Educational Investment in Spatial Equilibrium: Evidence from Indonesia

Allan Hsiao
University of Chicago

June 7, 2022

How does migration shape large-scale educational investment?

- Governments invest $3 trillion in education annually (World Bank 2022)
  - In Indonesia, 61,807 new primary schools (Sekolah Dasar INPRES program, 1973-1978)
- Educational investment targets students locally
  - But graduates migrate and seek employment nationally

This paper

- **Aggregate** and **distributional effects** of the INPRES program
  - Difference-in-difference with long-run outcomes (Duflo 2001)
  - Spatial heterogeneity in returns to education + implications for program design
- Spatial equilibrium model of **complementary education + migration** decisions
  - Returns to education: rural incentives depend on urban wages (if mobile)
  - Regional convergence: rural schools increase urban output (if mobile)
- **Results:** aggregate output ↑ (8%), regional inequality ↑ (12%)
  - Mobility magnifies both effects

Literature

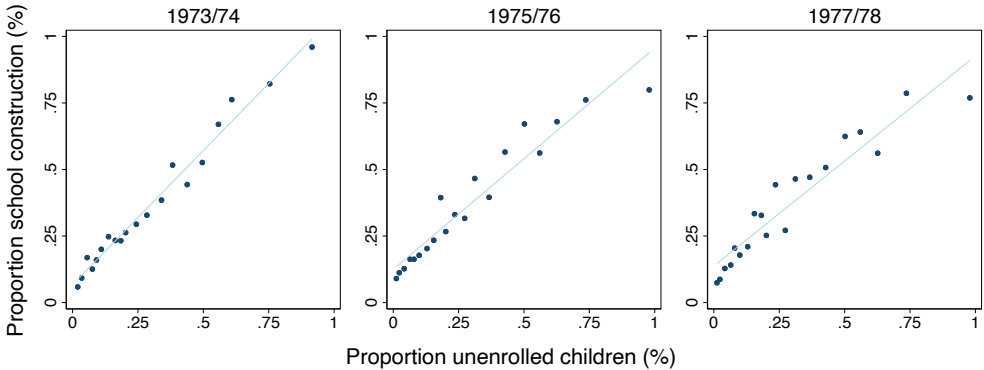
- **Education and migration** at scale in general equilibrium
  - Education: Khanna 2021, Dinerstein et al. 2022 (no migration)
  - Migration: Dahl 2002, Bryan et al. 2014, Bryan & Morten 2019 (no education)
  - Both: Eckert & Kleineberg 2021, Agostinelli et al. 2022 (no school construction)

- **INPRES program** evaluation with aggregate effects and counterfactuals
  - Duflo 2001/2004, Martinez-Bravo 2017, Ashraf et al. 2020, Bazzi et al. 2021

- **Place-based policy** with portable human capital benefits
  - Glaeser & Gottlieb 2008, Kline & Moretti 2014, Busso et al. 2013, Austin et al. 2018

Data and Stylized Facts

The INPRES program built 62,000 new primary school (1973-1978)




KEAMANAN
KEBERSIHAN
KETERTIBAN
KEINDAHAN
KEKELUARGAAN
KERINDANGAN
KESEHATAN
KEDISIPLINAN
KERAPIHAN
KEWASPADAAN
SD INPRES DAYA
KEC. BIRINGKANAYA
KOTA MAKASSAR

Data

- **Treatment** at district level
  - INPRES school construction (1973-1978)
  - Pre-program primary schools, child populations, enrollment rates
- **Long-run outcomes** at individual level
  - SUSENAS household surveys (2011-2014)
  - Districts of residence and birth, years of schooling, monthly wages

Difference-in-difference variation in school construction (Duflo 2001)

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

- **Young vs. old** students in age cohorts k
  - Young exposed to new schools, but old not
- **Many vs. few** new schools in origin districts j
  - More schools $\rightarrow$ bigger difference between young and old cohorts

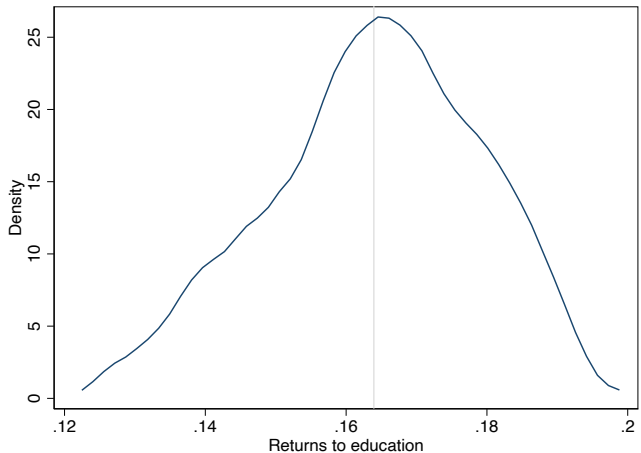
Long-term education and wage effects

	Treatment		
Outcomes	Estimate	SE	Obs
Years of schooling	0.103**	(0.0424)	233,517
— For wage earners	0.121**	(0.0495)	89,404
Log monthly wages	0.0195**	(0.00916)	89,404

