Development Economics: Lecture 3

The 'Big Push', Infrastructure and Spatial Development

Niclas Moneke

niclas.moneke@economics.ox.ac.uk

University of Oxford

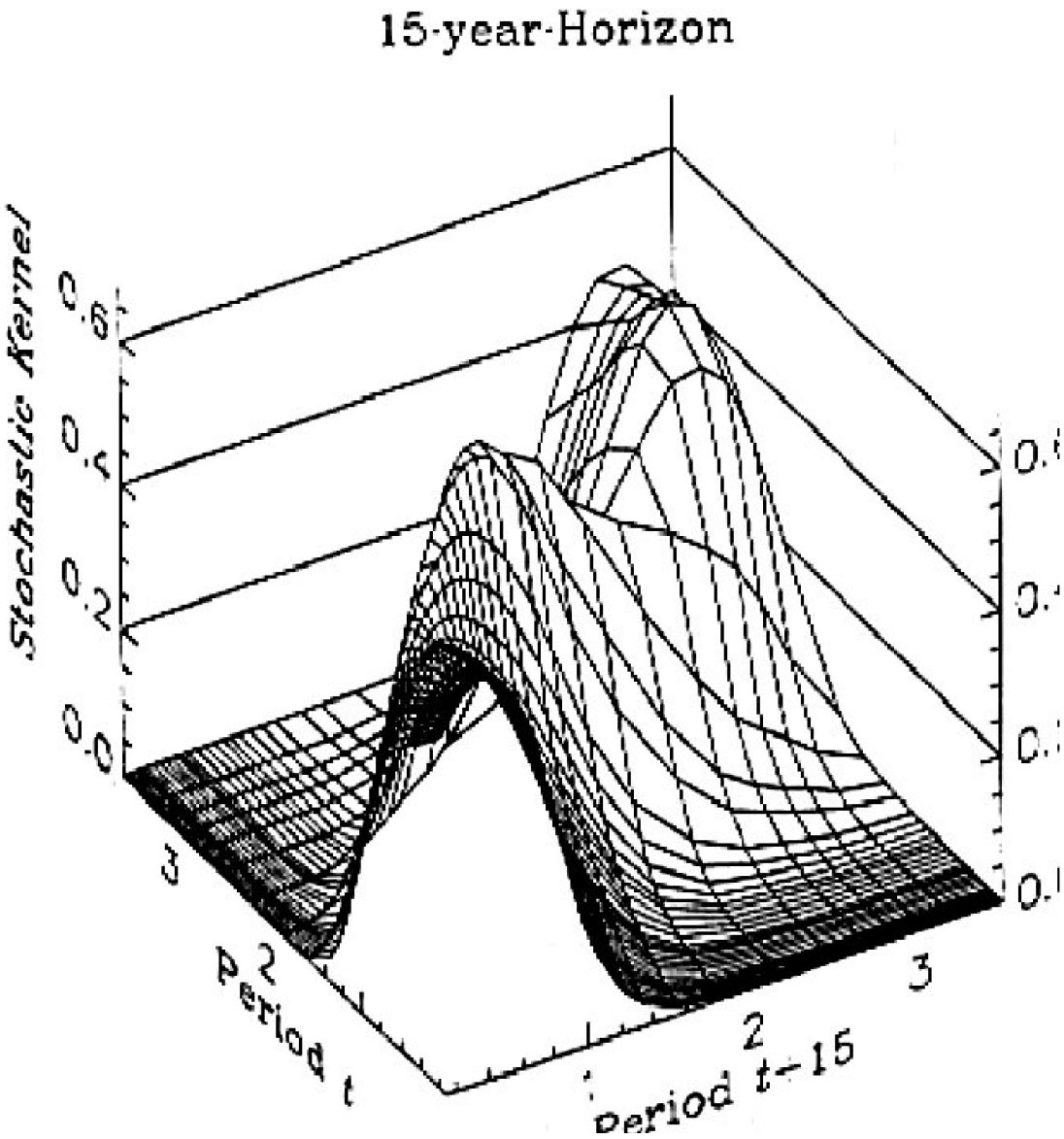
September 11, 2024

Structure of the course (days 1–5)

Topics 1-5 (Moneke)

- Topic 1 (Mon 09/09): Econ. Growth and Transformation
- Topic 2 (Tue 10/09): Poverty Traps and Policy Scale-up
- Topic 3 (Wed 11/09): Infrastructure and Spatial Development
- Topic 4 (Thu 12/09): Energy Access and Electrification Puzzle
- Topic 5 (Fri 13/09): Climate Change, Environment and Dev.

