Educational Investment in Spatial Equilibrium: Evidence from Indonesia

Allan Hsiao*

Princeton University

January 4, 2023

(most recent version here)

This paper quantifies the long-run aggregate and distributional effects of Indonesia's Sekolah Dasar INPRES program, one of the largest school construction programs in history. I do so with a spatial equilibrium model in which students invest in education, then migrate for employment after graduation. I find that the program increased aggregate output by 8%, with large gains for rural students but small gains for rural regions. Labor market integration magnifies each effect, as education and migration are complements: access to high urban wages raises the returns to education, but also encourages students to leave rural regions behind.

*Email: ajhsiao@princeton.edu. I am grateful for helpful feedback from Nikhil Agarwal, Milena Almagro, Abhijit Banerjee, Gharad Bryan, Eric Chyn, Michael Dinerstein, Jonathan Dingel, Dave Donaldson, Esther Duflo, James Fenske, Anders Humlum, Clément Imbert, Michael Kremer, Margaux Luflade, Melanie Morten, Ben Olken, Simon Quinn, Chris Roth, Tavneet Suri, Felix Tintelnot, and Rob Townsend. I acknowledge generous support from the National Science Foundation Graduate Research Fellowship and Jerry Hausman Graduate Dissertation Fellowship.

1 Introduction

Governments invest more than $3 trillion in education annually (World Bank 2022). This investment targets students locally, but graduates migrate and seek employment nationally. This paper studies how migration shapes educational investment in the context of Indonesia's Sekolah Dasar INPRES program, an unprecedented school construction effort that established 61,807 new primary schools from 1973 to 1978. Differences in mobility generate substantial spatial heterogeneity in the returns to education, and I show how these differences inform the design of the program.

I begin by analyzing the program with the difference-in-differences approach of Duflo (2001). In particular, I compare exposed (young) and unexposed (old) age cohorts in districts with high and low levels of school construction. National socioeconomic survey data from 2011 to 2014 capture a range of long-run education and employment outcomes, including years of schooling and monthly wages, and data on district of birth provide the link to school construction. I document two stylized facts.

First, the returns to education vary greatly over space. I estimate the program's impact on education and wages, and I find positive long-run effects. The ratio of the education and wage effects, which correspond to a first stage and reduced form, gives average returns to education. I complement this analysis with the change-in-changes approach of Athey and Imbens (2006) to estimate the full distribution of treatment effects. Average effects mask considerable heterogeneity for education and wages, which in turn reveal large variation in the returns to education across districts.

Second, variation in mobility explains much of the variation in returns to education. I measure mobility with labor market access, which I compute for each district as an inverse-distance-weighted average of pre-program population densities across nearby districts. This measure captures workers' proximity to high-wage urban labor markets, and I validate it by showing that migration rates are highest where market access is high. I find that districts with high market access drive the program's education and wage effects, and that they enjoy the highest returns to education.

I capture these stylized facts with a spatial equilibrium model in which individuals pursue education, then migrate for employment. Frictions include education costs and migration costs, and I interpret school construction in a given district as decreasing education costs in that district. Unlike typical place-based policies that provide only local benefits, schools build portable human capital. The model thus captures two margins of spatial interactions. First, the returns to education depend on labor market access. Mobility gives rural students access to high urban wages, which reward high human capital and thus raise the incentives to invest in education. Second, school construction has both local and non-local effects. Mobility implies that rural construction may not lead to regional convergence, as rural students leave after graduation and contribute to urban output.

To estimate the model, I use data on education, migration, and wages to form three moment conditions, which I log-linearize into regression equations and estimate sequentially. Each identifies a mutually exclusive set of parameters. I address the endogeneity of school construction, education, and human capital, which enter as independent variables, by again appealing to the INPRES program described by Duflo (2001). This difference-in-differences variation isolates the causal effect of school construction, while also providing instruments for education and human capital.

Having estimated the model, I can obtain counterfactual education, migration, and wages under a range of alternative scenarios. I solve the model by guessing wages in each location and iterating until convergence to a fixed point. Given initial wages, I compute human capital as individuals choose education and migration in response to these wages. Given human capital, I compute implied wages as firms set wages to reflect the marginal productivity of human capital. Initial wages must be consistent with implied wages in equilibrium, equalizing human capital supplied by individuals with that demanded by firms. Migration generates spatial interdependence, and so I solve jointly for wages across locations.

I use the model to quantify the aggregate and distributional effects of the program. In particular, I compare observed outcomes with outcomes under a counterfactual with zero school construction. The model then allows me to decompose the effects of mobility by mechanism, and to separate each from the general equilibrium effects generated by this large-scale program. The difference-in-differences analysis does not rely on the model, but it only captures net effects. Finally, I study the design of the program by simulating alternative allocations of school construction.

Quantifying aggregate effects, I find that the program increased output by eight percent. A decomposition exercise allows me to assess the impact of mobility. Without migration, the program has a direct effect of only three percent. Migration has three effects. First, holding education and wages fixed, allowing individuals to sort into high-productivity regions increases output by another two percentage point. Second, holding wages fixed, larger returns to education raise investment in education, increasing output by a further four percentage points. Third, diminishing marginal returns to human capital affect wages in equilibrium, decreasing output by one percentage point. Bryan et al. (2014) find large gains from sorting, but endogenizing education would raise them further, including in general equilibrium.

Quantifying distributional effects, I find that rural students benefit most. The program expanded opportunities for less-advantaged rural students with high marginal returns, and in doing so decreased inequality between rural and urban students by five percent. At the same time, the program explicitly aimed to encourage regional convergence, but mobility places convergence in tension with output gains. Without mobility, rural residents stay in rural regions but face low wages. Regional inequality falls, but so do output gains. With mobility, rural-to-urban migration fuels output gains, but rural regions gain little net of out-migration. Even so, they are better off than under zero construction, such that the program remains Pareto-improving. Regional inequality rises only because urban regions gain much more.

I conclude with guidance for Indonesian policy, which faces an equity-efficiency tradeoff under mobility. Rural school construction generates large returns, but also slows convergence between rural and urban regions. Investments in connected districts are especially effective, but these districts benefit least because most graduates leave. An alternative is to complement school construction with transportation infrastructure that improves mobility itself. Doing so boosts the effects of school construction, but not in a Pareto-improving way: rural regions suffer as out-migration rises. I illustrate these trade-offs by computing (ex-post) optimal allocations under a range of objective functions and tracing out the possibilities frontier.

My main contribution is to show how large-scale educational investment interacts with migration in general equilibrium. To this end, I build on a literature that studies educational infrastructure and student outcomes in developing countries (Burde and Linden 2013, Kazianga et al. 2013, Dinerstein et al. 2022, Khanna 2023), including work on the INPRES program itself (Duflo 2001, 2004, Martinez-Bravo 2017, Mazumder et al. 2019, Ashraf et al. 2020, Akresh et al. 2021, Bazzi et al. 2021).[1] I highlight meaningful spatial heterogeneity in the returns to education, and I quantify aggregate and distributional effects over the long run. Relative to Khanna (2023) and Dinerstein et al. (2022), who also study large-scale school construction programs, I focus on how mobility contributes to the returns to education in spatial equilibrium, as well as the implications of migration for program design.

I also build on a literature that applies quantitative spatial equilibrium models to studying the allocation of human capital over space, as reviewed by Redding and Turner (2015) and Redding and Rossi-Hansberg (2017). This work largely focuses on transportation, with recent examples in developing countries that include Tsivanidis (2019), Adukia et al. (2020), Moneke (2020), Balboni (2021), Zárate (2021), and Milsom (2022).[2] I show how spatial concerns apply to educational infrastructure via migration, and I provide new evidence on endogenous human capital formation in a spatial setting. Relative to Eckert and Kleineberg (2021) and Agostinelli et al. (2022), who also apply spatial frameworks to studying education, I quantify the effects of school construction at the national scale. The INPRES program provides quasi-experimental variation and allows me to study long-run labor market outcomes.

I evaluate the program with a spatial equilibrium model that captures individuals' education and migration decisions. The model builds on Bryan and Morten (2019) and Hsieh et al. (2019) 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). I emphasize the interaction between mobility and the returns to education, leverage quasi-experimental variation for estimation, and connect to infrastructure investment with an emphasis on distributional effects.

Finally, I engage with the literature on place-based policy, as reviewed by Glaeser and Gottlieb (2008), Kline and Moretti (2014a), Neumark and Simpson (2015), and Austin et al. (2018). Existing empirical work studies spatially targeted infrastructure

1 US-focused studies include Cellini et al. (2010), Neilson and Zimmerman (2014), Goncalves (2015), Hong and Zimmer (2016), and Conlin and Thompson (2017).

2 Other examples include work on roads (Fajgelbaum and Schaal 2020, Gertler et al. 2022, Graff 2022), highways (Allen and Arkolakis 2014, 2022, Faber 2014, Alder 2016, Yang 2017, Morten and Oliveira 2018), railroads (Donaldson and Hornbeck 2016, Donaldson 2018, Hornbeck and Rotemberg 2021, Fajgelbaum and Redding 2022), railways (Heblich et al. 2020, Severen 2022), and buses (Balboni et al. 2020).

investment (Kline and Moretti 2014b, Balboni et al. 2020) and enterprise subsidies (Neumark and Kolko 2010, Ham et al. 2011, Busso et al. 2013, Wang 2013, Criscuolo et al. 2019). These policies provide only local benefits, which in-migration can offset by increasing local prices or draining non-local productivity. By contrast, schools provide portable benefits that out-migration magnifies and distributes. I quantify these benefits for one of the largest school construction programs in history.