Long-term education and wage effects

	Placebo		
Outcomes	Estimate	SE	Obs
Years of schooling	-0.0176	(0.0318)	196,308
— For wage earners	0.0120	(0.0566)	55,091
Log monthly wages	-0.00765	(0.00890)	55,091

Heterogeneous returns to education by CIC (Athey & Imbens 2006)



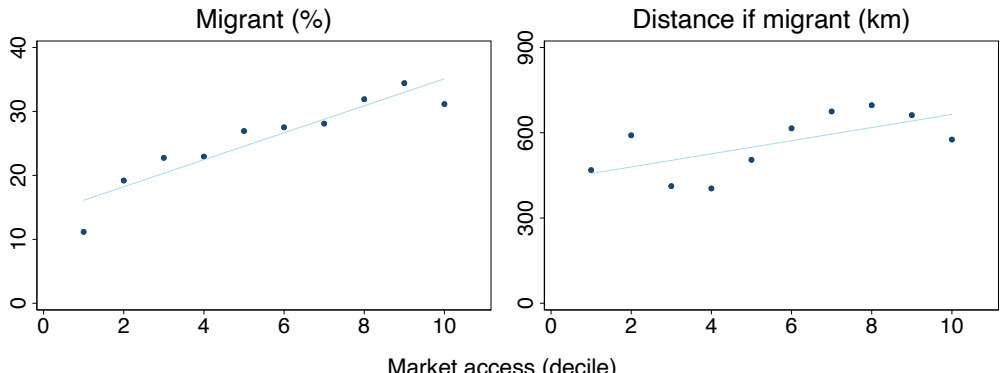
Migration levels are high and increasing in labor market access

- Average migration rate is 26%, and distance is 576 km
  - 16% cross-province vs. 31% cross-state in the US (ACS 2013-2014)
  - Many of those exposed to new schools migrate elsewhere

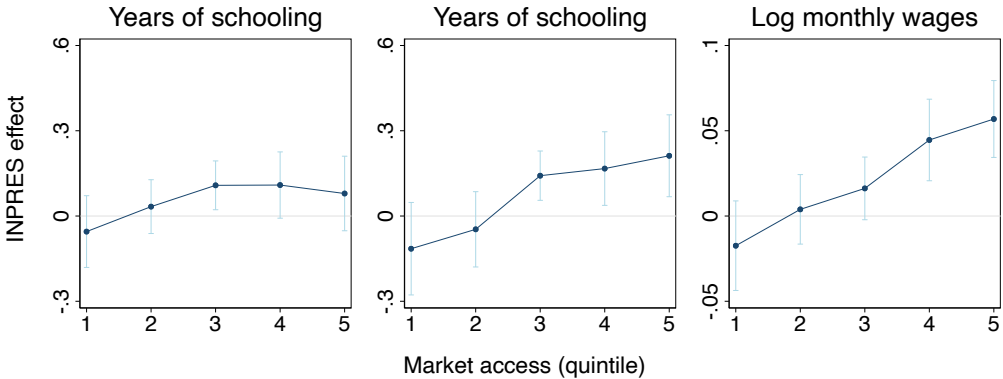
- Labor market access captures proximity to high urban wages

$$\text{MA}_d = \sum_{d'} \text{w}_{d'} \text{popden}_{d'} \quad \text{for} \quad \text{w}_{d'} \propto (1 + \text{dist}_{dd'})^{-2}$$

Migration levels are high and increasing in labor market access



INPRES effects are driven by labor market access



But INPRES does not change migration patterns

Outcomes	Estimate	SE	Obs
Migrant	0.0244	(0.0194)	244,793
Distance if migrant (km)	-5.097	(7.706)	62,717
Migrant to urban	0.0284	(0.0307)	242,646
Migrant to rural	0.0259	(0.0236)	244,793

- Consistent with model: INPRES affects costs of education, not migration
- In counterfactuals, INPRES effects under different migration costs

Model

Spatial equilibrium model

1. Government constructs schools
   - Build human capital that is portable
2. Individuals invest in education
   - In a district, more schools $\rightarrow$ better access $\rightarrow$ lower costs of education
3. Individuals migrate for work
   - Mobility gives rural students access to high urban wages
   - But also drains rural regions as students leave after graduation

School construction $a = \{a_\ell\}$, districts ℓ

- CES aggregate output

$$Y(a) = \left\{ \sum_{\ell} \left[A_\ell H_\ell(a_\ell) \right]^{\frac{\sigma-1}{\sigma}} \right\}^{\frac{\sigma}{\sigma-1}}$$