Poverty traps at macroeconomic scale? (Quah, 1993)



Poverty traps at macroeconomic scale? (Ray, 2014)

- Mobility matrix, 1982–2009

Cat 1: income $<$ 1/4 world av; Cat 2: between 1/4 and 1/2 world av; Cat 3: between 1/2 world av and world av; Cat 4: between world av and twice world av; Cat 5: income $>$ twice world av.

Obs	Cat	①	②	③	④	⑤
32	①	**84**	**13**	**3**	**0**	**0**
21	②	**43**	**43**	**14**	**0**	**0**
26	③	**0**	**27**	**50**	**23**	**0**
20	④	**0**	**0**	**20**	**70**	**10**
29	⑤	**0**	**0**	**0**	**3**	**97**

Rosenstein-Rodan's (1943) shoe factory

"Let us assume that 20,000 unemployed workers [...] are taken from the land and put into a large shoe factory. They receive wages substantially higher than their previous meagre income in natura. [...] If these workers spent all their wages on shoes, a market for the products of their enterprise would arise [...]. The trouble is that the workers will not spend all their wages on shoes.

If, instead, one million unemployed workers were taken from the land and put, not into one industry, but into a whole series of industries which produce the bulk of the goods on which the workers would spend their wages, what was not true in the case of one shoe factory would become true in the case of a whole system of industries: it would create its own additional market, thus realising an expansion of world output with the minimum disturbance of the world markets." (Rosenstein-Rodan (1943), pp. 205-206)

Smith and Young's complementarities

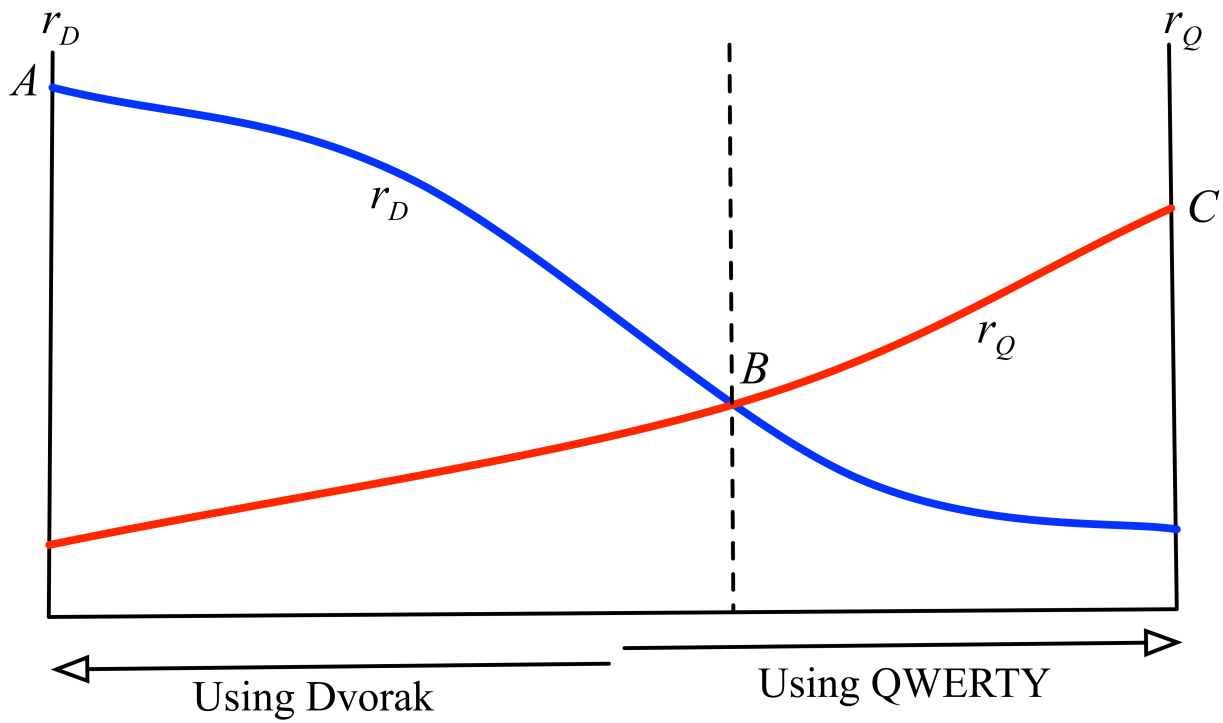
Smith: "The division of labor is limited by the size of the market."

