing experience, in the control zone. Earlier servicification in the control zone seems to be supported by greater progress in tertiary-level education, as indicated in Panel A.

5. Empirical Model

In this section, we first explain our empirical specifications for occupational choice and then those for human capital investment.

5.1 Occupational Choice

As our aim was to evaluate the impacts of modernization treatment, we adopted the DID framework for continuous treatment. We denote d_j as the distance from village j, where the original household (i.e., the ancestor) of individual i resided at the time of the 1977 survey, to its nearest entry point to the SLEx. The distance was measured in decimal degrees from the centroid of the village. To avoid selective migration, we use the distance from the original village, regardless of whether individual i is still in the same village. According to Figure 2, 48% of the sample individuals were in the same village, and another 17% were living in other villages in the same municipality. Appendix 3 verifies the robustness of our main results through a focus only on non-emigrants. Moreover, we adopted a binary treatment variable, which takes one if village j is located on the west side of Lake Laguna and zero otherwise (i.e., the distance is below the mean). The squared term d_j^2 is dropped when d_j is a binary treatment indicator. A_t is an indicator for individual i belonging to each after-treatment cohort; that is, those in their 30s, 40s, or 50s at the time of the 2017 tracking survey. The reference category included those aged 60–79 years old.

Under these settings, we employed the following form:

$$V^k_{ihjt} = \gamma^k A_t + \eta_1^k d_j + \eta_2^k d_j^2 + \theta^k (A_t d_j) + X_{ihjt}\lambda^k + \varepsilon^k_{ihjt}, \quad (2)$$

where V^k_{ihjt} is the probability of the outcome variables of the lifetime occupation k of individual i in family tree h of village j born in year t. The sector of individual occupation k is defined in two ways: fine and broad. The fine occupation category includes agriculture, manufacturing, industry, modern services, and traditional services. The broad category is binary, consisting of modern (manufacturing and modern services) and traditional occupations (all other services). Appendix Table B2 provides further data and our categorization for statistical analysis. X_{ihjt} is a vector of control variables, including gender, age, age squared, and a dummy variable that takes the value of one if the father is in the modern sector and zero otherwise. The standard error, ε^k_{ihjt}, is clustered at the village-cohort level as the variation in $A_t d_j$ occurs at this level. Because the number of clusters is small for the binary indicator (treated/control zones × 4 cohorts = 8 clusters), we additionally report p-values obtained using the wild bootstrap method (Cameron, Gelbach, and Miller, 2008). θ^k is a parameter of interest that can be interpreted as a cohort-specific heterogeneous treatment effect of modernization.

5.2 Education Choice

To further explore the potential mechanisms underlying the observed occupational decisions, we divided them into two channels: 1) the direct employment opportunity effect by modernization treatment captured by Equation (2), and 2) indirect channels promoting human capital investments.

We theorize the following model for years of schooling, S_{ihjt}, of individual i in family tree h of village j born in cohort t:

$$(3) \qquad S_{ihjt} = \mu E_{ihjt} + \gamma A_t + \eta_1 d_j + \eta_2 d_j^2 + \theta(A_t d_j) + Z_{ihjt}\lambda + \varepsilon_{ihjt},$$

where E is accessibility to a school at the school age of individual i. Z_{ihjt} is different from X_{ihjt} in Equation (2) in that it excludes the dummy variable that takes the value of one if the father is in the modern sector (zero otherwise) but includes the father's completed years of schooling as we are interested in the intergenerational dynamics of education. In Equation (3), we can quantify two treatment effects: one for school availability treatment, μ, and the other for modernization treatment, θ.

6. Estimation Results

First, we present the results of occupational choice in either broad (traditional or modern) or detailed categories (manufacturing, industry, modern services, traditional services, and others). Thereafter, we explore the effects of modernization treatment, school accessibility, and other determinants of human capital investment.

6.1 Results I: Occupational Choice

To examine whether an individual chose to work in the modern sector, we used a dummy variable that takes the value of one if an individual's lifetime primary occupation is in the modern sector as an outcome variable and zero otherwise. Table 2 reports the estimation results, where Columns (1) and (2) use the distance to the nearest highway exit/entrance and Columns (3) and (4) use the treatment indicator as the key independent variable.

[Insert Table 2 Here]

In Columns (1) and (2), the statistically significant and negative coefficients of the interaction term (i.e., θ in Equation (2)) are found for all of the treated age cohorts. This indicates that individuals who are young enough to benefit from modernization and live closer to highways are more likely to choose either manufacturing or modern services as their primary lifetime occupation. The point estimate ranges from -0.072 to -0.060. As the mean distances to the SLEx exit in decimal degrees among the treated and control villages are 0.28 and 2.60, respectively, the average gap is 2.32. The estimated coefficients indicate that individuals in the treated villages are 14–16 percentage points more likely to work in the modern sector. The same pattern was observed in Columns (3) and (4), where the binary treatment is used. The point estimate ranged from 0.135 to 0.167, which is consistent with the interpretation of the results observed with distance. The effects differed across age groups, with a slightly larger magnitude observed in younger individuals. We may interpret the impact on the younger cohort as the long-term impact of modernization treatment. Appendix Figure B1 graphically illustrates these estimated effects.

We also noted that having a father who is or has been employed in the modern sector significantly increases the likelihood of participation. However, the point estimate of 0.044 in Column (4) is smaller than that of the estimated treatment effects; this difference suggests that place-based policies can transcend intergenerational inequality dynamics. The pronounced effect of gender, specifically female gender, on the choice to engage in manufacturing or modern services is also evident. To determine the association between this finding and our primary findings, we estimated Equation (2) by gender.

Table 3 reports the heterogeneous effects by gender. Column (2), for instance, indicates that treated females in their 40s are as much as 17.6 percentage points more likely to engage in a modern sector job than their counterparts in the control villages. Appendix Figure B1 provides a graphical illustration. Overall, the magnitude of the estimated coefficient is only slightly larger among females than among males. However, a notable gender difference was observed in the intergenerational pattern; the significant and positive coefficient for males with fathers employed in the modern sector suggests that only males experienced an intergenerational effect from having a father working in the modern sector.

[Insert Table 3 Here]

Modernization positively affects the likelihood of individuals working in the modern sector, prompting us to examine its sectoral composition in detail.

Instead of dividing the six sectors into two groups, we categorized them into six distinct sectors: agriculture, manufacturing, industry, modern services, traditional services, and others, an approach that allowed us to discuss the differences in the evolution of the granular structural transformations. We used the multinomial logit model (MNL) to estimate Equation (2) and designated agriculture as the reference group. Table 4 presents the estimated coefficients derived from regressions for each sector. Compared to agriculture, manufacturing and modern services exhibited negative coefficients in Columns (1) and (3), indicating a shift to non-farm occupations in previously agrarian villages. Although the remaining sectors had negative coefficients, they were not statistically different from zero, and the point estimates were smaller than those of the modern sector. The same pattern was observed in Columns (6) to (10), where the binary treatment was used, which is consistent with the interpretation of the results observed with distance. A similar pattern was observed among the different age cohorts, indicating a greater effect in the younger age groups than in the older age groups. In terms of intergenerational effects, the more a father works in the modern sector, the higher the likelihood of his offspring working in the non-farm sector. As the coefficient of the female dummy reveals, females are more likely to work in modern services than in agriculture, presumably due to the infrequent participation of women in agriculture due to the physically demanding nature of the work (i.e., brawn-intensive work, as per Pitt et al., 2012) or limited access to land ownership.

Appendix Figure B2 illustrates the predicted probability for each occupation category across the age range of our sample to trace structural changes. Concentrating on the intersection of agriculture, manufacturing, and modern services, the transition point from agriculture to manufacturing is situated at an older age in the treatment zone: 55–60 years in the treatment zone (Panel A) and 45–50 years in the control zone (Panel B). This indicates that the structural shift from agriculture to manufacturing transpired earlier in villages that received the treatment. Conversely, during the same period, the control villages appear to have bypassed manufacturing and transitioned directly to modern services.

[Insert Table 4 Here]

Tables 5A and 5 B present the results of the MNL regression analyses by gender to examine heterogeneity. The first observation was that, comparing the estimated coefficients in both tables, males respond to modernization by increasing the likelihood of employment in manufacturing and modern services as opposed to agriculture. The transition from agricultural to non-agricultural work was more predominant in males than females. Conversely, younger females demonstrated a higher propensity to engage in modern services. Consistent with the findings of the overall sample, both genders exhibited significant and positive effects when their fathers were employed in the modern sector, which suggests that having a father engaged in the modern sector could motivate individuals to seek employment in non-agricultural fields.

[Insert Table 5 Here]

Appendix Figure B2 illustrates the predicted probability for each occupational category by gender. The earlier transition to manufacturing in the treated villages and the bypassing of manufacturing in the control villages seem evident for both males (Panels E and F) and females (Panels C and D), although the likelihood of employment in traditional services predominates for male workers. This phenomenon could be associated with the higher percentage of tertiary graduates in the control villages, as illustrated in Panel A of Figure 6.

6.2 Results II: Human Capital Investments

Having examined the direct labor “demand-side” channels of modernization on occupational choices, we now turn our attention to the indirect “supply side” channels through human capital investments. To that end, we applied the ordinary least squares model to estimate Equation (3) to determine the effect of modernization on educational attainment. The results of this estimation are detailed in Table 6 and illustrated graphically in Figure B3.

[Insert Table 6 Here]

Column (2) presents a negative and significant coefficient for females in their 30s and 40s, indicating that modernization treatment and the availability of modern sector jobs promoted school investment among females. As Heath and Mobarak (2015) and Lu et al. (2023) found in Bangladesh and China, respectively, younger females may have realized and responded to the increased return on education. Columns (5) and (6) present clear contrasts between the genders. The coefficient for females in their 30-50s was positive and significant, whereas that for males was negative and significant in the younger cohort. This negative impact on males may be attributed to the increased opportunity costs of investing in education. Indeed, the existing literature has revealed that increased job opportunities reduced educational investment in various settings, as seen in Atkin (2016) in Mexico after factory openings and in Rickman et al. (2017) in the U.S. after the shale boom. Likewise, our gendered results indicate that increased expected returns dominate the opportunity cost of schooling among females and vice versa among males.

To examine the complementarity between modernization treatment and school access, we delved deeper into how these treatment effects varied according to school accessibility by interacting the school access variable with the main DID term. As depicted in Table 7 and Appendix Figure B3 (Panels C and D), the statistically significant effect (negative) with the distance variable in Column (1) and the positive effect with the binary variable for the age cohort 40−59 in Column (4) indicate that exposure to modernized infrastructure positively impacts educational attainment when there is improved school access within the village, implying a potential complementarity between highways and school infrastructure. This trend is particularly noticeable among older age groups, suggesting that younger age groups may be less responsive to treatment despite having greater school accessibility.

[Insert Table 7 Here]

However, as indicated by the *p*-values in parentheses in Table 7 derived using a wild cluster bootstrapping procedure, these estimates were not statistically significant. Given that these estimated results are not robust when dealing with a small number of clusters, any interpretation of the complementarity between highways and school infrastructure should be approached with caution.

Separate regression results suggested that young women residing near highways were more inclined to attend school. Furthermore, female gender positively affected educational attainment. Parental education and school accessibility seem to be significant in both the complete sample and the gender-based subsample. Specifically, when controlling for access to school, the effect of modernization on education may be more closely tied to the demand side of education. Access to nearby schools is a key factor influencing education. School accessibility is set to one if a primary school is located within an individual's village (and zero otherwise). The treatment effect may be driven by the return to schooling, as corroborated by the higher returns for women in the Philippines. As Table 8 shows, data from the 2012 Philippines Family Income and Expenditure Survey in Laguna Province and Metro Manila indicate that returns on schooling for all workers and non-agricultural workers are marginally higher for females than males in the study area.

[Insert Table 8 Here]

6.3 Falsification Test using a Migrant Subsample

To assess the robustness of the findings, we conducted a falsification test using a subsample of migrants. The migrants in the falsification test are defined as individuals who have relocated from their original village on the east side of Lake Laguna to different locations, such as Metro Manila, provinces other than the Calabarzon region, or other countries. Table D1 in Appendix D presents the estimation results, replicating the analyses conducted in the previous sections. Generally, the results revealed that the main treatment effects were not significant when a subsample of migrants was used.