2 Data and Stylized Facts

This section describes the INPRES program and the data, then evaluates the program with a difference-in-differences approach.

2.1 The INPRES program

The program had the stated goal of constructing 62,000 primary schools nationwide: 6,000 in the fiscal year beginning in 1973, 6,000 in 1974, 10,000 in 1975, 10,000 in 1976, 15,000 in 1977, and 15,000 in 1978 (*Inpres* No. 10/1973, 6/1974, 6/1975, 3/1976, 3/1977, 6/1978). In 1973 and 1974, schools were distributed across districts in proportion to pre-program unenrollment rates for children of primary school age. From 1975 to 1978, unenrollment was instead defined relative to a 15% threshold, with no new schools for districts with unenrollment rates below 15%. Figure 1 shows that school construction is indeed proportional to unenrollment rates in the data, and appendix table A3 documents the resulting emphasis on rural, isolated districts. INPRES refers to the "presidential instructions" that established the program.

2.2 Data

District-level data on INPRES school construction come from Duflo (2001), which draws on data from the Ministry of National Development Planning (*Bappenas*) and the 1971 population census. The data record the number of primary schools constructed, the number of pre-program primary schools, 1971 child populations and enrollment rates, and INPRES water and sanitation spending per capita. I compute population densities by dividing 1971 populations by land area, and I use these population densities as a measure of ruralness. For each district, I compute labor market access as a weighted average of 1971 population densities across districts, where weights $(1 + \text{dist}_{dd'})^{-2}$ are inversely proportional to distance. Thus, districts that either contain or are close to urban centers have high market access, such that this measure captures proximity to high-wage urban labor markets.

The main individual-level data come from the 2011, 2012, 2013, and 2014 National Socioeconomic Surveys (SUSENAS). I observe districts of residence and birth, with the latter providing the link to INPRES program exposure. The data record educational and employment outcomes, including educational attainment and monthly wages. Self-employment activity is observed, but self-employment income is not. I restrict attention to male heads of household ages 2 to 24 in 1974 – when the first INPRES schools were completed – and I adjust districts to 1971 boundaries for consistency over time. I study male heads of household to avoid issues of intrahousehold bargaining that might otherwise constrain migration, which is the focus of this paper. "Districts" refer sub-provincial urban *kota* and rural *kabupaten*.

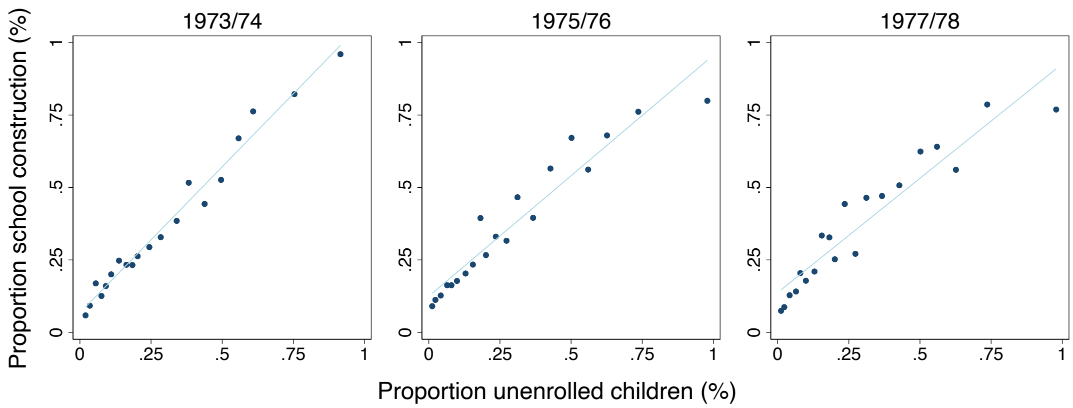
2.3 Education and wage effects

I estimate program effects by difference-in-differences as in Duflo (2001). Individuals ages 2 to 6 in 1974 – those young enough to benefit from new primary schools – form the treatment group, and those ages 12 to 17 in 1974 form the control group. I compare these groups in regions with high versus low levels of school construction.

$$Y_{ijk} = \delta_j + \delta_k + \beta S_j T_k + \boldsymbol{C}_j T_k \boldsymbol{\phi} + \varepsilon_{ijk}\,, \tag{1}$$

for individuals i born in district j and age cohort k. It includes outcome variable Y_{ijk}, district-of-birth fixed effect δ_j, year-of-birth fixed effect δ_k, school construction intensity S_j, treatment dummy T_k, district-of-birth controls $\boldsymbol{C}_j$, and error term ε_{ijk}. School construction intensity is the number of schools constructed per 1,000 children, and controls include 1971 child populations, 1971 enrollment rates, and INPRES spending on water and sanitation projects. I also include survey-year fixed effects because I pool SUSENAS data from multiple waves. The coefficient of interest is β, which captures the causal effect of school construction assuming common trends in high- and low-construction regions absent the program. As a placebo experiment, I compare two unexposed groups: those ages 12 to 17 and those ages 18 to 24 in 1974.

Figure 1: INPRES school construction vs. unenrollment rates by district



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 each district. The x-axis in 1973/1974 is the pre-program unenrollment rate among children of primary school age, and in other years is how much the rate exceeds 15%. I omit outliers by dropping the 5% of districts with extreme unenrollment rates.

Table 1: INPRES effects on education and labor

	Treatment			Placebo		
Outcomes	Estimate	SE	Obs	Estimate	SE	Obs
Years of schooling	0.103**	(0.0424)	233,517	-0.0176	(0.0318)	196,308
— For wage earners	0.121**	(0.0495)	89,404	0.0120	(0.0566)	55,091
Log monthly wages	0.0195**	(0.00916)	89,404	-0.00765	(0.00890)	55,091
Primary school completion	0.0585**	(0.0291)	233,517	-0.0134	(0.0167)	196,308
Middle school completion	0.0480**	(0.0207)	233,517	0.00573	(0.0156)	196,308
High school completion	0.0292	(0.0180)	233,517	-0.00167	(0.0140)	196,308
University completion	-0.0236	(0.0196)	233,517	-0.00792	(0.0214)	196,308
Employment	0.0304	(0.0278)	241,173	0.0309	(0.0216)	203,995
Wage employment	0.000376	(0.0131)	241,173	-0.0204	(0.0189)	203,995
Self-employment	-0.00219	(0.0119)	241,173	0.0140	(0.0142)	203,995
Weekly hours	-0.136	(0.102)	229,662	-0.00968	(0.109)	183,840

Each row is one treatment and one placebo regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Treatment compares individuals ages 2 to 6 and those ages 12 to 17 in 1974; placebo compares individuals ages 12 to 17 and those ages 18 to 24 in 1974. I run logit regressions for dummy outcomes. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 1 shows the long-run effects of the program on education and labor market outcomes. Consistent with the medium-run findings of Duflo (2001), school construction increases years of schooling, both in the full sample and for wage earners alone, and it increases log monthly wages. The education effects are driven by increased primary and middle school completion. The wage effects are not driven by increased employment, which suggests increased wage rates. These results also imply that the program does not meaningfully affect selection into the sample of wage earners. Placebo estimates are insignificant throughout.

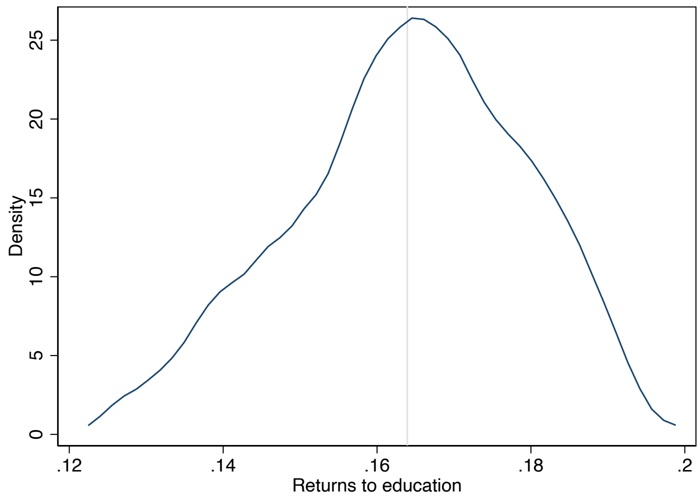
I then compute the implied returns to education by dividing the wage effect by the education effect. These effects correspond to the reduced form and first stage of a standard Wald estimator. In particular, I compute the proportional change in wages, as measured in log points, resulting from an additional year of education. I further consider the distribution of treatment effects in a change-in-changes framework, as formalized by Athey and Imbens (2006). Given a rank-invariance assumption, the empirical distributions of control and treatment outcomes reveal the full distributions of potential outcomes. I define districts with below-median school construction as control and those with above-median school construction as treatment. I then take the ratio of these estimates and obtain a distribution of returns to education. Figure 2 shows the result and reveals the considerable heterogeneity masked by the average.

2.4 Migration and labor market access

Figure 3 shows that baseline migration levels are high, particularly for districts with high labor market access, as individuals seek opportunities nationally. The average migration rate is 26%, and the average migration distance conditional on migration is 576 kilometers. The cross-province migration rate is 16%, compared to a cross-state migration rate of 31% in the United States, where mobility is relatively high.[3] Appendix figure A1 shows similar patterns across cohorts, with modestly higher levels of migration among younger, treated cohorts. Thus, spatial forces matter in equilibrium because many of those exposed to new schools migrate elsewhere.