Young (1928) noted that the converse also holds: The "division of labor" will determine the size of the market

$\rightarrow$ potential for virtuous cycle if coordination failure overcome

$\rightarrow$ economy with multiple equilibria: 'good' and 'bad' equilibrium

Strategic complements: technological externalities



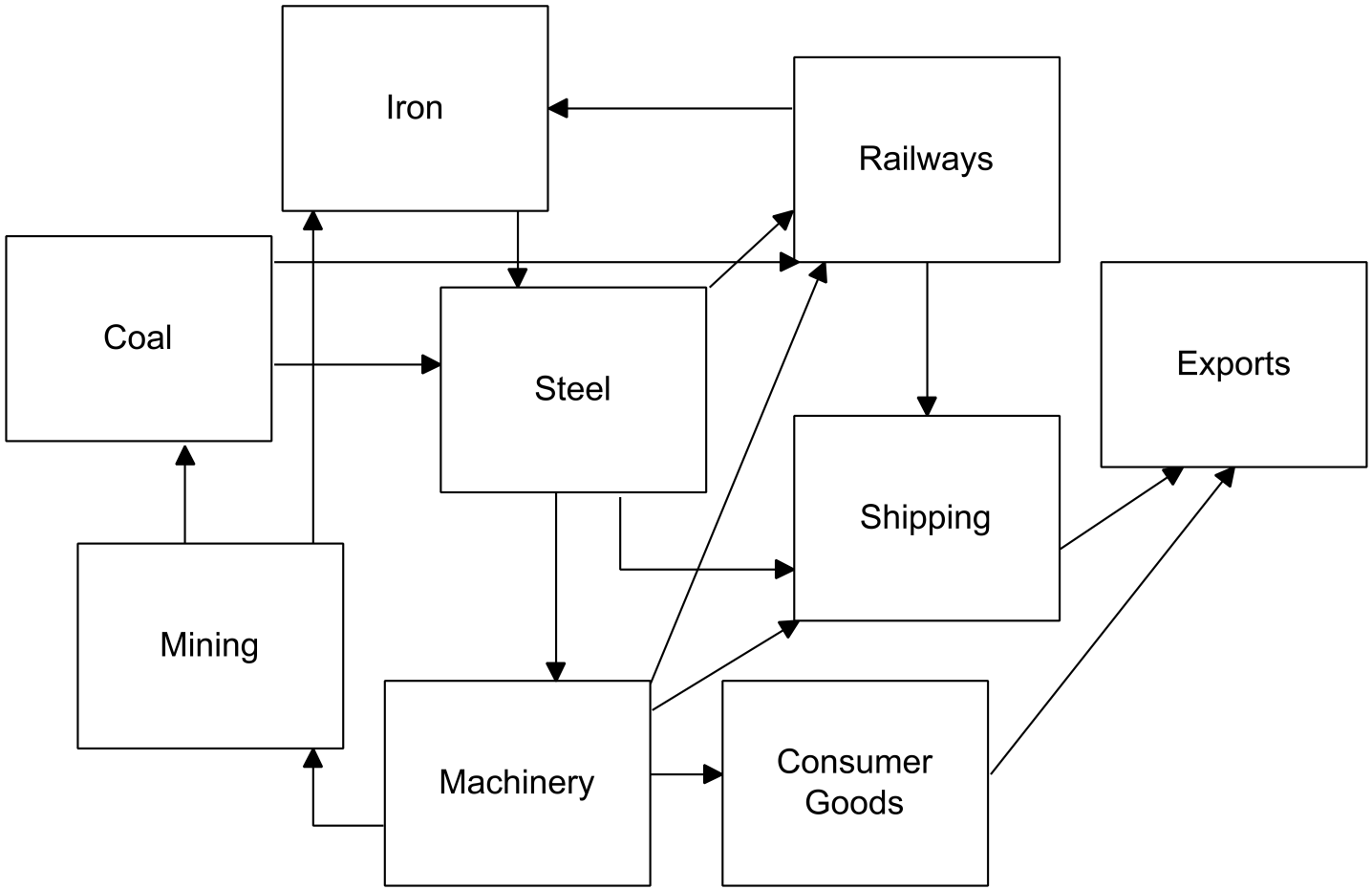
Coordination failure 1: demand complementarities