6.4 Robustness Check to Account for Spillover

To address the potential influence of spatial spillovers, we conducted a robustness check that controls for proximity to areas affected by the modernization treatment in Appendix E. Following Miguel and Kremer (2004), we incorporated ring terms, which capture the number of treated villages within specified distance intervals from each village, into our analysis. Table E1 presents an overview of the number of treated villages within three distance intervals: 0−5 km, 5−10 km, and 10−30 km from each sample village. Table E2 reports the estimation results that incorporate the ring terms. According to the table, the estimated treatment effects are larger than those reported in Table 2, especially for the youngest cohort age 30−39 in Full specification (4). These results suggest the existence of positive spillover effects that drive the downward bias in estimating the true treatment effect.

7. Concluding Remarks

This study examined the transformation of former agrarian villages in response to place-based policies aimed at enhancing connectivity and industrialization, with a particular focus on their effects on occupational dynamics and educational investment. By integrating three datasets—a household survey, satellite imagery, and administrative data—we found that increased exposure to infrastructure development, such as the construction of new highways and industrial parks, has led to structural transformation in Laguna Province of the Philippines. We leveraged the spatial and temporal variation in exposure to these place-based policies when individuals made decisions about schooling and occupation and employed a DID model to estimate the causal effect on years of schooling and lifetime primary occupation. Specifically, we used different age cohort dummies to capture the heterogeneous treatment effects. Additionally, we conducted a test to determine whether there were pre-existing trends in the proportion of built-up areas and population growth in the treatment villages compared to the control villages prior to the implementation of the policies. The event study plots, which showed no pre-trends, support the internal validity of our findings, and the non-significant estimated coefficients from the falsification test with a subsample of migrants support the parallel trend assumption.

The primary finding of this study is that investments in road infrastructure and the manufacturing sector influence occupational choices toward the modern sector. This result remains robust even when considering intergenerational effects and dummy treatment variables. A heterogeneity analysis involving gender indicated that male workers transitioned to the modern sector earlier than female workers. We also observed that individuals exposed to these infrastructural developments were more likely to work in manufacturing or modern services than in agriculture. This shift supports the transformation from agricultural to non-agricultural in previously agrarian villages, a change prompted by infrastructure development.

Although we found no statistically significant effect on human capital accumulation, an increase in years of schooling spurred by school construction suggests a complementarity between demand- and supply side interventions. The stronger treatment effect for female workers could be attributed to higher returns on schooling for females, particularly in non-farm occupations, as observed in the province.

REFERENCES

Adukia, Anjali, Sam Asher, and Paul Novosad. 2020. Educational investment responses to economic opportunity: Evidence from Indian road construction. *American Economic Journal: Applied Economics* 12 (1): 348–376.

Akresh, Richard, Daniel Halim, and Marieke Kleemans. 2023. Long-term and intergenerational effects of education: Evidence from school construction in Indonesia. *Economic Journal* 133 (650): 582–612.

Asher, Sam, and Paul Novosad. 2020. Rural roads and local economic development. *American Economic Review* 110 (3): 797–823.

Asian Development Bank (ADB). 2021. Completion report for Philippines: A road improvement and institutional development project. *ADB*.

Atkin, David. 2016. Endogenous skill acquisition and export manufacturing in Mexico. *American Economic Review* 106 (8): 2046−85.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2020. On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics* 145: 102442.

Baum-snow, Nathaniel, Brant J. Loren, Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. Roads, railroads and decentralization of Chinese cities. *Review of Economics and Statistics* 99 (3): 435–448.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2011. Migration and economic mobility in Tanzania: Evidence from a tracking survey. *Review of Economics and Statistics* 93 (3): 1010–1033.

Behrman, Jere R., Mark R. Rosenzweig, and Paul Taubman. 1994. Endowments and the allocation of schooling in the family and in the marriage market: The twins experiment. *Journal of Political Economy* 102 (6): 1131–1174.

Black, Sandra E., and Paul J. Devereux. 2011. Recent developments in intergenerational mobility. In *Handbook of labour economics volume 4b*, eds. D. Card, and O Ashenfelter. Amsterdam: Elsevier.

Black, Sandra, E., Paul J. Devereux, and Kjell G. Salvanes. 2005. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review* 95 (1): 437–449.

Brooks, Wyatt, and Kevin Donovan. 2020. Eliminating uncertainty in market access: The impact of new bridges in rural Nicaragua. *Econometrica* 88 (5): 1965–1997.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review* 103 (2): 897–947.

Bustos, Paula, Bruno Caprenttini, and Jacopo Ponticelli. 2016. Agricultural productivity and structural transformation: Evidence from Brazil. *American Economic Review* 106(6): 1320–1365.

Colin A. Cameron, Jonah B. Gelbach, and Douglas L. Miller. 2008. Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics* 90: 417–427.

Criscuolo, Chiara, Ralf Martin, Henry G. Overman, and John Van Reenen.

2019. Some causal effects of an industrial policy. *American Economic Review* 109 (1): 48–85.

Dalisay, Jose Y. 2010. *Builder of bridges: The Rudy Cuenca story.* Manila: Anvil Pub. ISBN 9789712724619.

Deaton, Angus. 2020. Randomization in the tropics revisited: A theme and eleven variations. In *Randomized control trials in the field of development: A critical perspective*, eds. Florent Bédécarrats, Isabelle Guérin, and François Roubaud. New York: Oxford University Press.

Dercon, Stefan, Pramila Krishnan, and Sofya Krutikova. 2013. Changing living standards in southern Indian villages 1975–2006: Revisiting the ICRISAT village level studies. *Journal of Development Studies* 49 (12): 1676–1693.

Card, David. 1999. The causal effect of education on earnings. In *Handbook of labor economics volume 3A*, eds. C. Ashenfelter and David Card. Amsterdam: Elsevier.

Doepke, Matthias and Tertilt, Michele. 2016. Families in macroeconomics. In *Handbook of macroeconomics* 2(23): 1789−1891, Amsterdam: Elsevier.

Donaldson, Dave. 2018. Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review* 108 (4–5): 899–934.

Duarte, Margarida, and Diego Restuccia. 2010. The role of the structural transformation in aggregate productivity. *Quarterly Journal of Economics* 125 (1): 129–173.

Duflo, Esther. 2001. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review* 91 (4): 795–813.

Ejrnæs, Mette, and Pörtner, Claus. 2004. Birth order and the intrahousehold allocation of time and education. *Review of Economics and Statistics* 86 (4): 1008–1019.

Ehrlich, Maximilian V., and Tobias Seidel. 2018. The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet. *American Economic Journal: Economic Policy* 10 (4): 344–74.

Evenson, Robert. 1980. *Note on the Laguna household survey in the Philippines*. Mimeograph.

Faber, Benjamin. 2014. Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. *Review of Economic Studies* 81 (3): 1046–1070.

Fan, Jingting, and Ben Zou. 2021. Industrialization from scratch: The construction of third front and local economic development in China's hinterland. *Journal of Development Economics* 152: 102698.

Fernández-Val, Iván, and Martin Weidner. 2018. Fixed effects estimation of large-T panel data models. *Annual Review of Economics* 10: 109–138.

Ghani, Ejaz., Arti G. Goswami, and William R. Kerr. 2016. Highway to success: The impact of the Golden Quadrilateral Project for the location and performance of Indian manufacturing. *The Economic Journal* 126: 317–357.

Glaeser, Edward L., and Joshua D. Gottlieb, 2008. The economics of place-making policies. *Brookings Papers on Economic Activity, Economic Studies Program* 39 (1): 155–253.

Gollin, Douglas, Casper W. Hansen, and Asger M. Wingender. 2021. Two blades of grass: The impact of the Green Revolution. *Journal of Political Economy* 129 (8): 2344–2384.

Gollin, Douglas, and Joseph P. Kaboski. 2023. New views of structural transformation: Insights from recent literature. *Oxford Development Studies.*

Harris, John R., and Michael P. Todaro. 1970. Migration, unemployment and development: A two-sector analysis. *American Economic Review* 60 (1): 126–142.

Hasan, Rana, Yi Jiang, and Radine Michelle Rafols. 2021. Place-based preferential tax policy and industrial development: Evidence from India's program on industrially backward districts. *Journal of Development Economics* 150: 102621.

Hayami, Yujiro, and Masao Kikuchi. 2000. *A rice village saga: Three decades of Green Revolution in the Philippines*. London: Macmillan Press.

Hayami, Yujiro, and Vernon W. Ruttan. 1985. *Agricultural development: An international perspective*. Maryland: Johns Hopkins University Press.

Hayashi, Fumio, and Edward C. Prescott. 2008. The depressing effect of agricultural institutions on the prewar Japanese economy. *Journal of Political Economy* 116 (4): 573–632.

Heath, Rachel A., and Mushfiq Mobarak. 2015. Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics* 115, 1–15.

Heblich, Stephan, Marlon Seror, Hao Xu, and Yanos Zylberberg. 2022. Industrial clusters in the long run: Evidence from million-rouble plants in China. *National Bureau of Economic Research Working Paper Series* 30744.

International Labour Organization. 2024. Labour Force Statistics database (LFS) ILOSTAT. http://ilostat.ilo.org/data.

Japan International Cooperation Agency (JICA) and National Economic and Development Authority (NEDA). 2014. Roadmap for transport infrastructure development for Metro Manila and its surrounding areas (Region III and Region IV-A). *JICA and NEDA*.

Jensen, Robert. 2012. Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *Quarterly Journal of Economics* 127 (2): 753–792.

Kline, Patrick, and Enrico Moretti. 2014a. People, places, and public policy: Some simple welfare economics of local economic development policies. *Annual Review of Economics* 6 (1): 629–662.

Kline, Patrick, and Enrico Moretti. 2014b. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics* 129 (1): 275–331.

Lanjouw, Peter, and Nicholas Stern. 2018. *Economic development in Palanpur over five decades*. Oxford: Oxford University Press.

Lewis, Arthur W. 1955. *Theory of economic growth*. London: Allen and Unwin Ltd..

Lu, Fangwen, Weizeng Sun, and Jianfeng Wu. 2023. Special economic zones and human capital investment: 30 years of evidence from China. *American Economic Journal: Economic Policy* 15 (3): 35-64.

Lu, Yi, Jin Wang, and Lianming Zhu. 2019. Place-based policies, creation, and agglomeration economies: Evidence from China's Economic Zone Program. *American Economic Journal: Economic Policy* 11 (3): 325–360.

Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan. 2011. Mechanism experiments and policy evaluations. *Journal of Economic Perspectives* 25(3): 17–38.

Manapat, Ricardo. 1991. *Some are smarter than others: The history of Marcos' crony capitalism.* New York: Aletheia Publications. ISBN 9719128704.

Matsuyama, Kiminori. 1992. Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory* 58 (2): 317–334.

Martincus, V. Christian, Jaeronimo Carballo, and Ana Cusolito. 2017. Roads, exports and employment: Evidence from a developing country. *Journal of Development Economics* 125 (2017): 21–39.

Neumark, David, and Helen Simpson. 2015. Place-based policies. In *Handbook of regional and urban economics,* eds. Gilles Duranton, Vernon Henderson, and William Strange. Amsterdam: Elsevier.

Pitt, Mark M., Mark R. Rosenzweig, and Mohammad Nazmul Hassan. 2012. Human capital investment and the gender division of labor in a brawn-based economy. *American Economic Review* 102 (7): 3531–60.

Popkin, M. Barry. 2020. Odyssey of a small-town midwestern boy to a scholarly path. *European Journal of Clinical Nutrition* 74: 979–982.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo. 2022. The human side of structural transformation. *America Economic Review* 122 (8): 2774–2814.

Ranis, Gustav, and John C. H. Fei. 1961. A theory of economic development. *American Economic Review* 51 (4): 533–565.

Ravallion, Martin. 2020. Should the randomistas (continue to) rule? *National Bureau of Economic Research Working Paper Series* 27554

Rickman, Dan S, Hongbo Wang, John V. Winters. 2017. Is shale development drilling holes in the human capital pipeline? *Energy Economics* 62: 283−290.

Rodrik, Dani. 2016. Premature deindustrialization. *Journal of Economic Growth* 21 (1): 1–33.