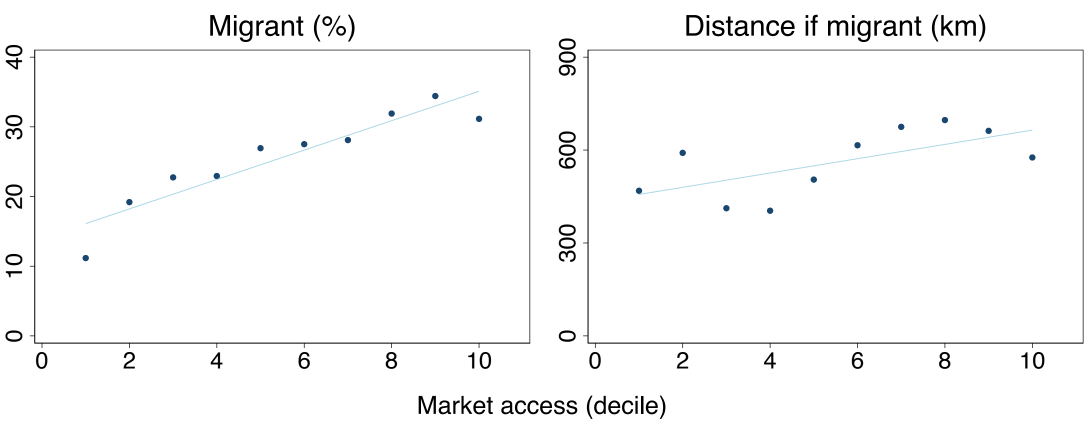
[3] I use 2013 and 2014 American Community Survey data to compute American migration rates. In doing so, I define migration as I do in the Indonesian context. Restricting attention to those born in the United States, which I take to include the 48 contiguous states plus the District of Columbia, I calculate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

Figure 2: Returns to education



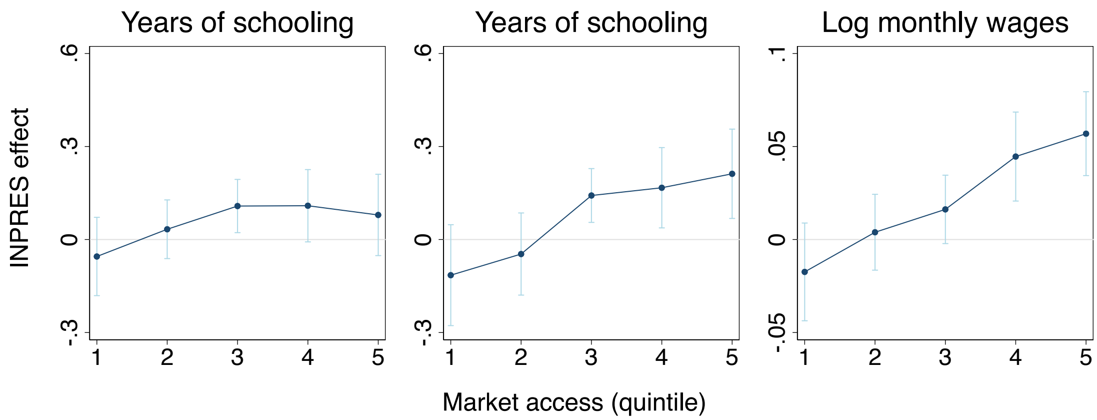
Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household ages 2 to 24 in 1974. I compute the distribution of returns to education by computing the distribution of education and wage treatment effects with change-in-changes and taking the ratio. The gray vertical lines shows the mean.

Figure 3: Migration and market access



Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household ages 2 to 24 in 1974. Migrants reside outside of their birth districts, and migration distances are Euclidean and between district centroids. Market access is an inverse-distance-weighted average of 1971 population densities across districts.

Figure 4: INPRES effects by market access



Each figure is one regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. I compare individuals ages 2 to 6 and those ages 12 to 17 in 1974. I report treatment effects by quartile of market access. Market access is an inverse-distance-weighted average of 1971 population densities across districts. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. Error bars shows 95% confidence bands.

Figure 4 shows that labor market access amplifies the INPRES treatment effect. I report interaction coefficients for quartiles $\boldsymbol{X_j}$ of birth-district market access.

$$Y_{ijk} = \delta_j + \delta_k + \boldsymbol{X_j} S_j T_k \boldsymbol{\beta} + \boldsymbol{C_j} T_k \boldsymbol{\phi} + \varepsilon_{ijk} \tag{2}$$

Effects increase in market access. Appendix figure A2 shows null effects in the placebo experiment, and appendix table A4 presents the regression table. Effects are indistinguishable from zero for districts with low market access, as barriers to migration limit the effective pool of job opportunities and thus the returns to education. I take market access as exogenous, as I construct the measure with 1971 populations that predate INPRES school construction and Euclidean distances that sidestep endogenous road networks. Neither quantity directly enters the allocation rule.

At the same time, table 2 shows that migration patterns do not themselves respond strongly to the program. Migration rates do not increase on the extensive margin, nor do migration distances on the intensive margin, and migration to both urban and rural destinations remains stable for urban and rural origins alike. This invariance is indeed consistent with the empirical model to come: in the model, school

Table 2: INPRES effects on migration

	Treatment			Placebo		
Outcomes	Estimate	SE	Obs	Estimate	SE	Obs
Migrant	0.0244	(0.0194)	244,793	-0.0249*	(0.0129)	210,543
Distance if migrant (km)	-5.097	(7.706)	62,717	-0.659	(6.656)	51,445
Migrant to urban	0.0284	(0.0307)	242,646	-0.0170	(0.0212)	207,096
Migrant to rural	0.0259	(0.0236)	244,793	-0.0169	(0.0170)	210,543
Migrant from urban to urban	0.0468	(0.0445)	116,594	0.0141	(0.0291)	105,664
Migrant from urban to rural	0.0449	(0.0276)	116,594	-0.0213	(0.0242)	105,664
Migrant from rural to urban	-0.00490	(0.0375)	126,052	-0.0484	(0.0335)	101,432
Migrant from rural to rural	-0.0113	(0.0260)	128,199	-0.0203	(0.0244)	104,879

Each row is one treatment and one placebo regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Treatment compares individuals ages 2 to 6 and those ages 12 to 17 in 1974; placebo compares individuals ages 12 to 17 and those ages 18 to 24 in 1974. I run logit regressions for dummy outcomes. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

construction lowers education costs but has no direct effect on either migration costs or migration itself. Moreover, even if the program changes neither migration nor labor market access over time, there remains large variation in the cross section that shapes the effects of school construction (as figures 3 and 4 show). Finally, this result invites the study of how school construction is affected when market access and migration do change, as I will emphasize in counterfactuals.

Indeed, consistent with mobility as a driver of wage gains, table 3 shows that people benefit more from school construction than places do. The first three columns show baseline estimates, as in table 1, that take 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, capturing effects on districts themselves. The latter estimates are indistinguishable from zero, suggesting that local gains dissipate as those who benefit most from the program eventually leave.[4]

[4] While the point estimates are consistently smaller and similarly precise, a significant difference from baseline would require even higher migration rates.

Table 3: INPRES effects for people vs. places

	People			Places		
	Years of schooling	Years of schooling	Log wages (month)	Years of schooling	Years of schooling	Log wages (month)
INPRES × young	0.103**	0.121**	0.0195**	0.0517	0.0260	0.0112
	(0.0424)	(0.0495)	(0.00916)	(0.0452)	(0.0506)	(0.00760)
Observations	233,517	89,404	89,404	232,915	89,252	89,252

Each row is one regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. I compare individuals ages 2 to 6 and those ages 12 to 17 in 1974. Treatment is school construction in the district of birth (people) or residence (places). I run logit regressions for dummy outcomes. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

3 Model

This section presents a spatial equilibrium model in which individuals invest in education, then migrate for work. The government invests in school construction with aggregate and distributional effects in mind.

3.1 Education and migration

Building on Bryan and Morten (2019) and Hsieh et al. (2019), individuals make education and migration decisions in two stages of life: (1) study and (2) work.[5] In the second stage, taking education as fixed, individuals i born in origin districts $j(i)$ and age cohorts $k(i)$ choose destinations ℓ to maximize utility

$$U_{i\ell}(e, \epsilon_{i\ell}) = u_{jk\ell} e^{\eta} \epsilon_{i\ell} \quad \text{for} \quad u_{jk\ell} \equiv \frac{a_\ell w_\ell s_{jk\ell}}{\tau^m_{jk\ell}} \tag{3}$$

with base utility $u_{jk\ell}$ and human capital $e^{\eta}\epsilon_{i\ell}$. These terms combine amenities a_ℓ, base wages w_ℓ, skill $s_{jk\ell}$, migration costs $\tau^m_{jk\ell}$, education e, human capital elasticity η, and skill draws $\epsilon_{i\ell}$. Base wages capture skill premia. Human capital thus amplifies base utility, as high-skill individuals benefit most from high base wages. These high-

[5] Relative to Bryan and Morten (2019), I endogenize education. Relative to Hsieh et al. (2019), I allow for imperfect labor-market information when choosing education.

wage individuals then benefit more from high amenities. Migration costs capture the financial, psychological, and opportunity costs of being away from home. In the first stage, individuals choose education e to maximize utility

$$U_i(e) = \mathbb{E}[\max_\ell U_{i\ell}(e, \epsilon_{i\ell})] - \tau^e_{jk} e \tag{4}$$

subject to education costs τ^e_{jk}. They do so knowing base utilities $u_{jk\ell}$ but not skill draws $\epsilon_{i\ell}$, which are not realized until the second stage. School construction reduces education costs, which capture difficulties in accessing education.