- industrial expansion raises income, generates demand for other industries (Rosenstein-Rodan, 1943)
- demand complementarity across producers of normal goods (cf. parable of Rosenstein-Rodan's (1943) shoe factory)
- complementarity pecuniary in nature, reflected in prices
- discussed prominently by other early development economists (cf. Fleming (1955), Hirschman (1958) and Nurkse (1953))

$\rightarrow$ demand complementarities foundation for policy of 'Big Push'

$\rightarrow$ one-off policy intervention to overcome coordination failure

Coordination failure 2: industry linkages (Hirschman, 1958)



Source: Hirschman, A. (1958). *The strategy of economic development*. Yale University Press.

Complementarities as source of multiple equilibria

- to obtain multiple equilibria, require a complementarity
- complementarity is a kind of externality
- for example:
  - two players, A and B
  - two actions, H high and L low
  - $u_B(H_B, H_A) - u_B(L_B, H_A) > u_B(H_B, L_A) - u_B(L_B, L_A)$
- interpretation:
  - "the gain to B of moving from low to high is larger when A plays high than when A plays low"
  - "the gain to B of moving from high to low is larger when A plays low than when A plays high"
  - actions H and L are 'strategic complements'

The nature of externalities

- at heart of strategic complements: pecuniary externalities
  (i.e. externalities that manifest themselves in prices)
- in contrast to technological externalities
  (e.g. direct, non-price effect, cf. QWERTY example)

$\rightarrow$ pecuniary externalities seem natural and compelling
(cf. Scitovsky (1954), who argued that they were dominant)

$\lightning$ however, pecuniary externalities hard to model with competitive markets

$\lightning$ violate First Fundamental Theorem of Welfare Economics

Recap: First Fundamental Theorem of Welfare

- person i has utility function u_i, endowment $x(i)$
- total endowment is $x \equiv \sum_i x(i)$
- vector of final consumption goods: $c \equiv \sum_i c(i)$
- production technology T converts x into c
- profit shares: $\theta(i)$ share of aggregate profits π for agent i
- prices: p for final goods, w for endowments

Recap, cont'd: First Fundamental Theorem of Welfare

Competitive equilibrium defined as:

1. $(p^*, w^*, c^*(i))$ exhibits profit maximization:

$$\pi \equiv p^* c^*(i) - w^* x \geq p^* c'(i) - w^* x' \ \forall (c', x' \in T)$$

2. and $c^*(i)$ maximises utility subject to budget constraint:

$$c^*(i) \text{ maximizes } u(c(i)) \text{ on } \{c(i) | p^* c(i) \leq w^* x(i) + \theta(i)\pi\}$$

Recap, cont'd: First Fundamental Theorem of Welfare

Competitive equilibrium defined as:

1. $(p^*, w^*, c^*(i))$ exhibits profit maximization:

$$\pi \equiv p^* c^*(i) - w^* x \geq p^* c'(i) - w^* x' \ \forall (c', x' \in T)$$

2. and $c^*(i)$ maximises utility subject to budget constraint:

$$c^*(i) \text{ maximizes } u(c(i)) \text{ on } \{c(i) | p^* c(i) \leq w^* x(i) + \theta(i)\pi\}$$

Theorem: a competitive equilibrium is Pareto optimal, so there can be no Pareto-ranked equilibria.