Rosenzweig, Mark R. 2003. Payoffs from panels in low-income countries: Economic development and economic mobility. *American Economic Review* 93 (2): 112–117.

Rosenzweig, Mark R. 2012. Thinking small: poor economics: A radical rethinking of the way to fight global poverty: Review essay. *Journal of Economic Literature* 50 (1): 115–127.

Rosenzweig, Mark R., and Kenneth I. Wolpin. 1986. Evaluating the effects of optimally distributed public programs: Child health and family planning interventions. *American Economic Review* 76 (3): 470–482.

Shenoy, Ajay. 2018. Regional development through place-based policies: Evidence from a spatial discontinuity. *Journal of Development Economics* 130: 173–189.

Solon Gary. 1999. Intergenerational mobility in the labor market. In *Handbook of labour economics volume 3A*, eds. O. Ashenfelter, and D. Card. Amsterdam: Elsevier.

The National Economic and Development Authority (NEDA). 2011. Calabarzon Regional Development Plan: 2011−2016. *NEDA*.

Walker, T.S., and J.G. Ryan. 1990. *Village and household economics in India's semi-arid tropics*. Baltimore: Johns Hopkins University Press.

Wang, Jin. 2013. The economic impact of special economic zones: Evidence from Chinese municipalities. *Journal of Development Economics* 101: 133–147.

Table 1

Descriptive Statistics in Study Villages of Laguna Province

	Mean	Min	Max	St.dv.	N
Age in 2017	44.47	30	79	11.05	4086
Female	0.49	0	1	0.50	4086
Younger cohort (30–49)	0.69	0	1	0.46	4086
Working in modern sector	0.42	0	1	0.49	4086
Working in manufacturing	0.20	0	1	0.40	4086
Working in modern services	0.23	0	1	0.42	4086
Father in modern sector	0.18	0	1	0.39	4086
Years of schooling	9.96	0	16	2.85	4075
Father's years of schooling	5.98	0	16	3.36	4071

Notes: Modern sector consists of manufacturing and modern services. Years of schooling is calculated by the highest completed grade.

Table 2

Lifetime Primary Occupation: Linear Probability Model

	Distance		Treatment indicator	
	(1)	(2)	(3)	(4)
Dependent variable: 1 [Modern Sector]				
Aged 30–39 × Distance	-0.071***	-0.072***		
	(0.018)	(0.018)		
Aged 40–49 × Distance	-0.060***	-0.063***		
	(0.018)	(0.019)		
Aged 50–59 × Distance	-0.061***	-0.061***		
	(0.022)	(0.022)		
Aged 30–39 × Treated			0.153***	0.154***
			(0.007)	(0.007)
			[0.110]	[0.102]
Aged 40–49 × Treated			0.161***	0.167***
			(0.010)	(0.011)
			[0.108]	[0.124]
Aged 50–59 × Treated			0.135***	0.137***
			(0.007)	(0.007)
			[0.102]	[0.102]
Female	0.199***	0.199***	0.199***	0.199***
	(0.021)	(0.021)	(0.028)	(0.028)
			[0.000]	[0.000]
Father in modern sector		0.043**		0.044***
		(0.020)		(0.011)
				[0.000]
Observations	4086	4086	4086	4086
Mean dependent value	0.424	0.424	0.424	0.424
R-squared	0.100	0.101	0.100	0.101

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (3) and (4), the p-value obtained using the wild bootstrap method is reported in brackets.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 3

Lifetime Primary Occupation by Gender: Linear Probability Model

	Female		Male	
	Distance	Treatment indicator	Distance	Treatment indicator
	(1)	(2)	(3)	(4)
Dependent variable: 1 [Modern Sector]				
Aged 30–39 × Distance	-0.085***		-0.042*	
	(0.026)		(0.022)	
Aged 40–49 × Distance	-0.054*		-0.057**	
	(0.029)		(0.022)	
Aged 50–59 × Distance	-0.058*		-0.046*	
	(0.030)		(0.024)	
Aged 30–39 × Treated		0.133***		0.117***
		(0.010)		(0.009)
		[0.000]		[0.000]
Aged 40–49 × Treated		0.176***		0.114***
		(0.012)		(0.007)
		[0.000]		[0.000]
Aged 50–59 × Treated		0.129***		0.089***
		(0.012)		(0.012)
		[0.000]		[0.000]
Father in modern sector	0.004	0.005	0.080**	0.080***
	(0.026)	(0.021)	(0.031)	(0.020)
		[0.825]		[0.005]
Observations	1990	1990	2096	2096
Mean Dependent Value	0.527	0.325	0.527	0.325
R-squared	0.071	0.074	0.070	0.073

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (2) and (4), the p-value obtained using the wild bootstrap method is reported in brackets.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4

Lifetime Primary Occupation: Multinomial Logit Model:

	Distance					Treatment indicator				
	Manufacturing	Industry	Modern Services	Traditional Services	Others	Manufacturing	Industry	Modern Services	Traditional Services	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Aged 30–39 × Distance	-0.499**	-0.124	-0.671***	-0.255	-0.392*					
	(0.232)	(0.259)	(0.189)	(0.206)	(0.235)					
Aged 40–49 × Distance	-0.473*	0.037	-0.611***	-0.274	-0.308					
	(0.260)	(0.224)	(0.188)	(0.200)	(0.251)					
Aged 50–59 × Distance	-0.434*	-0.082	-0.574**	-0.204	-0.318					
	(0.247)	(0.235)	(0.228)	(0.206)	(0.236)					
Aged 30–39 × Treated						1.031**	0.216	1.354**	0.400	0.932
						(0.515)	(0.841)	(0.545)	(0.676)	(0.684)
Aged 40–49 × Treated						1.323**	-0.106	1.531***	0.748	0.720
						(0.521)	(0.785)	(0.557)	(0.685)	(0.853)
Aged 50–59 × Treated						0.970*	0.414	1.124	0.382	0.396
						(0.570)	(0.777)	(0.704)	(0.660)	(0.704)
Female	1.606***	-2.701***	1.845***	0.743***	3.581***	1.624***	-2.722***	1.825***	0.734***	3.585***
	(0.167)	(0.377)	(0.188)	(0.161)	(0.251)	(0.172)	(0.375)	(0.189)	(0.160)	(0.251)
Father in modern sector	1.379***	1.015***	1.587***	1.312***	1.244***	1.353***	1.011***	1.622***	1.313***	1.233***
	(0.241)	(0.297)	(0.267)	(0.256)	(0.259)	1.624***	-2.722***	1.825***	0.734***	3.585***
Observations	4086	4086	4086	4086	4086	4086	4086	4086	4086	4086
Mean Dependent Value	0.198	0.079	0.226	0.286	0.106	0.198	0.079	0.226	0.286	0.106

Notes: Estimated coefficients are reported. The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (6) to (10), the p-value obtained using the wild bootstrap method is reported in brackets. The mean value of agriculture is 0.106.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5A

Lifetime Primary Occupation by Gender: Multinomial Logit Model (Female)

	Distance					Treatment indicator				
	Manufacturing	Industry	Modern Services	Traditional Services	Others	Manufacturing	Industry	Modern Services	Traditional Services	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Aged 30–39 × Distance	-0.319	0.144	-0.591**	-0.044	-0.223					
	(0.335)	(0.426)	(0.296)	(0.305)	(0.346)					
Aged 40–49 × Distance	-0.249	-0.385	-0.419	-0.088	-0.102					
	(0.259)	(0.316)	(0.263)	(0.298)	(0.319)					
Aged 50–59 × Distance	-0.038	-0.314	-0.223	0.241	0.027					
	(0.301)	(0.376)	(0.320)	(0.318)	(0.302)					
Aged 30–39 × Treated						0.653	-14.153***	1.458**	0.244	0.690
						(0.692)	(1.240)	(0.735)	(0.771)	(0.855)
Aged 40–49 × Treated						0.932	0.275	1.388	0.392	0.329
						(0.687)	(0.938)	(0.845)	(1.000)	(1.078)
Aged 50–59 × Treated						-0.149	0.209	0.381	-0.668	-0.443
						(0.665)	(1.077)	(0.831)	(0.772)	(0.733)
Father in modern sector	1.265**	2.578**	1.754***	1.389**	1.275**	1.234**	2.597**	1.781***	1.389**	1.259**
	(0.520)	(1.056)	(0.547)	(0.544)	(0.515)	(0.518)	(1.024)	(0.538)	(0.540)	(0.510)
Observations	1990	1990	1990	1990	1990	1990	1990	1990	1990	1990
Mean Dependent Value	0.234	0.003	0.293	0.219	0.197	0.234	0.003	0.293	0.219	0.197

Notes: Estimated coefficients are reported. The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (6) to (10), the p-value obtained using the wild bootstrap method is reported in brackets. The mean value of agriculture is 0.053.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 5B

Lifetime Primary Occupation by Gender: Multinomial Logit Model (Male)

	Distance					Treatment indicator				
	Manufacturing	Industry	Modern Services	Traditional Services	Others	Manufacturing	Industry	Modern Services	Traditional Services	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Aged 30–39 × Distance	-0.454*	-0.139	-0.551**	-0.331	-0.636*					
	(0.253)	(0.252)	(0.279)	(0.227)	(0.346)					
Aged 40–49 × Distance	-0.484*	0.006	-0.646**	-0.351*	-1.022*					
	(0.294)	(0.209)	(0.272)	(0.199)	(0.596)					
Aged 50–59 × Distance	-0.530	-0.197	-0.678**	-0.414**	-0.595*					
	(0.326)	(0.224)	(0.298)	(0.205)	(0.337)					
Aged 30–39 × Treated						1.119*	0.209	0.693	0.439	13.307***
						(0.646)	(0.843)	(0.628)	(0.764)	(1.198)
Aged 40–49 × Treated						1.327**	-0.097	1.030*	0.895	14.254***
						(0.658)	(0.746)	(0.575)	(0.605)	(1.460)
Aged 50–59 × Treated						1.630**	0.856	1.134	1.020	12.285***
						(0.681)	(0.747)	(0.697)	(0.647)	(1.762)
Father in modern sector	1.616***	0.989***	1.501***	1.329***	1.534***	1.572***	0.970***	1.511***	1.322***	1.494***
	(0.254)	(0.304)	(0.310)	(0.261)	(0.547)	(0.254)	(0.317)	(0.318)	(0.269)	(0.561)
Observations	2096	2096	2096	2096	2096	2096	2096	2096	2096	2096
Mean Dependent Value	0.163	0.150	0.162	0.349	0.019	0.163	0.150	0.162	0.349	0.019

Notes: Estimated coefficients are reported. The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (6) to (10), the p-value obtained using the wild bootstrap method is reported in brackets. The mean value of agriculture is 0.157.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6

Educational Attainment: Ordinary Least Squares Model

	Distance			Treatment indicator		
	Entire	Female	Male	Entire	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Years of schooling						
Aged 30–39 × Distance	-0.195	-0.407*	0.144			
	(0.175)	(0.211)	(0.159)			
Aged 40–49 × Distance	-0.143	-0.363*	0.210			
	(0.164)	(0.197)	(0.173)			
Aged 50–59 × Distance	-0.164	-0.157	-0.065			
	(0.191)	(0.203)	(0.210)			
Aged 30–39 × Treated				0.216***	0.632***	-0.687***
				(0.058)	(0.102)	(0.061)
				[0.016]	[0.122]	[0.102]
Aged 40–49 × Treated				0.265***	0.865***	-0.806***
				(0.074)	(0.103)	(0.035)
				[0.094]	[0.000]	[0.058]
Aged 50–59 × Treated				0.438***	0.400***	0.092
				(0.036)	(0.066)	(0.082)
				[0.090]	[0.030]	[0.046]
Female	0.613***			0.610**		
	(0.109)			(0.217)		
				[0.000]		
School Access	0.842**	0.542	1.316***	0.857***	0.568	1.291***
	(0.373)	(0.586)	(0.358)	(0.231)	(0.506)	(0.186)
				[0.116]	[0.556]	[0.072]
Father's years of schooling	0.204***	0.199***	0.215***	0.205***	0.202***	0.217***
	(0.020)	(0.026)	(0.023)	(0.015)	(0.031)	(0.027)
				[0.000]	[0.008]	[0.000]
Observations	4103	1999	2104	4103	1999	2104
Mean dependent value	9.960	10.281	9.656	9.960	10.281	9.656
R-squared	0.236	0.295	0.179	0.236	0.294	0.181

