Educational Investment in Spatial Equilibrium: Evidence from Indonesia

Allan Hsiao

Stanford University

October 3, 2024

(most recent version here)

This paper argues that mobility across space greatly amplifies the returns to education. Students in isolated rural labor markets have limited job opportunities and thus limited wage gains from additional years of schooling. Access to urban labor markets provides more job opportunities and thus larger gains. I capture this mechanism with a spatial equilibrium model in which students invest in education, then migrate for employment after graduation. I estimate the model with quasi-experimental variation from Indonesia's *Sekolah Dasar Inpres* program, one of the largest school construction programs in history. I find that mobility accounts for roughly half of the aggregate gains from the program. Improving mobility would increase these gains, but also widen regional inequality.

Email: ajhsiao@stanford.edu. I am grateful for helpful feedback from Francesco Agostinelli, Milena Almagro, Clare Balboni, Sam Bazzi, Gharad Bryan, Eric Chyn, Michael Dinerstein, Dave Donaldson, Esther Duflo, Clément Imbert, Gaurav Khanna, Tatjana Kleineberg, Margaux Luflade, Mónica Martinez-Bravo, Mushfiq Mobarak, Melanie Morten, Dávid Nagy, Ben Olken, Michael Peters, Tavneet Suri, and Román David Zárate.

1 Introduction

Governments spend more than $3 trillion each year on schools (World Bank 2022). Schools are immobile and serve students locally. But students are mobile and bring their education wherever they choose to work. Human capital is portable. I argue that mobility across space greatly amplifies the returns to education, which I define as the wage gains from additional schooling. Consider, for example, a rural Indonesian student who has invested heavily in computer programming skills. If this educated student can migrate freely to Jakarta, then they have access to programming jobs that reward their skills richly. Their returns to education are high. But if the student must stay close to home, then such jobs may not exist. They are educated but not compensated, and so their returns to education are low.

I pursue this example systematically with national socioeconomic data, quasi-experimental variation from a large-scale school construction program, and a quantitative spatial equilibrium model that isolates mechanisms. My goal is to illustrate how mobility affects the returns to education and, in turn, the aggregate and distributional impacts of school construction at scale. My setting is Indonesia, where the *Sekolah Dasar Inpres* program constructed nearly 62,000 new primary schools from 1973 to 1978. I measure education and wage outcomes four decades later with data from the 2011 to 2014 waves of the National Socioeconomic Survey (*Susenas*).

I evaluate the *Inpres* program with a difference-in-differences approach, and I find evidence consistent with my simple example. As in Duflo (2001), I estimate program effects by comparing exposed (young) and unexposed (old) individuals born in districts with high and low levels of school construction. I document positive and significant effects on education and wages. I then consider effect heterogeneity by labor market access, applying the market access approach of Donaldson and Hornbeck (2016). I measure market access as proximity to high-wage urban markets, and I estimate its interaction with the difference-in-differences variation. I instrument with faraway variation to address the potential correlation between local market access and economic shocks. I find that high market access roughly doubles the baseline effects of the program. The rural Indonesian student is too far from Jakarta to access high-paying programming jobs – even if they were to invest in programming skills – and so *Inpres* schools have limited impact.

I capture this intuition with a quantitative spatial equilibrium model in which individuals pursue education, then migrate for employment. Education and migration costs present frictions, and school construction decreases education costs. Individuals supply human capital that firms use for production, subject to diminishing marginal returns within locations. These diminishing returns act as a congestion force. Space enters on three counts. First, individuals consider both local and non-local labor markets. Mobility increases access to high urban wages, which raise the incentives to invest in education. Second, schools have both local and non-local impacts. Rural school construction may not lead to regional convergence, as rural students can leave after graduation. Third, equilibrium wage rates clear labor markets in each location. Markets are interdependent across locations and cannot be evaluated in isolation.

I estimate the model with the same data and identifying variation used to evaluate the *Inpres* program. Estimation is simple. I show that the model yields two key moment equations, which I can estimate sequentially by linear regression. The challenge is that education and migration enter as endogenous independent variables. The solution is that difference-in-differences variation provides an instrument for education, while faraway labor market access provides an instrument for migration. These instruments are sufficient for recovering all model parameters except for the congestion parameter, which I instead calibrate with estimates from the literature. Congestion is akin to a peer effect, and peer effects encounter a "reflection" problem that makes identification challenging (Manski 1993).

My model-based approach has several advantages relative to the difference-in-differences analysis, despite relying on the same identifying variation. Of course, difference-in-differences avoids the structure and assumptions of the economic model. But while difference-in-differences can only quantify net impacts, the model specifies mechanisms that allow me to decompose these impacts. While difference-in-differences can only study the *Inpres* program as implemented, the model provides a means to extrapolate from actual to hypothetical allocations of school construction. And while difference-in-differences cannot identify equilibrium effects experienced by the control group, as these effects difference out, the model features wage rates that adjust in equilibrium for treatment and control groups alike. Drawing on these advantages, I conduct two counterfactual exercises that highlight the impacts of mobility.

The first exercise finds that mobility accounts for roughly half of the aggregate gains achieved by the *Inpres* program. I isolate the direct effect of school construction by restricting students to their home locations. Most have low local returns to education, and so *Inpres* schools increase aggregate output by only 3%. I then allow for mobility, which has three effects: matching, motives, and market. The matching effect is that individuals sort into locations where they are more productive. I measure it by allowing for mobility but holding education and wage rates fixed. Output gains rise from 3 to 4%. The motives effect is that higher returns to education raise investment in education. Allowing education to adjust but holding wage rates fixed, output gains rise from 4 to 7%. The market effect is that diminishing marginal returns to human capital reduce equilibrium wage rates. Allowing education and wage rates to adjust, output gains fall from 7 to 6%. The total impact of mobility is thus to raise output gains from 3 to 6%. Bryan et al. (2014) find large gains from matching alone, but accounting for motives would raise them further, including in equilibrium.

The second exercise reveals a distributional tension. On one hand, the program decreased wage inequality between rural and urban *students* by 4%. The program expanded opportunities for less-advantaged rural students with high marginal returns, thereby delivering large gains for rural students relative to urban students. On the other hand, the program increased wage inequality between rural and urban *regions* by 9%. Rural students gain precisely by leaving to pursue high urban wages, contributing to urban growth but not to rural growth. The policymaker concerned with regional inequality must balance these opposing impacts. For Indonesia, regional convergence has long been an explicit policy goal, including for the *Inpres* program. Indeed, I invert the model to back our government preferences, and I find *Inpres* construction to be consistent with an objective function that places equal weight on aggregate output and regional equity.

In terms of policy, the first exercise suggests that a government can maximize aggregate gains with a "big push" that combines school construction with transportation investments aimed at improving mobility. Greater mobility raises the returns to education and thus the impact of educational infrastructure. At the same time, the second exercise emphasizes the potential for "brain drain" as greater mobility encourages rural students to leave home. And while economists often prioritize people over places, regional inequality may remain a salient concern for policymakers and politicians. Aggregate gains come with distributional costs.

My main contribution is to show how mobility affects educational investment at scale. I add endogenous education to a large literature, reviewed by Redding and Turner (2015) and Redding and Rossi-Hansberg (2017), that uses quantitative spatial equilibrium models to capture mobility. While this work largely focuses on transportation infrastructure, I show how spatial concerns also apply to educational infrastructure, and I provide new evidence on human capital formation in a spatial setting.[1] I also add geography and space to a literature that evaluates large-scale educational interventions. In comparison to Khanna (2023), Dinerstein et al. (2023), and Fujimoto et al. (2024), who also study national educational policy, I underscore the central importance of mobility and spatial heterogeneity for the returns on government investment. A small set of papers works at a similar intersection. In comparison to Agostinelli et al. (2024) and Eckert and Kleineberg (2024), who also apply spatial frameworks to studying education, I offer a tractable framework that is identified with quasi-experiment variation and a setting that allows me to evaluate the national, long-run impacts of school spending.

Methodologically, I build most directly on Duflo (2001) and Bryan and Morten (2019). The former highlights the *Inpres* variation, and the latter quantifies migration frictions with a spatial equilibrium model.[2] In combining the two, my approach relates to a broader theme in empirical work that well-identified effects can inform and distinguish between economic models, and that economic models can broaden the external validity of well-identified effects (Nakamura and Steinsson 2018). Relative to Bryan and Morten (2019), I endogenize education and use quasi-experimental variation to estimate the model. Relative to Duflo (2001), I emphasize mobility and use the model to analyze the long-run benefits of school construction. I quantify these benefits for one of the largest school construction programs in history.

[1] Recent examples for transportation include work on roads (Adukia et al. 2020, Fajgelbaum and Schaal 2020, Milsom 2023, Balboni 2024, Gertler et al. 2024, Graff 2024, Moneke 2024), highways (Allen and Arkolakis 2014, 2022, Faber 2014, Yang 2018, Alder 2023, Morten and Oliveira 2023), railroads (Donaldson and Hornbeck 2016, Donaldson 2018, Fajgelbaum and Redding 2022, Hornbeck and Rotemberg 2024), railways (Heblich et al. 2020, Severen 2023), subways (Gonzalez-Navarro and Turner 2018, Zárate 2024), and buses (Balboni et al. 2020, Tsivanidis 2023).

[2] Duflo (2001) spurs further work on the *Inpres* program that includes Duflo (2004), Martinez-Bravo (2017), Mazumder et al. (2019), Ashraf et al. (2020), Akresh et al. (2023), Bazzi et al. (2023). Bryan and Morten (2019) builds on Hsieh et al. (2019) and lies within a broader literature on selection into occupations (Roy 1951, Heckman 1974, Heckman and Sedlacek 1985, Keane and Wolpin 1997) and migration (Dahl 2002, Kennan and Walker 2011, Moretti 2011, Young 2013).

2 Program and Data

Indonesia's *Sekolah Dasar Inpres* program built primary schools at national scale in the 1970s. I describe the program and data. Appendix A provides additional detail on data sources and construction.

2.1 Background

Inpres refers to the presidential instructions (*Instruksi Presiden*) that established the *Inpres* program across two five-year development plans (*Repelita*). *Repelita I* focused on infrastructure investment from 1969 to 1974, and *Repelita II* placed additional priority on education from 1974 to 1979. The latter particularly emphasized rural development and regional equity. Coupled with windfall oil revenues, these government priorities enabled school construction of unprecedented scale. The stated goal was 62,000 new primary schools nationwide: 6,000 per year in 1973 and 1974, 10,000 per year in 1975 and 1976, and 15,000 per year in 1977 and 1978.[3] The program also funded teacher recruitment and compensation, as well as concurrent improvements in water and sanitation systems. Large-scale investment in physical infrastructure is typical of Indonesian development policy, with similar efforts for roads (Gertler et al. 2024), healthcare (Hsiao 2022), and flood protection (Hsiao 2023, 2024).

The program targeted low-enrollment districts with an allocation rule aimed at promoting regional equity. In 1973 and 1974, *Inpres* school construction was allocated across districts in proportion to pre-program unenrollment rates among children of primary school age. From 1975 to 1978, unenrollment was defined relative to a 15% threshold. Districts with unenrollment rates below 15% received no new schools, while those above 15% received new schools in proportion to how much their pre-program unenrollment rate exceeded 15%. Appendix figure A1 shows that school construction across districts is indeed proportional to child unenrollment rates in the data. "Districts" refer collectively to rural regencies (*kabupaten*) and urban municipalities (*kota*), each of which are subprovincial administrative units.

[3] *Inpres* No. 10/1973, 6/1974, 6/1975, 3/1976, 3/1977, and 6/1978 document these goals.

2.2 Treatment data

District-level data on *Inpres* school construction come from Duflo (2001), which draws on Ministry of National Development Planning (*Bappenas*) reports from 1973 to 1978 and population census reports from 1971. I observe the number of *Inpres* primary schools constructed from 1973 to 1978, *Inpres* water and sanitation spending from 1973 to 1978, child populations in 1971, and child enrollment rates in 1971. The data record planned school construction, which Duflo (2001) argues aligns well with actual school construction.

I verify the data by consulting the original primary sources. For school construction, I use the *Bappenas* reports to correct a modest number of inconsistencies. For child populations, I transcribe the census reports and again make limited corrections. Children are defined as those of age 5 to 14. For child enrollment, I construct child enrollment rates with a random sample of the census microdata. I draw total populations from the census reports, and I take the water and sanitation spending data as given in Duflo (2001). Appendix table A2 compares the resulting data to those in Duflo (2001) and shows that the data remain similar.

I construct a measure of labor market access that captures proximity to cities. Appendix figure A2 shows that human capital clusters in cities: districts with higher population densities have higher levels of education and wages, controlling for province fixed effects. This pattern of concentration suggests that urban areas offer higher returns to human capital than rural areas do, and it motivates studying treatment effect heterogeneity along this margin. For each district j in the set of districts $\mathcal{J}$, I use total population and land area to compute population density popden_j in 1971. I use district centroids to calculate Euclidean distances $\text{dist}_{jj'}$ to all other districts j' in hundreds of kilometers. Labor market access is an inverse distance weighted average of population densities across districts, and I take quadratic weights $a = 2$ as baseline.

$$MA_j^a = \sum_{j'} \frac{\text{popden}_{j'}}{(\text{dist}_{jj'} + 1)^a}. \tag{1}$$

Districts with high labor market access either contain or are close to cities. Population densities may be endogenous, and so I exclude own and nearby population densities within b kilometers to isolate faraway variation, which will serve as an instrument. I

Table 1: Treatment data

	Mean	SD	Min	Max	N
Inpres schools per 1000 children	2.26	1.09	0.69	8.29	282
Inpres school construction	219	174	16.0	824	282
Population (1000s)	422	362	12.4	1,974	282
Child population (1000s)	114	97.9	3.80	542	282
Child enrollment rates	0.51	0.14	0.02	0.84	282
Log population density	5.06	1.96	0.04	10.1	282
Log market access	8.63	1.32	5.25	10.8	282

Each observation is one district. Columns show means, standard deviations, minima, maxima, and the number of observations. Rows summarize the number of *Inpres* schools constructed from 1973 to 1978 and population and enrollment in 1971. Children are those of age 5 to 14. Market access is an inverse distance weighted sum of 1971 population densities across districts. Data: *Bappenas* reports (1973-78), census reports (1971).

take distance bands $b = 50$ kilometers as baseline.

$$MA_j^{ab} = \sum_{j'} \frac{\text{popden}_{j'}}{(\text{dist}_{jj'} + 1)^a} \cdot \mathbb{1}(\text{dist}_{jj'} > b) \tag{2}$$

Table 1 summarizes the data. The average district received 219 new primary schools from the *Inpres* program, yielding 2.26 new primary schools per 1,000 children recorded in 1971. This measure is the main treatment variable. Child enrollment rates in 1971 were low at 51% on average and did not exceed 84%. For the district with the lowest level of child enrollment, only 2% of school-aged children attended school. The data include districts with very high population densities in 1971, exceeding 24,000 people per square kilometer, as well as districts with far lower densities. The average value in levels is 742 people per square kilometer. Labor market access is tighter in distribution because it averages over population densities. Districts may themselves be sparsely populated, but still have high market access based on proximity to densely populated areas.

2.3 Outcome data

Individual-level data on socioeconomic outcomes come from the 2011, 2012, 2013, and 2014 National Socioeconomic Surveys (*Susenas*). That is, I measure long-run outcomes four decades after the first *Inpres* schools were completed in 1974. The analysis focuses on these long-run outcomes, but I use similar data from the 1976 Intercensal Population Surveys (*Supas*) to define a pre-*Inpres* baseline for counterfactuals. Following Duflo (2001), I restrict attention to male individuals of age 2 to 24 in 1974. These age groups correspond to individuals of age 39 to 64 in the *Susenas* data. I observe districts of residence and birth, and I use the latter to determine exposure to *Inpres* school construction. The underlying assumption is that individuals pursue primary schooling where they are born. I adjust districts to 1973 boundaries for consistency with the treatment data. I define migrants as those who reside outside of their birth districts, and I code whether migration occurs across provinces or to urban destinations. I compute Euclidean migration distances between district centroids.

The data record education. I observe the highest level of schooling attended and the number of years within each level. I compute the years of schooling completed, counting primary school as six years, middle school as three, high school as three, community college as two, and university as four. I pool vocational, religious, and traditional schooling for middle school and high school. I subtract one from the number of years attended if individuals report not having completed a given level of schooling. Tertiary schooling is subject to measurement error, noting that community college programs range from one to three years and that master's and doctoral programs exceed four years. The baseline analysis takes twelve as the maximum years of schooling to avoid this measurement issue, given that the impacts of *Inpres* primary schools likely concentrate at lower levels.

The data also record wages. I measure monthly net wages from an individual's main job, inclusive of both money and goods, as well as hours worked in the past week. I use these measures to compute log hourly wages. I define the employed as those who report working as their primary activity in the past week. I observe whether individuals work as employees, employers, own-account workers, or unpaid family workers. I code employees as wage-employed and others as self-employed. The data measure income for some of the self-employed, but appendix table A3 shows these data to be incomplete. I therefore focus on wages in my analysis. Duflo (2001) discusses the potential selection issue at length.

Table 2 summarizes the data. I observe large improvements in education and wages. Years of schooling have doubled from 3.06 in the 1970s to 7.30 in the 2010s.

Table 2: Outcome data

	Supas 1976			*Susenas* 2011-2014		
	Mean	SD	N	Mean	SD	N
Male	1.00	0.00	22,969	1.00	0.00	483,503
Age	48.4	6.98	22,969	49.5	6.48	483,503
Years of schooling	3.06	3.52	22,969	7.30	3.92	483,503
Years of schooling, wage > 0	4.95	4.24	6,237	8.79	3.69	164,046
Log hourly wages	7.96	0.93	6,237	9.12	0.88	164,046
Completed primary school	0.29	0.45	22,969	0.75	0.43	483,503
Completed middle school	0.11	0.31	22,969	0.44	0.50	483,503
Completed high school	0.06	0.23	22,969	0.31	0.46	483,503
Employed	0.89	0.32	22,969	0.91	0.29	483,503
Wage-employed	0.33	0.47	20,363	0.41	0.49	438,909
Self-employed	0.67	0.47	20,363	0.59	0.49	438,909
Weekly hours	43.3	15.6	20,363	42.7	15.9	438,909
Migrant	0.28	0.45	22,969	0.25	0.43	483,503
Provincial migrant	0.46	0.50	6,351	0.64	0.48	122,103
Urban migrant	0.79	0.41	6,351	0.60	0.49	122,103
Distance	400	634	6,351	572	623	122,103

Each observation is one male individual of age 39 to 64. Columns show means, standard deviations, and the number of observations. Rows summarize demographics, including education and wages. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. The employed are those who report working as their primary activity. Conditional on employment, I observe if they are wage- or self-employed and their weekly hours worked. Migrants are those whose districts of birth and residence do not coincide. Conditional on migration, I observe if they migrate across provinces, if they migrate to an urban destination, and their distance migrated. Distances are Euclidean and between district centroids. Districts are based on 1973 boundaries. Data: *Supas* survey (1976), *Susenas* surveys (2011-14).