Fréchet skill draws facilitate solving in closed form. Following McFadden (1974) and Eaton and Kortum (2002),

$$F(\epsilon_1, \ldots, \epsilon_L) = \exp\left\{ - \sum_\ell \epsilon_\ell^{-\theta} \right\}$$

with high θ implying low skill dispersion. Migration probabilities are

$$m_{jk\ell} = \frac{u^\theta_{jk\ell}}{MA_{jk}} \quad \text{for} \quad MA_{jk} \equiv \sum_\ell u^\theta_{jk\ell} \tag{5}$$

with labor market access MA_{jk}. Education is fixed across destination choices and thus does not enter directly. Before realizing skill draws $\epsilon_{i\ell}$, expected utility is

$$\mathbb{E}[\max_\ell U_{i\ell}(e, \epsilon_{i\ell})] = \sum_\ell \mathbb{E}[U_{i\ell}(e, \epsilon_{i\ell}) \,|\, \text{choose } \ell] m_{jk\ell} = e^\eta MA^{\frac{1}{\theta}}_{jk} \gamma$$

given $\mathbb{E}[\epsilon_{i\ell} \,|\, \text{choose } \ell] = m^{-\frac{1}{\theta}}_{jk\ell} \gamma$ for $\gamma \equiv \Gamma(1 - \frac{1}{\theta})$. Education follows from equation 4, while wages combine base wages, skill, and human capital.

$$e_{jk} = \arg\max_e \left\{ U_i(e) \right\} = \left(\frac{\eta}{\tau^e_{jk}} \right)^{\frac{1}{1-\eta}} \left(MA^{\frac{1}{\theta}}_{jk} \gamma \right)^{\frac{1}{1-\eta}}, \tag{6}$$

$$w_{jk\ell} = \mathbb{E}[w_\ell s_{jk\ell} e^\eta_{jk} \epsilon_{i\ell} \,|\, \text{choose } \ell] = \frac{\tau^m_{jk\ell}}{a_\ell} \left(\frac{\eta}{\tau^e_{jk}} \right)^{\frac{\eta}{1-\eta}} \left(MA^{\frac{1}{\theta}}_{jk} \gamma \right)^{\frac{1}{1-\eta}} \tag{7}$$

I observe migration $m_{jk\ell}$, education e_{jk}, and wages $w_{jk\ell}$ (but not base wages w_ℓ).[6] I

[6] Unobserved skill draws $\epsilon_{i\ell}$ allow individual-level variation in wages but not education, which is chosen before skill draws are realized. Log-normal expectational errors υ_i would allow education

separate $(a_\ell, \tau^m_{jk\ell})$ from $(w_\ell, s_{jk\ell})$ assuming observed wages only capture the latter.

Comparative statics are as follows. Education costs decrease education and thus wages. Migration costs increase wages, as those who overcome barriers to entry are positively selected. Amenities decrease wages given compensating differentials. Labor market access amplifies the returns to education, increasing education, wages, and the gains from school construction. At the same time, market access encourages out-migration, which limits the local gains from school construction. School construction thus has smaller effects on places than it does on people.

The model focuses on internal migration for employment after graduation. First, employment is after education, but child labor may increase the opportunity cost of education in some locations.[7] Education costs τ^e_{jk} capture these costs. Second, migration is for employment, but parents may choose destinations based on the schools available to their children.[8] Amenities a_ℓ capture this availability. Third, migration is internal. Indeed, international out-migration is limited to less than 0.5% of the total population (World Bank 2022).[9]

The model also imposes several simplifications. First, the human capital returns to education are homogeneous, given a common parameter η. But the total returns to education are not, as utility combines human capital with heterogeneous base utility $u_{jk\ell}$. Second, individuals know base utilities across all destinations when choosing their education. The unrealized skill draws can accommodate additional uncertainty, in which case $u_{jk\ell}$ should be interpreted as known base utility. Third, individuals realize these skill draws across all destinations before choosing where to migrate, but informational frictions may apply for faraway destinations. Migration costs $\tau^m_{jk\ell}$ capture such frictions. Fourth, I abstract from sequential migration with skill and information accumulation over the life cycle, but I capture much of these effects because I observe long-run wages net of such activity.

---

to vary individually without affecting estimation, which relies only on group-level variation.

7 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).

8 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).

9 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.

3.2 Output and equilibrium

In each location, firms use human capital to produce goods with diminishing marginal returns. On the supply side, diseconomies of scale may arise as increased production places pressure on local factor markets. On the demand side, consumers may view products across destinations as imperfect substitutes, with equilibrium prices that decline in the quantities supplied. Perfect competition implies that base wages reflect marginal productivity. Output and wages are

$$Y_\ell = A_\ell H_\ell^\kappa \,, \quad w_\ell = \kappa A_\ell H_\ell^{\kappa-1} \tag{8}$$

for productivity A_ℓ, human capital H_ℓ, and production elasticity $0 < \kappa < 1$. Applying $Y_\ell / w_\ell = H_\ell / \kappa$, I obtain aggregate output as a sum of wages.

$$Y = \sum_\ell Y_\ell = \frac{1}{\kappa} \sum_{j,k,\ell} N_{jk} m_{jk\ell} w_{jk\ell} \tag{9}$$

for labor force N_{jk}. School construction raises output by raising human capital.

Base wages clear markets for human capital. Individuals supply human capital in exchange for wages, which affect labor market access and thus education and migration as in equations 5 and 6.

$$H_\ell^{\text{supply}} = \sum_{j,\,k} N_{jk} m_{jk\ell} \mathbb{E}[s_{jk\ell} e_{jk}^\eta \epsilon_{i\ell} \,|\, \text{choose } \ell] \,. \tag{10}$$

Firms demand human capital to produce goods. They set wages equal to diminishing marginal returns, which imply downward-sloping demand. Rearranging equation 8,

$$H_\ell^{\text{demand}} = \left(\frac{\kappa A_\ell}{w_\ell} \right)^{\frac{1}{1-\kappa}} . \tag{11}$$

In equilibrium, $H_\ell^{\text{supply}}(w_\ell) = H_\ell^{\text{demand}}(w_\ell)$ for base wages w_ℓ across locations ℓ. School construction thus has general equilibrium effects, as increased human capital affects wages in the labor market. Migration generates spatial equilibrium effects, as human capital in one location affects wages in all other locations.

3.3 School construction

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 $\bar{C}$,

$$\max_S \; \lambda_0 Y(S) - \lambda_1 D_1(S) - \lambda_2 D_2(S) \quad \text{s.t.} \quad \lambda_0 + \lambda_1 + \lambda_2 = 1\,, \quad C(S) \leq \bar{C}\,. \qquad (12)$$

Aggregate output is given by equation 9. Distributional concerns focus on wage gaps across people and places. For people, wage gap D_1 between urban and rural origins captures differences in opportunity for individuals. For places, wage gap D_2 between urban and rural destinations captures regional disparities net of migration.

$$D_1 = \frac{1}{\kappa} \sum_{j,k,\ell} (U_j - R_j) N_{jk} m_{jk\ell} w_{jk\ell}\,, \quad D_2 = \frac{1}{\kappa} \sum_{j,k,\ell} (U_\ell - R_\ell) N_{jk} m_{jk\ell} w_{jk\ell}$$

for urbanness U_j and ruralness R_j. Absent migration, $D_1 = D_2$.

4 Estimation

This section describes estimation and identification, then presents estimates.

4.1 Frictions and amenities

In a first step, I consider the wage premium for education. I divide equation 7 by equation 6 to eliminate the endogenous labor market access term.

$$\frac{w_{jk\ell}}{e_{jk}} = \frac{\tau^e_{jk} \tau^m_{jk\ell}}{\eta a_\ell}$$

I parameterize education and migration costs as functions of schools and distances. Let S_{jk} and $D_{j\ell}$ each denote schools and distances plus one.

$$\tau^e_{jk} = S_{jk}^{-\sigma} \tau^{e'}_{jk}\,, \quad \tau^m_{jk\ell} = D_{j\ell}^{\delta} \tau^{m'}_{jk\ell}$$

Substituting, decomposing $\tau_{jk}^{e'}\tau_{jk\ell}^{m'} = \tau_j\tau_k\varepsilon_{jk\ell}^{\tau}$, and taking logs with $\tilde{x} \equiv \log x$,

$$\tilde{w}_{jk\ell} - \tilde{e}_{jk} = -\tilde{\eta} - \sigma\tilde{S}_{jk} + \delta\tilde{D}_{j\ell} - \tilde{a}_{\ell} + \tilde{\tau}_j + \tilde{\tau}_k + \tilde{\varepsilon}_{jk\ell}^{\tau}\,. \tag{13}$$

This equation identifies education and migration cost parameters (σ, δ) and relative amenities $\frac{a_\ell}{a_0}$. Since schools are not randomly assigned, I again appeal to the INPRES school construction program as in Duflo (2001). Interaction term $S_{jk} = S_jT_k$ for schools S_j and age cohort exposure T_k, alongside origin and cohort fixed effects, captures the difference-in-differences comparison between young and old cohorts in districts with high and low levels of INPRES construction.[10] Euclidean distances represent bilateral resistance and yield the gravity equation typical of quantitative spatial models (Redding and Rossi-Hansberg 2017). In this setting, gravity leads to positive selection and thus high wages among individuals who overcome long distances, with higher δ implying stronger gravity and higher wages.[11] Origin and destination fixed effects isolate the effect of distance by controlling for systematic differences in remote places, including school quality and economic conditions.