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (4), (5), and (6), the *p*-value obtained using the wild bootstrap method is reported in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7

Educational Attainment with Treatment and School Access: Ordinary Least Squares Model

	Distance			Treatment indicator		
	Entire	Female	Male	Entire	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Years of schooling						
Aged 30–39 × Distance	0.321	0.162	0.779**			
	(0.418)	(0.565)	(0.314)			
Aged 40–49 × Distance	0.634	0.651	0.821**			
	(0.397)	(0.512)	(0.322)			
Aged 50–59 × Distance	0.379	0.687	0.506			
	(0.409)	(0.547)	(0.313)			
Aged 30–39 × Distance × School Access	-0.498	-0.506	-0.647*			
	(0.443)	(0.592)	(0.347)			
Aged 40–49 × Distance × School Access	-0.797*	-1.020*	-0.626*			
	(0.421)	(0.537)	(0.364)			
Aged 50–59 × Distance × School Access	-0.565	-0.839	-0.620			
	(0.441)	(0.571)	(0.375)			
Distance × School Access	-0.167	-0.627	0.052			
	(0.354)	(0.512)	(0.397)			
Aged 30–39 × Treated				0.174	1.113	-1.210*
				(0.657)	(0.752)	(0.547)
				[0.798]	[0.182]	[0.063]
Aged 40–49 × Treated				-0.763	-0.438	-1.364*
				(0.914)	(1.223)	(0.705)
				[0.431]	[0.731]	[0.094]
Aged 50–59 × Treated				-0.410	-0.851	-0.854
				(0.742)	(0.949)	(0.617)
				[0.598]	[0.400]	[0.209]
Aged 30–39 × Treated × School Access				-0.050	-0.725	0.508
				(0.600)	(0.663)	(0.558)
				[0.935]	[0.310]	[0.393]
Aged 40–49 × Treated × School Access				1.062	1.294	0.552
				(0.844)	(1.134)	(0.701)
				[0.249]	[0.291]	[0.457]

Aged 50–59 × Treated × School Access				0.940	1.242	1.117
				(0.716)	(0.894)	(0.654)
				[0.231]	[0.207]	[0.131]
Treated × School Access				-0.078	1.118**	-0.400
				(0.222)	(0.333)	(0.266)
				[0.735]	[0.012]	[0.176]
Aged 30–39 × School Access	0.018	-0.049	0.143	-0.648	-0.220	-1.089*
	(0.552)	(0.654)	(0.580)	(0.620)	(0.690)	(0.564)
				[0.331]	[0.759]	[0.095]
Aged 40–49 × School Access	0.550	0.917	0.202	-1.158	-1.118	-1.040
	(0.554)	(0.640)	(0.666)	(0.856)	(1.161)	(0.693)
				[0.218]	[0.368]	[0.177]
Aged 50–59 × School Access	0.933*	1.038	1.119*	-0.348	-0.741	-0.317
	(0.561)	(0.656)	(0.601)	(0.712)	(0.920)	(0.627)
				[0.640]	[0.447]	[0.629]
Female	0.609***			0.604**		
	(0.109)			(0.219)		
				[0.028]		
Father’s years of schooling	0.205***	0.199***	0.216***	0.206***	0.203***	0.217***
	(0.020)	(0.026)	(0.023)	(0.016)	(0.032)	(0.026)
				[0.000]	[0.000]	[0.000]
School Access	1.405**	2.014***	1.482	1.105***	0.494	1.679***
	(0.592)	(0.758)	(0.942)	(0.189)	(0.275)	(0.196)
				[0.000]	[0.000]	[0.000]
Observations	4103	1999	2104	4103	1999	2104
Mean dependent value	9.960	10.281	9.656	9.960	10.281	9.656
R-squared	0.236	0.295	0.179	0.236	0.294	0.181

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a “before” or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (4), (5), and (6), the p-value obtained using the wild bootstrap method is reported in brackets.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8

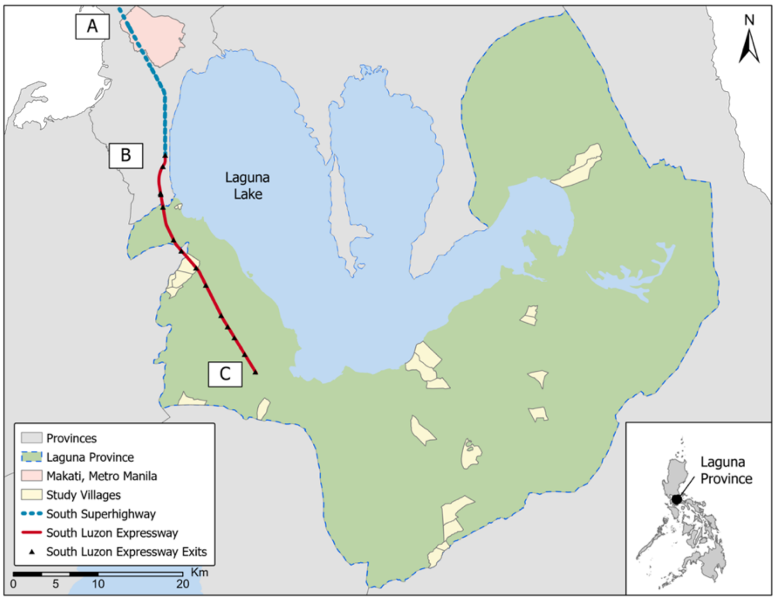
Returns to Schooling by Sector: Family Income and Expenditure Survey 2012

	Entire	Agriculture	Non-agriculture
Dependent variables: Wages			
Panel A: Female			
Years of Schooling	0.118***	0.065	0.119***
	(0.010)	(0.184)	(0.010)
Observations	1030	19	1024
R-squared	0.237	0.822	0.225
Panel B: Male			
Years of Schooling	0.100***	0.006	0.102***
	(0.005)	(0.048)	(0.006)
Observations	3227	99	3202
R-squared	0.196	0.389	0.196

Notes: *** p<0.01, ** p<0.05, * p<0.10. Robust standard errors are in parentheses. The coefficients of experience and experience squared are excluded. Municipality fixed effect is controlled for.

Figure 1

Study Villages and Constructed Highway



Notes: The locations of the South Luzon Expressway (SLEx) were plotted based on geo-referenced data provided by the South Luzon Tollway Corporation. The route from Metro Manila (A) to Alabang (B) started in 1969. The expansion from Alabang (B) to Calamba (C) began in 1978.

Figure 2

Tracked Households and Current Locations

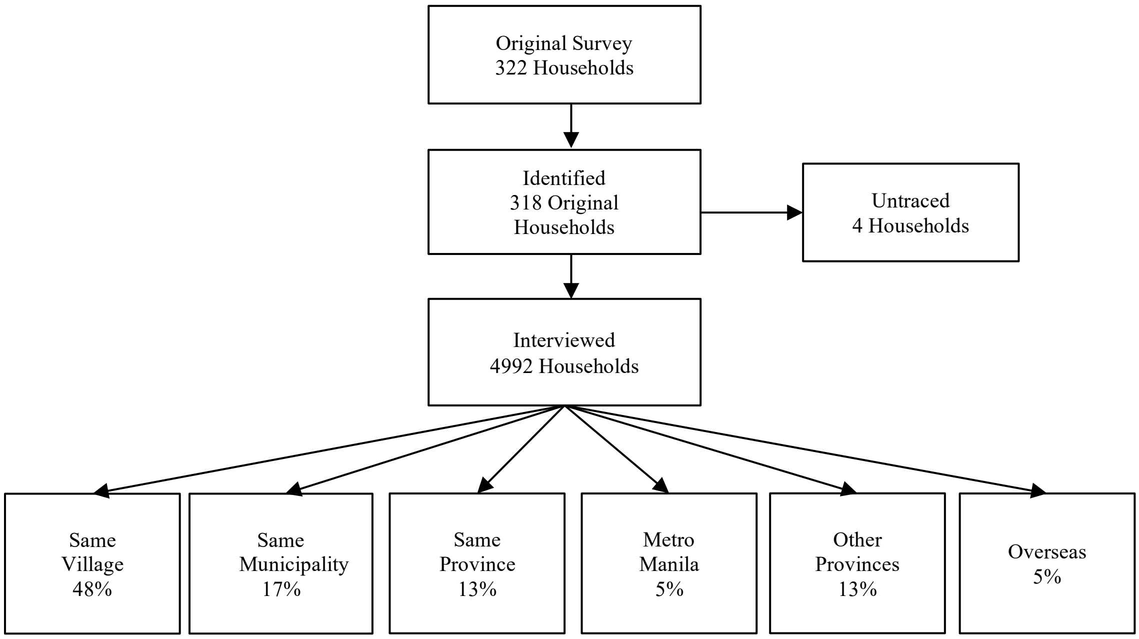


Figure 3

Distance between Sample Villages and Highway in 2017

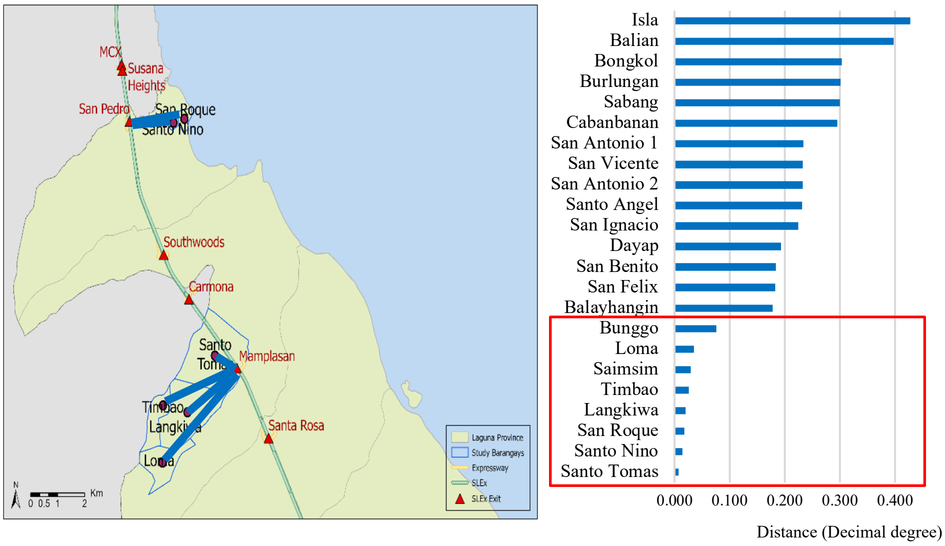
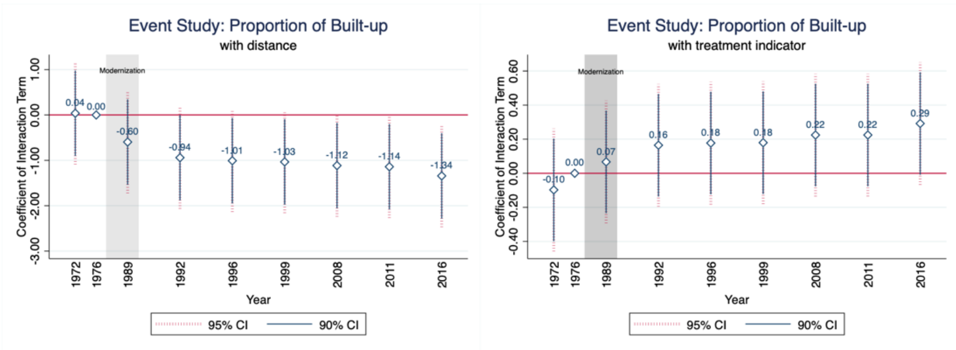