The data cover individuals of age 39 to 64 in each case. Wage earners have one to two more years of schooling. Real wages have more than tripled: average (natural) log hourly wages, as measured in year-2011 dollars, rise from 7.96 to 9.12. Log hourly wages of 9.12 correspond to roughly \$1 USD per hour based on year-2011 exchange rates. By comparison, the World Bank used a threshold of \$1.90 USD per day to define extreme poverty in 2011. Completion rates for primary, middle, and high school have increased by similar magnitudes, although high school completion remains low at 31%

in the *Susenas* data.

Employment and migration outcomes are more stable over time. Employment rates are around 90%. Among the employed, the wage-employment rate is 33 to 41%, and average working hours are around 43 hours per week. There is meaningful migration at 25 to 28% of individuals, with most migrants crossing provincial boundaries and migrating to urban destinations. The unconditional cross-province migration rate is 13 to 16%. By comparison, the cross-state migration rate is 31% in the United States, a high-income economy with presumably lower migration costs.[4]

3 Evaluation

I evaluate the long-run effects of the *Inpres* program with a difference-in-differences approach. I focus on education and wages, and I document heterogeneity by labor market access. Appendix B discusses additional outcomes and robustness.

3.1 Identification

As in Duflo (2001), I compare treated and control age groups in districts with high and low levels of program exposure. The first *Inpres* primary schools were completed in 1974. The treatment age group is individuals of age 2 to 6 in 1974, as they were young enough to benefit from the program. The control age group is individuals of age 12 to 17 in 1974, as they were too old to benefit. Age groups are thus analogous to time in the typical difference-in-differences setting. Program exposure is given by the intensity of school construction in a district. The regression specification is

$$Y_{ijkt} = \alpha_j + \alpha_k + \alpha_t + \beta S_j T_k + \boldsymbol{C}_j T_k \boldsymbol{\phi} + \varepsilon_{ijkt} \tag{3}$$

for individuals i born in district j and age group k, as measured in survey year t. I study how the interaction between program intensity S_j and treatment dummy T_k affects outcomes Y_{ijkt}, controlling for birth district fixed effects α_j, age group fixed effects α_k, survey year fixed effects α_t, and birth district controls $\boldsymbol{C}_j$. The error term

[4] I use 2013 and 2014 American Community Survey data to compute American migration rates. In doing so, I define migration as I do in the Indonesian context. Restricting attention to those born in the United States, which I take to include the 48 contiguous states plus the District of Columbia, I calculate the proportion of individuals residing outside of their state of birth.

is ε_{ijkt}, and I cluster standard errors by district of birth.

Coefficient of interest β captures the causal effect of *Inpres* program intensity. Intensity is the number of *Inpres* schools constructed from 1973 to 1978 per 1,000 children in 1971. The identifying assumption is parallel trends in high- and low-construction districts absent the program. I test this assumption with a placebo experiment, comparing individuals of age 12 to 17 and those of age 18 to 24 in 1974. Both groups were too old to benefit from *Inpres* school construction. Controls include child populations in 1971, child enrollment rates in 1971, and *Inpres* water and sanitation spending from 1973 to 1978. The first and second controls address the concern that *Inpres* schools were allocated as a function of these measures. That is, high- and low-construction districts differed along these dimensions to begin with. The third control addresses the concern that high-construction districts received concurrent investment in water and sanitation systems. These effects should not be attributed to school construction.

I also consider treatment effect heterogeneity with respect to labor market access MA_j, as defined by equation 1.

$$Y_{ijkt} = \alpha_j + \alpha_k + \alpha_t + \beta S_j T_k + \gamma MA_j S_j T_k + \delta MA_j T_k + \boldsymbol{C_j} T_k \boldsymbol{\phi} + \varepsilon_{ijkt} \tag{4}$$

Coefficient γ on the triple interaction term $MA_j S_j T_k$ captures treatment effect heterogeneity. I control for all two-way interactions. The original difference-in-difference term $S_j T_k$ is the common treatment effect, and the additional term $MA_j T_k$ allows for differential trends among districts with high and low market access. Birth district fixed effects α_j absorb the remaining two-way interaction $MA_j S_j$. To facilitate interpretation, I estimate equation 4 with the z-score of market access. Coefficient β is the *Inpres* effect for districts with average market access, while coefficient γ is the added effect for districts with market access that is one standard deviation higher.

A potential concern is the endogeneity of market access. On one hand, I measure market access with 1971 populations that predate *Inpres* school construction and Euclidean distances that sidestep endogenous road networks. Neither quantity enters the allocation rule or responds to the program, possibly lessening the endogeneity concern. On the other hand, my measure of market access may still be correlated with omitted variables. Own-district population density is one such confounder: positive economic shocks may attract migrants, affecting educational composition and wages alongside population density and market access. If positive shocks are more likely in urban areas, then I will misattribute these shocks to market access. Similar concerns apply to nearby districts, as population densities and economic shocks may each be spatially correlated. I thus turn to faraway variation, as in Donaldson and Hornbeck (2016) and Jedwab and Storeygard (2022). I instrument for market access with the faraway measure of equation 2, omitting both own- and nearby-district population densities.[5] In defining "faraway," the trade-off is between exclusion and relevance. The farthest districts are likely orthogonal to own-district unobservables. But market access downweights these districts, and so they are also weak instruments.

3.2 Long-run effects

Inpres school construction in the 1970s improves education and wages in the 2010s. For education, table 3 shows that years of schooling rise by nearly 0.1 years for each *Inpres* school per 1,000 children. The average district received 2.26 *Inpres* schools per 1,000 children (table 1), yielding a total average effect of approximately 0.2 years of schooling. For the full sample of individuals, this effect is relative to a control mean of 6.7 years. For the subsample of wage earners, I estimate an effect size that is larger but nonetheless similar. The larger effect is relative to a larger control mean. Although the subsample of wage earners is less than half of the full sample, the similarity of the estimates provides some reassurance that sample selection does not strongly bias the results. I will present both sets of estimates for all regressions that take years of schooling as the outcome.

Appendix figure B1 presents effects on completion at each level. I estimate impacts on dummy variables that encode whether individuals have completed one or more years of schooling, two or more years, three or more years, and so on. Completion of six, nine, and twelve years of schooling correspond to primary, middle, and high school completion. *Inpres* effects are concentrated at the primary and middle school levels. Each *Inpres* school per 1,000 children increases primary and middle school completion rates by more than 1% each, implying increases of more than 2% for the average district. These impacts are relative to control means of 69% and 37% for primary and middle school completion. *Inpres* primary schools do not have significant

[5] I instrument for $MA_j S_j T_k$ and $MA_j T_k$ with $MA_j^b S_j T_k$ and $MA_j^b T_k$. MA_j^b is as in equation 2.

Table 3: Education and wages

	Treatment			Placebo		
	Years of schooling	Years of schooling	Log wages	Years of schooling	Years of schooling	Log wages
Inpres effect	0.0847**	0.0994**	0.0256***	0.00779	0.0366	-0.0108
	(0.0356)	(0.0385)	(0.00943)	(0.0291)	(0.0486)	(0.00926)
Age group FE	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	264,307	98,413	98,413	222,194	59,727	59,727
Control mean	6.695	8.434	9.195	6.140	7.054	8.892

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. Treatment regressions compare individuals of age 2 to 6 versus 12 to 17, and placebo regressions compare those of age 12 to 17 versus 18 to 24. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

impacts on high school education.

For wages, table 3 shows that log hourly wages rise by nearly 0.03 log points for each *Inpres* school per 1,000 children, with an effect size of nearly 0.06 log points for the average district and a control mean of 9.2. These increased wage rates are consistent with human capital gains from additional education. Appendix table B1 suggests that the effect is not instead driven by other margins of employment, as I find no significant impact on overall, wage-, or self-employment status. I also find no significant impact on weekly hours worked. Moreover, the lack of an effect on wage employment is consistent with the similarity of the years of schooling estimates between the full sample and the subsample of wage earners. Strong effects on wage employment would otherwise induce stronger selection into the latter sample.

I obtain insignificant placebo estimates throughout. Table 3 estimates insignificant effects for education and wages, appendix figure B1 for educational completion, and appendix table B1 for employment outcomes. The only exception is appendix table B2, where the placebo effect on migration is weakly significant at the 10% level.

Moreover, my long-run estimates are largely consistent with the medium-run estimates in table 4 of Duflo (2001). My estimates for years of schooling are roughly half as large, although they remain positive and statistically significant. The discrepancy seems to arise from the 2011-14 *Susenas* data. Appendix table B3 replicates Duflo (2001) with 1995 *Supas* data and my baseline specification, with magnitudes squarely within the range presented in Duflo (2001). Although I follow the same age cohorts in both cases – those of age 2 to 24 in 1974 – differences may arise because the *Susenas* surveys require recall of educational attainment at a much later date. The *Supas* and *Susenas* surveys also differ in frequency, scale, and intention.[6] My estimates for log wages are on the higher end of Duflo (2001), but it is natural to expect different wage effects later in the life cycle. Taken together, smaller education and larger wage effects may imply larger returns to education in my long-run dataset.

3.3 Labor market access

Inpres school construction has stronger education and wage effects for districts with high labor market access. Intuitively, high market access increases the pool of job opportunities and thus the returns to education. Conversely, low market access limits the pool of opportunities, and so schooling has low returns – even if *Inpres* construction makes schooling more available. Table 4 estimates this interaction, with market access in z-score units, and finds it to be large and statistically significant. I obtain similar but slightly larger estimates with the IV specification, which I take as my baseline. I find the interacted effects to be as large as the mean effects.[7] Increasing market access beyond the mean by one standard deviation leads to a doubling of the *Inpres* effects. That is, individuals born in districts with high labor market access benefit most from *Inpres* school construction. In targeting these districts, the program

[6] *Supas* is decadal – five years after each census – while *Susenas* is annual. *Supas* 2015 included 652,000 households, while *Susenas* 2014 included 143,000. *Supas* targets population counts across districts, while *Susenas* targets socioeconomic conditions and how they change over time. I also note that the most direct comparisons are to columns 3 and 6 in table 4 of Duflo (2001). These columns include the full set of controls, as I do for all of my specifications.

[7] I estimate mean effects that are themselves larger than those in the uninteracted regressions of table 3. Indeed, market access is right-skewed, and so a district with mean market access remains proximate to major labor markets. The larger magnitudes are well in line with those in Duflo (2001), although I again estimate stronger wage effects.

Table 4: Labor market access

	OLS			IV: 50 km		
	Years of schooling	Years of schooling	Log wages	Years of schooling	Years of schooling	Log wages
Inpres effect	0.132***	0.165***	0.0395***	0.133***	0.169***	0.0410***
	(0.0376)	(0.0394)	(0.00969)	(0.0384)	(0.0405)	(0.00971)
Inpres effect × *MA*	0.112***	0.152***	0.0247***	0.107**	0.154***	0.0273**
	(0.0402)	(0.0434)	(0.00911)	(0.0498)	(0.0561)	(0.0117)
Age group FE	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	264,307	98,413	98,413	264,307	98,413	98,413
Control mean	6.695	8.434	9.195	6.695	8.434	9.195
F-statistic				162.7	99.52	99.52

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. OLS regressions do not instrument for labor market access, while IV regressions instrument with faraway variation that excludes districts within 50 km. I compare individuals of age 2 to 6 versus 12 to 17. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. Market access is an inverse distance weighted sum of 1971 population densities across districts. I convert market access to a z-score and interpret units in standard deviations. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

maximizes its aggregate impact.

At the same time, high labor market access can encourage out-migration, particularly for rural districts with limited local opportunities. In targeting these districts, the program helps rural individuals who eventually leave, with little net benefit to the districts themselves. Figure 1a confirms that out-migration is highest for districts with high market access. I estimate a positive and significant coefficient β for the specification

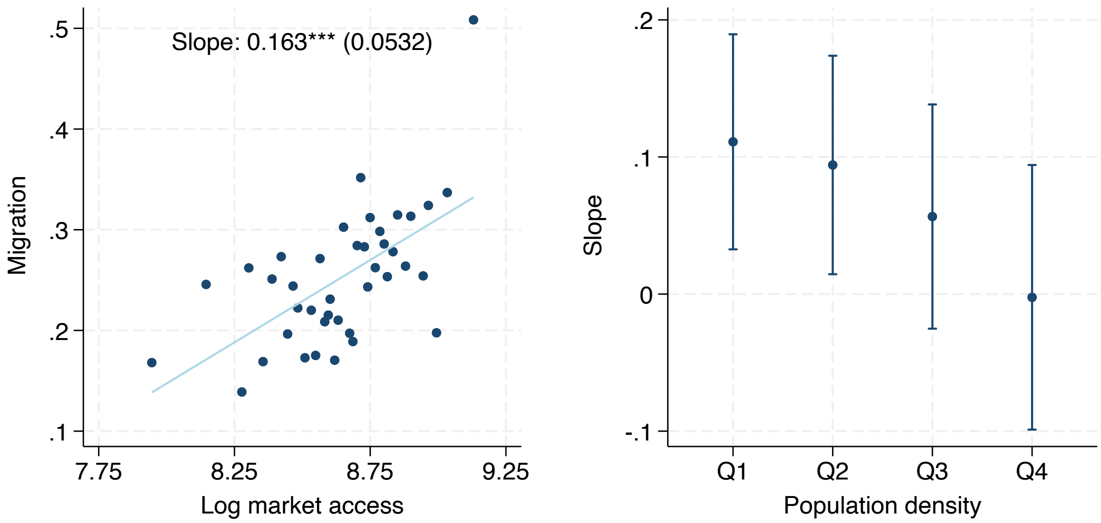
$$Y_{ijt} = \alpha_P + \alpha_t + \beta \log MA_j + \varepsilon_{ijt}$$

with individuals i born in districts j of provinces $P(j)$, as measured in survey year t. I regress out-migration in 2011 to 2014 on log labor market access in 1971, con-

Figure 1: Migration and market access

(a) Full sample

(b) By population density



The left figure is a binned scatterplot, and each observation is one male individual of age 39 to 64. I report the slope coefficient and standard error. The right figure plots the analogous slope coefficients with 95% confidence bands for each quartile of the population density distribution. Migrants are those whose districts of birth and residence do not coincide. Population density is as measured in 1971. Market access is an inverse distance weighted sum of 1971 population densities across districts. Population density and market access are for individiuals' districts of birth. I control for province and survey-year fixed effects, and I cluster standard errors by province. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Susenas* surveys (2011-14), census reports (1971).

trolling for province and survey year fixed effects and clustering standard errors by province. Figure 1b shows that rural districts drive the correlation. For quartiles $Q(j)$ of population density, I estimate the specification

$$Y_{ijt} = \alpha_P + \alpha_Q + \alpha_t + \beta_Q \log MA_j + \varepsilon_{ijt}.$$

I obtain positive and significant for rural districts in the lower quartiles of population density, with insignificant estimates for urban districts.[8] Rural-born individuals benefit from market access but leave, while urban-born individuals benefit and stay. Labor market access amplifies *Inpres* effects at some distributional cost.

Consistent with out-migration, appendix table B4 documents attenuated *Inpres* effects for places relative to people. The first three columns show baseline estimates,

[8] Coefficients β^q are all smaller than coefficient β of figure 1a because quartile fixed effects absorb level differences between urban and rural districts.

as in table 3, taking birth-district school construction as treatment. They capture effects on individuals, inclusive of those who migrate away. The last three columns instead take current-district construction as treatment. They show the smaller effects on districts' own long-run outcomes, net of out-migration that dissipates the local gains from school construction. People benefit more than places do.

One nuance, however, is that this distributional tension arises even if migration patterns do not themselves respond strongly to the program. Indeed, appendix table B2 shows that migration patterns do not respond on either the extensive or intensive margins. I detect no impact on migration rates, migration across provinces, migration to urban destinations, or migration distances. But large and heterogenous migration rates in the cross-section – as seen in figure 1a – still lead to gains that escape rural districts. If *Inpres* schools were to increase migration rates, then they would amplify the concern: rural individuals would leave at even higher rates.

Finally, appendix B considers robustness. The main results are stable with respect to the chosen distance weights and bands, the censoring of tertiary education, the inclusion of flexible spatial controls, and the exclusion of the largest labor markets.

4 Model

I build an economic model for further analysis. In the model, individuals invest in education and migrate for work. School construction encourages education. Firms use labor to produce, and wage rates clear labor markets in equilibrium. Appendix C provides derivations.

4.1 Choices and frictions

Building on Bryan and Morten (2019) and Hsieh et al. (2019), individuals make education and migration choices in two stages of life: school and work.[9] First, they realize schooling shocks and choose education. Second, they realize skill shocks and choose whether and where to migrate. These choices determine the supply of human capital across locations. Education and migration costs present frictions.

[9] Relative to Bryan and Morten (2019), I endogenize education. Relative to Hsieh et al. (2019), I allow individuals to choose education without perfect information on migration.

Consider an individual i of origin $j(i)$ and age group $k(i)$. In the first stage, they realize schooling shock ϵ and choose education e to maximize utility.

$$u_{jk}(\epsilon) = \max_e \left\{ \bar{v}_{jk}(e) - c_{jk}(e, \epsilon) \right\} \tag{5}$$

Utility includes two components: labor utility and education costs.

$$\bar{v}_{jk}(e) = \mathbb{E}[v_{jk}(e, \varepsilon) \,|\, e], \quad c_{jk}(e, \epsilon) = e\tau^e_{jk}\epsilon \tag{6}$$

Labor utility $\bar{v}_{jk}$ is the future utility from work, subject to skill shocks ε that are not realized until the second stage. Costs c_{jk} include regional costs and individual shocks. Education costs τ^e_{jk} capture systematic barriers to education, which school construction can help to alleviate. Schooling shocks ϵ capture idiosyncratic barriers and are independent and identically distributed.