4.2 Human capital, skill, and base wages

In a second step, I consider wages as a function of education and migration. Applying equations 5 and 6 to equation 7,

$$w_{jk\ell} = w_\ell s_{jk\ell} e_{jk}^{\eta} m_{jk\ell}^{-\frac{1}{\theta}}\gamma\,.$$

Decomposing $s_{jk\ell} = s_j s_k \varepsilon_{jk\ell}^{s}$ and taking logs,

$$\tilde{w}_{jk\ell} = \tilde{\gamma} + \eta\tilde{e}_{jk} - \frac{1}{\theta}\tilde{m}_{jk\ell} + \tilde{w}_\ell + \tilde{s}_j + \tilde{s}_k + \tilde{\varepsilon}_{jk\ell}^{s}\,. \tag{14}$$

This equation identifies human capital elasticity η, skill draw dispersion θ, and base wages w_ℓ, alongside skill $s_{jk\ell}$ as a residual. I identify base wages in levels given observed wages $w_{jk\ell}$ as numeraire and $\gamma = \Gamma(1 - \frac{1}{\theta})$ pinned down by θ.[12] Endogeneity

[10] Estimation can use total or INPRES schools for S_j, as origin fixed effects absorb initial stocks $\bar{S}_j$. Within $\tau_{jk\ell}$, I can also include interaction term $\boldsymbol{C}_jT_k$ with controls $\boldsymbol{C}_j$ to match specification 1.

[11] The Fréchet assumption characterizes selection and thus identifies δ from observed $w_{jk\ell}/e_{jk}$.

[12] Identifying base wages w_ℓ in levels also requires log-normal skill $s_{jk\ell}$ of mean one, otherwise $w_\ell s_{jk\ell}$ and $(cw_\ell)(\frac{1}{c}s_{jk\ell})$ are observationally equivalent for constant c. In equation 13, amenities

arises because unobserved skill $\varepsilon^s_{jk\ell}$ is mechanically correlated with both education e_{jk} and migration $m_{jk\ell}$, as in equations 5 and 6, through base utility $u_{jk\ell}(s_{jk\ell})$. For education, unobserved skill echoes the typical concern over ability as an omitted variable. I instrument for education with INPRES exposure S_{jk}, again leveraging the difference-in-differences variation. I thus isolate the causal effect of education on wages as in Duflo (2001). I instrument for migration with distances $D_{j\ell'}$ from origins j to destinations $\ell' \neq \ell$. These alternative destinations affect migration to ℓ, but are arguably uncorrelated and excludable with respect to ℓ. Faraway alternatives are more likely to be uncorrelated and excludable, but may also be less relevant.

4.3 Production

In a third step, I consider output as a function of human capital. By equation 8,

$$Y'_\ell = \kappa A_\ell H_\ell^\kappa$$

for output Y'_ℓ and human capital H_ℓ.

$$Y'_\ell \equiv \kappa Y_\ell = \sum_{j,k} N_{jk} m_{jk\ell} w_{jk\ell}\,, \quad H_\ell = \sum_{j,k} N_{jk} m_{jk\ell}^{1-\frac{1}{\theta}} s_{jk\ell} e_{jk}^{\eta} \gamma$$

I construct dollar-denominated Y'_ℓ from data alone, and I compute H_ℓ from data combined with estimates of θ, η, and $s_{jk\ell}$ from step two of estimation.[13] Taking logs,

$$\tilde{Y}'_\ell = \tilde{\kappa} + \kappa \tilde{H}_\ell + \tilde{A}_\ell\,. \tag{15}$$

This equation identifies production elasticity κ and productivity A_ℓ.[14] I identify productivity in levels given output as numeraire and a constant term pinned down by κ. Endogeneity arises because productivity A_ℓ is mechanically correlated with human capital H_ℓ through both education and migration, as in equations 5 and 6. I instrument for H_ℓ with INPRES school construction in alternative locations $\ell' \neq \ell$.

are not identified in levels because they are not captured by observed wages. Scaling amenities and education costs in equation 4 affects neither education nor wages, and so observed wages cannot serve as numeraire. Furthermore, friction $\tau_{jk\ell}$ need not be log-normal with mean one.

[13] Neither this step nor step two rely on step one of estimation.

[14] This κ parameter that governs general equilibrium forces is indeed not identified by difference-in-differences equation 14, as Khanna (2023) argues. Equation 15 shows that identifying κ requires variation across locations, but this variation is differenced out in equation 14.

Table 4: Parameter estimates

	Estimate	SE
Education costs (σ)	0.110**	(0.047)
Migration costs (δ)	0.042***	(0.004)
Human capital (η)	0.688**	(0.311)
Skill dispersion (θ)	21.31**	(10.52)
Production (κ)	0.767***	(0.101)

Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

This construction increases human capital in ℓ through migration from ℓ'. But it does not directly target ℓ and thus is arguably uncorrelated and excludable. As before, construction in faraway alternatives ℓ' may be more plausibly uncorrelated and excludable, but less relevant.

4.4 Estimates

Table 4 present parameter estimates. The σ parameter captures the relationship between school construction and education costs, and the δ parameter captures the relationship between distance and migration costs. Both parameters are positive and statistically significant, suggesting that school construction decreases education costs while distance increases migration costs. Fréchet parameter θ is similar to that estimated in Bryan and Morten (2019). Production parameter κ is less than one, suggesting diminishing marginal returns. Bryan and Morten (2019) calibrate this parameter to a value of 1.05 based on estimates from the literature, but their model includes congestion forces that offset this parameter. In my model, this parameter is net of congestion and thus has a smaller value.

For human capital elasticity η, table 5 presents an IV estimate of 0.7 compared to an OLS estimate of 0.4. The IV estimate is larger than the OLS estimate, as is the case in Duflo (2001). There is a relatively strong first stage that indeed disappears in the placebo experiment. For the US, Hsieh et al. (2019) choose a 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. I take my IV estimate of $\eta = 0.7$ as a baseline value, but also I consider robustness to the

Table 5: Human capital elasticity η

	Treatment			Placebo
	OLS	IV	First stage	First stage
Log years of schooling	0.393***	0.688**		
	(0.00721)	(0.311)		
INPRES $\times$ young			0.0284***	0.00564
			(0.00899)	(0.0110)
Observations	89,404	89,404	89,404	55,091
F-statistic			9.97	0.26

Each column is one regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Treatment estimates compare individuals ages 2 to 6 and those ages 12 to 17 in 1974; placebo estimates compare individuals ages 12 to 17 and those ages 18 to 24 in 1974. The outcome variable is log monthly wages, and the instrument for log years of schooling is the interaction of INPRES program intensity and treatment cohort. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

OLS value of 0.4 and the Hsieh et al. (2019) value of 0.1.

5 Counterfactuals

This section quantifies the program's long-run aggregate and distributional effects, highlighting the equity-efficiency trade-off facing the policymaker.

5.1 Solving the model

I compute aggregate output Y given schools S_j by solving the model as follows.

1. Given schools S_j, compute education costs τ^e_{jk}.
2. Guess base wages w_ℓ.
3. Given w_ℓ, compute migration $m_{jk\ell}$ with equation 5.
4. Given S_j and w_ℓ, compute education e_{jk} with equation 6.
5. Given $m_{jk\ell}$ and e_{jk}, compute human capital H_ℓ with equation 10.
6. Given H_ℓ, recompute base wages w_ℓ with equation 11.
7. Repeat from step (2) until convergence.

This algorithm solves for base wages w_ℓ as a fixed point, with migration, education, and wages all adjusting in equilibrium. I can also solve the model while changing parameters other than education costs in step one. However, the baseline analysis focuses on school construction and its effects on education costs, holding fixed other parameters like migration costs, amenities, and productivities. Note that relative amenities are sufficient for solving the model, as the normalization cancels.

5.2 Evaluating the program

Table 6 presents the aggregate and distributional effects of the program. The program increases aggregate output by eight percent relative to zero construction. Students from rural regions experience the largest gains, as new schools bring greater benefits to people from less-educated rural regions relative to more-educated urban ones. In increasing the opportunities available to rural students, the program decreases inequality between people from rural and urban regions by five percent. 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 regions themselves, net of out-migration. Reducing inequality across places was an explicit motivation for targeting INPRES school construction to low-enrollment regions, and both equity and political economy considerations can rationalize such a policy goal.[15] I find that the program increases inequality between rural and urban places by twelve percent. 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 regions. The program remains a Pareto improvement relative to zero school construction because rural regions still benefit from modest output gains and higher human capital. But regional inequality rises because urban regions benefit even more.

Mobility drives both aggregate and distributional effects. Shutting down mobility entirely by setting migration costs to one, I find that the direct effect of school construction is to increase output by only three percent. Inequality among people and inequality among places coincide and are similar to inequality under zero construction. Although rural students have higher marginal returns to education, they are confined

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

Table 6: INPRES aggregate and distributional effects

	Aggregate output	Inequality (people)	Inequality (places)
Evaluating the program			
Zero construction	1.00	1.00	1.00
Actual INPRES allocation	1.08	0.95	1.12
Decomposing migration effects			
Direct effect of construction	1.03	0.99	0.99
+ Migration	1.05	0.97	1.04
+ Migration-induced education	1.09	0.94	1.14
+ New equilibrium wages	1.08	0.95	1.12

Each row is one counterfactual. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Values are ratios relative to zero construction. 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 wages fixed. I then allow education to adjust but hold wages fixed. Finally, I allow wages to adjust.

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.

Allowing migration affects output in three ways. First, conditional on education and wages, individuals sort into high-productivity districts with larger returns to education, raising output by two percentage points. Second, conditional on wages, these larger returns in turn increase investment in education, raising output by four percentage points. Third, wages adjust in general equilibrium, decreasing output by one percentage point. Marginal productivity falls as total urban human capital rises, which depresses urban wages. Previous work has focused on sorting, but the education effect dominates here even net of general equilibrium effects (Bryan et al. 2014). At the same time, output gains are driven by rural students migrating to high urban wages, which increases inequality across places.

