Educational Investment in Spatial Equilibrium: Evidence from Indonesia

Allan Hsiao
Princeton University

June 27, 2022

How does migration shape educational investment?

- Governments invest $3 trillion in education annually (World Bank 2022)
  - In Indonesia, 61,807 new primary schools (INPRES 1973-1978)

- Schools serve students locally, but graduates seek employment nationally
  1. Non-local incentives for individual investment
  2. Non-local effects of collective investment (at scale)

This paper

- Aggregate and distributional effects of the **INPRES program**
  - Difference-in-difference with long-run outcomes (Duflo 2001)
  - Spatial equilibrium model to decompose effects and redesign program

- **Complementarity between education and migration**
  1. Rural schooling depends on urban wages (non-local incentives)
  2. Rural schools increase urban output (non-local effects)

- Results: aggregate output $\uparrow$ (8%), inequality $\updownarrow$ (people $\downarrow$ 5%, places $\uparrow$ 12%)
  - Tension between returns to education and regional convergence

This paper

- Aggregate and distributional effects of the **INPRES program**
  - Difference-in-difference with long-run outcomes (Duflo 2001)
  - Spatial equilibrium model to decompose effects and redesign program

- **Complementarity between education and migration**
  1. Rural schooling depends on urban wages (non-local incentives)
  2. Rural schools increase urban output (non-local effects)

- Results: aggregate output $\uparrow$ (8%), inequality $\updownarrow$ (people $\downarrow$ 5%, places $\uparrow$ 12%)
  - Tension between returns to education and regional convergence

This paper

- Aggregate and distributional effects of the **INPRES program**
  - Difference-in-difference with long-run outcomes (Duflo 2001)
  - Spatial equilibrium model to decompose effects and redesign program

- **Complementarity between education and migration**
  1. Rural schooling depends on urban wages (non-local incentives)
  2. Rural schools increase urban output (non-local effects)

- Results: aggregate output $\uparrow$ (8%), inequality $\updownarrow$ (people $\downarrow$ 5%, places $\uparrow$ 12%)
  - Tension between returns to education and regional convergence

Literature

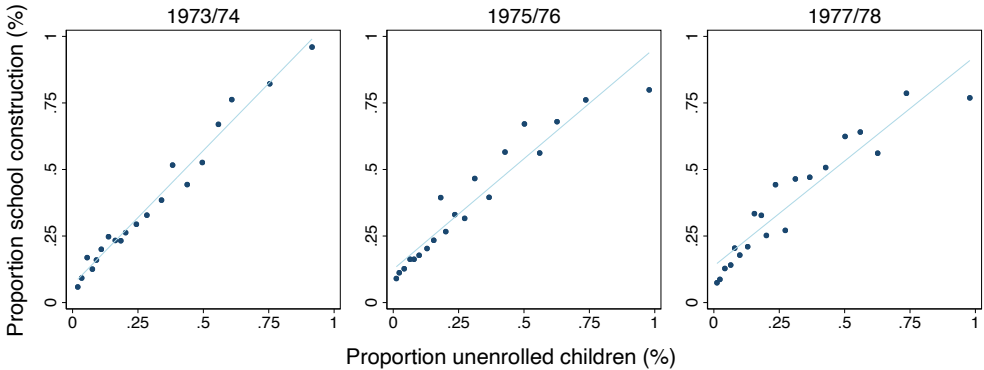
- **Education and migration** at scale in general equilibrium
  - Education: Khanna 2021, Dinerstein et al. 2022 (no migration)
  - Migration: Dahl 2002, Bryan et al. 2014, Bryan & Morten 2019 (no education)
  - Both: Eckert & Kleineberg 2021, Agostinelli et al. 2022 (no school construction)