In the second stage, the individual considers locations ℓ. They take education e as given, and they realize skill shocks $\varepsilon = \{\varepsilon_\ell\}$ across locations. They choose a destination to maximize labor utility.

$$v_{jk}(e, \varepsilon) = \max_\ell \left\{ v_{jk\ell}(e, \varepsilon_\ell) \right\} \tag{7}$$

Labor utility includes three components: amenities, wages, and costs.

$$v_{jk\ell}(e, \varepsilon) = \frac{a_\ell w_{jk\ell}(e, \varepsilon_\ell)}{\tau^m_{jk\ell}}, \quad w_{jk\ell}(e, \varepsilon_\ell) = r_\ell h_{jk\ell}(e, \varepsilon_\ell), \quad h_{jk\ell}(e, \varepsilon_\ell) = e^\eta s_{jk\ell}\varepsilon_\ell \tag{8}$$

For each location ℓ, amenities a_ℓ capture unobserved destination qualities. Wages $w_{jk\ell}$ depend on wage rates r_ℓ per unit of human capital $h_{jk\ell}$. Human capital is perfectly substitutable, and so individuals face common wage rates within locations. Human capital increases in education e, subject to elasticity $\eta < 1$ that captures diminishing marginal returns to education. Human capital also increases in skill $s_{jk\ell}$ and skill shock ε_ℓ. As in Eaton and Kortum (2002), skill shocks follow Fréchet distribution

$$F(\varepsilon_1, \ldots, \varepsilon_L) = \exp\{-\sum_\ell \varepsilon_\ell^{-\theta}\}.$$

A high value for Fréchet parameter θ implies low skill dispersion. Skill shocks are uncorrelated with schooling shocks. Migration costs $\tau^m_{jk\ell}$ are the financial, psychological, and opportunity costs of moving and working away from home.

It is convenient to define location utility $\tilde{v}_{jk\ell}$, collecting elements that do not vary individually. I also define labor market access MA_{jk} as a power sum of location utilities, with weighting given by Fréchet parameter θ.

$$\tilde{v}_{jk\ell} = \frac{a_\ell r_\ell s_{jk\ell}}{\tau^m_{jk\ell}}, \quad MA_{jk} = \sum_\ell \tilde{v}^\theta_{jk\ell}$$

This measure of market access is correlated with the previous measure, as defined by equation 1, because both sum wages across locations and downweight by distance. For this measure, wages enter through wage rates r_ℓ and skill $s_{jk\ell}$ in the numerator, distance enters through migration costs $\tau^m_{jk\ell}$ in the denominator, and the functional form follows from the Fréchet distributional assumption. For the previous measure, the functional form does not depend on any particular modeling assumption, and so its interpretation is independent of the modeling exercise that follows.

4.2 Returns to education

Labor market access magnifies the returns to education. Evaluating expected wages $\bar{w}_{jk}$ and utility $\bar{v}_{jk}$ from labor,

$$\begin{aligned}
\bar{w}_{jk}(e) &= \mathbb{E}[\max_\ell w_{jk\ell}(e, \varepsilon_\ell) \,|\, e] = \gamma e^\eta MA_{jk}^{\frac{1}{\theta}} \left(\frac{\sum_\ell (\frac{\tau^m_{jk\ell}}{a_\ell}) \tilde{v}^\theta_{jk\ell}}{MA_{jk}} \right), \\
\bar{v}_{jk}(e) &= \mathbb{E}[\max_\ell v_{jk\ell}(e, \varepsilon_\ell) \,|\, e] = \gamma e^\eta MA_{jk}^{\frac{1}{\theta}}.
\end{aligned}$$

The γ term is a scalar given by $\gamma = \Gamma(1 - \frac{1}{\theta})$ and gamma function $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$. It accounts for the positive expected utility derived from skill shocks ε. The marginal wage and welfare gains from education e are

$$\frac{\partial \bar{w}_{jk}}{\partial e} = MA_{jk}^{\frac{1}{\theta}} \left(\frac{\gamma\eta}{e^{1-\eta}} \right) \left(\frac{\sum_\ell (\frac{\tau^m_{jk\ell}}{a_\ell}) \tilde{v}^\theta_{jk\ell}}{MA_{jk}} \right), \quad \frac{\partial \bar{v}_{jk}}{\partial e} = MA_{jk}^{\frac{1}{\theta}} \left(\frac{\gamma\eta}{e^{1-\eta}} \right).$$

The key observation is that the wage and welfare gains from education are increasing in market access MA_{jk}. This comparative static is seen most clearly when

$\tau^m_{jk\ell} = a_\ell$ for all ℓ, such that wage and welfare gains coincide. Education is more useful when students can access labor markets that reward human capital richly. When $\tau^m_{jk\ell} = a_\ell$ does not hold for all ℓ, wage gains embed additional variation in amenities a_ℓ and migration costs $\tau^m_{jk\ell}$. Utility depends on wages and amenities in combination, and so high amenities allow for lower wage gains because of compensating differentials. Similarly, utility depends on wages net of migration costs, and so high migration costs call for high wage gains. For both amenities and migration costs, this complication arises because each is an unobserved quantity. Finally, wage and welfare gains are decreasing in education when $\eta < 1$ because of diminishing marginal returns. The empirical challenge is that the difference-in-differences analysis conflates the above forces. The model provides conceptual guidance and a basis for disentangling them.

4.3 Outcomes

I characterize education, wage, and migration outcomes, highlighting comparative statics that are consistent with the *Inpres* effects documented in section 3. Education follows from equation 5.

$$\bar{e}_{jk} = \mathbb{E}[e] = \left(\frac{\gamma \eta M A_{jk}^{\frac{1}{\theta}}}{\tau^e_{jk}} \right)^{\frac{1}{1-\eta}} \bar{\epsilon} \tag{9}$$

for scalars $\bar{\epsilon} = \mathbb{E}[\epsilon^{-\frac{1}{1-\eta}}]$ and γ as previously defined. The expression does not depend on destination ℓ because individuals choose education before realizing their skill shocks and choosing their destination. Location characteristics enter only collectively through labor market access MA_{jk}. The first comparative static is that education costs τ^e_{jk} reduce education, as is natural. In reducing education costs, school construction should encourage education. The second comparative static is that labor market access MA_{jk} amplifies the impact of education costs. At the extreme, $MA_{jk} = 0$ implies that $e = 0$ for all $\tau^e_{jk} > 0$. Reducing education costs has zero effect. Conversely, high market access expands the pool of job opportunities and raises the returns to education. Reducing education costs should thus have large effects. Tables 3 and 4 support both predictions: *Inpres* school construction increases years of schooling completed, particularly where market access is high.

Wages follow from equations 8, evaluated conditionally for observed destinations.

$$\bar{w}_{jk\ell} = \mathbb{E}[w \,|\, \text{choose } \ell] = \left(\frac{\gamma \tau^m_{jk\ell} MA^{\frac{1}{\theta}}_{jk}}{a_\ell}\right)\left(\frac{\bar{e}_{jk}}{\bar{\epsilon}}\right)^{\eta}\tilde{\epsilon} \tag{10}$$

for scalars $\bar{\epsilon} = \mathbb{E}[\epsilon^{-\frac{1}{1-\eta}}]$, $\tilde{\epsilon} = \mathbb{E}[\epsilon^{-\frac{\eta}{1-\eta}}]$, and γ as previously defined. The conditional expectation captures positive selection on skill shocks ε_ℓ, whereby those observed in destination ℓ have likely realized high skill shocks for that destination. Otherwise they would have selected into another destination instead. Fréchet shocks allow me to characterize this selection in closed form. The comparative static is that wages are increasing in education. School construction should thus have directionally similar effects on wages and education. Tables 3 and 4 again support this prediction: *Inpres* construction increases wages, particularly where labor market access is high – just as it does for education.

Migration follows from equation 7. Substituting equations 8, it becomes

$$\bar{m}_{jk\ell} = \mathbb{P}[\tilde{v}_{jk\ell}\varepsilon_\ell \geq \tilde{v}_{jk\ell'}\varepsilon_{\ell'} \text{ for all } \ell'] = \frac{\tilde{v}^{\theta}_{jk\ell}}{\sum_{\ell'} \tilde{v}^{\theta}_{jk\ell'}} \tag{11}$$

Fréchet skill shocks ε_ℓ give choice probabilities of familiar closed form. Migration choices depend on location utilities $\tilde{v}_{jk\ell}$ for every location in the choice set, but not on education e. Education does not vary by location, and so it drops from the maximand and does not affect destination choice. The comparative static is thus that migration is invariant to education. School construction should not affect migration, and table B2 supports this prediction: *Inpres* school construction has no significant impact by any measure.[10]

4.4 Output

Aggregate output is given by national production Y, which sums across locations. In each location, perfectly competitive firms produce with total human capital H_ℓ as

[10] To capture migration effects, I could interpret school construction as affecting both education and migration costs. Fixed costs of migration would offer a microfoundation for the latter. School construction increases education and thus wages. Higher wages then lower the relative burden of these fixed costs, thereby encouraging migration. In my setting, however, the null results of table B2 are not consistent with such an extension.

the sole input, subject to productivity A_ℓ and production elasticity κ. The production function determines demand for human capital.

$$Y = \sum_\ell Y_\ell, \quad Y_\ell = A_\ell H_\ell^\kappa \tag{12}$$

Productivity captures spatial heterogeneity in the productive use of human capital, including through differences in technology, physical capital, and sectoral composition. Production elasticity $0 < \kappa < 1$ captures diminishing marginal returns within locations. I interpret these diminishing returns as a net congestion force.[11]

By perfect competition, firms hire until the marginal product of total human capital H_ℓ meets wage rate r_ℓ.

$$r_\ell = \kappa A_\ell H_\ell^{\kappa-1} = \frac{\partial Y_\ell}{\partial H_\ell} \tag{13}$$

Total human capital sums over individuals in each location, as do total wages.

$$H_\ell = \sum_{j,k} N_{jk}\bar{m}_{jk\ell}\bar{h}_{jk\ell}, \quad W_\ell = \sum_{j,k} N_{jk}\bar{m}_{jk\ell}\bar{w}_{jk\ell}$$

Populations N_{jk} count working individuals from origins j and age groups k, of which proportion $\bar{m}_{jk\ell}$ choose destination ℓ. These individuals have average human capital $\bar{h}_{jk\ell} = \mathbb{E}[h \,|\, \text{choose } \ell]$ and average wages $\bar{w}_{jk\ell} = \mathbb{E}[w \,|\, \text{choose } \ell]$.

The above expressions allow for straightforward interpretation of aggregate output. Combined with equations 8, which imply $\bar{w}_{jk\ell} = r_\ell \bar{h}_{jk\ell}$, it follows that

$$Y_\ell = \frac{r_\ell H_\ell}{\kappa} = \frac{W_\ell}{\kappa}. \tag{14}$$

The first equality decomposes production as the product of wage rates and total human capital. The second states that production is given by total wages. Both are subject to a scalar κ, which cancels when considering percentage changes. That is, the model provides convenient microfoundation by which changes in aggregate output

[11] A natural interpretation is that production congests local factor markets. Appendix C demonstrates a similar force with goods markets: if production is imperfectly substitutable across locations, then added production suppresses goods prices and thus marginal productivity. An opposing force arises from agglomeration through innovation or economies of scale. I capture these forces collectively, as it is difficult to separate them empirically.

coincide with changes in total wages.

4.5 Equilibrium

I define equilibrium as the set of wage rates $r = \{r_\ell\}$ that clear human capital markets in every location.

$$H_\ell^D(r_\ell) = H_\ell^S(r) \quad \forall\, \ell \tag{15}$$

Firms demand human capital. Production is local, and so demand can be evaluated separately for each location. Inverting equation 13,

$$H_\ell^D(r_\ell) = \left(\frac{\kappa A_\ell}{r_\ell}\right)^{\frac{1}{1-\kappa}}.$$

Demand is downward-sloping, as marginal returns are diminishing for $\kappa < 1$. Individuals supply human capital. Migration is national, and so supply must be evaluated jointly across locations. By equation 14,

$$H_\ell^S(r) = \frac{W_\ell(r)}{r_\ell}.$$

Supply is upward-sloping, as high wage rates attract human capital and increase total wages. Wage rates enter total wages W_ℓ through migration and average wages, as given by equations 10 and 11.

4.6 Discussion

The model focuses on internal migration for employment after graduation. First, migration is internal. Indeed, international out-migration is limited to less than 0.5% of the total population (World Bank 2022).[12] Second, migration is for employment. I abstract from migration for schooling, whereby parents move so that their children can attend better schools.[13] However, amenities a_ℓ capture this mechanism to the

[12] From 1980 to 2015, Indonesia experienced net out-migration of 905,000. Foreign-born individuals account for part of this out-migration, as their population fell by 428,000 over the same period. The total population was 259,000,000 in 2015.

[13] The literature on moving to opportunity, as reviewed by Chyn and Katz (2021), finds that children experience positive education effects after moving to better neighborhoods, subject to disruption effects that sometimes dominate (Chetty et al. 2016, Chetty and Hendren 2018, Chyn 2018, Laliberté 2021, Nakamura et al. 2021, Rojas-Ampuero and Carrera 2021).

extent that good schools act as amenities in certain locations. Third, employment is after graduation. I abstract from child labor, which entails employment before graduation. However, education costs τ^e_{jk} capture this force to the extent that child labor increases the opportunity costs of education in certain locations.[14]

The model also imposes several simplifications. First, I assume a common elasticity η of human capital with respect to education. But I note that the returns to education remain heterogeneous: wage rates r_ℓ differ across destinations, skill $s_{jk\ell}$ differs across origins and age groups, and skill shock ϵ differs across individuals. Each affects how an additional year of schooling translates into added wages. Second, individuals know location utilities $\tilde{v}_{jk\ell}$ across locations when choosing education. But I can accommodate uncertainty to the extent that it is Fréchet distributed and thus absorbed by skill shocks ε_ℓ. Third, individuals know these skill shocks across destinations when choosing migration, even to faraway locations. But migration costs $\tau^m_{jk\ell}$ capture information frictions to the extent that they are increasing in distance. Fourth, I abstract from sequential migration with skill accumulation over the life cycle. These dynamics require substantially more modeling. I will estimate the model with long-run wages and interpret counterfactuals as being net of these dynamics.

I also note a contrast between school construction and "place-based policy" in the form of spatially targeted infrastructure investment.[15] School construction is place-based because schools serve students locally. But other place-based policies provide only local benefits. In-migration offsets these gains by increasing local prices and draining non-local productivity. By contrast, schools provide portable benefits. Out-migration magnifies and distributes these gains.

5 Estimation

I describe estimation and identification, and I present estimates. Estimation reduces to linear regression with instruments. Appendix C provides derivations.

[14] A literature on child labor and education finds that higher returns to child labor reduce investment in schooling (Atkin 2016, Shah and Steinberg 2017, Bau et al. 2021, Shah and Steinberg 2021).

[15] Glaeser and Gottlieb (2008), Neumark and Simpson (2015), and Austin et al. (2018) review the literature on place-based policy.

5.1 Frictions and amenities

I estimate education costs, migration costs, and relative amenities with aggregate data by origin j, age group k, and destination ℓ. I define origins as birth districts, age groups as those of age 2 to 6 versus 12 to 17 in 1974, and destinations as current districts of residence. By aggregating, I eliminate unobserved individual shocks (ϵ, ε) for schooling and skill, and I obtain a simple log-linear expression to take to data.

$$\log \bar{w}_{jk\ell} - \log \bar{e}_{jk} = \log \frac{\tilde{\epsilon}}{\bar{\epsilon}\eta} - \log a_\ell + \log \tau^e_{jk} + \log \tau^m_{jk\ell} \tag{16}$$

Estimation is a linear regression with fixed effects. On the left-hand side, I construct each term from data. The first term is the log of average hourly wages by origin, age group, and destination. I only observe wages for individuals who select into a given destination ℓ, and so the model accounts explicitly for selection on unobserved skill shocks ε_ℓ. The second term corresponds to average years of schooling. On the right-hand side, I estimate each term. The first term is a constant that combines scalars $(\tilde{\epsilon}, \bar{\epsilon})$, which depend on the distribution of schooling shocks ϵ, and parameter η, which is the human capital elasticity. The second term is a fixed effect by destination ℓ, and it identifies amenities a_ℓ. The third term is a fixed effect by origin and age group jk, and it identifies education costs τ^e_{jk}. The fourth term is the residual, and it has a structural interpretation as migration costs $\tau^m_{jk\ell}$. I thus obtain estimates $(a_\ell, \tau^e_{jk}, \tau^m_{jk\ell})$, noting that each is only identified in relative terms: fixed effects give estimates only relative to the base case, and residuals are mechanically centered at zero. These relative estimates will suffice for counterfactuals.

Having recovered education costs τ^e_{jk}, I can estimate how they respond to *Inpres* school construction from 1973 to 1978. The endogeneity problem is that schools are not randomly assigned. In particular, construction targeted low-enrollment districts and thus is likely correlated with unobserved education costs. I again appeal to difference-in-differences as in Duflo (2001): for school construction S_j per 1,000 children and age group exposure T_k, I compare young and old cohorts in districts with high and low levels of school construction.

$$\log \tau^e_{jk} = \alpha_j + \alpha_k + \beta S_j T_k + \boldsymbol{C}_j T_k \boldsymbol{\phi} + \varepsilon_{jk} \tag{17}$$

The coefficient of interest is parameter β, which measures the impact of school construction on education costs. I match specification 3 with origin fixed effects α_j, age group fixed effects α_k, and controls $\boldsymbol{C}_j$ for child populations in 1971, child enrollment rates in 1971, and *Inpres* water and sanitation spending from 1973 to 1978. The fixed effects isolate the difference-in-differences comparison, while also absorbing the base case value for education costs. The difference with specification 3 is that estimation relies on aggregate and not individual data. I pool across survey years to minimize noise in computing aggregate education and wages, and so I cannot include survey year fixed effects α_t.

5.2 Wage rates and human capital

I estimate wage rates, the human capital parameter, the Fréchet parameter, and skill. I again use aggregate data by origin j, age group k, and destination ℓ to estimate a linear regression equation with fixed effects.

$$\log \bar{w}_{jk\ell} = \log \frac{\gamma \tilde{\epsilon}}{\bar{\epsilon}^{\eta}} + \log r_\ell + \eta \log \bar{e}_{jk} - \frac{1}{\theta} \log \bar{m}_{jk\ell} + \log s_{jk\ell} \tag{18}$$

On the left-hand side, I construct average log hourly wages from data. On the right-hand side, I estimate each term. The first term is a constant that combines scalar $\gamma = \Gamma(1-\frac{1}{\theta})$, which is pinned down by Fréchet parameter θ, and scalars $(\tilde{\epsilon}, \bar{\epsilon})$ as above. The second term is a fixed effect by destination ℓ, and it identifies wage rates r_ℓ. I note the distinction between wages $\bar{w}_{jk\ell}$ and wage rates r_ℓ: the former are wages per hour as observed in the data, while the latter are wages per hour per unit of human capital as estimated here. I do not observe wage rates because human capital includes unobserved components. The third term measures average years of schooling by origin and age group, and it identifies the elasticity η of human capital with respect to years of schooling. The fourth term measures migration choice probabilities by origin, age group, and destination, and it identifies Fréchet skill dispersion θ. I compute probabilities from data on individual migration decisions using a simple frequency estimator. The fifth term is the residual, and it has a structural interpretation as skill $s_{jk\ell}$. I thus obtain estimates $(\eta, \theta, r_\ell, s_{jk\ell})$. As before, $(r_\ell, s_{jk\ell})$ are identified only in relative terms, but these relative estimates will suffice for counterfactuals.