5.3 Redesigning the program

I study the design of the program by considering alternative allocations of school construction, subject to a budget constraint. I do so by maximizing the set of objective functions described in equation 12. 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, thereby characterizing the policymaker's possibilities frontier and quantifying 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, as spatial interdependence demands considering locations jointly. Indeed, school construction in one district depends on and affects labor markets in all other districts. 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. This cutoff was 0% for 1973-1974 construction and 15% for 1975-1978.

The analysis proceeds as follows. First, I vary the allocation rule cutoff over a grid of unenrollment cutoffs, and I obtain the resulting allocations. High cutoffs concentrate school construction in low-enrollment regions, which by appendix table A3 tend to be rural and isolated. Second, for each allocation I used the estimated model to compute effects on aggregate output, person-based inequality, and place-based inequality. Third, I consider an objective function focused on aggregate output ($\lambda_0 = 1$), and I select the allocation that maximizes this objective. This step amounts to maximization by grid search over the one-dimensional cutoff grid, greatly simplifying optimization because I search only over cutoffs and not over the full set of possible allocations. Fourth, I repeat the previous step 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.

In the first step, note that each allocation rule is subject to the observed budget constraint. I use total expenditures to define the budget, and indeed the INPRES program specified 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. I then distribute the total expenditures 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 many new schools as district one does. Future work can consider how aggregate and distributional effects vary with the budget, which I currently treat as fixed.

Figure 5 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 at the cost of elevated place-based inequality. Person-based inequality is positively correlated with aggregate output and thus avoids a trade-off. I also consider a policymaker balancing person- and place-based inequality, which are negatively correlated, in appendix figure B3.

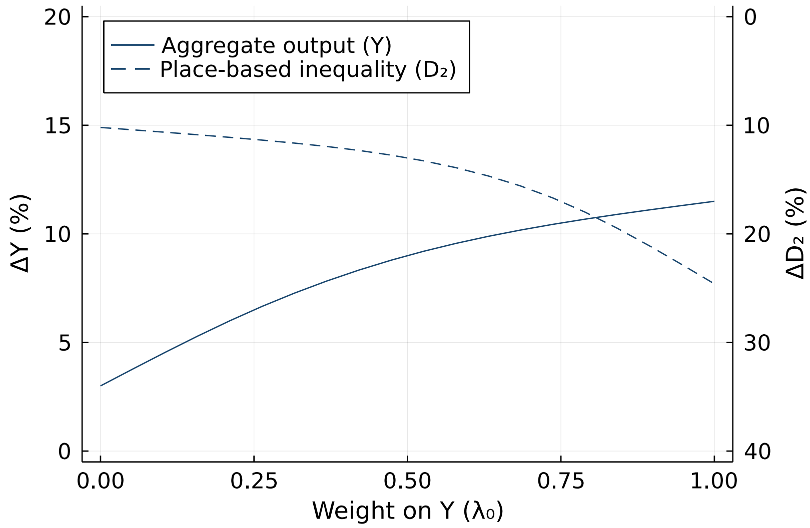
In figure 5a, more weight on aggregate output implies a higher cutoff, which increases construction in low-enrollment regions and thus both aggregate output and place-based inequality. I highlight the trade-off by reversing the axis for place-based inequality, which enters the objective function negatively. Conversely, more weight on place-based inequality implies a lower cutoff, which generates the opposite effects. Targeting low-enrollment regions thus raises aggregate output gains by enhancing opportunities for underserved students, but at the cost of increased regional disparities. Furthermore, figure 5a 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 8% and 12% effects produced by the actual allocation.[16]

In figure 5b, 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, lowering migration costs – such as by investing in roads – greatly amplifies 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 coordinated investment in schools and roads is substantially more effective than school construction alone, it is also no longer Pareto-improving relative to zero construction.

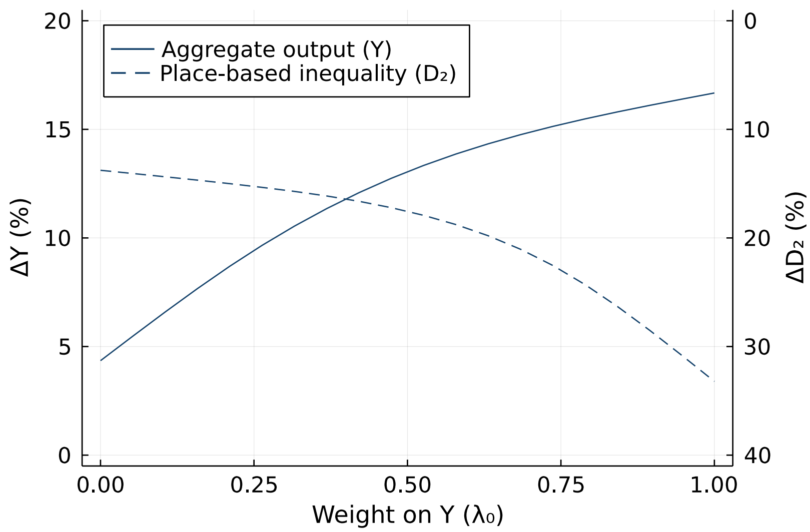
[16] The actual allocation falls below each curve because it departs slightly from the cutoff rule.

Figure 5: Effects on aggregate output vs. place-based inequality

(a) Baseline



(b) Halved migration costs



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 each y-axis, 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.

6 Conclusion

Spatial effects are crucial for evaluating large-scale educational investment because graduates migrate for employment. Mobility amplifies the returns to education, increasing output but draining rural regions. This paper capture these forces with a spatial equilibrium model and uses it to quantify 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 eight percent, but also increased regional inequality by twelve percent. Migration accounts for nearly all of each effect.

Several lines of inquiry are left for future work. First, future work might quantify the complementary effects of joint investment in schools and roads. Such an approach may be valuable given the interaction between education and migration costs. 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. 2022.

Akresh, Richard, Daniel Halim, and Marieke Kleemans. Long-term and Intergenerational Effects of Education: Evidence from School Construction in Indonesia. 2021.

Alder, Simon. Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development. 2016.

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.

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

Athey, Susan and Guido Imbens. Identification and Inference in Nonlinear Difference-in-Differences Models. Econometrica, 74(2):431–497, 2006.

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. 2021.

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. Islam and the State: Religious Education in the Age of Mass Schooling. 2021.

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. Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. Econometrica, 82(5):1671–1748, 2014.

Burde, Dana and Leigh Linden. Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools. American Economic Journal: Applied Economics, 5(3):27–40, 2013.

Busso, Matias, Jesse Gregory, and Patrick Kline. Assessing the Incidence and Efficiency of a Prominent Place Based Policy. American Economic Review, 103(2): 897–947, 2013.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. Quarterly Journal of Economics, 125(1):215–261, 2010.

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.

Conlin, Michael and Paul Thompson. Impacts of New School Facility Construction: An Analysis of a State-financed Capital Subsidy Program in Ohio. Economics of Education Review, 59:13–28, 2017.

Criscuolo, Chiara, Ralf Martin, Henry Overman, and John Van Reenan. Some Causal Effects of an Industrial Policy. American Economic Review, 109(1):48–85, 2019.

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

Dinerstein, Michael, Christopher Neilson, and Sebastián Otero. The Equilibrium Effects of Public Provision in Education Markets: Evidence from a Public School Expansion Policy. 2022.

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. Saving the American Dream? Education Policies in Spatial General Equilibrium. 2021.

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, 2022.

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

Gertler, Paul, Marco Gonzalez-Navarro, Tadeja Gračner, and Alexander Rothenberg. Road Maintenance and Local Economic Development: Evidence from Indonesia's Highways. 2022.

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

Goncalves, Felipe. The Effects of School Construction on Student and District Outcomes: Evidence from a State-Funded Program in Ohio. 2015.

Graff, Tilman. Spatial Inefficiencies in Africa's Trade Network. 2022.

Ham, John, Charles Swenson, Ayse Imrohoroglu, and Heonjae Song. Government Programs Can Improve Local Labor Markets: Evidence from State Enterprise Zones, Federal Empowerment Zones and Federal Enterprise Community. Journal of Public Economics, 95:779–797, 2011.

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.

Hong, Kai and Ron Zimmer. Does Investing in School Capital Infrastructure Improve Student Achievement? Economics of Education Review, 53:143–158, 2016.

Hornbeck, Richard and Martin Rotemberg. Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies. 2021.

Hsiao, Allan. Democratization and Infrastructure Investment: Evidence from Healthcare in Indonesia. 2021.

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.

Kazianga, Harounan, Dan Levy, Leigh Linden, and Matt Sloan. The Effects of "Girl-Friendly" Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso. American Economic Journal: Applied Economics, 5(3):41–62, 2013.

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.

Kline, Patrick and Enrico Moretti. People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs. Annual Review of Economics, 6:629–662, 2014a.

Kline, Patrick and Enrico Moretti. Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. Quarterly Journal of Economics, 129(1):275–331, 2014b.

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

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.

McFadden, Daniel. Conditional Logit Analysis of Qualitative Choice Behavior. In Zarembka, Paul, editor, Frontiers in Econometrics, volume 1, chapter 4, pages 105–142. Academic Press, 1974.

Milsom, Luke. Moving Opportunity: Local Connectivity and Spatial Inequality. 2022.

Moneke, Niclas. Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia. 2020.

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. 2018.

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

Neilson, Christopher and Seth Zimmerman. The Effect of School Construction on Test Scores, School Enrollment, and Home Prices. Journal of Public Economics, 120:18–31, 2014.