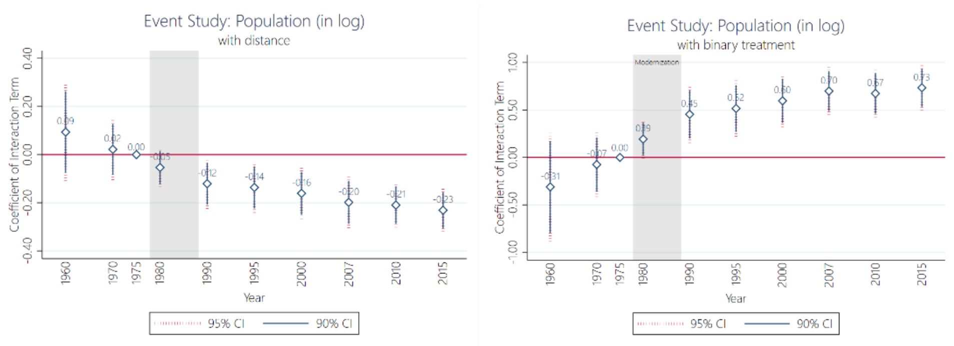
Figure 4

Checks for Parallel Trend

Panel A: Event study Plots: Changes in the Proportion of Built-up



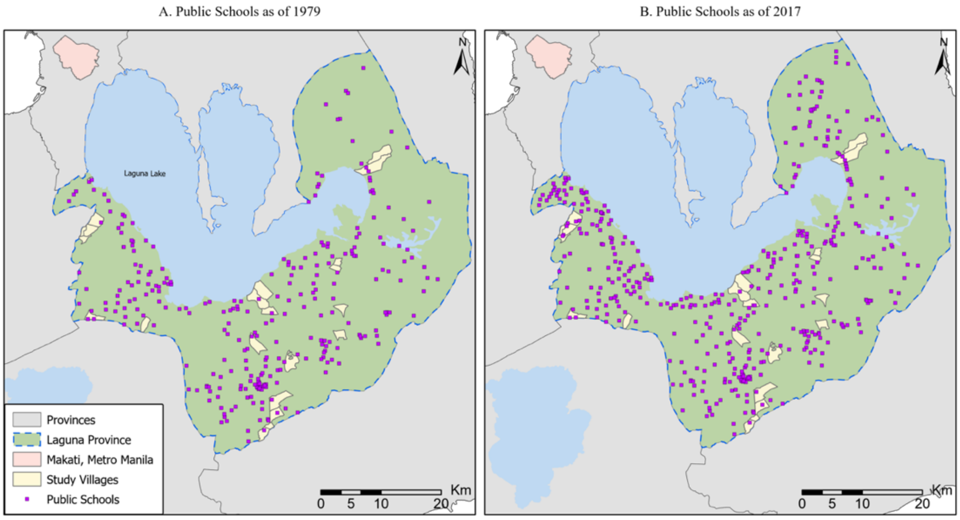
Panel B: Event study Plots: Changes in the Population



Notes: Solid navy and broken cranberry bars indicate 90% and 95% confidence intervals, respectively.

Figure 5

Location of Public Primary Schools in 1979 and 2019

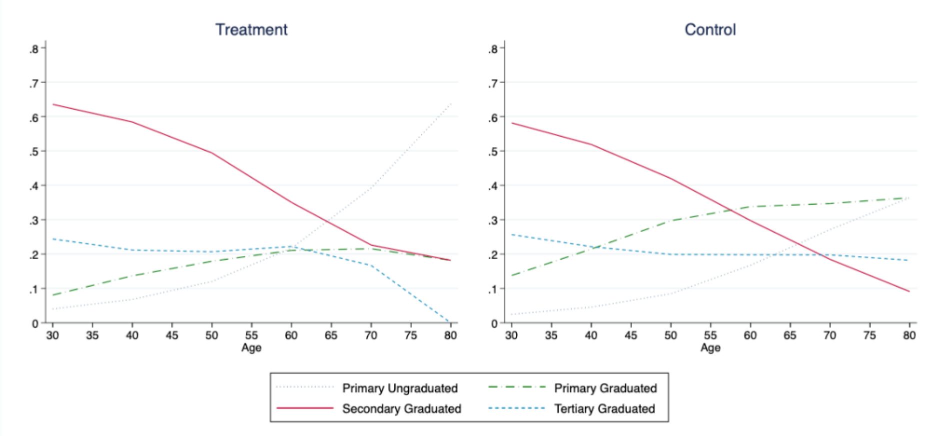


Notes: The locations of schools are plotted based on the address in the Master List of Schools for the 2017–2018 academic year obtained from the website of the Department of Education (DepEd).

Figure 6

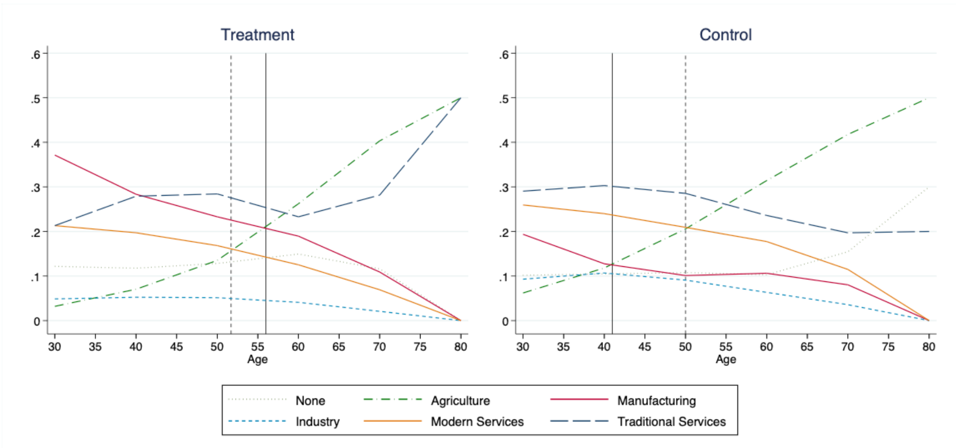
Descriptive Analysis

Panel A: Age-Specific Educational Attainment



Note: Locally Weighted Scatterplot Smoothing (LOWESS) is used with bandwidth=0.8.

Panel B: Age-Specific Employment Share



Note: Locally Weighted Scatterplot Smoothing (LOWESS) is used with bandwidth=0.8.

Appendix A: Study Area and Survey

A.1 The Study Area

Figure A1 shows the geographical position of Manila, the capital city, northwest of the lake, indicating its developmental status as of 1976 through the prevalence of built-up areas. Situated to the south of Lake Laguna is the Laguna Province, where a comparison of the southwest and southeast sides of the lake reveals a similar scarcity of built-up areas. We classified the area into four distinct categories using a machine learning algorithm: water bodies (blue), vegetation (green), bare land (orange), and built-up areas (red). There appears to be an even distribution of development on both sides, suggesting that during the 1970s, primary industry was the economic backbone of Laguna Province.

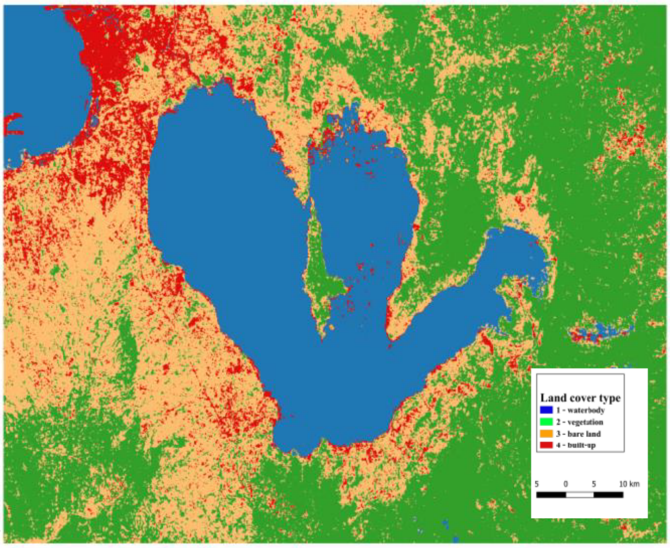


FIGURE A1. THE STUDY AREA IN 1976

Notes: The data sources are Landsat images, courtesy of the NASA Goddard Space Flight Center and the U.S. Geological Survey. We classified the area into four distinct categories using a machine learning algorithm: water bodies (blue), vegetation (green), bare land (orange), and built-up areas (red). To create these classifications, we generated land cover and land use maps from historical Landsat satellite image data. The Landsat satellite images used as inputs were preprocessed to ensure that they were free of cloud cover, achieved by composing observation data from the dry season over a three-year period, which included the years before and after the target years. We employed the random forest algorithm, a machine learning technique, to classify the pixels in the Landsat images into four categories. The classification model was trained using a dataset acquired through visual interpretation of false-color composite Landsat images.

The expressway, which opened in 1978, links the western part of the province to Manila. This enhanced connectivity drew investments, leading to the establishment of several industrial parks in the 1980s. Consequently, owing to expressways and industrial parks, the western side underwent swift modernization.

Figure A2, which uses land use information derived from Landsat satellite images, depicts the evolution of land use from 1972 to 2016. [1] We can discern a transition in land cover from agricultural use, represented by green vegetation and orange bare land, to built-up areas, denoted in red, on the lake's southwest side, where the expressway and industrial parks are located. Conversely, minimal changes in the built-up areas are evident on the southeast side.

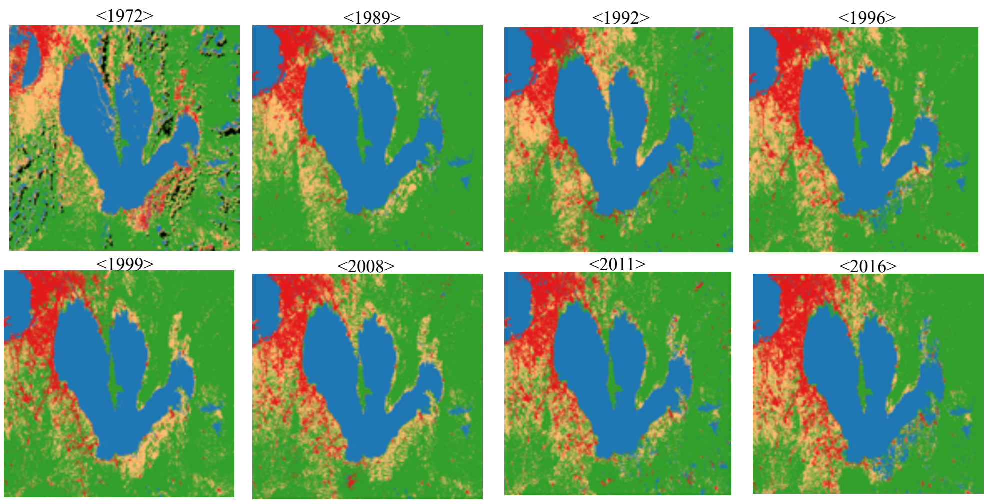


FIGURE A2. LAND USE CHANGES OBSERVED BY SATELLITE IMAGES

Notes: The same as Figure A2.

[1] We construct a land-use data set using satellite images, each with a pixel size of 269 m × 269 m. Because we combine all of the images, including those from both dry and wet seasons, we concurrently refine and implement quality control measures to address the low resolution of the older data set from the 1970s, seasonal variations, and a mechanical failure that occurred in the 2000s.

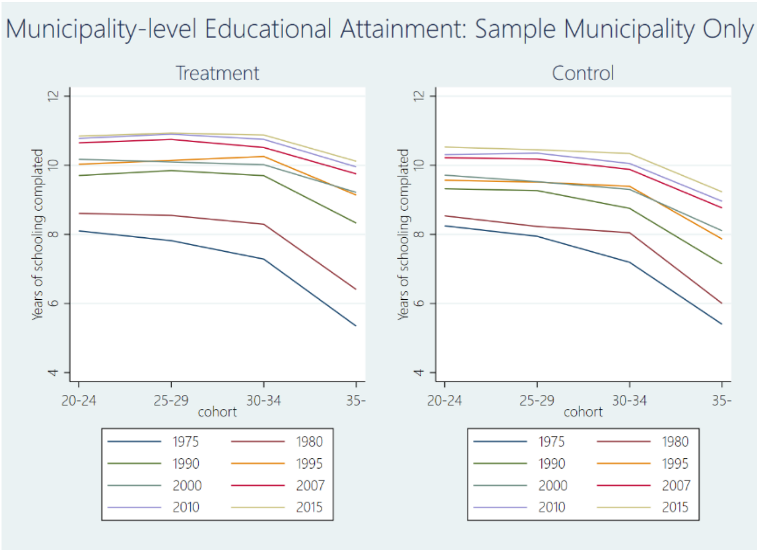
A.2 Additional Checks for Pre-Trends

As an additional check for parallel trends before SLEx construction, we used municipality-cohort level data on education. We digitized information from the population and housing Census for Laguna Province from 1970 to 2015. The municipality-cohort level data illustrate how human capital accumulation has progressed in the treated and control municipalities, although the data are based on the existing population at the time of each Census and do not consider migration. Our sample consisted of 23 villages located in ten municipalities, including three treated municipalities closer to the SLEx and seven farther away.