The endogeneity problem is that education and migration are not randomly assigned. In particular, by equations 9 and 11, unobserved skill $s_{jk\ell}$ is mechanically correlated with each through labor market access MA_{jk}. For education, this correlation echoes the typical concern over ability as an omitted variable. I again appeal to difference-in-differences and match specification 3.

$$\log \bar{e}_{jk} = \tilde{\alpha}_j + \tilde{\alpha}_k + \tilde{\beta} S_j T_k + \boldsymbol{C}_j T_k \tilde{\boldsymbol{\phi}} + \tilde{\varepsilon}_{jk} \tag{19}$$

As in Duflo (2001), difference-in-differences term $S_j T_k$ offers an instrument for education, conditional on origin fixed effects, age group fixed effects, and controls. I thus isolate the causal effect of education on wages in equation 18. For migration, I instrument with labor market access $MA_{j\ell}$ interacted with age group exposure T_k.

$$MA_{j\ell}^{ab} = \sum_{j'} \frac{\text{popden}_{j'}}{(\text{dist}_{jj'} + 1)^a} \cdot \mathbb{1}(\text{dist}_{jj'} > b) \cdot \mathbb{1}(\text{dist}_{j'\ell} > b)$$

I isolate faraway variation with quadratic $a = 2$ and cutoff $b = 50$ kilometers as baseline. I construct this measure directly from data as in section 3. I cannot use the model-based measure because it depends on Fréchet parameter θ, which I seek to estimate here. Like in section 3, I omit locations close to origin j. Unlike section 3, I also omit locations close to destination ℓ. Unobservables in each of these locations are likely to be correlated with skill $s_{jk\ell}$, which constitutes the error term.

5.3 Production

I calibrate the production parameter, which governs congestion. An attempt at estimation projects wage rates r_ℓ, as recovered above, onto total human capital H_ℓ. By equation 13,

$$\log r_\ell = \log \kappa + (\kappa - 1) \log H_\ell + \log A_\ell.$$

The identification argument is that the relationship between total human capital and wage rates is informative of congestion. For $\kappa < 1$, congestion depresses wage rates as human capital increases. For $\kappa = 1$, a lack of congestion implies no relationship between human capital and wage rates.

However, this approach is subject to the reflection problem and the identification challenges that accompany it (Manski 1993). Indeed, the reflection problem is typical when estimating group effects, and congestion amounts to a group effect. The problem is revealed by substituting into equation 18 and rewriting with total wages $W_\ell = r_\ell H_\ell$.

$$\log \bar{w}_{jk\ell} = \log \frac{\gamma \tilde{\epsilon} \kappa^{\frac{1}{\kappa}}}{\bar{\epsilon}^{\eta}} + \frac{\kappa - 1}{\kappa} \log W_\ell + \frac{1}{\kappa} \log A_\ell + \eta \log \bar{e}_{jk} - \frac{1}{\theta} \log \bar{m}_{jk\ell} + \log s_{jk\ell},$$

noting that total wages W_ℓ can be computed from data. Productivities A_ℓ are identified by destination fixed effects, but I do not need these productivities for counterfactuals. Instead, the parameter of interest is κ, which I must identify by regressing disaggregate outcome $\bar{w}_{jk\ell}$ on aggregate outcome W_ℓ.

Identifying variation must therefore "manipulate peer characteristics in a manner unrelated to individual characteristics" (Angrist 2014). In the specification above, I require variation that manipulates wages $\bar{w}_{jk\ell}$ for individuals who come from different origins and age groups, but who select into the same destination. *Inpres* variation is insufficient: it generates variation in education, but individuals still self-select into particular destinations. Peer group assignment is not random. Market access instruments are also insufficient: they generate variation in peer group assignment, but this variation is muted at the group level. Mechanically, for a given destination, negative shocks for some origins offset positive shocks for other origins. Angrist (2014) highlights this issue of weak instruments and the bias that follows. Zárate (2023) nicely summarizes the problem, particularly in online appendix B, and describes how stratified randomization of peers can overcome it – a high bar for my observational setting. I therefore turn to calibration, taking $\kappa = 0.925$ as baseline. This value corresponds to the central congestion scenario in Bryan and Morten (2019).

5.4 Estimates

I implement estimation sequentially. First, I start with frictions and amenities. I estimate equations 16 and 17 jointly, substituting 17 into 16 to obtain a single regression equation. I obtain estimates $(\beta, a_\ell, \tau^e_{jk}, \tau^m_{jk\ell})$. Second, I consider wage rates and human capital, estimating equations 18 and 19 jointly with two-stage least squares. I accommodate the controls needed for instrument validity in equation 19 by including these controls in equation 18, rewriting the latter with $\log s_{jk\ell} = \log \tilde{s}_{jk\ell} + \tilde{\alpha}_j + \tilde{\alpha}_k + \boldsymbol{C}_j T_k \tilde{\boldsymbol{\phi}}$. I obtain estimates $(\eta, \theta, r_\ell, s_{jk\ell})$. Third, I calibrate production parameter κ. Each step is independent, as none depends in any way on the results of

Table 5: Parameters

			Estimate	SE
β	*Inpres* effect	DD	-0.056**	(0.028)
η	Human capital	IV	0.224**	(0.106)
		OLS	0.142***	(0.049)
θ	Skill dispersion	IV	19.96**	(10.13)
		OLS	18.23**	(8.98)
κ	Production	–	0.925	–

Each panel is one step of estimation. *Inpres* effect β captures the extent to which *Inpres* school construction per 1,000 children reduces education costs τ^e_{jk}. I estimate it by difference-in-differences (DD). Human capital parameter η is the elasticity of human capital with respect to years of schooling, while Fréchet parameter θ captures skill dispersion. I estimate them both with and without instruments (IV and OLS). Production parameter κ is the elasticity of production with respect to total human capital in each location. I calibrate this parameter. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

the others. Proceeding sequentially is thus without loss.

Table 5 presents the common parameters. First, *Inpres* effect β captures the relationship between school construction and education costs. It is negative and statistically significant, suggesting that school construction decreases education costs – consistent with the positive *Inpres* effects on education and wages that I document in section 3. An additional school per 1,000 children reduces education costs by 6%, implying reductions of 12% for the average district. For counterfactuals, this parameter has the most immediate influence on the gains from school construction. If schools greatly reduce education costs, then school construction leads to large gains. Identification benefits from direct reliance on well-studied *Inpres* variation.

Second, human capital parameter η is the elasticity of human capital with respect to education. Wages are proportional to human capital, and so this parameter corresponds to returns to education as typically specified in a large body of empirical work. I compute and report the marginal effects of an additional year of schooling on log wages. My IV estimate of 0.22 is larger than my OLS estimate of 0.14, although omitted variable bias from unobserved ability would suggest otherwise. Instead, my estimates are consistent with measurement error in education that attenuates OLS estimates. The magnitudes I obtain are somewhat larger than the returns to education estimated in Duflo (2001).[16] This parameter plays a crucial role in counterfactuals, as higher returns to education imply greater gains from school construction. A large body of estimates dating back to Mincer (1958) allows me to benchmark this parameter or even to calibrate it. Psacharopoulos (2024) provides a recent survey.

Third, Fréchet parameter θ is roughly 20, with similar estimates between IV and OLS. This magnitude is similar to that estimated in Bryan and Morten (2019). This parameter is the only one for which identification relies on instrumenting with faraway market access. An alternative is to calibrate this parameter and move $\frac{1}{\theta}\bar{m}_{jk\ell}$ to the left-hand side in estimating equation 18. Estimation would then rely solely on the variation generated by *Inpres* school construction. For counterfactuals, this parameter enters when computing supply of human capital. Large skill dispersion increases the value of labor market access because individuals choose their destinations, and so they can select their best skill shocks rather than their average skill shocks. Market access then amplifies the positive impacts of school construction.

Fourth, production parameter κ is calibrated to $\kappa = 0.925$, capturing marginal returns that diminish at modest rate. I draw this value from the central congestion scenario of Bryan and Morten (2019), which also incorporates agglomeration and goods markets with calibrated values from the literature. These additional forces combine to generate congestion on net, and so I focus on congestion rather than attempting to separate each force empirically. For counterfactuals, this parameter enters when computing demand for human capital. Higher values of κ correspond to flatter demand curves. For $\kappa = 1$, the demand curve is perfectly flat because a lack of congestion allows firms to scale without experiencing productivity losses. Perfectly elastic demand then eliminates equilibrium adjustments in wage rates, even as school construction raises human capital levels nationally.

6 Counterfactuals

I quantify how mobility shapes the long-run aggregate and distributional effects of the *Inpres* program, and I highlight the equity-efficiency trade-off facing the poli-

[16] For the US, Hsieh et al. (2019) choose a lower value of 0.1 that corresponds to the fraction of output spent on human capital accumulation. They obtain this value by dividing education spending by the labor share of GDP.

cymaker. Appendix D details how I solve the model.

6.1 School construction

School construction reduces education costs. I characterize this effect by differencing equation 17. Changes in school construction S_j – from baseline S_j to counterfactual S_j' – induce changes in education costs τ_{jk}^e for treated age groups T_k.

$$\log \tau_{jk}^{e'} - \log \tau_{jk}^{e} = \beta(S_j' - S_j)T_k$$

for $\beta < 0$. I rewrite as changes $\hat{x} = \frac{x'}{x}$ for counterfactual x' and baseline x.

$$\hat{\tau}_{jk}^e = \exp\left[\beta(S_j' - S_j)T_k\right]$$

To reach this point, I only need to estimate parameter β. Other terms do not need to be estimated: counterfactual S_j' is as given by the counterfactual scenario of interest, and observed (S_j, T_k) are observed as data.

Education costs then affect production. I characterize this effect with "exact-hat" algebra as in Dekle et al. (2008), which allows me to express counterfactuals as changes relative to observed data. Changes in production are given simply by changes in wage rates and human capital by location.

$$\hat{Y}_\ell = \hat{r}_\ell \hat{H}_\ell$$

I solve for equilibrium responses in wage rates r_ℓ and human capital H_ℓ with equilibrium conditions 15, which state that human capital markets clear across space. Computation is lightweight because I solve using exact-hat expressions for human capital demand and supply. I characterize how each responds to any given changes $(\hat{\tau}_{jk}^e, \hat{\tau}_{jk\ell}^m)$ in education and migration costs, fixing common parameters $(\beta, \eta, \theta, \kappa)$, productivities A_ℓ, amenities a_ℓ, skill $s_{jk\ell}$, and shocks (ϵ, ε). To reach this point, I must estimate common parameters (κ, η, θ) and baseline values $(a_\ell, r_\ell, s_{jk\ell}, \tau_{jk\ell}^m)$. I do not need to estimate baseline values in levels, as any normalizations in levels cancel when I work in changes. I can therefore quantify how school construction affects production by location, with the flexibility to consider how migration costs influence this response.

I now define criteria for evaluation. Suppose a government allocates schools S across locations, building human capital to maximize a combination of aggregate output Y and distributional concerns (D_1, D_2). For non-negative weights λ, costs C, and budget constraint $\bar{C}$,

$$\max_S \ \lambda_0 Y(S) - \lambda_1 D_1(S) - \lambda_2 D_2(S) \quad \text{s.t.} \quad \lambda_0 + \lambda_1 + \lambda_2 = 1\,, \quad C(S) \leq \bar{C}\,. \tag{20}$$

Weights λ are a reduced form for the complex political motives that govern the balance between equity and efficiency. Production in each location follows from the exact-hat expression above. Aggregate output is production summed across locations, as in equation 12. Distributional concerns focus on rural-urban wage gaps, distinguishing between wage gaps across people and across places. For urbanness (U_j, U_ℓ) and ruralness (R_j, R_ℓ),

$$D_1 = \frac{1}{\kappa} \sum_{j,k,\ell} (U_j - R_j) N_{jk} \bar{m}_{jk\ell} \bar{w}_{jk\ell}, \quad D_2 = \frac{1}{\kappa} \sum_{\ell} (U_\ell - R_\ell) Y_\ell.$$

Wage gap D_1 across people of rural versus urban origin captures differences in opportunity for individuals, while wage gap D_2 across rural versus urban destinations captures regional disparities net of migration. Absent migration, $D_1 = D_2$.

6.2 Evaluating the program

Table 6 evaluates the aggregate and distributional effects of the program. Starting with total effects, the program increases aggregate output by 6% relative to zero construction. Students from rural districts experience the largest gains, as new schools bring greater benefits to people from less-educated rural districts relative to more-educated urban ones. In increasing the opportunities available to rural students, the program decreases inequality between people from rural and urban districts by 4%. That is, inequality across people falls as urban and rural students converge following nationwide school construction.

The government may also value convergence between rural and urban districts themselves, net of out-migration. Reducing inequality across places was an explicit motivation for targeting *Inpres* school construction to low-enrollment districts, and

Table 6: Aggregate and distributional effects

	Aggregate output	Inequality (people)	Inequality (places)
Total effects			
Zero construction	1.00	1.00	1.00
Actual *Inpres* allocation	1.06	0.96	1.09
Decomposition			
Zero construction	1.00	1.00	1.00
+ School construction	1.03	0.99	0.99
+ Migration	1.04	0.98	1.03
+ Migration-induced education	1.07	0.95	1.11
+ New equilibrium wage rates	1.06	0.96	1.09

Each row is one counterfactual. Values are ratios relative to zero construction. Aggregate output is given by aggregate wages. Person-based inequality is between people born in urban versus rural places. Place-based inequality is between people living in urban versus rural places. In the second panel, I start with *Inpres* school construction under infinite migration costs. Next, I lower migration costs to those estimated but hold education and wage rates fixed. I then allow education to adjust but continue to hold wage rates fixed. Finally, I allow wage rates to adjust. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

both equity and political economy considerations can rationalize such a policy goal.[17] I find that the program increases inequality between rural and urban places by 9%. Rural-to-urban migration fuels output gains by connecting rural human capital to high urban wages, but it does so at the expense of rural districts. The program remains a Pareto improvement relative to zero school construction because rural districts still benefit from modest output gains and higher human capital. But regional inequality rises because urban districts benefit even more.

Turning to the decomposition, mobility drives both aggregate and distributional effects. I isolate the direct impact of school construction by raising migration costs to infinity. In shutting mobility down, I find that aggregate output increases by only 3%. Low mobility leads to low labor market access and low benefits from investing in education, even when school construction lowers the costs of doing so. For inequality among people and places, the two measures coincide because there is no migration. Everyone born in a place stays in the place, and so people and places align. Moreover,

[17] Indonesia's transmigration program of the 1980s is another example of a policy aimed at developing "lagging" regions.

inequality on both counts is similar to inequality under zero construction. Although rural students have higher marginal returns to education, they are confined to low-wage labor markets and thus invest little in education. Urban students have access to high urban wages and thus larger incentives to invest in education, but they also face lower marginal returns given higher baseline levels of education. These offsetting forces mute inequality relative to the total effects.

Next, I allow for migration by reducing migration costs from infinity to their baseline estimated values. I do so in three phases to isolate three effects: matching, motives, and market. First, I allow individuals to realize their skill shocks and sort into their preferred destinations, but I do not allow education or wage rates to adjust. Individuals maintain the lower educational investment that they choose under infinite migration costs, and wage rates do not adjust in equilibrium. Improved matching increases the aggregate output effect by one percentage point (from 3% to 4%). Second, I allow education but not wage rates to adjust. Reduced migration costs increase labor market access and thus the motives to invest in education. Greater education increases the aggregate output effect by three percentage points (from 4% to 7%). Third, I allow wages to adjust. Greater education raises human capital, generating congestion that affects markets by causing wage rates to fall in equilibrium. Lower wages dampen the aggregate output effect by one percentage point (from 7% to 6%). Previous work has focused on the matching effect (Bryan et al. 2014), but I find that the motives effect dominates – even net of market effects. At the same time, output gains are driven by rural students migrating to high urban wages, which increases inequality across places.

6.3 Redesigning the program

I redesign the *Inpres* program by considering alternative allocations of school construction with the objective functions of equation 20, subject to the observed budget constraint. In particular, I search over allocations to maximize aggregate output ($\lambda_0 = 1$), person-based inequality ($\lambda_1 = 1$), place-based inequality ($\lambda_2 = 1$), and combinations of the three (for $\lambda_0 + \lambda_1 + \lambda_2 = 1$). I then compute each allocation's effect on output and inequality, characterizing both the policymaker's possibilities frontier and the implied equity-efficiency trade-off.

The challenge is that, for each objective function, the optimization problem is difficult to solve. Computing each optimal allocation requires solving a combinatorial problem. I must consider locations jointly – rather than one at a time – because migration generates spatial independence. That is, school construction in one district affects labor markets in all other districts, and so I cannot evaluate districts individually. The result is a severe curse of dimensionality. I therefore simplify the problem by focusing on allocation rules similar to the one used in reality. The actual rule allocated schools in proportion to 1971 child unenrollment in excess of a cutoff level, with a cutoff of 0% for 1973-1974 construction and 15% for 1975-1978.

The analysis proceeds as follows. First, I consider an objective function focused on aggregate output ($\lambda_0 = 1$). Second, I select the allocation rule that maximizes this objective. I do so by searching over a grid of allocation rule cutoff values. Optimization is greatly simplified because I search over the one-dimensional set of allocation rule cutoffs and not over the high-dimensional set of allocations themselves. Third, I evaluate each cutoff by computing the resulting allocation of school construction, then using the estimated model to compute effects on aggregate output, person-based inequality, and place-based inequality. High cutoffs concentrate school construction in low-enrollment districts, which tend to be rural and isolated. Fourth, I repeat these steps for alternative objective functions, which I obtain by varying the λ weights. The model thus captures the possibilities frontier facing the policymaker. It also generates policy prescriptions: for any given objective function, the model delivers the optimal cutoff rule and the resulting aggregate and distributional effects.