Neumark, David and Jed Kolko. Do Enterprise Zones Create Jobs? Evidence from California's Enterprise Zone Program. Journal of Urban Economics, 68(1):1–19, 2010.

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

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.

Severen, Christopher. Commuting, Labor, and Housing Market Effects of Mass Transportation: Welfare and Identification. 2022.

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 Bogota's TransMilenio. 2019.

Wang, Jin. The Economic Impact of Special Economic Zones: Evidence from Chinese Municipalities. Journal of Development Economics, 101:133–147, 2013.

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. 2017.

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

Zárate, Román David. Spatial Misallocation, Informality and Transit Improvements: Evidence from Mexico City. 2021.

APPENDIX

A Data and Stylized Facts

The 1976 and 1995 Intercensal Population Surveys (SUPAS) allow for additional placebo experiments with individual-level data for earlier periods. As in the baseline placebo experiment, I compare age cohorts that are both unexposed. In the 1995 SUPAS data, I compare individuals ages 12 to 17 and those ages 18 to 24 in 1974 – the same cohorts in the primary placebo experiment. This comparison replicates the placebo experiments in Duflo (2001). In the 1976 SUPAS data, I compare individuals ages 12 to 17 and those ages 18 to 24 in 1955. I focus on these earlier cohorts because those in the primary placebo experiment are not yet of working age. Tables A1 and A2 show that these experiments largely produce insignificant estimates for the outcomes considered in the main analysis.

Table A1: INPRES effects on education and labor (placebo)

	SUPAS 1976			SUPAS 1995		
Outcomes	Estimate	SE	Obs	Estimate	SE	Obs
Years of schooling	-0.0175	(0.0703)	18,173	0.0411	(0.0470)	64,392
— For wage earners	0.169	(0.164)	6,461	-0.0185	(0.0765)	25,159
Log monthly wages	0.00219	(0.0280)	6,461	0.000511	(0.00808)	25,159
Primary school completion	-0.0489	(0.0479)	18,061	0.00535	(0.0252)	64,392
Middle school completion	-0.01000	(0.0535)	18,135	0.0213	(0.0219)	64,392
High school completion	0.103	(0.0738)	17,838	0.0408	(0.0287)	64,392
University completion	0.239	(0.151)	12,598	-0.0357	(0.0520)	63,828
Employment	0.0900	(0.119)	16,135	-0.0125	(0.0487)	68,595
Wage employment	-0.000908	(0.0526)	18,098	0.00264	(0.0171)	69,114
Self-employment	0.00755	(0.0568)	18,067	-0.00622	(0.0180)	69,114
Weekly hours	-0.250	(0.270)	16,903	0.0277	(0.128)	66,423

Each row is two placebo regressions. Data focus on male heads of household. SUPAS 1976 compares individuals ages 12 to 17 and those ages 18 to 24 in 1955; SUPAS 1995 compares individuals ages 12 to 17 and those ages 18 to 24 in 1974. I run logit regressions for dummy outcomes. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2: INPRES effects on migration (placebo)

	SUPAS 1976			SUPAS 1995		
Outcomes	Estimate	SE	Obs	Estimate	SE	Obs
Migrant	0.0480	(0.0590)	16,860	-0.0224	(0.0232)	69,114
Distance if migrant (km)	11.57	(22.65)	5,584	25.47**	(12.32)	20,639
Migrant to urban	0.0544	(0.0625)	17,058	-0.0128	(0.0321)	66,207
Migrant to rural	-0.0395	(0.0950)	15,967	-0.0167	(0.0280)	69,114
Migrant from urban to urban	-0.0532	(0.0780)	11,751	0.0352	(0.0419)	38,415
Migrant from urban to rural	-0.0571	(0.123)	10,735	-0.00784	(0.0460)	38,415
Migrant from rural to urban	0.238**	(0.110)	5,307	-0.0908*	(0.0507)	27,792
Migrant from rural to rural	-0.0758	(0.154)	5,232	0.00545	(0.0377)	30,699

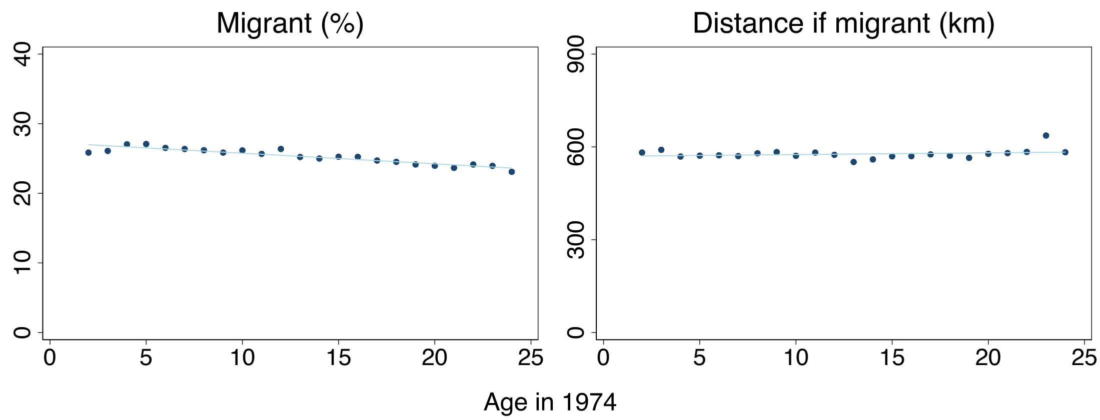
Each row is two placebo regressions. Data focus on male heads of household. SUPAS 1976 compares individuals ages 12 to 17 and those ages 18 to 24 in 1955; SUPAS 1995 compares individuals ages 12 to 17 and those ages 18 to 24 in 1974. I run logit regressions for dummy outcomes. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3: INPRES school construction vs. ruralness and isolation

	School construction	School construction	School construction
Population density (ruralness)	-0.0748***		-0.0309***
	(0.0155)		(0.0118)
Labor market access (isolation)		-0.417***	-0.359***
		(0.0759)	(0.0851)
Observations	282	282	282

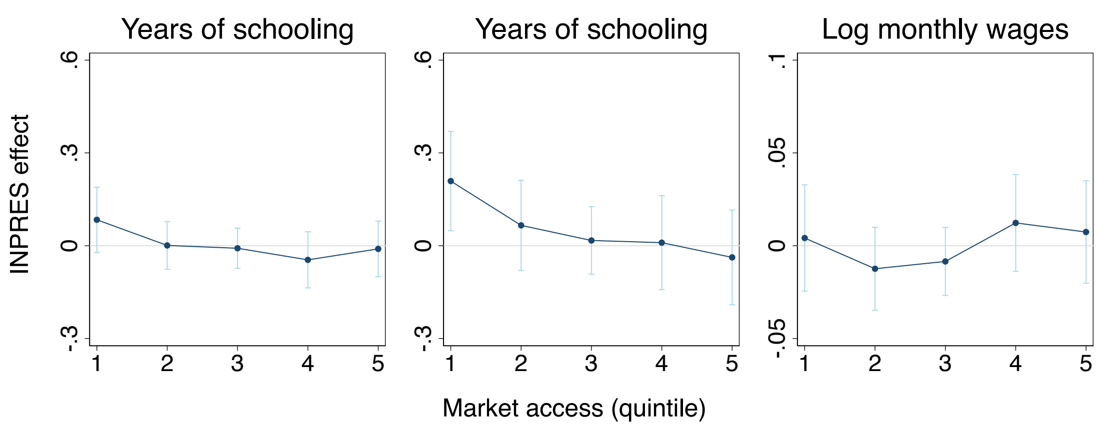
Each row is one regression, and each observation is a district. Population density is 1971 population divided by land area. Market access is an inverse-distance-weighted average of 1971 population densities across districts. School construction is INPRES schools built per million children in 1971. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure A1: Migration and age cohort



Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household ages 2 to 24 in 1974. Migrants reside outside of their birth districts, and migration distances are Euclidean and between district centroids.

Figure A2: Placebo INPRES effects by market access



Each figure is one regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. I compare individuals ages 12 to 17 and those ages 18 to 24 in 1974. I report treatment effects by quartile of market access. Market access is an inverse-distance-weighted average of 1971 population densities across districts. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. Error bars shows 95% confidence bands.

Table A4: INPRES effects by market access

	Treatment			Placebo		
	Years of schooling	Years of schooling	Log wages (month)	Years of schooling	Years of schooling	Log wages (month)
INPRES × young	0.0309	-0.0457	-0.0167	0.0157	0.0817	-0.0152
	(0.0489)	(0.0612)	(0.0103)	(0.0367)	(0.0679)	(0.0106)
—— × MA	0.0899**	0.207***	0.0449***	-0.0412	-0.0855*	0.00923
	(0.0350)	(0.0427)	(0.00624)	(0.0256)	(0.0451)	(0.00733)
Observations	233,517	89,404	89,404	196,308	55,091	55,091

Each column is one regression. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Treatment compares individuals ages 2 to 6 and those ages 12 to 17 in 1974; placebo compares individuals ages 12 to 17 and those ages 18 to 24 in 1974. Market access is an inverse-distance-weighted average of 1971 population densities across districts. Regressions control for birth district, birth year, and survey year fixed effects, as well as 1971 child population, 1971 enrollment rates, and INPRES spending on water and sanitation projects. Standard errors are clustered by birth district. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

B Counterfactuals

I consider three extensions. First, spatial spillovers imply that investment should be centralized. Table B1 illustrates that construction is greatly reduced if districts must fund school construction themselves. In this case, the program increases aggregate output by only one percent. The reason is that districts realize only part of the benefits of local construction, and they do not internalize benefits for other districts. I determine construction levels by computing the marginal social benefit of construction in each district under the observed allocation, then reducing construction in each district until the marginal district benefit of construction matches the computed social benefit. I do so taking other districts' investment to be zero. Centralized investment increases investment by internalizing cross-district spillover effects, raising aggregate output by another three percentage points. It also takes advantage of complementarities in investment. Taking other districts' investment to instead be at observed levels, aggregate output increases by a further three percentage points.