- **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

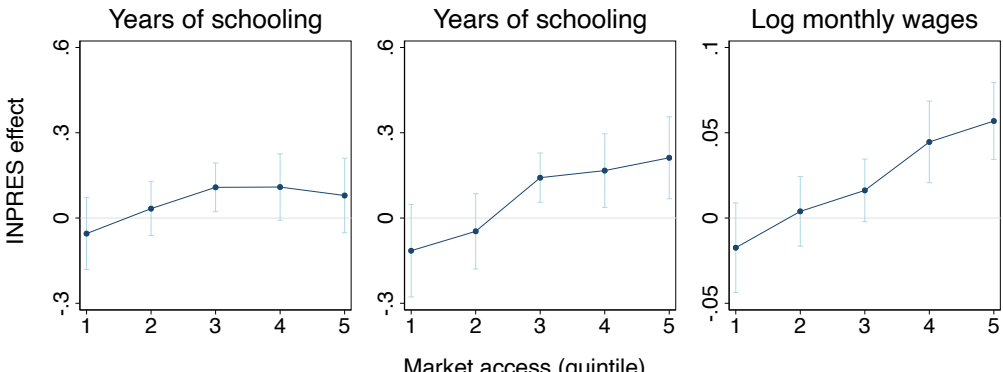
	Placebo		
Outcomes	Estimate	SE	Obs
Years of schooling	-0.0176	(0.0318)	196,308
— For wage earners	0.0120	(0.0566)	55,091
Log monthly wages	-0.00765	(0.00890)	55,091

Driven by labor market access (non-local incentives)

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

- Captures access to high urban wages
  - Population density in 1971 with inverse-distance weights

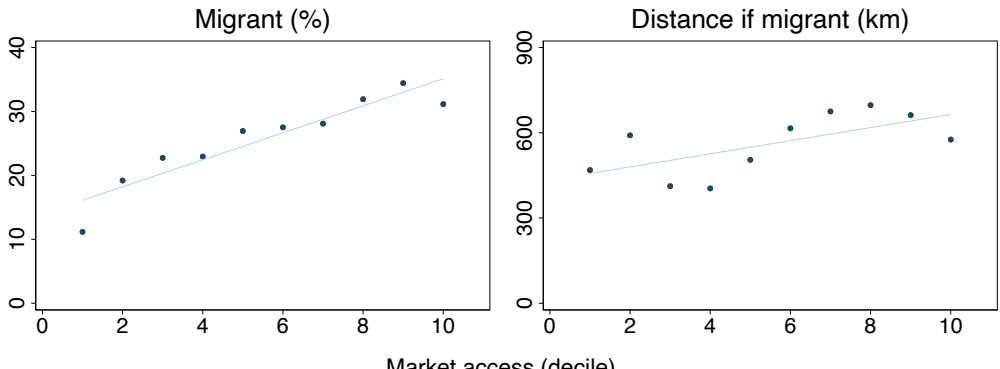
Driven by labor market access (non-local incentives)



Migration rates are high (non-local effects)

- Average migration rate of 26% and distance of 576 km
  - Increasing in labor market access
- Cross-province 16% vs. cross-state 31% in US (ACS 2013-2014)

Migration rates are high (non-local effects)



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

- Non-local incentives and effects still apply
- Consistent with model: INPRES affects costs of education, not migration
- Counterfactuals: INPRES effects under different migration costs

Model

Spatial equilibrium model

1. Government constructs schools
   - Build human capital that is portable (aggregate output)
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 (person-based inequality)
   - But rural students leave after graduation (place-based inequality)

Education and migration frictions

$$U(e,\epsilon) = \alpha_\ell \varepsilon_{jk\ell}^{\alpha} [\underbrace{(1-\widehat{\tau_{j\ell}^{m}}) w_\ell h_{jk} \varepsilon_{jk\ell}^{h} e^{\eta} \epsilon}_{\text{net labor income}} - \underbrace{(1+\widehat{\tau_{jk}^{e}}) c \varepsilon_{jk\ell}^{c} e}_{\text{cost of education}}]$$

(migration: $\tau_{j\ell}^{m}$; education: τ_{jk}^{e})

- Individual i, origin j, age cohort k, destinations ℓ
  - Each destination has $e_\ell^*(\epsilon_\ell) = \max_e U_\ell(e,\epsilon_\ell) \;\Rightarrow\; U_\ell(\epsilon_\ell) = U_\ell(e_\ell^*,\epsilon_\ell)$

- Compare destinations, then pick best to get choice probabilities

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

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

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

Aggregate and distributional effects

- 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)

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

Estimation

Human capital function (INPRES as IV)

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

$$\Downarrow$$

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

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

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

Education and migration costs (INPRES as DD)

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

$$\Downarrow$$

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

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)

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 $\sigma = 8$

Counterfactuals

Goals

- **Evaluate** relative to zero-construction counterfactual
  - Aggregate and distributional effects
- **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
  - Indonesia's INPRES program built 62,000 primary schools in 1970s
  - Aggregate output $\uparrow$ (8%), person-based inequality $\downarrow$ (5%), place-based $\uparrow$ (12%)
- Education and migration are **complementary**
  - Big gains for rural students who leave rural regions behind