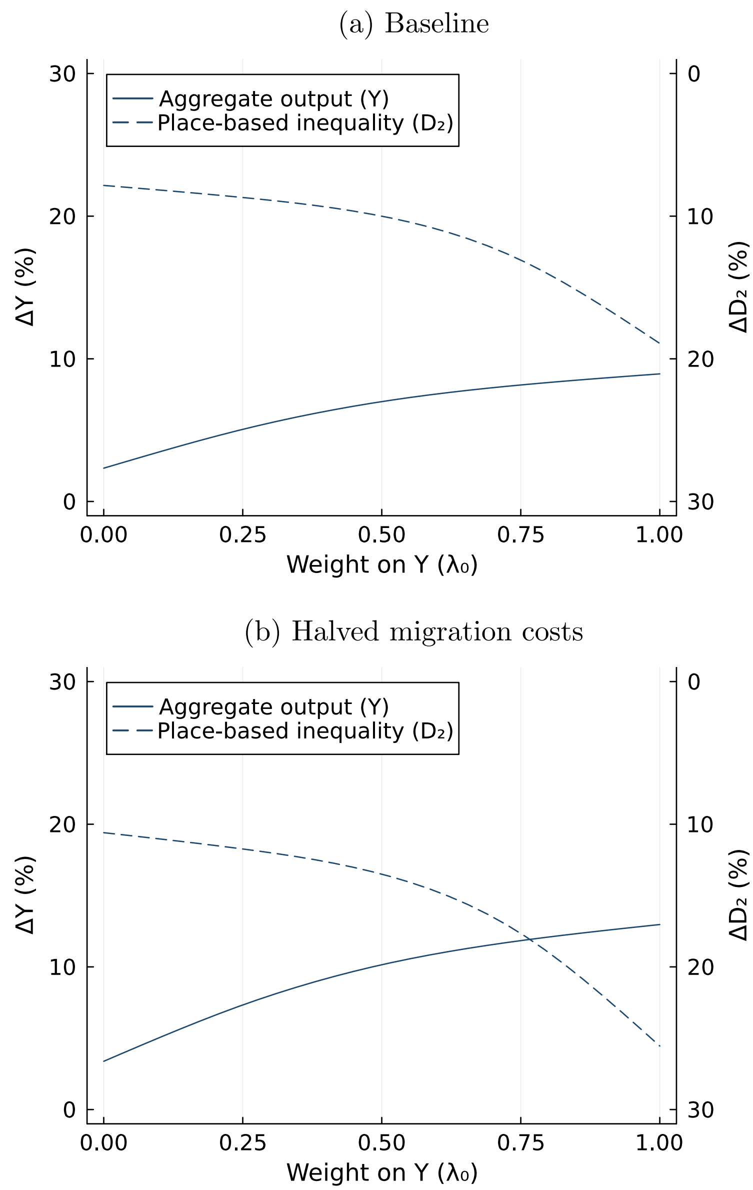
Every allocation I consider is subject to the observed budget constraint. I use total expenditures to define the budget, and indeed the *Bappenas* reports specify school construction costs by district. For 1973, these costs range from 2.5M IDR for non-urban districts in Sumatra, Java, Bali, and Kalimantan to 7M IDR for districts in Greater Jakarta. Given a cutoff level, for each district I compute 1971 child unenrollment in excess of this level, then I distribute the budget across districts in proportion to excess unenrollment. For example, if excess unenrollment is 10% in district one and 20% in district two, then district two receives twice as much funding as district one does. I treat this budget as fixed, although a public finance approach might consider feedback between production and local resources.

Figure 2 illustrates the results, plotting policymaker preferences alongside the resulting impacts on aggregate output and place-based inequality. For policymaker preferences, I increase weight λ_0 on aggregate output at the expense of weight $\lambda_2 = 1 - \lambda_0$ on place-based inequality, holding fixed weight $\lambda_1 = 0$ on person-based inequality. I focus on aggregate output and place-based inequality to capture the equity-efficiency trade-off, as raising aggregate output comes only at the cost of increased place-based inequality. Appendix figure D1 considers a policymaker that considers person- and place-based inequality, which invoke a similar trade-off. By contrast, there is no trade-off between aggregate output and person-based inequality. When school construction targets individuals from less-educated places with higher marginal returns, aggregate output is high and person-based inequality is low – positive outcomes on both counts.

For figure 2a, I begin on the right. When the weight on aggregate output is high ($\lambda_0 = 1$ on the x-axis), the policymaker concentrates construction in districts with high unenrollment, achieving high aggregate output and high place-based inequality. The policymaker does so with an allocation rule that uses a high unenrollment cutoff to avoid allocating schools to districts with low unenrollment. Moving leftward on figure 2a, the opposite holds when the weight on aggregate output is low ($\lambda_0 = 0$ on the x-axis). The policymaker concentrates construction in districts with low unenrollment, achieving low aggregate output and low place-based inequality. I highlight the trade-off by reversing the axis for place-based inequality, which enters the objective function negatively. Moving from right to left, one line rises while the other line falls. The reason is that targeting low-enrollment regions raises aggregate output gains by enhancing opportunities for underserved students. But these students access opportunities precisely by leaving for urban centers, thereby worsening regional inequality. Furthermore, figure 2a suggests a government objective function with approximately equal weights on aggregate output and place-based inequality, as $\lambda_0 = 0.5$ roughly corresponds to the 6% and 9% effects produced by the actual allocation.

In figure 2b, I repeat the analysis under reduced migration costs, which increase mobility and magnify the equity-efficiency tradeoff. Given the interaction between migration and education costs, reduced migration costs – perhaps through road investments – greatly amplify the output gains from school construction. At the same time, place-based inequality also rises. But unlike the baseline scenario, in which rural regions experience small but nonetheless positive gains, these counterfactuals involve meaningful losses to rural regions. The reason is that lower migration costs increase rural-to-urban migration, which drains rural populations. Thus, although co-

Figure 2: Aggregate output vs. place-based inequality



I vary the objective function holding fixed weight $\lambda_1 = 0$ on person-based inequality D_1. I thus vary weight $\lambda_0 \in [0, 1]$ on aggregate output Y, which in turn affects weight $\lambda_2 = 1 - \lambda_0$ on place-based inequality D_2. For both y-axes, higher is better. The left axes are percentage increases in Y relative to zero construction, with Y entering the objective function positively. The right axes are percentage increases in D_2 relative to zero construction, with D_2 entering the objective function negatively and thus flipped axes in the figures. The bottom figure repeats the exercise of the top figure under 50% lower migration costs. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

ordinated investment in schools and roads is substantially more effective than school construction alone, it is also no longer Pareto-improving relative to zero construction.

7 Conclusion

Spatial effects are crucial for evaluating large-scale educational interventions because graduates migrate for employment. Mobility amplifies the returns to education, increasing output but draining rural regions. I capture these forces with a spatial equilibrium model, and I use the model to illustrate how mobility shapes the aggregate and distributional effects of Indonesia's *Sekolah Dasar Inpres* program, which constructed 62,000 primary schools in the mid-1970s. I find that the program increased long-run aggregate output by 6%, and that mobility accounts for half of this effect. The importance of mobility suggests gains from coordinated investment in education and transportation, although such investment can also widen regional inequality.

Several lines of inquiry are left for future work. First, I highlight regional equity concerns that stem from spillover impacts of local investment on non-local outcomes. The political economy of these linkages has important implications, particularly when investment is chosen and funded locally. Second, I assume that school construction lowers education costs by increasing physical access, but the effects of new schools might also depend on factors like school quality and interactions with existing schools. Third, public school construction may prompt equilibrium responses by the private sector that affect aggregate outcomes. Such work informs policymakers' ongoing efforts to invest in education, which remains fundamental to economic development.

References

Adukia, Anjali, Sam Asher, and Paul Novosad. Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction. *American Economic Journal: Applied Economics*, 12(1):348–376, 2020.

Agostinelli, Francesco, Margaux Luflade, and Paolo Martellini. On the Spatial Determinants of Educational Access. 2024.

Akresh, Richard, Daniel Halim, and Marieke Kleemans. Long-Term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia. *The Economic Journal*, 133(650):582–612, 2023.

Alder, Simon. Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development. 2023.

Allen, Treb and Costas Arkolakis. Trade and the Topography of the Spatial Economy. *Quarterly Journal of Economics*, 129(3):1085–1140, 2014.

Allen, Treb and Costas Arkolakis. The Welfare Effects of Transportation Infrastructure Improvements. *Review of Economic Studies*, 2022.

Angrist, Joshua. The Perils of Peer Effects. *Labour Economics*, 30:98–108, 2014.

Ashraf, Nava, Natalie Bau, Nathan Nunn, and Alessandra Voena. Bride Price and Female Education. *Journal of Political Economy*, 128(2):591–641, 2020.

Atkin, David. Endogenous Skill Acquisition and Export Manufacturing in Mexico. *American Economic Review*, 106(8):2046–2085, 2016.

Austin, Benjamin, Edward Glaeser, and Lawrence Summers. Jobs for the Heartland: Place-Based Policies in 21st-Century America. *Brookings Papers on Economic Activity*, 1:151–255, 2018.

Balboni, Clare. In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities. 2024.

Balboni, Clare, Gharad Bryan, Melanie Morten, and Bilal Siddiqi. Transportation, Gentrification, and Urban Mobility: The Inequality Effects of Place-Based Policies. 2020.

Bau, Natalie, Martin Rotemberg, Manisha Shah, and Bryce Millett Steinberg. Human Capital Investment in the Presence of Child Labor. 2021.

Bazzi, Samuel, Masyhur Hilmy, and Benjamin Marx. Religion, Education, and the State. 2023.

Bryan, Gharad and Melanie Morten. The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):2229–2268, 2019.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. Under-investment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica*, 82(5):1671–1748, 2014.

Chetty, Raj and Nathaniel Hendren. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *Quarterly Journal of Economics*, 133(3):1107–1162, 2018.

Chetty, Raj, Nathaniel Hendren, and Lawrence Katz. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):855–902, 2016.

Chyn, Eric. Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children. *American Economic Review*, 108(10):3028–3056, 2018.

Chyn, Eric and Lawrence Katz. Neighborhoods Matter: Assessing the Evidence for Place Effects. *Journal of Economic Perspectives*, 35(4):197–222, 2021.

Dahl, Gordon. Mobility and the Return to Education: Testing a Roy Model with Multiple Markets. *Econometrica*, 70(6):2367–2420, 2002.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum. Global Rebalancing with Gravity: Measuring the Burden of Adjustment. *IMF Staff Papers*, 55(3):511–540, 2008.

Dinerstein, Michael, Daniel Morales, Christopher Neilson, and Sebastián Otero. Equilibrium Effects of Public Provision in Education Markets. 2023.

Donaldson, Dave. Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, 108(4-5):899–934, 2018.

Donaldson, Dave and Richard Hornbeck. Railroads and American Economic Growth: A "Market Access" Approach. *Quarterly Journal of Economics*, 131(2):799–858, 2016.

Duflo, Esther. Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *American Economic Review*, 91(4):795–813, 2001.

Duflo, Esther. The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia. *Journal of Development Economics*, 74:163–197, 2004.

Eaton, Jonathan and Samuel Kortum. Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779, 2002.

Eckert, Fabian and Tatjana Kleineberg. The Geography of Opportunity: Education, Work, and Intergenerational Mobility Across US Counties. 2024.

Faber, Benjamin. Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, 81(3):1046–1070, 2014.

Fajgelbaum, Pablo and Stephen Redding. Trade, Structural Transformation, and Development: Evidence from Argentina 1869-1914. *Journal of Political Economy*, 130(5):1249–1318, 2022.

Fajgelbaum, Pablo and Edouard Schaal. Optimal Transport Networks in Spatial Equilibrium. *Econometrica*, 88(4):1411–1452, 2020.

Fujimoto, Junichi, David Lagakos, and Mitchell VanVuren. Macroeconomic Effects of "Free" Secondary Schooling in the Developing World. 2024.

Gertler, Paul, Marco Gonzalez-Navarro, Tadeja Gračner, and Alexander Rothenberg. Road Maintenance and Local Economic Development: Evidence from Indonesia's Highways. *Journal of Urban Economics*, 143:103687, 2024.

Glaeser, Edward and Joshua Gottlieb. The Economics of Place-Making Policies. *Brookings Papers on Economic Activity*, 39(1):155–253, 2008.

Gonzalez-Navarro, Marco and Matthew Turner. Subways and Urban Growth: Evidence from Earth. *Journal of Urban Economics*, 108:85–106, 2018.

Graff, Tilman. Spatial Inefficiencies in Africa's Trade Network. *Journal of Development Economics*, 171:103319, 2024.

Heblich, Stephan, Stephen Redding, and Daniel Sturm. The Making of the Modern Metropolis: Evidence from London. *Quarterly Journal of Economics*, 135(4):2059–2133, 2020.

Heckman, James. Shadow Prices, Market Wages, and Labor Supply. *Econometrica*, 42(4):679–694, 1974.

Heckman, James and Guilherme Sedlacek. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93(6):1077–1125, 1985.

Hornbeck, Richard and Martin Rotemberg. Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies. 2024.

Hsiao, Allan. Democratization and Infrastructure Investment: Evidence from Healthcare in Indonesia. 2022.

Hsiao, Allan. Sea Level Rise and Urban Adaptation in Jakarta. 2023.

Hsiao, Allan. Sea Level Rise and Urban Inequality. *AEA Papers and Proceedings*, 114:47–51, 2024.

Hsieh, Chang-Tai, Erik Hurst, Charles Jones, and Peter Klenow. The Allocation of Talent and U.S. Economic Growth. *Econometrica*, 87(5):1439–1474, 2019.

Jedwab, Rémi and Adam Storeygard. The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010. *Journal of the European Economic Association*, 20(1):1–38, 2022.

Keane, Michael and Kenneth Wolpin. The Career Decisions of Young Men. *Journal of Political Economy*, 105(3):473–522, 1997.

Kennan, John and James Walker. The Effect of Expected Income on Individual Migration Decisions. *Econometrica*, 79(1):211–251, 2011.

Khanna, Gaurav. Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India. *Journal of Political Economy*, 131(2), 2023.

Laliberté, Jean-William. Long-Term Contextual Effects in Education: Schools and Neighborhoods. *American Economic Journal: Economic Policy*, 13(2):336–377, 2021.

Manski, Charles. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*, 60(3):531–542, 1993.

Martinez-Bravo, Monica. The Local Political Economy Effects of School Construction in Indonesia. *American Economic Journal: Applied Economics*, 9(2):256–289, 2017.

Mazumder, Bhashkar, Maria Rosales-Rueda, and Margaret Triyana. Intergenerational Human Capital Spillovers: Indonesia's School Construction and Its Effects on the Next Generation. *AEA Papers and Proceedings*, 109:243–249, 2019.

Milsom, Luke. Moving Opportunity: Local Connectivity and Spatial Inequality. 2023.

Mincer, Jacob. Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66(4):281–302, 1958.

Moneke, Niclas. Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia. 2024.

Moretti, Enrico. Local Labor Markets. *Handbook of Labor Economics*, 4b:1237–1313, 2011.

Morten, Melanie and Jaqueline Oliveira. The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City. 2023.

Nakamura, Emi and Jón Steinsson. Identification in Macroeconomics. *Journal of Economic Perspectives*, 32(3):59–86, 2018.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson. The Gift of Moving: Intergenerational Consequences of a Mobility Shock. *Review of Economic Studies*, 2021.

Neumark, David and Helen Simpson. Place-Based Policies. *Handbook of Regional and Urban Economics*, 5:1197–1287, 2015.

Psacharopoulos, George. Returns to Education: A Brief History and an Assessment. *Education Economics*, 32(5):561–565, 2024.

Redding, Stephen and Esteban Rossi-Hansberg. Quantitative Spatial Economics. *Annual Review of Economics*, 9:21–58, 2017.

Redding, Stephen and Matthew Turner. Transportation Costs and the Spatial Organization of Economic Activity. *Handbook of Regional and Urban Economics*, 5: 1339–1398, 2015.

Rojas-Ampuero, Fernanda and Felipe Carrera. Sent Away: The Long-Term Effects of Slum Clearance on Children and Families. 2021.

Roy, Andrew. Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2):135–146, 1951.

Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Sarkar, Matthew Sobek, Derek Burk, Dan Ehrlich, Quinn Heimann, and Jane Lee. Integrated Public Use Microdata Series, International: Version 7.5, 2024. URL https://international.ipums.org.

Severen, Christopher. Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification. *Review of Economics and Statistics*, 105(5): 1073–1091, 2023.

Shah, Manisha and Bryce Millett Steinberg. Drought of Opportunities: Contemporaneous and Long-Term Impacts of Rainfall Shocks on Human Capital. *Journal of Political Economy*, 125(2):527–261, 2017.

Shah, Manisha and Bryce Millett Steinberg. Workfare and Human Capital Investment: Evidence from India. *Journal of Human Resources*, 56(2):380–405, 2021.

Tsivanidis, Nick. Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio. 2023.

World Bank. World Bank Open Data, 2022. URL https://data.worldbank.org.

Yang, Yang. Transport Infrastructure, City Productivity Growth, and Sectoral Reallocation: Evidence from China. 2018.

Young, Alwyn. Inequality, the Urban-Rural Gap, and Migration. *Quarterly Journal of Economics*, 128(4):1727–1785, 2013.

Zárate, Román Andrés. Uncovering Peer Effects in Social and Academic Skills. *American Economic Journal: Applied Economics*, 15(3):35–79, 2023.

Zárate, Román David. Spatial Misallocation, Informality and Transit Improvements: Evidence from Mexico City. 2024.

Appendix

A Data

I describe the data and data construction. Table A1 lists data sources.

A.1 Treatment data

Data on the *Inpres* program and district demographics come from Duflo (2001), which draws on Ministry of National Development Planning (*Bappenas*) reports from 1973 to 1978 and population census reports from 1971. For each district, the data record the planned number of *Inpres* schools by year from 1973 to 1978, *Inpres* water and sanitation spending from 1973 to 1978, total and child populations in 1971, and child enrollment rates in 1971. From the original dataset, I drop duplicate and missing observations to obtain measures for 282 unique districts. Districts include rural regencies (*kabupaten*) and urban municipalities (*kota*). I verify the data and make adjustments as follows.