- Inequality (people vs. places)

$$D(a) = Y^U(a) - Y^R(a) \quad \text{for} \quad Y_\ell^U(a) = \lim_{\sigma \to \infty} U_\ell Y_\ell(a)$$

- Costs $C(a)$ from data

Frictions: education and migration costs

$$U(e,\epsilon) = \alpha_\ell \varepsilon^{\alpha}_{jk\ell} [\underbrace{(1-\tau^m_{j\ell}) w_\ell h_{jk} \varepsilon^h_{jk\ell} e^\eta \epsilon}_{\text{net labor income}} - \underbrace{(1+\tau^e_{jk}) c \varepsilon^c_{jk\ell} e}_{\text{cost of education}}]$$

- Individual i, origin j, age cohort k, destinations ℓ
  - Each destination has education choice e^* and utility $U(e^*, \epsilon)$
  - Given amenities α_ℓ, migration costs $\tau^m_{j\ell}$, base wages w_ℓ, Fréchet draw ϵ
  - And human capital h_{jk}, education costs τ^e_{jk} across destinations

- Compare destinations, then pick best to get choice probabilities

$$\pi_{jk\ell} = \frac{\tilde{w}^\theta_{jk\ell}}{\sum_{\hat{\ell}} \tilde{w}^\theta_{jk\hat{\ell}}} \quad \text{for} \quad \tilde{w}_{jk\ell} \equiv \alpha_\ell^{1-\eta} (1-\tau^m_{j\ell}) w_\ell \tilde{\varepsilon}_{jk\ell}$$

Choice probabilities, education, and wages

- LHS variables observed in data

$$\pi_{jk\ell} = \tilde{w}^{\theta}_{jk\ell} / \sum_{\hat{\ell}} \tilde{w}^{\theta}_{jk\hat{\ell}}$$

$$\overline{\text{educ}}_{jk\ell} = \mathbb{E}\left[e^* \mid \text{individuals choose } \ell\right]$$

$$\overline{\text{wage}}_{jk\ell} = \mathbb{E}\left[w_\ell h_{jk} \varepsilon^h_{jk\ell} e^\eta \epsilon \mid \text{individuals choose } \ell,\, e = e^*\right]$$

- Education and wages are increasing in labor market access

$$\overline{\text{educ}}_{jk\ell},\, \overline{\text{wage}}_{jk\ell} \propto \left(\sum_{\hat{\ell}} \tilde{w}^{\theta}_{jk\hat{\ell}} \right)^{\frac{1}{\theta(1-\eta)}} \equiv \text{MA}_{jk}$$

Equilibrium

- Base wages w_ℓ and prices p_ℓ

$$w_\ell = p_\ell A_\ell\,, \quad p_\ell = \left(\frac{Y}{Y_\ell}\right)^{\frac{1}{\sigma}}$$

- Agglomeration κ and congestion μ

$$A_\ell = \bar{A}_\ell H_\ell^\kappa\,, \quad \alpha_\ell = \bar{\alpha}_\ell \left(\sum_{j,k} N_{jk}\pi_{jk\ell}\right)^{-\mu}$$

Estimation

Human capital function (INPRES as IV)

$$\text{wage}_i \propto \text{hcap}_i = (\text{educ}_i)^{\eta}$$

$$\Downarrow$$

$$\log \text{wage}_{ijk} = \delta_j + \delta_k + \eta \log \text{educ}_{ijk} + \boldsymbol{C}_j T_k \boldsymbol{\phi} + \varepsilon_{ijk}$$

$$\log \text{educ}_{ijk} = \delta_j + \delta_k + \beta S_j T_k + \boldsymbol{C}_j T_k \boldsymbol{\phi} + \varepsilon_{ijk}$$

Education and migration costs (INPRES as DD)

$$1 + \tau^{e}_{jk} = (1 + S_j T_k)^{-\beta} \delta_j \delta_k (1 + \boldsymbol{C}_j T_k)^{\boldsymbol{\phi}}$$
$$1 - \tau^{m}_{j\ell} = (1 + d^{P}_{j\ell})^{-\varphi_1} (1 + d^{D}_{j\ell})^{-\varphi_2}$$

$$\Downarrow$$

$$\log \overline{\mathsf{educ}}_{jk\ell} - \log \overline{\mathsf{wage}}_{jk\ell} = \beta \log(1 + S_j T_k) - \log \delta_j - \log \delta_k - \boldsymbol{\phi} \log(1 + \boldsymbol{C}_j T_k) - \varphi_1 \log(1 + d^{P}_{j\ell}) - \varphi_2 \log(1 + d^{D}_{j\ell}) + \log \frac{\eta}{c} - \log \varepsilon^{c}_{jk\ell}$$