Figure A3 shows the mean years of education. The data is available for 1975, 1980, 1990, 1995, 2000, 2007, 2010, and 2015. As the defined cohort differs by the Census round, we plotted the four cohorts for which the data were available throughout the years between 1975 and 2015; that is, 20−24, 25−29, 30−34, and 35 and over. Panel A plots the data for the ten municipalities included in the sample. Panel B plots the data for all 30 municipalities in Laguna Province, for which treatment is defined by the distance to the SLEX. The 15 municipalities within the median distance to the SLEX are defined as the treated municipalities and those farther as the control ones.

In both panels, the data for 1975, conducted before the SLEX construction, present similar patterns in the treatment and control municipalities. Educational attainment increased as the Census rounds proceeded. In recent years, the treated municipalities have completed slightly more years of schooling, which is consistent with our main conclusions. Indeed, the *t*-test using municipality-cohort level data does not reject the null hypothesis of the presence of a difference between the treated and control municipalities in 1975 and 1980 but rejects it in 1990 and all subsequent years. The same pattern was consistently observed when we used ten sample municipalities (3 treated and 3 control) and 30 municipalities (15 treated and 15 control).

Panel A: Sample Municipality Only



Panel B: All Municipalities

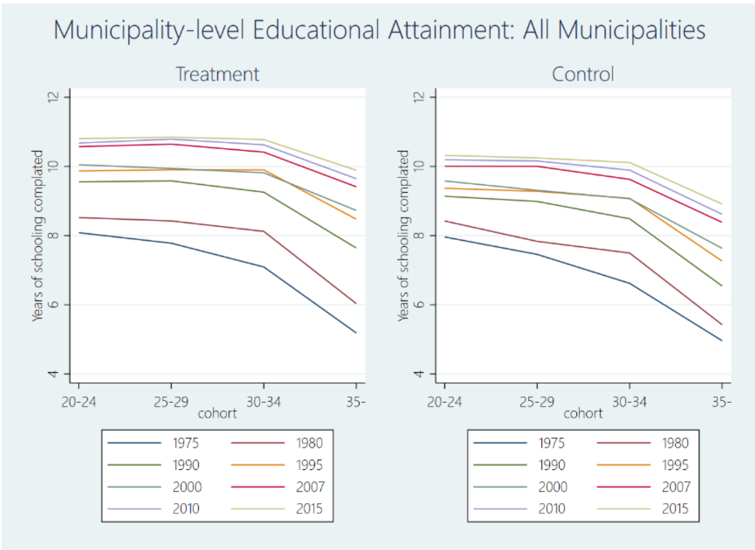


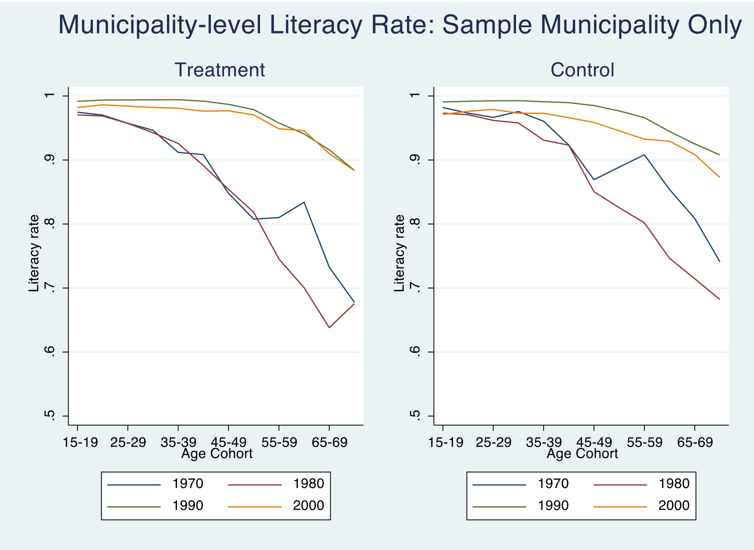
FIGURE A3. COHORT-LEVEL EDUCATIONAL ATTAINMENT BY TREATMENT MUNICIPALITY

Notes: The source is the Census of population and housing for various years.

Figure A4 plots the literacy rate for the 5-year cohort for the treated and control municipalities. The data is available for 1970, 1980, 1990, and 2000, and we plot data for the consistently available twelve 5-year cohorts, including

15−19, 20−24, etc., up to 70−74. There are fluctuations since the literacy rate is not based on the total enumeration but is computed based on 5%−20% of the population. Importantly, the data for 1970, which were acquired before the SLEx construction, present a similar pattern. When we use ten sample municipalities, the *t*-test does not reject the null hypothesis of the presence of a difference between the treated and control municipalities in 1970, 1980, and 1990 but rejects it in 2000. When we use all 30 municipalities, the *t*-test also rejects the null hypothesis in 1970, but the mean value is higher among the control municipalities (0.89 for the treated 15 municipalities and 0.86 for the control 15 municipalities).

Panel A: Sample Municipality Only



Panel B: All Municipalities

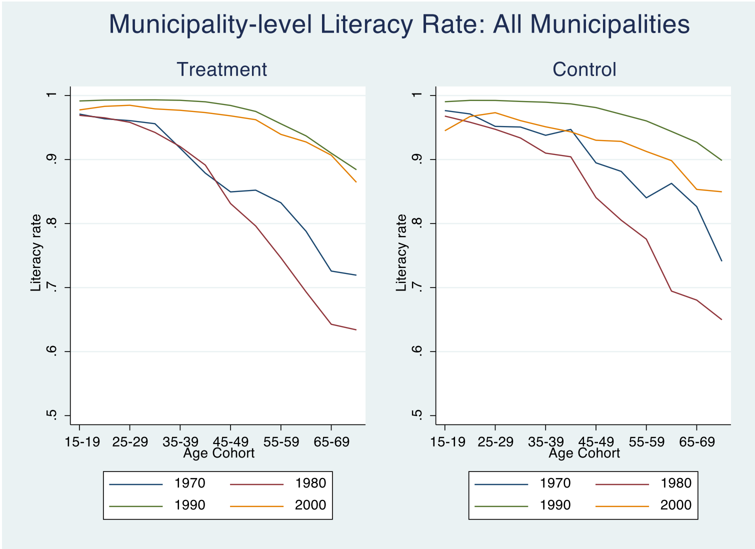


FIGURE A4. COHORT-LEVEL LITERACY RATE BY TREATMENT MUNICIPALITY

Notes: The source is the Census of population and housing for various years.

We used municipality-cohort level Census data to estimate a DID model to analyze the impact of SLEx construction on educational attainment. Our unit of observation was the municipal cohort, and we estimated the model using 10 sample municipalities or all 30 municipalities in Laguna Province. Since our definition of treatment was those aged 20 years or younger in 1978 in the main analysis, we defined the cohort 20−24 and younger in 1980 as the after-treatment cohort. In the 1990 Census, for instance, cohorts 30−34 and younger were considered the post-treatment cohort. We controlled for municipality, cohort, and Census year fixed effects. The results are shown in Table A1. The coefficients of the treated municipality are all positive and significant, except in Column 3. These results support our main conclusion that the SLEx construction facilitates investments in human capital.

Table A1

Educational Attainment Using the Aggregated Census Data

	Years of Schooling Completed		Literacy Rate	
	Sample only	All	Sample only	All
	(1)	(2)	(3)	(4)
Treated municipality * After-treatment cohort	0.27**	0.30***	0.011	0.038***
	(0.13)	(0.065)	(0.016)	(0.011)
Census year = 1970	(No data)		(Reference)	
Census year = 1975	(Reference)		(No data)	
Census year = 1980	0.58***	0.60***	-0.028**	-0.064***
	(0.11)	(0.070)	(0.013)	(0.0094)
Census year = 1990	1.66***	1.62***	0.075***	0.072***
	(0.11)	(0.066)	(0.012)	(0.0085)
Census year = 1995	2.12***	2.09***	(No data)	
	(0.10)	(0.066)		
Census year = 2000	2.21***	2.22***	0.054***	0.038***
	(0.10)	(0.065)	(0.012)	(0.0087)
Census year = 2007	2.81***	2.85***	(No data)	
	(0.10)	(0.065)		
Census year = 2010	2.99***	3.05***	(No data)	
	(0.10)	(0.066)		
Census year = 2015	3.16***	3.23***	(No data)	
	(0.10)	(0.067)		
Observations	320	960	480	1438
R-squared	0.946	0.940	0.597	0.608

Notes: The after-treatment cohort is 20–24 and younger as of 1980. The coefficient for the municipality and cohort fixed effects are not reported for brevity. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

A.3 The Tracking Survey

The Laguna Multipurpose Household Survey (LMHS) was originally conceived and implemented by Professors Robert Evenson and Barry M. Popkin in 1975. This survey encompassed 34 villages and 576 households within the Laguna Province of the Philippines, an area spanning 1,795 km^2 and housing a population of 803,750 (Evenson et al., 1980). The initial 34 villages were selected using stratified random sampling. Thirteen of these villages, which are representative of lowland rice-farming communities, were selected from a previous survey conducted by the University of the Philippines Los Baños, known as the Farm and Home Development Office Survey. The remaining villages were randomly selected from a comprehensive list of all villages within each of the three other categories: six from uplands, three from fishing villages, and 12 from semi-urban villages. These 34 sample villages were chosen to accurately represent the socioeconomic conditions across Laguna Province. Within each of these 34 villages, 16 households were randomly selected from the village Census (an additional 9 households were selected from each of the three fishing villages). This sampling framework resulted in 576 households (with one missing household) surveyed in the 1975 wave of the LMHS.

Subsequent surveys were administered consistently in 1977, 1979, 1982, 1985, 1990, 1992, and 1998 (Ejrnæs and Pörtner, 2004; Rosenzweig and Wolpin, 1986). The original data files and respondent lists from the initial 1975 survey were not accessible. Therefore, we used data from the second wave of the 1977 LMHS as the baseline dataset. These data were gathered immediately before SLEX completion. The second wave targeted a subsample of respondents in the first wave and consisted of 322 households, which we assumed represented the socioeconomic conditions of households in Laguna Province in 1977.

We revisited the village in which the households surveyed in 1977 were located. We consulted with village leaders, senior citizens, and other knowledgeable individuals, whom we referred to as "informants," to determine the current location of the original household. In instances where the original household head and spouse were deceased, which was often the case, we gathered information about the residences of the children from the original household. If the original household was still in the same village and some of the original members were still residing there, the tracking process was relatively straightforward. We endeavored to identify at least one, but preferably more, member or descendant of the original household from the information provided by the informants. If the informants were unable to identify any members of the original household, we approached neighbors, 1977 landowners of the housing lot, current occupants of the lot, relatives, friends, schoolmates, or coworkers of the original household members to identify the current residences of the head, spouse, and other existing members of the original household.

After identifying the original household members and their descendants, we conducted visits to a selection of these individuals, referred to herein as "respondents," to gather data on the family tree of the original household via our tracking module. To amass as much information as possible from all descendants, we employed proxy reports. Subsequently, we requested that these respondents introduce us to other potential respondents who could provide any missing information. This data collection process was repeated until comprehensive information on the original household members and their descendants was gathered. When we had exhausted the pool of available respondents within the original village or its immediate surroundings, we made phone calls to potential respondents residing in distant areas to collect the necessary information.

Appendix B: Additional Tables and Figures

This section presents additional tables to provide supplementary information on the main analysis and graphically illustrates the main estimation results.