I consult the original *Bappenas* reports to verify the school construction data, and I correct a modest number of data entry errors and inconsistencies: 1 for 1973, 4 for 1974, 12 for 1975, 5 for 1976, 3 for 1977, and 2 for 1978. Among these changes, the corrected values are on average 85% larger or smaller than their original values. In the corrected data, construction in 1976 and 1978 matches that in 1975 and 1977, respectively, as stipulated by the presidential instructions.[18] I take water and sanitation spending to be as constructed by Duflo (2001), noting that this spending seems to be proportional to school construction in the *Bappenas* reports, and I fill six missing observations to avoid dropping these districts.[19] For total and child populations, I again consult the original census reports and make a small number of corrections: 29 for total population and 16 for child population. The average magnitudes of these

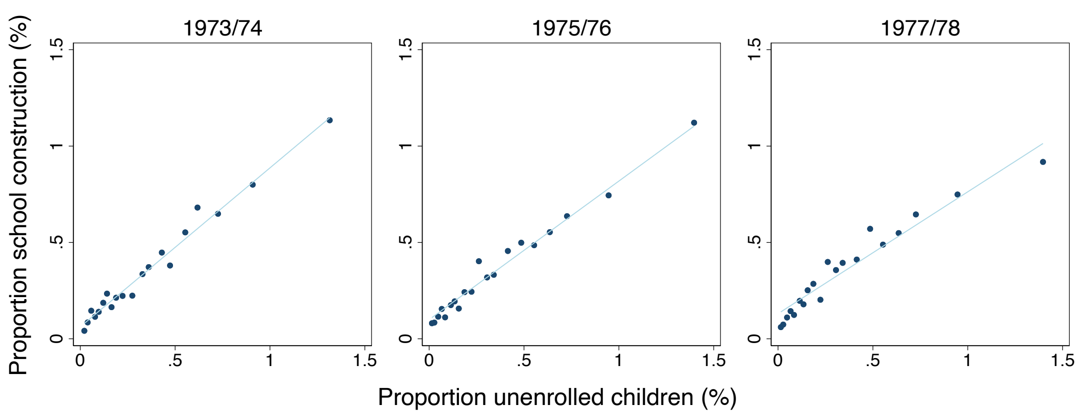
[18] In reference to *Inpres* No. 6/1975, *Inpres* No. 3/1976 states: "*Pembagian jumlah gedung Sekolah Dasar [...] adalah sama dengan jumlah gedung Sekolah Dasar tahap pertama yang dibangun [...] berdasarkan Instruksi Presiden Nomor 6 Tahun 1975.*" *Inpres* No. 6/1978 contains analogous text in reference to *Inpres* No. 3/1977.

[19] I impute one missing value for Southeast Aceh with the average value among other rural districts in the province of Aceh. I fill five missing values for the Special Capital Region of Jakarta with zero, noting that the main results are robust to excluding Jakarta (table B8).

Table A1: Data sources

Period	Format	Source	Description
Treatment data			
1973-1978	Reports	Ministry of National Development Planning (*Bappenas*), Duflo (2001)	*Inpres* school construction and water and sanitation spending
1971	Reports	Population Census (*SP*), Duflo (2001)	Total and child populations
1971	Microdata	Population Census (*SP*), Ruggles et al. (2024)	Child enrollment
2011	Shapefiles	Global Administrative Areas Database (GADM)	Administrative boundaries (level 2)
1993-2014	Crosswalk	World Bank	Redistricting over time
Outcome data			
1976	Microdata	Intercensal Population Survey (*Supas*), Ruggles et al. (2024)	Education, wages, employment, and migration
2011-2014	Microdata	National Socioeconomic Survey (*Susenas*)	Education, wages, employment, and migration

Figure A1: Program allocation



Each figure is a binned scatter plot, and each observation is one district. The y-axis is the proportion of total school construction allocated to a district. The x-axis for 1973/1974 is unenrollment rate in 1971 among children of primary school age. The x-axis for other years is how much the rate exceeds 15%. Data: *Bappenas* reports (1973-78), census reports (1971).

changes are 65% and 38% of the original values. Child populations are defined to include individuals of age 5 to 14. With *Inpres* school construction from 1973 to 1978 and child populations for 1971, I recompute the main treatment variable as the total number of *Inpres* schools constructed per 1,000 children.

I construct child enrollment rates with census microdata from 1971, drawing on a 0.54% sample from IPUMS International (Ruggles et al. 2024). These microdata record whether respondents are currently in school, and I compute enrollment rates among individuals of age 5 to 14 to match those captured by the child population data. I do so for the 275 districts I observe in the IPUMS data. Redistricting between 1971 and 1973 explains the difference from the 282 districts of the school construction data. I fill the seven missing observations for (offshoot) child districts by applying the enrollment values of the corresponding (original) parent districts. I obtain rather different measures of child enrollment, and the average magnitude of the changes is 221% of the original values.

Table A2 summarizes the data adjustments relative to the original data from Duflo (2001). The baseline data are balanced with a slightly smaller number of districts. *Inpres* school construction and populations remain similar on average. I obtain higher child enrollment rates, as the Duflo (2001) data seem to define child enrollment rates based on school attendance for individuals of age 5 and above, rather than individuals of age 5 to 14. I obtain similar magnitudes when I construct enrollment rates based on the former sample of individuals. At the same time, I note that child enrollment enters the analysis only as a control variable. I also note that the regressions in Duflo (2001) may not rely on this particular measure, as the data also record the number of children attending school.

I also construct a measure of labor market access that captures proximity to cities. I use district boundary shapefiles from the Global Administrative Area Database (GADM) to compute land area by district, as well as district centroids. Because of redistricting over time, some district polygons must be merged to obtain the district boundaries of the treatment data. Merged land areas sum over individual land areas, and merged centroids are area-weighted averages of individual centroids. I calculate population densities from total populations and land areas, and I compute Euclidean distances between centroids for all pairs of districts. For each district, I define market access as an inverse distance weighted sum of own and neighboring population den-

Table A2: Treatment data adjustments

	Baseline			Duflo (2001)		
	Mean	SD	N	Mean	SD	N
Inpres schools per 1000 children	2.26	1.09	282	2.34	1.26	287
Inpres school construction	219	174	282	221	172	293
1973 construction	21.2	20.0	282	21.4	20.1	286
1974 construction	21.2	18.5	282	21.4	18.5	286
1975 construction	35.4	29.8	282	35.5	29.7	286
1976 construction	35.4	29.8	282	35.5	29.7	286
1977 construction	53.1	44.5	282	53.1	44.2	285
1978 construction	53.1	44.5	282	53.2	44.3	285
Population (1000s)	422	362	282	420	365	287
Child population (1000s)	114	97.9	282	115	97.9	287
Child enrollment rates	0.51	0.14	282	0.18	0.10	287

Each observation is one district. Columns show means, standard deviations, and the number of observations in the baseline data and the Duflo (2001) data. Rows summarize the number of *Inpres* schools constructed from 1973 to 1978 and population and enrollment in 1971. Children are those of age 5 to 14. Data: *Bappenas* reports (1973-78), census reports (1971).

sities. My baseline measure uses quadratic inverse distance weighting, with distances measured in hundreds of kilometers.

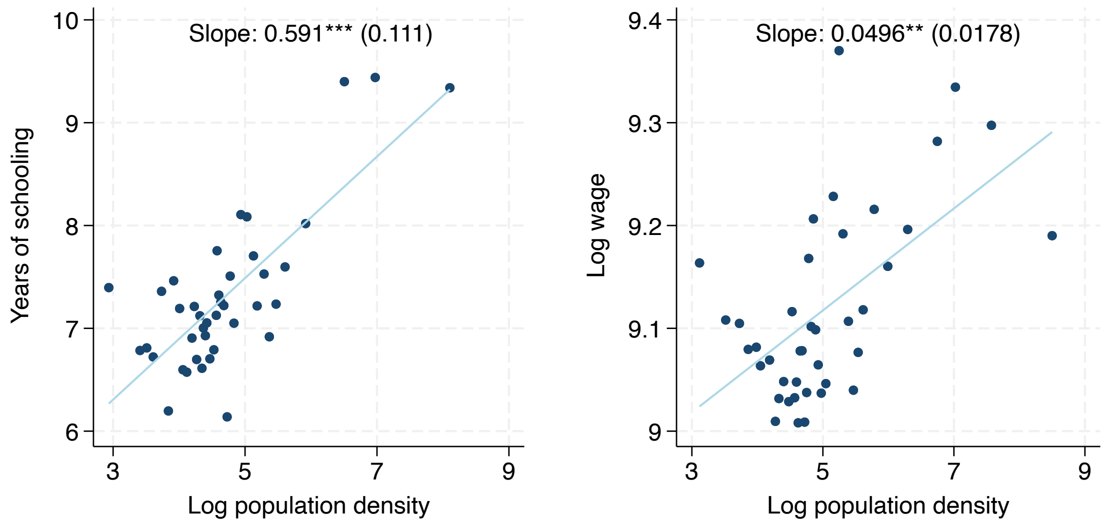
Figure A2 motivates my focus on proximity to cities. With microdata on education and wages from 2011 to 2014, which I describe in the following section, I estimate the specification

$$Y_{ijt} = \alpha_P + \alpha_t + \beta \log \text{popden}_j + \varepsilon_{ijt}$$

for individuals i living in districts j of provinces $P(j)$, as measured in survey year t. I regress education and wages in 2011 to 2014 on log population densities in 1971, controlling for province and survey year fixed effects and clustering standard errors by province. Figure A2 plots the data as binned scatterplots, and it reports positive and significant coefficients β. Education and wages are higher in cities. This pattern is consistent with the idea that cities offer higher returns to human capital, which may impact the treatment effects of *Inpres* school construction. Population densities may be endogenous because of omitted variable bias, and so the main analysis excludes own-district population density and relies instead on faraway variation in labor market access. My baseline measure defines "faraway" as farther than 50 kilometers.

Figure A2: Population density

(a) Education (b) Wages



Each figure is one binned scatterplot, and each observation is one male individual of age 39 to 64. Education is years of schooling completed, and wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. Education and wages are as measured in 2011 to 2014, and population density is as measured in 1971. I control for province and survey-year fixed effects, and I cluster standard errors by province. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Susenas* surveys (2011-14), census reports (1971).

A.2 Outcome data

Data on individual outcomes come from the *Supas* and *Susenas* socioeconomic surveys. The *Supas* data are for the survey year 1976, and the *Susenas* data are for 2011, 2012, 2013, and 2014. I rely primarily on the *Susenas* data to study the long-run impacts of *Inpres* school construction, but I also estimate the model on the *Supas* data to establish baseline parameter values for the pre-*Inpres* period.

I observe demographics, education, and wages. Demographics include gender, age, district of birth, and district of residence. Following Duflo (2001), I use gender and age to restrict attention to male individuals of age 39 to 64. For the *Susenas* data, this sample corresponds to those of age 2 to 24 when the first *Inpres* schools were completed in 1974. I use district of birth to link individuals to *Inpres* school construction, assuming that individuals pursue primary schooling where they are born. I use district of residence to construct migration outcomes, with migrants defined as those whose districts of birth and residence do not coincide. For migrants, I observe whether they migrate across provinces, whether they migrate to urban districts. I compute Euclidean migration distances based on district centroids.

I encounter the typical challenge of redistricting over time, as decentralization has prompted many cases of larger districts that split into smaller, independent districts. I merge districts as coded in the *Supas* and *Susenas* data to match those coded in the *Inpres* school construction data. I do so with a district proliferation crosswalk from the World Bank, combined with painful manual inspection. I obtain a consistent set of boundaries over time, covering 26 provinces and 282 districts.

For education, I observe the highest level of education attended and either completion or the number of years attended at that level. I pool vocational, religious, and traditional schooling for middle school and high school. At the tertiary level, I distinguish between community colleges that offer vocational associate's degrees (*Diploma*) and universities that offer traditional bachelor's degrees (*Sarjana*). For the *Susenas* data, it is important to code zero years of schooling for those who report never having attended schooling in a preliminary question, otherwise these observations are treated as missing. I compute the years of schooling completed. Primary school is six years, middle school is three, high school is three, community college is two, and university is four. For those who attend but do not complete a given level of schooling, I subtract one from the number of years attended at that level.[20]

Tertiary schooling is a challenge. Three-year vocational programs exist (*D3*), but the *Susenas* data pool these programs with associate's degrees (*Sarjana Muda* in older parlance) that may not be three-year programs. The data also pool master's and doctoral programs (*S2* and S3), which typically require different periods of study. Attendance without completion is more common at this level, with recall bias serving to worsen measurement error. Furthermore, the *Inpres* program built primary schools, which in any case may have limited impacts on tertiary schooling outcomes. The baseline analysis takes twelve years of schooling as the maximum, sidestepping the potential measurement error from tertiary schooling. Figure B6 shows that the main results are robust to incorporating the variation from tertiary schooling.

For employment, I observe employment status, type, income, and hours. I code individuals as employed if they report working as their primary activity in the past

[20] For someone who reports attending three years of middle school, I record eight years of schooling completed rather than nine years of schooling attained.

Table A3: Income data

	Supas 1976			*Susenas* 2011-2014		
	Log income	Nonzero income	N	Log income	Nonzero income	N
Wage-employed	7.96	0.94	6,639	9.12	0.90	181,393
Self-employed	7.68	0.01	13,724	8.86	0.59	257,516

Each observation is one male individual of age 39 to 64. Columns show mean log income conditional on reporting nonzero income, the proportion of individuals who report nonzero income, and the number of observations. Income is log hourly net income in year-2011 Indonesian rupiah from an individual's main job. Rows summarize these statistics for the wage- and self-employed. Data: *Supas* survey (1976), *Susenas* surveys (2011-14).

week. Individuals also report whether they work as own-account workers, employers with temporary or permanent employees, temporary or permanent employees, or unpaid family workers.[21] I categorize employees as wage-employed and others as self-employed. I observe typical monthly income from an individual's main job, inclusive of both money and goods, and I adjust for inflation to obtain income in year-2011 dollars. I also observe the number of hours worked in the past week at an individual's main job, which I combine with monthly income to compute log hourly income.

A data limitation is that the income data are incomplete for self-employed individuals. Table A3 shows this incompleteness by tabulating the income data by employment type. For the wage-employed, I record nonzero incomes for 90% of individuals in the *Susenas* data and 94% in the *Supas* data. These data are nearly complete. But for the self-employed, I record nonzero incomes for only 59% of individuals in the *Susenas* data and 1% in the *Supas* data. I observe that reported incomes are lower among the self-employed, and I speculate that self-employed income is inherently more variable and prone to reporting bias. My analysis focuses on income for the wage-employed, and I refer to wages in place of income.

[21] In the *Supas* data, the categories are instead own-account workers, employers, employees, and unpaid family workers.

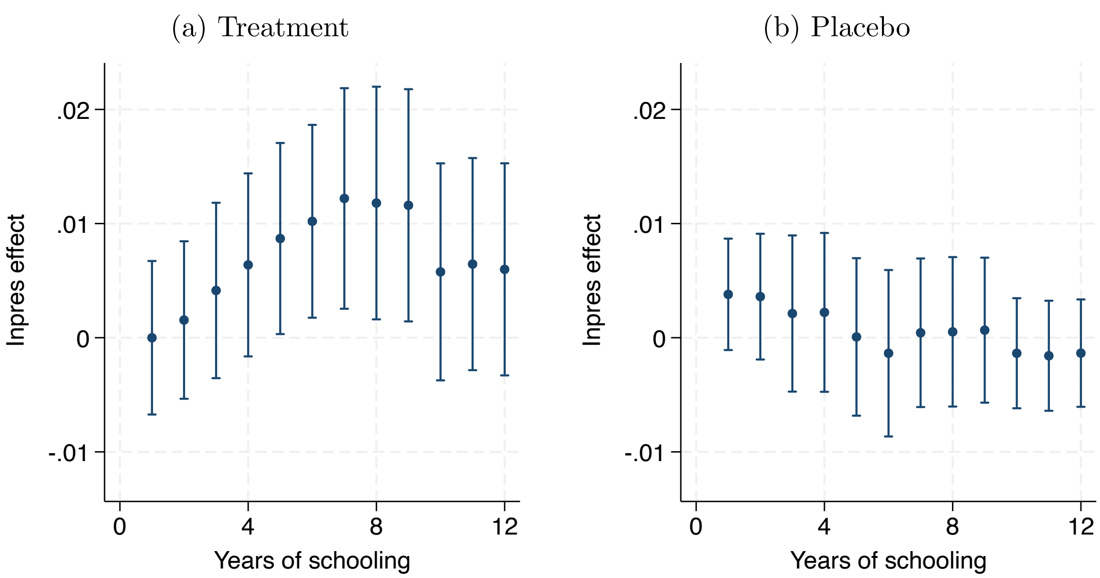
B Evaluation

I present additional results and robustness checks.

B.1 Additional results

I document *Inpres* effects on educational completion, employment, and migration outcomes. I also consider medium-run and place effects.

Figure B1: Educational completion



Each dot is one regression, and each observation is one male individual of age 2 to 24 in 1974. Each figure plots *Inpres* effect estimates with 95% confidence bands, evaluating completion of s or more years of schooling for $s \in \{1, 2, \ldots, 12\}$. Treatment regressions compare individuals of age 2 to 6 versus 12 to 17, and placebo regressions compare those of age 12 to 17 versus 18 to 24. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

Table B1: Employment

	Treatment				Placebo			
	Employed	Wage-employed	Self-employed	Weekly hours	Employed	Wage-employed	Self-employed	Weekly hours
Inpres effect	0.000631	0.000907	-0.000907	-0.151	0.00328	-0.000461	0.000461	0.0216
	(0.00142)	(0.00321)	(0.00321)	(0.106)	(0.00331)	(0.00453)	(0.00453)	(0.104)
Age group FE	x	x	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x	x	x
Observations	264,307	246,031	246,031	246,031	222,194	191,996	191,996	191,996
Control mean	0.910	0.403	0.597	42.35	0.808	0.275	0.725	40.13

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. Treatment regressions compare individuals of age 2 to 6 versus 12 to 17, and placebo regressions compare those of age 12 to 17 versus 18 to 24. The employed are those who report working as their primary activity. Conditional on employment, I observe if they are wage- or self-employed and their weekly hours worked. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

Table B2: Migration

	Treatment				Placebo			
	Migrant	Provincial migrant	Urban migrant	Distance	Migrant	Provincial migrant	Urban migrant	Distance
Inpres effect	0.00315	0.00772	-0.00343	-7.805	-0.00437*	-0.00401	-0.00376	0.0217
	(0.00316)	(0.00722)	(0.00564)	(7.784)	(0.00224)	(0.00588)	(0.00527)	(7.209)
Age group FE	x	x	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x	x	x
Observations	264,307	67,618	67,618	67,618	222,194	54,572	54,572	54,572
Control mean	0.252	0.622	0.612	563.8	0.238	0.640	0.581	583.9

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. Treatment regressions compare individuals of age 2 to 6 versus 12 to 17, and placebo regressions compare those of age 12 to 17 versus 18 to 24. Migrants are those whose districts of birth and residence do not coincide. Conditional on migration, I observe if they migrate across provinces, if they migrate to an urban destination, and their distance migrated. Distances are Euclidean and between district centroids. *Inpres* effects are coefficient estimates for difference-in-difference term S_jT_k, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

Table B3: Medium-run effects

	Treatment			Placebo		
	Years of schooling	Years of schooling	Log wages	Years of schooling	Years of schooling	Log wages
Inpres effect	0.137**	0.180***	0.0179*	-0.0112	0.0134	-0.00405
	(0.0530)	(0.0596)	(0.0107)	(0.0392)	(0.0560)	(0.00894)
Age group FE	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	77,890	30,581	30,581	77,849	29,928	29,928
Control mean	6.909	8.567	8.687	6.448	8.391	8.776

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. I replicate Duflo (2001) with 1995 data on medium-run outcomes. Treatment regressions compare individuals of age 2 to 6 versus 12 to 17, and placebo regressions compare those of age 12 to 17 versus 18 to 24. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Data: *Supas* survey (1995), *Bappenas* reports (1973-78), census reports (1971).