Second, more sophisticated allocation rules lead to larger aggregate output effects. Allocation rules can vary in the weights and observables considered. For a

Table B1: INPRES effects on aggregate output

	Aggregate output
Zero construction	1.00
Uniform construction	1.02
Actual INPRES allocation	1.08
District-based investment	1.01
+ Internalizing spillovers	1.04
+ Internalizing complementarities	1.08

Each row is one counterfactual. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Values are ratios relative to zero construction.

given weighting scheme $\mathcal{P}$, set of observables X, and budget A, I choose weights

$$\rho^* = \arg\max_{\rho \in \mathcal{P}} \left\{ W\big(a(\rho; X, A)\big) \right\}$$

in order to maximize a given objective function $W(a)$. That is, optimization is over weights ρ that correspond to allocations $a(\rho)$ as follows.

$$a_\ell(\rho; X, A) = A \left(\frac{X_\ell \rho}{\sum_\ell X_\ell \rho} \right)$$

Proportional weighting schemes use uniform weights, avoiding optimization altogether ($\mathcal{P}_{\text{prop}} = \{\rho \mid \rho = \mathbf{1}\}$). The actual allocation rule was proportional to unenrollment and thus falls within this set. Linear and quadratic schemes offer more flexibility by admitting weights parameter to optimize over.

Table B2 presents allocation rules of varying complexity. The actual rule captures diminishing marginal returns by conditioning on unenrollment. Indeed, this rule is more effective than rules conditioned on other single observables. It is also more effective than a uniform rule that neglects observables entirely (table B1). More flexible weighting schemes increase effectiveness, but larger gains come from combining unenrollment and other observables. Ruralness is a rough proxy for market access that is already commonly considered in regional planning, and distance to the nearest urban district is an even better proxy. Rules that combine unenrollment and urban distance – even a simple proportional one – substantially outperform those that con-

Table B2: Effects of sophisticated allocation rules on aggregate output

	Weighting scheme			
Observables	Proportional	Cutoff	Linear	Quadratic
Child population	1.05	1.05	1.05	1.06
Unenrollment	1.07	1.07	1.08	1.08
Ruralness	1.04	1.04	1.04	1.05
Urban distance	1.04	1.04	1.05	1.05
Child population + ruralness	1.07	1.07	1.08	1.08
Child population + urban distance	1.07	1.08	1.08	1.09
Unenrollment + ruralness	1.08	1.09	1.09	1.09
Unenrollment + urban distance	1.09	1.09	1.09	1.10

Each row is one counterfactual. Data come from SUSENAS 2011, 2012, 2013, and 2014 and focus on male heads of household. Values are ratios relative to aggregate output under zero construction. Unenrollment is unenrolled school-age child population, and urban distance is Euclidean distance to the nearest urban district.

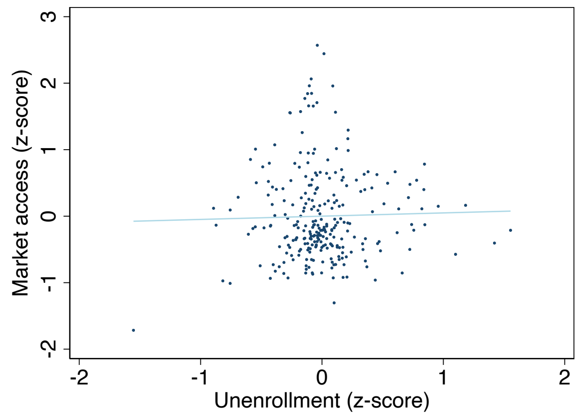
sider unenrollment alone. Unenrollment alone is insufficient because it is uncorrelated with market access (figure B1), and thus misses an important force in the full model.

Third, uncertainty can rationalize the use of simple allocation rules. Long-run migration costs are uncertain at the time of allocating school construction, and these migration costs have important effects on schooling and wages as previously discussed. I therefore consider expected aggregate output

$$\mathbb{E}_{\upsilon}[Y(a;\upsilon)] = \int Y(a;\upsilon) f(\upsilon) d\upsilon \,,$$

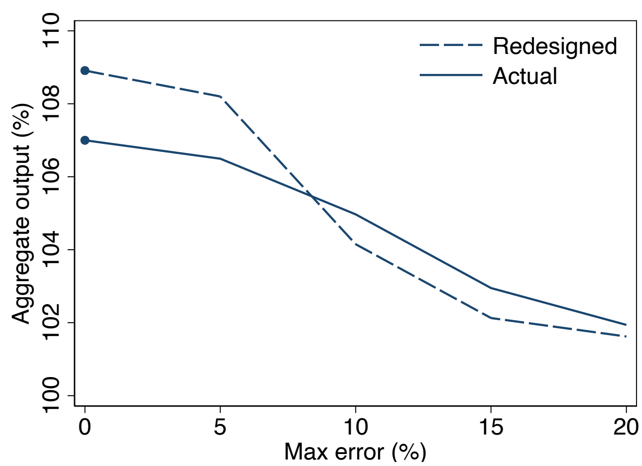
subject to multiplicative distortions υ to migration costs. Figure B2 shows that a redesigned rule – one proportional to unenrollment and urban distance – dominates when uncertainty is low, as it allows more precise targeting. But it also involves weight parameters fit in expectation, effectively overfitting to mean error scenarios. As such, the actual rule – one proportional to unenrollment alone – dominates when uncertainty is high. That is, complex rules can be more effective, but simpler rules are more robust. Indeed, policymakers often employ simple rules in complex environments, including population cutoffs and ranked lists for public investment. Future work can consider a similar exercise with uncertainty in the effects of school construction on education costs. More broadly, rationalizing the actual allocation rule

Figure B1: Pre-INPRES unenrollment vs. market access



Each observation is one district. Market access is an inverse-distance-weighted average of 1971 population densities across districts. Unenrollment is total unenrolled school-age child population. The figure controls for 1971 population.

Figure B2: Aggregate output by allocation rule under uncertainty

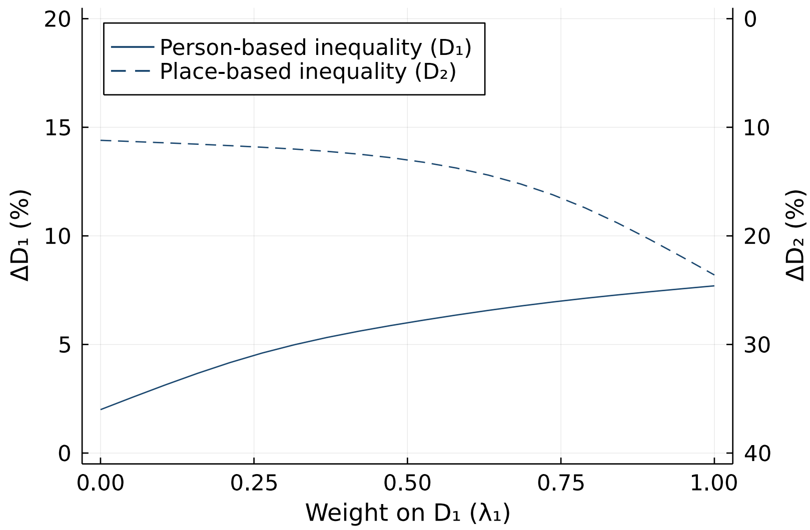


The redesigned rule is proportional to unenrollment and urban distance, and the actual rule is proportional to unenrollment. Aggregate output values are relative to zero construction. Max error $\bar{\upsilon}$ implies multiplicative distortions $\upsilon \sim U[1-\bar{\upsilon}, 1+\bar{\upsilon}]$ to migration costs.

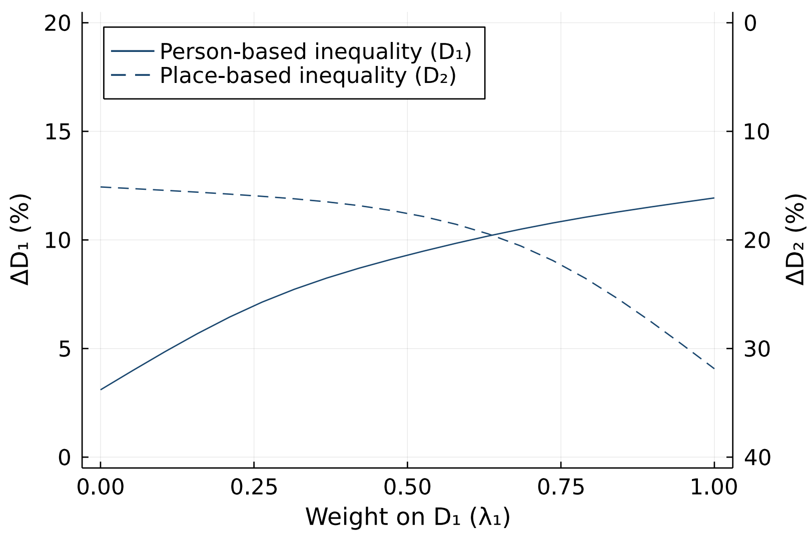
is possible with other objective functions as well. Political concerns are one example, and I pursue this line of inquiry in related work on healthcare infrastructure in Indonesia (Hsiao 2021).

Figure B3: Effects on 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 each 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.