Recap, cont'd: First Fundamental Theorem of Welfare

Competitive equilibrium defined as:

1. $(p^*, w^*, c^*(i))$ exhibits profit maximization:
$$\pi \equiv p^* c^*(i) - w^* x \geq p^* c'(i) - w^* x' \ \forall (c', x' \in T)$$
2. and $c^*(i)$ maximises utility subject to budget constraint:
$$c^*(i) \text{ maximizes } u(c(i)) \text{ on } \{c(i) | p^* c(i) \leq w^* x(i) + \theta(i)\pi\}$$

Theorem: a competitive equilibrium is Pareto optimal, so there can be no Pareto-ranked equilibria.

- Proof: suppose not, then there exists allocation $c(i)$ such that:
  $c \equiv \sum_i c(i)$ is feasible, $(c, x) \in T$
  $u(c(j)) \geq u(c^*(j)) \ \forall j$, with strict inequality for some j

$\rightarrow$ $p^* c(j) \geq p^* c^*(j) \ \forall j$, with strict inequality for some j

$\rightarrow$ $p^* c - w^* x \geq p^* c^* - w^* x$, contradicting profit max. $\square$

Murphy et al. (1989): objective

- 'Big Push' lit. envisions multiple, Pareto-ranked equilibria
- requires economy to sustain two possible levels of industrialisation

$\rightarrow$ industrialisation must be individually unprofitable at low levels of aggregate industrialisation, but individual industrialisation profitable at high levels of aggregate industrialisation

Murphy et al. (1989): objective

- 'Big Push' lit. envisions multiple, Pareto-ranked equilibria
- requires economy to sustain two possible levels of industrialisation

$\rightarrow$ industrialisation must be individually unprofitable at low levels of aggregate industrialisation, but individual industrialisation profitable at high levels of aggregate industrialisation

$\lightning$ even individually unprofitable industrialisation must have externalities on other sectors that make industrialisation in other sectors more profitable

Murphy et al. (1989): objective

- 'Big Push' lit. envisions multiple, Pareto-ranked equilibria
- requires economy to sustain two possible levels of industrialisation

$\rightarrow$ industrialisation must be individually unprofitable at low levels of aggregate industrialisation, but individual industrialisation profitable at high levels of aggregate industrialisation

$\lightning$ even individually unprofitable industrialisation must have externalities on other sectors that make industrialisation in other sectors more profitable

$\rightarrow$ for now, this more stringent requirement envisioned by Rosenstein-Rodan (1943) et al. dropped

$\rightarrow$ focus only on own industrialisation with positive externalities

MSV (1989) with demand externality: setup

- model of demand-side externalities
- continuum of sectors, $i \in [0, 1]$
- identical individuals with labour endowment L and utility function:

$$\int_0^1 lnx(i)di$$

- if income is y, then y spent on every good i
- normalise wage to 1, $y = \pi + L$, (profits + labour income)

- each sector has two technologies, cottage and industrialised
- cottage: $x = l$, no setup cost
- industrialised: $x = \alpha l$, where $\alpha > 1$, and setup cost $F(i)$ for i

MSV (1989) with demand externality: setup cont'd

- CRS cottage technology and unit demand elasticity imply $p = 1$
- profit from industrialisation in sector i given by:

$$py - \frac{y}{\alpha} - F(i) = \frac{\alpha - 1}{\alpha} y - F(i) \equiv ay - F(i)$$

$\rightarrow$ larger income y conducive to industrialisation

- arrange sectors in increasing order of $F(i)$:
  - $F(0) = 0$
  - $F(1) = \infty$