Table B4: Place effects

	People: birth district effects			Places: current district effects		
	Years of schooling	Years of schooling	Log wages	Years of schooling	Years of schooling	Log wages
Inpres effect	0.0847**	0.0994**	0.0256***	0.0554	0.0254	0.0146*
	(0.0356)	(0.0385)	(0.00943)	(0.0380)	(0.0428)	(0.00840)
Age group FE	x	x	x	x	x	x
District FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	264,307	98,413	98,413	264,307	98,413	98,413
Control mean	6.695	8.434	9.195	6.695	8.434	9.195

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. People regressions consider *Inpres* exposure by birth district as in table 3, while place regressions are by current district of residence. I compare individuals of age 2 to 6 versus 12 to 17. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district j and treatment dummy T_k for age group k. I control for age group, district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

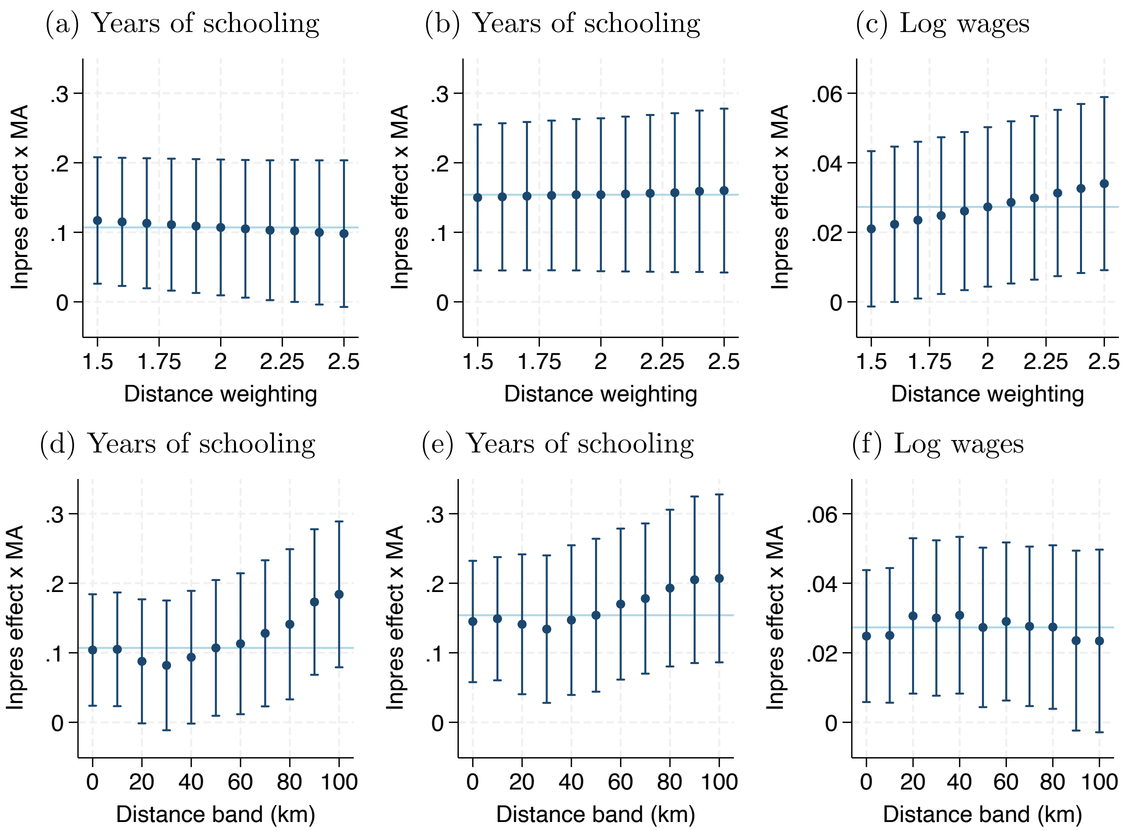
B.2 Robustness

There are several ways to construct labor market access and its instrument from the data. Both depend directly on the choices of distance weights and bands. Figure B5 shows that the results of table 4 are stable with respect to these choices. Each subfigure plots the market access interaction coefficient for the main education and wage outcomes. The top three subfigures vary the distance weight in the market access measure of equation 1. Relative to inverse quadratic weighting with $a = 2$ as baseline, I consider the range $\mathcal{A} = [1.5, 2.5]$. The bottom three subfigures vary the distance band in the market access instrument of equation 2. Relative to defining faraway variation with $b = 50$ kilometers as baseline, I consider the range $\mathcal{B} = [0, 100]$. In each case, I find the baseline estimates – marked with light blue horizontal lines – to lie well within the confidence bands of the alternative estimates. The next section will present a model and a particular expression for market access, but at this point I do not rely on the structure of this model.

I conduct several additional tests. Table B6 uses the uncensored measure of education, which introduces measurement error from tertiary education. The results are robust, although the non-interacted effects are somewhat attenuated. Table B7 includes spatial controls in the form of cubic polynomials of district coordinates. These controls flexibly absorb unobserved spatial variation that may be confounded with market access. I obtain similar estimates, although the absorbed spatial variation also weakens the strength of the instrument. Estimates remain similar with a lower distance band of 20 kilometers, which strengthens the instrument at cost to the plausibility of exclusion.[22] Table B8 excludes Jakarta and the greater metropolitan area (*Jabodetabek*), which form the largest labor market in Indonesia. The results are not driven by this subset of districts, where unobservables are perhaps more likely confounders.

[22] Donaldson and Hornbeck (2016) and Jedwab and Storeygard (2022) also include spatial controls. The former does not instrument with faraway market access – instead including nearby market access as a control – and thus does not encounter this issue. The latter does instrument and encounter the same issue. The authors emphasize cutoffs of 55 and 111 kilometers, which are comparable to my baseline distance band of 50 kilometers but larger than my lower distance band of 20 kilometers. The authors benefit from data covering most of the African continent, with total land area nearly 13 times that of Indonesia.

Table B5: Distance weights and bands



Each dot is one regression, and each observation is one male individual of age 2 to 24 in 1974. Each figure plots market access interaction estimates with 95% confidence bands. I vary distance weights a and distance bands b, as defined in equations 1 and 2, around the baseline values of $a = 2$ and $b = 50$. I compare individuals of age 2 to 6 versus 12 to 17. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. Market access is an inverse distance weighted sum of 1971 population densities across districts. I convert market access to a z-score and interpret units in standard deviations. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

Table B6: Uncensored education

	No MA		MA OLS		MA IV: 50 km	
	Years of schooling	Years of schooling	Years of schooling	Years of schooling	Years of schooling	Years of schooling
Inpres effect	0.0684*	0.0751*	0.110***	0.137***	0.117***	0.149***
	(0.0360)	(0.0438)	(0.0376)	(0.0445)	(0.0386)	(0.0455)
Inpres effect × *MA*			0.0916**	0.130***	0.107**	0.172***
			(0.0393)	(0.0490)	(0.0503)	(0.0602)
Age group FE	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	264,307	98,413	264,307	98,413	264,307	98,413
Control mean	7.009	9.203	7.009	9.203	7.009	9.203
F-statistic					162.7	99.52

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. I include tertiary education when computing years of schooling, rather than taking twelve years to be the maximum. MA regressions include an interaction with labor market access. OLS regressions do not instrument, while IV regressions instrument with faraway variation that excludes districts within 50 km. I compare individuals of age 2 to 6 versus 12 to 17. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. *Inpres* effects are coefficient estimates for difference-in-difference term S_jT_k, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. Market access is an inverse distance weighted sum of 1971 population densities across districts. I convert market access to a z-score and interpret units in standard deviations. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

Table B7: Spatial controls

	IV: 20 km			IV: 50 km		
	Years of schooling	Years of schooling	Log wages	Years of schooling	Years of schooling	Log wages
Inpres effect	0.0912**	0.0928*	0.0283***	0.0735*	0.0634	0.0401***
	(0.0405)	(0.0475)	(0.0104)	(0.0444)	(0.0629)	(0.0129)
Inpres effect $\times$ *MA*	0.0816*	0.153***	0.0378***	0.101**	0.169***	0.0375***
	(0.0465)	(0.0546)	(0.0115)	(0.0503)	(0.0596)	(0.0118)
Age group FE	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	264,307	98,413	98,413	264,307	98,413	98,413
Control mean	6.695	8.434	9.195	6.695	8.434	9.195
F-statistic	119.6	54.09	54.09	17.59	6.264	6.264

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. I include spatial controls in the form of cubic polynomials of district centroid coordinates. IV regressions instrument for labor market access with faraway variation, excluding districts within 20 and 50 km. I compare individuals of age 2 to 6 versus 12 to 17. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. Market access is an inverse distance weighted sum of 1971 population densities across districts. I convert market access to a z-score and interpret units in standard deviations. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

Table B8: Largest markets

	Excluding Jakarta			Excluding *Jabodetabek*		
	Years of schooling	Years of schooling	Log wages	Years of schooling	Years of schooling	Log wages
Inpres effect	0.130***	0.170***	0.0421***	0.133***	0.173***	0.0419***
	(0.0395)	(0.0425)	(0.00995)	(0.0396)	(0.0426)	(0.00998)
Inpres effect × *MA*	0.106**	0.156***	0.0278**	0.108**	0.157***	0.0281**
	(0.0507)	(0.0578)	(0.0121)	(0.0511)	(0.0579)	(0.0121)
Age group FE	x	x	x	x	x	x
Birth district FE	x	x	x	x	x	x
Survey year FE	x	x	x	x	x	x
Observations	260,321	96,050	96,050	256,400	94,024	94,024
Control mean	6.656	8.398	9.188	6.662	8.419	9.191
F-statistic	290.2	246.1	246.1	559.8	519.6	519.6

Each column is one regression, and each observation is one male individual of age 2 to 24 in 1974. Jakarta regressions exclude individuals born in the five districts of the Special Capital Region of Jakarta. *Jabodetabek* regressions additionally exclude the greater metropolitan area, covering modern-day Bogor, Depok, Tangerang, and Bekasi. I compare individuals of age 2 to 6 versus 12 to 17. Schooling is years completed, both unconditionally and conditional on observing nonzero wages. Wages are log hourly net wages in year-2011 Indonesian rupiah from an individual's main job. *Inpres* effects are coefficient estimates for difference-in-difference term $S_j T_k$, given program intensity S_j for district of birth j and treatment dummy T_k for age group k. Market access is an inverse distance weighted sum of 1971 population densities across districts. I convert market access to a z-score and interpret units in standard deviations. I control for age group, birth district, and survey year fixed effects; child populations and enrollment rates in 1971; and *Inpres* water and sanitation spending from 1973 to 1978. I cluster standard errors by birth district based on 1973 boundaries. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).

C Model and Estimation

I derive expressions for migration, education, wages, production, and human capital. I also derive the key estimating equations. I then extend the model to include endogenous goods prices.

C.1 Preliminaries

I collect several expressions before proceeding to derivations. First, recall the definitions of labor market access and location utility.

$$MA_{jk} = \sum_{\ell} \tilde{v}_{jk\ell}^{\theta}, \quad \tilde{v}_{jk\ell} = \frac{a_\ell r_\ell s_{jk\ell}}{\tau_{jk\ell}^{m}} \tag{21}$$

Second, I define the following expected values of schooling shock ϵ.

$$\bar{\epsilon} = \mathbb{E}[\epsilon^{-\frac{1}{1-\eta}}], \quad \tilde{\epsilon} = \mathbb{E}[\epsilon^{-\frac{\eta}{1-\eta}}] \tag{22}$$

for human capital elasticity η. Each is a constant, as schooling shocks are IID. Third, conditional on choosing destination ℓ, the expected value of skill shock ε_ℓ is

$$\mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell] = \gamma \bar{m}_{jk\ell}^{-\frac{1}{\theta}} \tag{23}$$

for migration probability $\bar{m}_{jk\ell}$, scalar $\gamma = \Gamma(1 - \frac{1}{\theta})$, and gamma function $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$. Hsieh et al. (2019) derive this expression on appendix page 4 for choice shocks that follow Fréchet distribution $F(\varepsilon_1, \ldots, \varepsilon_L) = \exp\{-\sum_\ell \varepsilon_\ell^{-\theta}\}$. Positive selection arises because high values of ε_ℓ increase the likelihood of observing choice ℓ, especially for choices with otherwise low probabilities $\bar{m}_{jk\ell}$.

C.2 Labor utility and wages

Education e yields labor utility $\bar{v}_{jk}(e)$ in expectation of skill shocks $\varepsilon = \{\varepsilon_\ell\}$ and destination choice ℓ. I evaluate this utility as folllows.

$$\bar{v}_{jk}(e) = \mathbb{E}[\max_{\ell} v_{jk\ell}(e, \varepsilon_\ell) \,|\, e] = e^{\eta} \sum_{\ell} \bar{m}_{jk\ell} \tilde{v}_{jk\ell} \mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell] = \gamma e^{\eta} MA_{jk}^{\frac{1}{\theta}}$$

The first equality starts with definition 6 for expected labor utility $\bar{v}_{jk}(e)$ and substitutes definitions 7, 8, and 21. The second applies the Law of Iterated Expectations over destinations ℓ. The third simplifies, noting that

$$\bar{m}_{jk\ell}\tilde{v}_{jk\ell}\mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell] = \gamma\tilde{v}_{jk\ell}\bar{m}_{jk\ell}^{1-\frac{1}{\theta}} = \gamma\tilde{v}_{jk\ell}^{\theta} MA_{jk}^{\frac{1}{\theta}-1}.$$

The first equality substitutes equation 23 for conditional skill shock $\mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell]$, and the second applies equations 11 and 21. I can similarly evaluate the expected wages $\bar{w}_{jk}(e)$ from labor.[23]

$$\bar{w}_{jk}(e) = \mathbb{E}[\max_\ell w_{jk\ell}(e,\varepsilon_\ell) \,|\, e] = e^\eta \sum_\ell \bar{m}_{jk\ell} r_\ell s_{jk\ell}\mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell] = \gamma e^\eta MA_{jk}^{\frac{1}{\theta}} \left(\frac{\sum_\ell (\frac{\tau^m_{jk\ell}}{a_\ell})\tilde{v}_{jk\ell}^\theta}{MA_{jk}} \right)$$

Labor utility and wages coincide when $\tau^m_{jk\ell} = a_\ell$ for all ℓ. Perfect mobility and homogeneous amenities yield one such case.

C.3 Education

Education maximizes utility $u_{jk}(\epsilon)$. The trade-off is that education raises expected labor utility tomorrow, but it incurs education costs today. The maximization problem is given by definition 5 for utility $u_{jk}(\epsilon)$. I substitute the above expression for expected labor utility $\bar{v}_{jk}(e)$, and I solve the first order condition.

$$e = \arg\max_e \left\{ \gamma e^\eta MA_{jk}^{\frac{1}{\theta}} - e\tau^e_{jk}\epsilon \right\} = \left(\frac{\gamma\eta MA_{jk}^{\frac{1}{\theta}}}{\tau^e_{jk}\epsilon} \right)^{\frac{1}{1-\eta}}$$

Schooling shocks ϵ allow education to vary individually. Aggregating over individuals,

$$\bar{e}_{jk} = \mathbb{E}[e] = \left(\frac{\gamma\eta MA_{jk}^{\frac{1}{\theta}}}{\tau^e_{jk}} \right)^{\frac{1}{1-\eta}} \mathbb{E}[\epsilon^{-\frac{1}{1-\eta}}] = \left(\frac{\gamma\eta MA_{jk}^{\frac{1}{\theta}}}{\tau^e_{jk}} \right)^{\frac{1}{1-\eta}} \bar{\epsilon}$$

The first equality defines the expectation. The second substitutes education e and applies the expectation. The third applies definition 22 for $\bar{\epsilon}$. Taking logs,

$$(1-\eta)\log\bar{e}_{jk} = \log\gamma + \log\eta + (1-\eta)\log\bar{\epsilon} + \frac{1}{\theta}\log MA_{jk} - \log\tau^e_{jk}. \quad (24)$$

[23] For completeness, I note that $\bar{m}_{jk\ell} r_\ell s_{jk\ell}\mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell] = \gamma r_\ell s_{jk\ell}\bar{m}_{jk\ell}^{1-\frac{1}{\theta}} = \gamma(\frac{\tau^m_{jk\ell}}{a_\ell})\tilde{v}_{jk\ell}^\theta MA_{jk}^{\frac{1}{\theta}-1}$.