Table B1

List of Selected Industrial Parks

Code	Name	Year of Establishment	Land area (ha)	Number of companies	Filipino (%)
1	AG&P Special Economic Zone	N/A	40.3	N/A	96
2	Cocochem Agro-Industrial Park	2013	42.0	3	100
3	First Philippine Industrial Park	1997	331.9	13	70
4	First Philippine Industrial Park II	2013	91.8	N/A	70
5	Keppel Philippines Marine SEZ	2007	22.9	N/A	100
6	Laguna Technopark SEZ	1989	314.9	241	61
7	Laguna Technopark Annex	1989	29	N/A	61
8	Light Industry & Science Park III	N/A	110.5	N/A	100
9	Lima Technology Center	1997	280.2	11	60
10	Philtown Technology Park	2006	66.6	N/A	100
11	Tabangao Special Economic Zone	N/A	86.0	1	100
12	Cavite Economic Zone	1980	278.5	382	100
13	Cavite Economic Zone II	N/A	53.7	N/A	60
14	Daiichi Industrial Park	1996	55.0	4	100
15	EMI Special Economic Zone	N/A	12.2	1	60
16	First Cavite Industrial Estate	1991	71.8	63	60
17	Gateway Business Park	1989	110.1	19	80
18	Golden Mile Business Park	2002	45.1	38	64
19	People's Technology Complex	2000	59.0	14	100
20	Suntrust Ecotown Tanza	2014	116.2	N/A	100
21	Calamba Premiere International Park	1999	65.6	18	60
22	Carmelray Industrial Park I	1992	80.0	22	100

Table B2

Sector Classification of Occupation

Code	Industry of Primary Occupation (Lifetime)	Sector
A	Agriculture, forestry, and fishing	Agriculture
C	Manufacturing	Manufacturing
B	Mining and quarrying	Industry
D	Electricity, gas, steam, and air-conditioning supply	
E	Water supply, sewerage, waste management, and remediation activities	
F	Construction	
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	Traditional Services
H	Transportation and storage	
I	Accommodation and food service activities	
J	Information and communication	Modern Services
K	Financial and insurance activities	
M	Professional, scientific, and technical services	
N	Administrative and support service activities	
P	Education	
Q	Human health and social work activities	
L	Real estate activities	
O	Public administrative and defense; compulsory social security	
R	Arts, entertainment, and recreation	
S	Other service activities	
U	Activities of extraterritorial organizations and bodies	
T	Activities of private households as employers and undifferentiated goods producing activities of households for own use	Others
V	Full-time student	
W	Housewife/husband	
X	No job	

Note: The industry code in the survey questionnaire follows the classification of industries in the 2009 Philippine Standard Industrial Classification.

Table B3

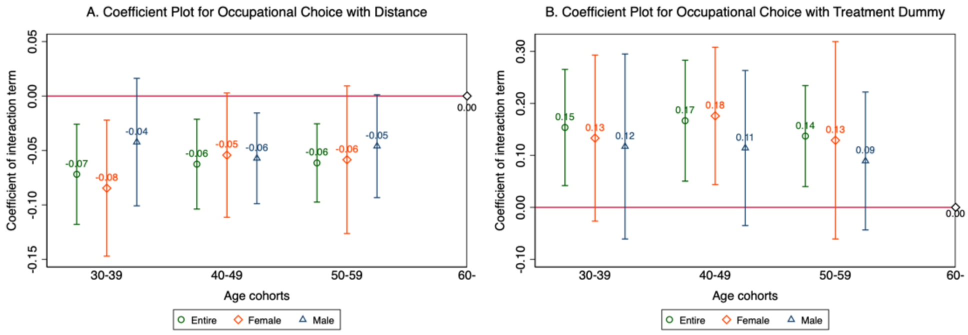
Descriptive Statistics among the Non-migrant Sample

	Mean	Min	Max	St.dv.	N
Age in 2017	45.41	30	79	11.33	2681
Female	0.47	0	1	0.50	2681
Younger cohort (30–49)	0.66	0	1	0.48	2681
Working in modern sector	0.40	0	1	0.49	2681
Working in manufacturing	0.20	0	1	0.40	2681
Working in modern services	0.20	0	1	0.40	2681
Father in modern sector	0.15	0	1	0.36	2681
Years of schooling	9.68	0	16	2.85	2678
Father's years of schooling	5.58	0	14	3.25	2668

Notes: The modern sector consists of manufacturing and modern services. Years of schooling is calculated by the highest completed grade. The sample consists of people who still reside in the original municipality.

Figure B1

Coefficient Plots of Main Results

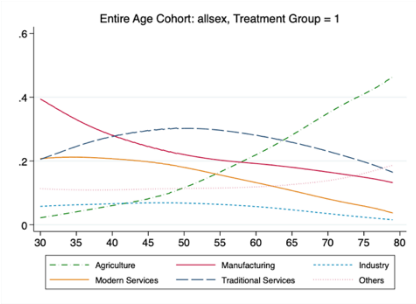


Note: These figures are based on the estimated coefficients reported in Tables 2 and 3.

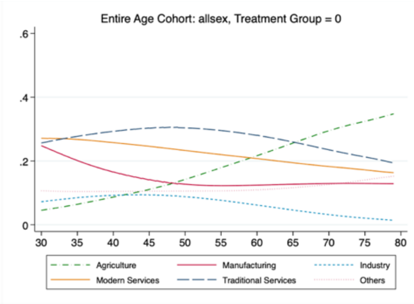
FigureB2

Predicted Average Probability of Occupational Choice by Occupation and Age

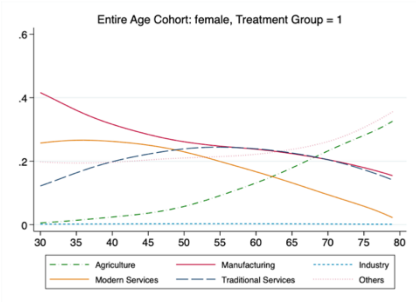
(a) All treatment zones.



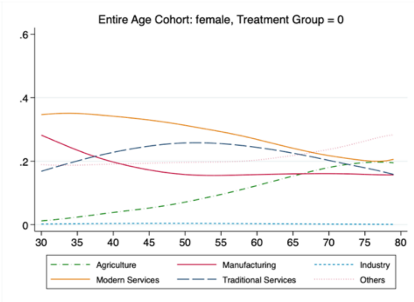
(b) All control zones.



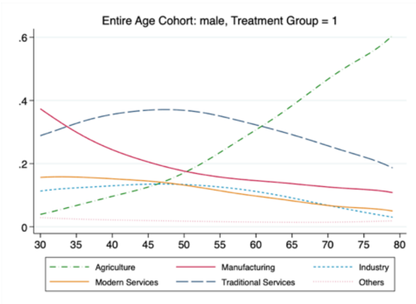
(c) Female, Treatment Zone



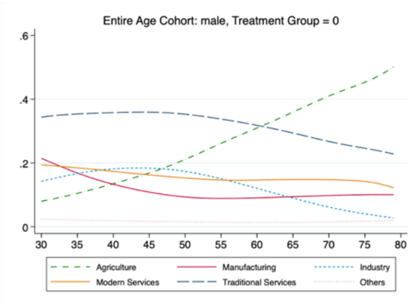
(d) Female, Control Zone



(e) Male, Treatment Zone



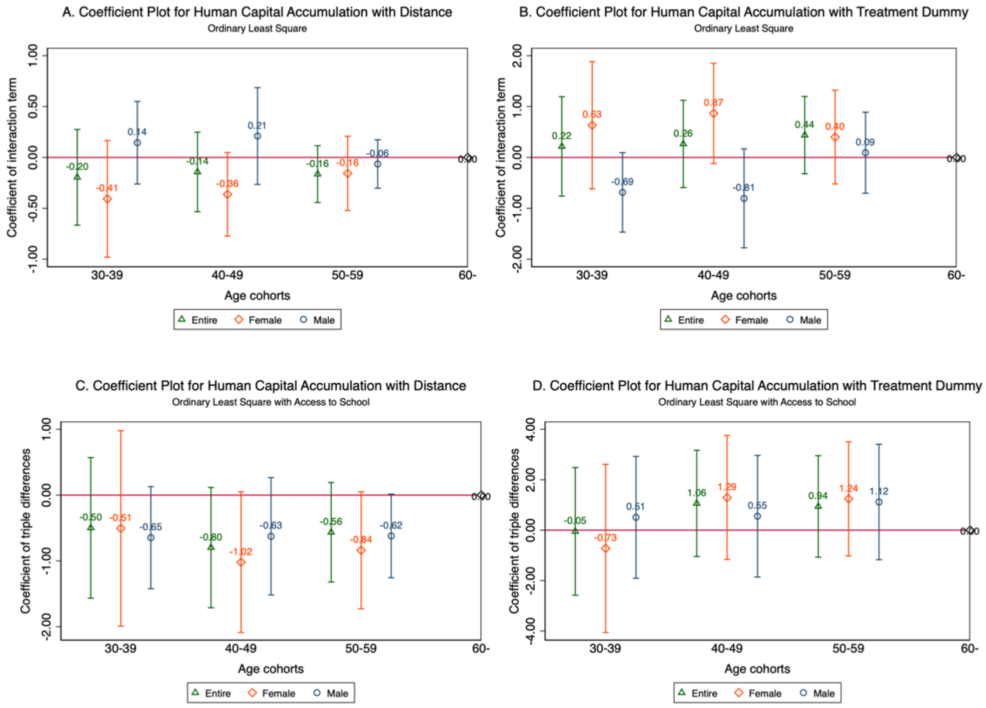
(f) Male, Control Zone



Note: These figures are based on the estimated results of the multinomial logit model, as shown in Table 5.

Figure B3

Coefficient Plots of Treatment Zone Effects



Note: These figures are based on the estimated coefficients reported in Tables 6 and 7.

Appendix C: Subsample Analysis with Non-migrants

As a robustness check, this section presents the results estimated using only a subsample of households that remained in the same village or another village in the same municipality as the ancestor's original household. The subsample constituted 65% of the sample (see Figure 2). Table C1 corresponds to Table 2 (lifetime occupation), Table C2 to Table 3 (lifetime occupation by gender), Tables C3 to Table 4 (lifetime occupation using the multinomial logit model), Tables C4 and C5 to Table 6 (education).

Table C1

Linear Probability Model: Lifetime Primary Occupation (Non-migrants)	Female		Male	
	Distance	Treatment indicator	Distance	Treatment indicator
	(1)	(2)	(3)	(4)
Dependent variable: 1 [Modern Sector]				
Aged 30–39 × Distance	-0.098***	-0.101***		
	(0.024)	(0.025)		
Aged 40–49 × Distance	-0.063***	-0.068***		
	(0.019)	(0.019)		
Aged 50–59 × Distance	-0.056**	-0.057**		
	(0.027)	(0.027)		
Aged 30–39 × Treated			0.214***	0.218***
			(0.010)	(0.011)
			[0.068]	[0.056]
Aged 40–49 × Treated			0.142***	0.152***
			(0.011)	(0.013)
			[0.052]	[0.062]
Aged 50–59 × Treated			0.117***	0.123***
			(0.012)	(0.013)
			[0.048]	[0.038]
	0.212***	0.211***	0.212***	0.212***
	(0.022)	(0.022)	(0.025)	(0.025)
			[0.000]	[0.000]
Father in modern sector		0.074**		0.070**
		(0.031)		(0.025)
				[0.034]
Observations	2681	2681	2681	2681
Mean Dependent Value	0.404	0.404	0.404	0.404
R-squared	0.121	0.123	0.120	0.122

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (3) and (4), the p-value obtained using the wild bootstrap method is reported in brackets. The sample consists of people who still reside in the original municipality.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C2

Linear Probability Model: Lifetime Primary Occupation by Gender (Non-migrants)

	Female		Male	
	Distance	Treatment indicator	Distance	Treatment indicator
	(1)	(2)	(3)	(4)
Dependent variable: 1 [Modern Sector]				
Aged 30–39 × Distance	-0.085***		-0.047	
	(0.031)		(0.031)	
Aged 40–49 × Distance	-0.056*		-0.053*	
	(0.032)		(0.029)	
Aged 50–59 × Distance	-0.081**		-0.017	
	(0.038)		(0.033)	
Aged 30–39 × Treated		0.124***		0.143***
		(0.011)		(0.011)
		[0.000]		[0.016]
Aged 40–49 × Treated		0.159***		0.121***
		(0.018)		(0.006)
		[0.088]		[0.010]
Aged 50–59 × Treated		0.150***		0.041**
		(0.016)		(0.015)
		[0.024]		[0.000]
Father in modern sector	-0.008	-0.005	0.083**	0.082***
	(0.032)	(0.014)	(0.037)	(0.013)
		[0.114]		[0.000]
Observations	1409	1409	1535	1535
Mean Dependent Value	0.540	0.540	0.319	0.319
R-squared	0.073	0.071	0.086	0.086