Other parameters (INPRES as moments)

$$\sum_{i=1}^{n}[y_i - \exp(x_i\hat{\beta})]x_i = 0$$

- Poisson pseudo-maximum likelihood (Santos Silva & Tenreyro 2006)
  - Common in spatial models to accommodate zeros in choice probabilities

$$\log \overline{\text{educ}}_{jk\ell} - \log \overline{\text{wage}}_{jk\ell}$$

$$\Delta_\ell \log \overline{\text{educ}}_{jk\ell}, \quad \Delta_\ell \log \overline{\text{wage}}_{jk\ell}, \quad \Delta_\ell \log \pi_{jk\ell}$$

INPRES treatment effects

Calibrated parameters (Bryan & Morten 2019)

- Agglomeration $\kappa = 0.05$
- Congestion $\mu = 0.075$
- Elasticity of substitution $\sigma = 8$

Estimated human capital function

	Treatment			Placebo		
	OLS	IV	First stage	OLS	IV	First stage
Log years of schooling	0.393***	0.688**		0.394***	-1.357	
	(0.00721)	(0.311)		(0.00678)	(3.523)	
INPRES × young			0.0284***			0.00564
			(0.00899)			(0.0110)
Observations	89,404	89,404	89,404	55,091	55,091	55,091
F-statistic			9.97			0.26

Estimated education and migration costs

	Treatment		Placebo	
	Estimate	SE	Estimate	SE
β	0.110**	(0.0467)	0.0514	(0.0457)
φ_1	0.0415***	(0.00353)	0.0388***	(0.00423)
φ_2	0.0184	(0.0500)	-0.0299	(0.0658)

Counterfactuals

Quantifying aggregate and distributional effects

- **Evaluate** relative to zero-construction counterfactual
- **Decompose** effects of mobility by mechanism
  - And separate each from the general equilibrium effects
  - Diff-in-diff avoids model but only captures net effects
- Study program **design**
  - By simulating alternative allocations of school construction

Computing aggregate output

- New schools $\rightarrow$ new prices, productivities, migration (algorithm in paper)
  - Adjustments to observed quantities, like in exact-hat algebra (Dekle et al. 2008)

$$Y_\ell(a) = \frac{1}{p_\ell} \sum_{j,k} N_{jk} \pi_{jk\ell}(a) \overline{\text{wage}}_{jk\ell}(a)$$

- Special case: zero agglomeration + perfect substitution ($\kappa = 0$, $\sigma \to \infty$)
  - Parameter β is enough! No need to estimate others.

$$Y'_\ell = \sum_{j,k} N_{jk} \pi_{jk\ell} \overline{\text{wage}}_{jk\ell} \left(\frac{1 + S'_j T_k}{1 + S_j T_k} \right)^{\frac{\beta\eta}{1-\eta}}$$

The program increased aggregate output by 8%

	Aggregate output
Zero construction	**1.00**
+ Direct effect of construction	1.02
+ Migration	1.03
+ Migration-induced schooling	1.07
+ New equilibrium wages	**1.08**

- Small gains without migration (direct effect) or without education (sorting)
  - Complementarity between education and migration
  - Gains from sorting are already large (Bryan et al. 2014)

With especially large benefits for rural students

	Inequality (people)
Zero construction	**1.00**
+ Direct effect of construction	0.99
+ Migration	0.98
+ Migration-induced schooling	0.96
+ New equilibrium wages	**0.95**

- Expanded opportunity for rural students with high marginal returns
  - Decreased inequality between rural and urban students by 5%

But also increased inequality across places by 12%

	Inequality (places)
Zero construction	**1.00**
+ Direct effect of construction	0.99
+ Migration	1.02
+ Migration-induced schooling	1.11
+ New equilibrium wages	**1.12**

- The program explicitly aimed to encourage regional convergence
  - But mobility places convergence in tension with output gains
  - Rural regions still enjoy net gains, but urban regions gain more

Equity-efficiency tradeoff under mobility

	Aggregate output	Inequality (people)	Inequality (places)
Actual INPRES allocation	1.08	0.95	1.12
Prioritizing rural regions	**1.09**	0.93	**1.14**
+ Halving migration costs	**1.13**	0.90	**1.18**
Prioritizing urban regions	1.04	0.97	1.06
+ Halving migration costs	1.08	0.93	1.12

- Rural construction generates large returns but widens rural-urban gap (still Pareto)
- Alternative: schools + roads, although rural out-migration will rise (not Pareto)

Conclusion

Summary

- Evaluating large-scale educational investment in spatial equilibrium
  - 62,000 primary schools built in 1970s Indonesia
- Aggregate output $\uparrow$ (8%), regional inequality $\uparrow$ (12%)
  - Big gains for rural students who leave rural regions behind