C.4 Wages

Wages reward human capital. By equations 8 for wages and human capital,

$$w = w_{jk\ell}(e, \varepsilon_\ell) = r_\ell e^\eta s_{jk\ell} \varepsilon_\ell.$$

Skill shocks ε_ℓ allow wages to vary individually. Aggregating over individuals,

$$\bar{w}_{jk\ell} = \mathbb{E}[w \,|\, \text{choose } \ell] = r_\ell \mathbb{E}[e^\eta] s_{jk\ell} \mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell] = \left(\frac{\gamma \tau^m_{jk\ell} MA_{jk}^{\frac{1}{\theta}}}{a_\ell} \right) \left(\frac{\bar{e}_{jk}}{\bar{\epsilon}} \right)^\eta \tilde{\epsilon}.$$

The first equality defines the expectation. The second substitutes wages w and applies the expectation, noting that schooling and skill shocks are uncorrelated. The third substitutes equation 23 for conditional skill shocks $\mathbb{E}[\varepsilon_\ell \,|\, \text{choose } \ell]$ and rewrites with equation 11 for migration probabilities $\bar{m}_{jk\ell}$. It also substitutes

$$\mathbb{E}[e^\eta] = \left(\frac{\gamma \eta MA_{jk}^{\frac{1}{\theta}}}{\tau^e_{jk}} \right)^{\frac{\eta}{1-\eta}} \mathbb{E}[\epsilon^{-\frac{\eta}{1-\eta}}] = \left(\frac{\bar{e}_{jk}}{\bar{\epsilon}} \right)^\eta \tilde{\epsilon}$$

for $\tilde{\epsilon}$ as given by definition 22. Taking logs,

$$\log \bar{w}_{jk\ell} = \log \gamma + \log \tilde{\epsilon} - \eta \log \bar{\epsilon} - \log a_\ell + \eta \log \bar{e}_{jk} + \frac{1}{\theta} \log MA_{jk} + \log \tau^m_{jk\ell}. \tag{25}$$

C.5 Migration

Migration maximizes labor utility $U_{jk\ell}$ among the choice set of destinations ℓ.

$$m = \arg\max_\ell \tilde{v}_{jk\ell} \varepsilon_\ell$$

Skill shocks ε_ℓ across locations allow migration choices to vary individually. I note that skill shocks allow for individual variation in both wages and migration because there are independent shocks for each location. For an individual observed in destination ℓ, skill shock ε_ℓ for that destination can rationalize any observed wage. Fixing ε_ℓ, skill shocks $\{\varepsilon_{\ell'}\}$ in other destinations $\ell' \neq \ell$ can rationalize the observed migration to ℓ. That is, even if ε_ℓ is low, lower values for $\{\varepsilon_{\ell'}\}$ will rationalize choosing ℓ. Aggregating over individuals,

$$
\begin{aligned}
\bar{m}_{jk\ell} &= \mathbb{P}[\tilde{v}_{jk\ell'}\varepsilon_{\ell'} \leq \tilde{v}_{jk\ell}\varepsilon_\ell] \quad \forall \ell' \\
&= \int_0^\infty \frac{\partial}{\partial \varepsilon_\ell} \exp\left\{ -\sum_{\ell'} \left(\frac{\tilde{v}_{jk\ell}\varepsilon_\ell}{\tilde{v}_{jk\ell'}}\right)^{-\theta} \right\} d\varepsilon_\ell \\
&= \int_0^\infty \exp\left\{ -\varepsilon_\ell^{-\theta} \left(\frac{\sum_{\ell'} \tilde{v}^\theta_{jk\ell'}}{\tilde{v}^\theta_{jk\ell}}\right) \right\} \theta \varepsilon_\ell^{-\theta-1} d\varepsilon_\ell \\
&= \frac{\tilde{v}^\theta_{jk\ell}}{\sum_{\ell'} \tilde{v}^\theta_{jk\ell'}} \left[\exp\left\{ -\sum_{\ell'} \left(\frac{\tilde{v}_{jk\ell}\varepsilon_\ell}{\tilde{v}_{jk\ell'}}\right)^{-\theta} \right\} \right]_0^\infty = \frac{\tilde{v}^\theta_{jk\ell}}{\sum_{\ell'} \tilde{v}^\theta_{jk\ell'}}
\end{aligned}
$$

Fréchet shocks give choice probabilities of familiar form. The first equality defines the probability. The second computes the probability with Fréchet CDF $F(\varepsilon_1, \ldots, \varepsilon_L) = \exp\{-\sum_\ell \varepsilon_\ell^{-\theta}\}$ over candidate values of ε_ℓ. The third computes the Fréchet PDF, the fourth integrates, and the fifth simplifies. Note that $F = 1$ for $\varepsilon_\ell \to \infty$ and $F = 0$ for $\varepsilon_\ell \to 0$ given support $(0, \infty)$. I take logs and substitute definitions 21 for location utility $\tilde{v}_{jk\ell}$ and labor market access MA_{jk}.

$$
\log \bar{m}_{jk\ell} = \theta \log a_\ell + \theta \log r_\ell + \theta \log s_{jk\ell} - \theta \log \tau^m_{jk\ell} - \log MA_{jk} \tag{26}
$$

C.6 Production and human capital

By equations 12 and 13 for production Y_ℓ and wage rates r_ℓ, I can write production as a function of wage rates and total human capital H_ℓ in a location.

$$
Y_\ell = \frac{r_\ell H_\ell}{\kappa}
$$

In each location, production depends on the total human capital H_ℓ in that location. Total human capital sums over individuals, and I define total wages simlarly.

$$
H_\ell = \sum_{j,k} N_{jk}\bar{m}_{jk\ell}\bar{h}_{jk\ell}, \quad W_\ell = \sum_{j,k} N_{jk}\bar{m}_{jk\ell}\bar{w}_{jk\ell}
$$

for populations N_{jk}, migration probabilities $\bar{m}_{jk\ell}$, average human capital $\bar{h}_{jk\ell}$, and average wages $w_{jk\ell}$.

Human capital is proportional to wages at all levels of aggregation: individual, average, and total.

$$h = h_{jk\ell}(e, \varepsilon_\ell) = \frac{w}{r_\ell}, \quad \bar{h}_{jk\ell} = \mathbb{E}[h \,|\, \text{choose } \ell] = \frac{\bar{w}_{jk\ell}}{r_\ell}, \quad H_\ell = \frac{W_\ell}{r_\ell}$$

The left expression holds at the individual level. It follows from equations 8. The middle expression holds on average. It aggregates with conditional expectations over skill shocks ε_ℓ, with the first equality defining the expectation and the second applying that wages $\bar{w}_{jk\ell} = \mathbb{E}[w \,|\, \text{choose } \ell]$. The right expression holds in total. It follows from the definitions of H_ℓ and W_ℓ. Combining the above, I can express production in terms of total wages.

$$Y_\ell = \frac{W_\ell}{\kappa}$$

C.7 Estimating equations

I derive two main estimating equations. First, I subtract education equation 24 from wage equation 25.

$$\log \bar{w}_{jk\ell} - \log \bar{e}_{jk} = \log \tilde{\epsilon} - \log \bar{\epsilon} - \log \eta - \log a_\ell + \log \tau^e_{jk} + \log \tau^m_{jk\ell}$$

Second, I apply migration equation 26 to wage equation 25. I rewrite equation 26 to express labor market access MA_{jk} in terms of migration $\bar{m}_{jk\ell}$, then I substitute into equation 25.

$$\log \bar{w}_{jk\ell} = \log \gamma + \log \tilde{\epsilon} - \eta \log \bar{\epsilon} + \log r_\ell + \eta \log \bar{e}_{jk} - \frac{1}{\theta} \log \bar{m}_{jk\ell} + \log s_{jk\ell}$$

C.8 Extension: goods markets

I can extend the model to allow for goods markets alongside labor markets with endogenous goods prices. Setting $\kappa = 1$ for simplicity, output Y_ℓ depends on goods prices p_ℓ and production G_ℓ, where production combines productivity A_ℓ and human capital H_ℓ.

$$Y_\ell = p_\ell G_\ell, \quad G_\ell = A_\ell H_\ell$$

Competitive labor markets imply that wage rates r_ℓ reflect marginal productivity.

$$r_\ell = \frac{\partial Y_\ell}{\partial H_\ell} = p_\ell A_\ell$$

Perfectly competitive manufacturers produce final good G from location-specific goods G_ℓ with constant elasticity of substitution $\sigma > 1$. They sell at fixed price p.

$$Y = pG, \quad G = \left(\sum_\ell G_\ell^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

Manufacturers purchase location-specific goods at prices p_ℓ with costless trade across locations. Competitive goods markets imply that prices reflect marginal productivity.

$$p_\ell = \frac{\partial Y}{\partial G_\ell} = p \left(\frac{Y}{Y_\ell} \right)^{\frac{1}{\sigma}}$$

Thus, for a given location ℓ, human capital H_ℓ has two effects on output Y_ℓ. Rewriting output as $Y_\ell = r_\ell H_\ell$ shows these effects neatly. The direct effect of human capital is to increase output. The indirect effect is to decrease wage rates r_ℓ, which fall as added output suppresses prices p_ℓ. The indirect effect attenuates the direct effect.

D Counterfactuals

I derive exact-hat expressions with notation $\hat{x} = \frac{x'}{x}$ in changes for counterfactual x' and baseline x. The goal is to characterize how production responds to changes $(\hat{\tau}^e_{jk}, \hat{\tau}^m_{jk\ell})$ in education and migration costs, accounting for equilibrium changes $\hat{r}_\ell$ in wage rates. To this end, consider candidate changes $\hat{r}_\ell$ and take them as given.

On the demand side, recall $H^D_\ell = (\frac{\kappa A_\ell}{r_\ell})^{\frac{1}{1-\kappa}}$ for production elasticity κ, productivities A_ℓ, and wage rates r_ℓ. Changes in demand are

$$\hat{H}^D_\ell = \frac{H^{D'}_\ell}{H^D_\ell} = \hat{r}_\ell^{-\frac{1}{1-\kappa}}.$$

Productivities are fixed, and so they cancel. Changes $\hat{r}_\ell$ are as given, and parameter κ is as previously estimated.

On the supply side, recall $H^S_\ell = \frac{W_\ell}{r_\ell}$ for total wages W_ℓ and wage rates r_ℓ. Changes in supply are

$$\hat{H}^S_\ell = \frac{H^{S'}_\ell}{H^S_\ell} = \frac{\hat{W}_\ell}{\hat{r}_\ell}.$$

Changes $\hat{r}_\ell$ are again as given. For $\hat{W}_\ell = \frac{\sum_{j,\,k} N_{jk}\bar{m}'_{jk\ell}\bar{w}'_{jk\ell}}{\sum_{j,\,k} N_{jk}\bar{m}_{jk\ell}\bar{w}_{jk\ell}}$, I require counterfactual terms

$$\bar{m}'_{jk\ell}\bar{w}'_{jk\ell} = \bar{m}_{jk\ell}\bar{w}_{jk\ell}\hat{m}_{jk\ell}\hat{w}_{jk\ell},$$

where I write changes $(\hat{m}_{jk\ell}, \hat{w}_{jk\ell})$ rather than $(\hat{\bar{m}}_{jk\ell}, \hat{\bar{w}}_{jk\ell})$. I compute baseline $\bar{m}_{jk\ell}\bar{w}_{jk\ell}$ from observed $(\bar{m}_{jk\ell}, \bar{w}_{jk\ell})$. Substituting equations 10 and 11, these terms satisfy

$$\bar{m}_{jk\ell}\bar{w}_{jk\ell} = \frac{\tilde{\epsilon}(\gamma\eta^\eta)^{\frac{1}{1-\eta}} a_\ell^{\theta-1} r_\ell^\theta s^\theta_{jk\ell}}{(\tau^e_{jk})^{\frac{\eta}{1-\eta}}(\tau^m_{jk\ell})^{\theta-1}(MA_{jk})^{1-\frac{1}{\theta(1-\eta)}}}.$$

In changes, it follows that

$$\hat{m}_{jk\ell}\hat{w}_{jk\ell} = \frac{\hat{r}_\ell^\theta}{(\hat{\tau}^e_{jk})^{\frac{\eta}{1-\eta}}(\hat{\tau}^m_{jk\ell})^{\theta-1}(\hat{MA}_{jk})^{1-\frac{1}{\theta(1-\eta)}}}.$$

Scalars $(\tilde{\epsilon}, \gamma, \eta)$, amenities a_ℓ, and skill $s_{jk\ell}$ are fixed, and so they cancel. Changes $(\hat{r}_\ell, \hat{\tau}^e_{jk}, \hat{\tau}^m_{jk\ell})$ are as given, and parameters (η, θ) are as previously estimated. I compute changes $\hat{MA}_{jk}$ from equation 21 for labor market access.

$$\hat{MA}_{jk} = \frac{MA'_{jk}}{MA_{jk}}, \quad MA'_{jk} = \sum_{\ell} \left(\frac{a_\ell r'_\ell s_{jk\ell}}{\tau^{m'}_{jk\ell}} \right)^\theta, \quad MA_{jk} = \sum_{\ell} \left(\frac{a_\ell r_\ell s_{jk\ell}}{\tau^{m}_{jk\ell}} \right)^\theta$$

Amenities a_ℓ and skill $s_{jk\ell}$ are fixed, but they do not cancel because of the summation. Baseline $(a_\ell, r_\ell, s_{jk\ell}, \tau^m_{jk\ell})$ and parameter θ are as estimated, and counterfactual $(r'_\ell, \tau^{m'}_{jk\ell})$ are as given by $r'_\ell = r_\ell \hat{r}_\ell$ and $\tau^{m'}_{jk\ell} = \tau^m_{jk\ell} \hat{\tau}^m_{jk\ell}$. Baseline $(a_\ell, r_\ell, s_{jk\ell}, \tau^m_{jk\ell})$ need not be estimated in levels, as normalizations cancel when computing $\hat{MA}_{jk}$. With $\hat{MA}_{jk}$ in hand, I work backward to obtain $\hat{m}_{jk\ell}\hat{w}_{jk\ell}$, $\bar{m}'_{jk\ell}\bar{w}'_{jk\ell}$, $\hat{W}_\ell$, and lastly $\hat{H}^S_\ell$.

Taking stock, I begin with counterfactual changes $(\hat{\tau}^e_{jk}, \hat{\tau}^m_{jk\ell})$ for education and migration costs, I consider candidate changes $\hat{r}_\ell$ for wage rates, and I compute changes $(\hat{H}^D_\ell, \hat{H}^S_\ell)$ as above for human capital demand and supply. Fixing $(\hat{\tau}^e_{jk}, \hat{\tau}^m_{jk\ell})$, I consider many candidates $\hat{r}_\ell$, and I compute many changes $(\hat{H}^D_\ell, \hat{H}^S_\ell)$ across locations. I stop when I find the set $\hat{r} = \{\hat{r}_\ell\}$ that satisfy equilibrium conditions

$$\hat{H}^D_\ell = \hat{H}^S_\ell \quad \forall \, \ell.$$

These conditions in changes are equivalent to conditions 15 in levels because markets also clear in baseline equilibrium.[24] Thus, for any changes $(\hat{\tau}^e_{jk}, \hat{\tau}^m_{jk\ell})$ in education and migration costs, I obtain equilibrium changes $(\hat{r}_\ell, \hat{H}_\ell)$ in wage rates and human capital for each location. Recalling that production $Y_\ell = \frac{r_\ell H_\ell}{\kappa}$, in changes I obtain

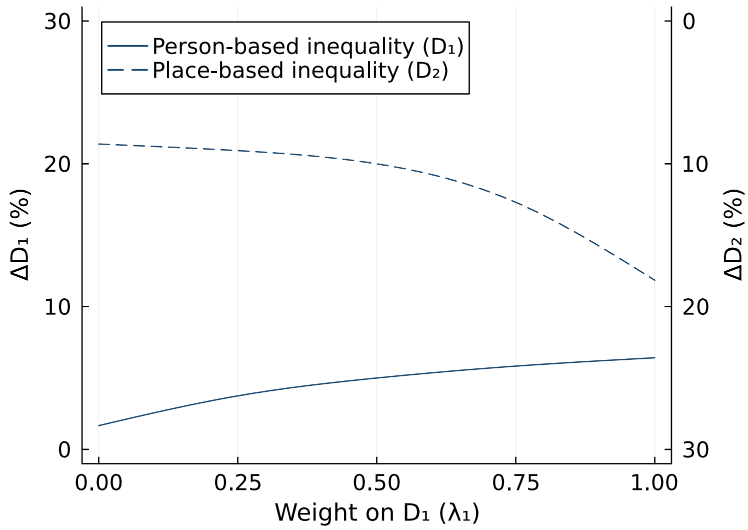
$$\hat{Y}_\ell = \hat{r}_\ell \hat{H}_\ell.$$

To reach this point, I require data $(\bar{m}_{jk\ell}, \bar{w}_{jk\ell})$, parameters (κ, η, θ), and estimates $(a_\ell, r_\ell, s_{jk\ell}, \tau^m_{jk\ell})$. Estimates need not be identified in levels.

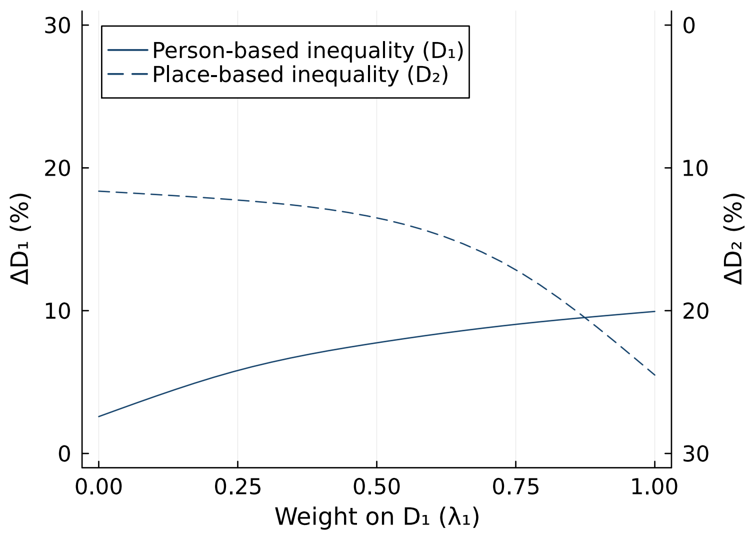
[24] In levels, equilibrium conditions 15 require that $H^{D'}_\ell = H^{S'}_\ell$ for counterfactual $(H^{D'}_\ell, H^{S'}_\ell)$. The condition becomes $H^D_\ell \hat{H}^D_\ell = H^S_\ell \hat{H}^S_\ell$ because, by definition, $H^{D'}_\ell = H^D_\ell \hat{H}^D_\ell$ and $H^{S'}_\ell = H^S_\ell \hat{H}^S_\ell$ for baseline (H^D_ℓ, H^S_ℓ) and changes $(\hat{H}^D_\ell, \hat{H}^S_\ell)$. But $H^D_\ell = H^S_\ell$ in baseline equilibrium, and so the condition simplifies to $\hat{H}^D_\ell = \hat{H}^S_\ell$ in changes.

Figure D1: Person- vs. place-based inequality

(a) Baseline



(b) Halved migration costs



I vary the objective function holding fixed weight $\lambda_0 = 0$ on aggregate output Y. I thus vary weight $\lambda_1 \in [0, 1]$ on person-based inequality D_1, which in turn affects weight $\lambda_2 = 1 - \lambda_1$ on place-based inequality D_2. For both y-axis, higher is better. The left axes are percentage decreases in D_1 relative to zero construction, with D_1 entering the objective function negatively. The right axes are percentage increases in D_2 relative to zero construction, with D_2 entering the objective function negatively and thus flipped axes in the figures. The bottom figure repeats the exercise of the top figure under 50% lower migration costs. Data: *Susenas* surveys (2011-14), *Bappenas* reports (1973-78), census reports (1971).