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (2) and (4), the p-value obtained using the wild bootstrap method is reported in brackets. The sample consists of people who still reside in the original municipality.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C3

Multinomial Logit Model: Lifetime Primary Occupation (Non-migrants)

	Distance					Treatment indicator				
	Manufacturing	Industry	Modern Services	Traditional Services	Others	Manufacturing	Industry	Modern Services	Traditional Services	Others
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Aged 30–39 × Distance	-0.674***	0.135	-0.774***	-0.273	-0.364					
	(0.261)	(0.295)	(0.223)	(0.269)	(0.295)					
Aged 40–49 × Distance	-0.590**	0.185	-0.577***	-0.248	-0.332					
	(0.255)	(0.253)	(0.201)	(0.238)	(0.333)					
Aged 50–59 × Distance	-0.411	0.097	-0.458*	-0.140	-0.117					
	(0.268)	(0.274)	(0.263)	(0.257)	(0.276)					
Aged 30–39 × Treated						2.357***	-0.196	2.010***	0.919	1.255**
						(0.470)	(0.734)	(0.624)	(0.735)	(0.632)
Aged 40–49 × Treated						2.277***	-0.388	1.797***	1.005	1.298
						(0.479)	(0.568)	(0.639)	(0.688)	(0.831)
Aged 50–59 × Treated						1.735***	0.106	1.242	0.536	0.325
						(0.610)	(0.649)	(0.799)	(0.771)	(0.613)
Female	1.768***	-2.893***	1.886***	0.736***	3.645***	1.805***	-2.900***	1.905***	0.743***	3.661***
	(0.201)	(0.555)	(0.217)	(0.202)	(0.315)	(0.218)	(0.554)	(0.229)	(0.207)	(0.321)
Father in modern sector	1.205***	0.656**	1.268***	0.923***	0.712**	-0.098	0.056	-0.005	0.062	-0.079
	(0.293)	(0.314)	(0.316)	(0.279)	(0.302)	(0.079)	(0.109)	(0.078)	(0.078)	(0.119)
Observations	2681	2681	2681	2681	2681	2681	2681	2681	2681	2681
Mean Dependent Value	0.129	0.203	0.081	0.201	0.284	0.129	0.203	0.081	0.201	0.284

Notes: Estimated coefficients are reported. The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. The mean value of agriculture is 0.103. The ample consists of people who still reside in the original municipality.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table C4

Ordinary Least Square: Educational Attainment with Treatment (Non-migrants)

	Distance			Treatment indicator		
	Entire	Female	Male	Entire	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Years of schooling						
Aged 30–39 × Distance	-0.272	-0.487**	0.080			
	(0.190)	(0.210)	(0.215)			
Aged 40–49 × Distance	-0.138	-0.362*	0.244			
	(0.163)	(0.186)	(0.204)			
Aged 50–59 × Distance	-0.161	-0.249	0.039			
	(0.191)	(0.195)	(0.229)			
Aged 30–39 × Treated				0.482***	0.789***	-0.300***
				(0.071)	(0.149)	(0.080)
				[0.012]	[0.138]	[0.100]
Aged 40–49 × Treated				0.365***	0.816***	-0.586***
				(0.077)	(0.178)	(0.032)
				[0.054]	[0.028]	[0.058]
Aged 50–59 × Treated				0.420***	0.448**	0.041
				(0.059)	(0.142)	(0.106)
				[0.050]	[0.034]	[0.084]
Female	0.647***			0.646**		
	(0.126)			(0.198)		
				[0.000]		
School Access	0.554*	0.396	0.785	0.568**	0.450	0.767**
	(0.280)	(0.446)	(0.484)	(0.206)	(0.564)	(0.322)
				[0.138]	[0.350]	[0.066]
Father's years of schooling	0.187***	0.159***	0.217***	0.188***	0.160***	0.217***
	(0.022)	(0.035)	(0.026)	(0.016)	(0.039)	(0.029)
				[0.000]	[0.102]	[0.000]
Observations	2684	1250	1434	2684	1250	1434
Mean dependent value	9.668	10.014	9.366	9.668	10.014	9.366
R-squared	0.241	0.310	0.185	0.240	0.308	0.185

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (4), (5) and (6), the p-value obtained using the wild bootstrap method is reported in brackets. The sample consists of people who still reside in the original municipality.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

Table C5

Ordinary Least Square: Educational Attainment with Treatment and School Access (Non-migrants)

		Distance			Treatment indicator	
	Entire	Female	Male	Entire	Female	Male
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Years of schooling						
Aged 30–39 × Distance × School Access	-0.180	-0.675	0.100			
	(0.434)	(0.587)	(0.449)			
Aged 40–49 × Distance × School Access	-0.906**	-1.122**	-0.901*			
	(0.413)	(0.531)	(0.460)			
Aged 50–59 × Distance × School Access	-0.426	-0.646	-0.677			
	(0.436)	(0.585)	(0.450)			
Aged 30–39 × Treated × School Access				-0.242	-0.144	-0.347
				(0.423)	(0.592)	(0.260)
				[0.585]	[0.814]	[0.224]
Aged 40–49 × Treated × School Access				1.774**	1.617	1.951***
				(0.625)	(0.990)	(0.410)
				[0.025]	[0.146]	[0.002]
Aged 50–59 × Treated × School Access				0.963*	0.876	1.766***
				(0.498)	(0.736)	(0.348)
				[0.094]	[0.273]	[0.001]
Female	0.643***			0.642**		
	(0.127)			(0.199)		
				[0.015]		
Father's years of schooling	0.188***	0.160***	0.216***	0.190***	0.163***	0.218***
	(0.022)	(0.035)	(0.026)	(0.016)	(0.040)	(0.028)
				[0.000]	[0.005]	[0.000]
Observations	2684	1250	1434	2684	1250	1434
Mean dependent value	9.668	10.014	9.366	9.668	10.014	9.366
R-squared	0.242	0.308	0.186	0.242	0.306	0.187

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance and the zone-cohort level for the treatment indicator. In Columns (4), (5) and (6), the p-value obtained using the wild bootstrap method is reported in brackets. The sample consists of people who still reside in the original municipality.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix D: Placebo Analysis with Control Villages

As a falsification test, this section presents the results estimated using only a subsample of the migrated households in the control villages. The rationale is that households in the control villages are unlikely to be affected by the treatment. Table D1 corresponds to Table 2 (Lifetime occupation).

Table D1

Linear Probability Model for Falsification Test: Lifetime Primary Occupation

(Control villages)

	(1)	(2)	(3)	(4)
	Dependent variable: 1 [Modern Sector]			
Aged 30–39 × Distance	-0.182	-0.179	-0.133	-0.133
	(0.114)	(0.114)	(0.089)	(0.086)
Aged 40–49 × Distance	-0.090	-0.089	0.003	0.004
	(0.112)	(0.111)	(0.086)	(0.083)
Aged 50–59 × Distance	-0.176	-0.166	-0.120	-0.117
	(0.122)	(0.123)	(0.095)	(0.093)
Female	0.254***	0.251***	0.227***	0.225***
	(0.053)	(0.054)	(0.042)	(0.041)
Father in modern sector		-0.076		-0.041
		(0.069)		(0.075)
Location of current residence:				
Metro Manila			X	X
Other provinces	X	X	X	X
Overseas	X	X	X	X
Observations	239	239	394	394
Mean dependent value	0.536	0.536	0.500	0.500
R-squared	0.116	0.116	0.105	0.104

Notes: The cohort aged 60 or above at the time of our survey in 2017 is taken as a "before" or comparison group. The coefficient for age, age squared, and village and cohort fixed effects are not reported for brevity. Standard errors in parentheses are clustered at the village-cohort level for distance.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Appendix E: Spatial Spillovers

To account for the possibility of spatial spillovers, this section presents the results after controlling for the number of treated villages in the vicinity of each sample village. Specifically, we follow Miguel and Kremer (2004) and control for ring terms. Table E1 presents the number of treated villages within the buffer of 0−5 km, 5−10 km, and 10−30 km, and Table E2 uses these variables as additional controls.

Table E1

Number of Treated Villages within the Village Buffer

Village code	Village name	0−5 km	5−10 km	10−30 km
1	Sto. Nino	2	4	2
2	San Roque	2	4	2
3	Langkiwa	4	2	2
4	Sto. Tomas	4	2	2
5	Timbao	4	2	2
6	Loma	4	2	2
7	Bunggo	1	1	6
8	Saimsim	1	1	6
9	San Benito	0	0	6
10	Balayhangin	0	0	2
11	San Felix	0	0	6
12	Dayap	0	0	2
13	Sto. Angel	0	0	0
14	San Ignacio	0	0	1
15	San Vicente	0	0	1
16	San Antonio 1	0	0	1
17	San Antonio 2	0	0	1
18	Bongkol	0	0	1
19	Burlungan	0	0	0
20	Cabanbanan	0	0	0
21	Sabang	0	0	0
22	Balian	0	0	0
23	Isla	0	0	0

Table E2

Lifetime Primary Occupation: Linear Probability Model

	(1)	(2)	(3)	(4)
Aged 30–39 × Treated	0.246***	0.086	0.450***	0.273*
	(0.045)	(0.081)	(0.122)	(0.147)
Aged 40–49 × Treated	0.236***	0.143	0.152	-0.010
	(0.071)	(0.107)	(0.133)	(0.164)
Aged 50–59 × Treated	0.087	0.071	0.100	0.102
	(0.068)	(0.076)	(0.140)	(0.153)
Aged 30–39 × Number of treated villages within 5 km	0.050	0.075**	-0.013	0.020
	(0.032)	(0.034)	(0.046)	(0.048)
Aged 40–49 × Number of treated villages within 5km	0.049	0.063*	0.072	0.101**
	(0.033)	(0.036)	(0.047)	(0.050)
Aged 50–59 × Number of treated villages within 5km	0.045	0.047	0.048	0.047
	(0.035)	(0.036)	(0.051)	(0.053)
Aged 30–39 × Number of treated villages within 5–10 km	-0.121***	-0.096***	-0.134***	-0.106**
	(0.029)	(0.032)	(0.043)	(0.047)
Aged 40–49 × Number of treated villages within 5–10 km	-0.106***	-0.092**	-0.095*	-0.069
	(0.039)	(0.042)	(0.052)	(0.057)
Aged 50–59 × Number of treated villages within 5–10 km	-0.034	-0.032	-0.050	-0.052
	(0.035)	(0.037)	(0.048)	(0.051)
Aged 30–39 × Number of treated villages within 10–30 km		0.024**	0.049	0.045
		(0.012)	(0.040)	(0.039)
Aged 40–49 × Number of treated villages within 10–30 km		0.014	-0.012	-0.016
		(0.014)	(0.042)	(0.040)
Aged 50–59 × Number of treated villages within 10–30 km		0.002	-0.019	-0.020
		(0.010)	(0.044)	(0.042)
Aged 30–39 × Number of all villages within 5 km			0.014	0.005
			(0.032)	(0.031)
Aged 40–49 × Number of all villages within 5km			-0.011	-0.018
			(0.035)	(0.035)
Aged 50–59 × Number of all villages within 5km			0.016	0.018
			(0.033)	(0.032)
Aged 30–39 × Number of all villages within 5–10 km				0.020*
				(0.011)
Aged 40–49 × Number of all villages within 5–10 km				0.018
				(0.013)
Aged 50–59 × Number of all villages within 5–10 km				-0.000
				(0.009)
Female	0.196***	0.196***	0.194***	0.194***
	(0.022)	(0.022)	(0.022)	(0.022)
Father in modern sector	0.035*	0.038*	0.039*	0.041*
	(0.020)	(0.020)	(0.021)	(0.021)
Observations	3951	3951	3951	3951
Mean dependent value	0.424	0.424	0.424	0.424
R-squared	0.102	0.103	0.104	0.104