$\rightarrow$ all sectors in $[0, n]$ industrialise

MSV (1989) with demand externality: equilibrium

- if n sectors industrialise at aggregate income $y(n)$, zero profit:

$$ay(n) - F(n) = 0$$

MSV (1989) with demand externality: equilibrium

- if n sectors industrialise at aggregate income $y(n)$, zero profit:

$$ay(n) - F(n) = 0$$

- then national income $y(n)$ given by:

$$y(n) = \int_0^n \pi(i)di + L = \int_0^n [ay(n) - F(i)]\, di + L = any(n) - nA(n) + L$$

where $A(n)$ is average of all fixed costs on $[0, n]$

MSV (1989) with demand externality: equilibrium

- if n sectors industrialise at aggregate income $y(n)$, zero profit:

$$ay(n) - F(n) = 0$$

- then national income $y(n)$ given by:

$$y(n) = \int_0^n \pi(i)di + L = \int_0^n [ay(n) - F(i)]\, di + L = any(n) - nA(n) + L$$

where $A(n)$ is average of all fixed costs on $[0, n]$

$\rightarrow$ therefore: $y(n) = \frac{L - nA(n)}{1 - an}$

- substitute in to zero profit: $[1 - an]F(n) + anA(n) = aL$

$\rightarrow$ derivative of LHS is $[1 - an]F'(n) > 0$, so solution unique

MSV (1989) with demand externality: equilibrium

- if n sectors industrialise at aggregate income $y(n)$, zero profit:

$$ay(n) - F(n) = 0$$

- then national income $y(n)$ given by:

$$y(n) = \int_0^n \pi(i)di + L = \int_0^n [ay(n) - F(i)]\, di + L = any(n) - nA(n) + L$$

where $A(n)$ is average of all fixed costs on $[0, n]$

$\rightarrow$ therefore: $y(n) = \frac{L - nA(n)}{1 - an}$

- substitute in to zero profit: $[1 - an]F(n) + anA(n) = aL$

$\rightarrow$ derivative of LHS is $[1 - an]F'(n) > 0$, so solution unique

$\lightning$ complementarity, but no multiplicity?

MSV (1989) with demand externality: conclusion

- ↯ complementarity, but no multiplicity?
- externality generated via payoffs alone
- if firm's payoff is positive, so is the externality → firm invests
- firm does not internalize the externality, but does not need to
- likewise for the case in which profits are negative

→ 'Big Push' literature around Rosenstein-Rodan (1943) requires source of complementarity beyond private profit alone

MSV (1989) with industry wage premium: setup

- assume wage premium in industry: $w = 1 + v$
- simplify fixed cost: $F(i) = F \ \forall i$ (easy to relax)
- profit from industrialising in any sector (when demand is y):

$$\pi = y - \frac{1+v}{\alpha} y - F(1+v)$$

MSV (1989) with industry wage premium: setup

- assume wage premium in industry: $w = 1 + v$
- simplify fixed cost: $F(i) = F \ \forall i$ (easy to relax)
- profit from industrialising in any sector (when demand is y):

$$\pi = y - \frac{1+v}{\alpha} y - F(1+v)$$

$\rightarrow$ no industrialisation equilibrium: $y(0) = L$, if

$$L\left(1 - \frac{1+v}{\alpha}\right) - F(1+v) \leq 0$$

MSV (1989) with industry wage premium: setup

- assume wage premium in industry: $w = 1 + v$
- simplify fixed cost: $F(i) = F \ \forall i$ (easy to relax)
- profit from industrialising in any sector (when demand is y):

$$\pi = y - \frac{1+v}{\alpha} y - F(1+v)$$

$\rightarrow$ no industrialisation equilibrium: $y(0) = L$, if

$$L\left(1 - \frac{1+v}{\alpha}\right) - F(1+v) \leq 0$$

$\rightarrow$ industrialisation equilibrium: $y(1) = \alpha(L - F)$, if

$$L\left(1 - \frac{1+v}{\alpha}\right) - F \geq 0$$

MSV (1989) with industry wage premium: setup

- assume wage premium in industry: $w = 1 + v$
- simplify fixed cost: $F(i) = F \ \forall i$ (easy to relax)
- profit from industrialising in any sector (when demand is y):

$$\pi = y - \frac{1+v}{\alpha} y - F(1+v)$$

$\rightarrow$ no industrialisation equilibrium: $y(0) = L$, if

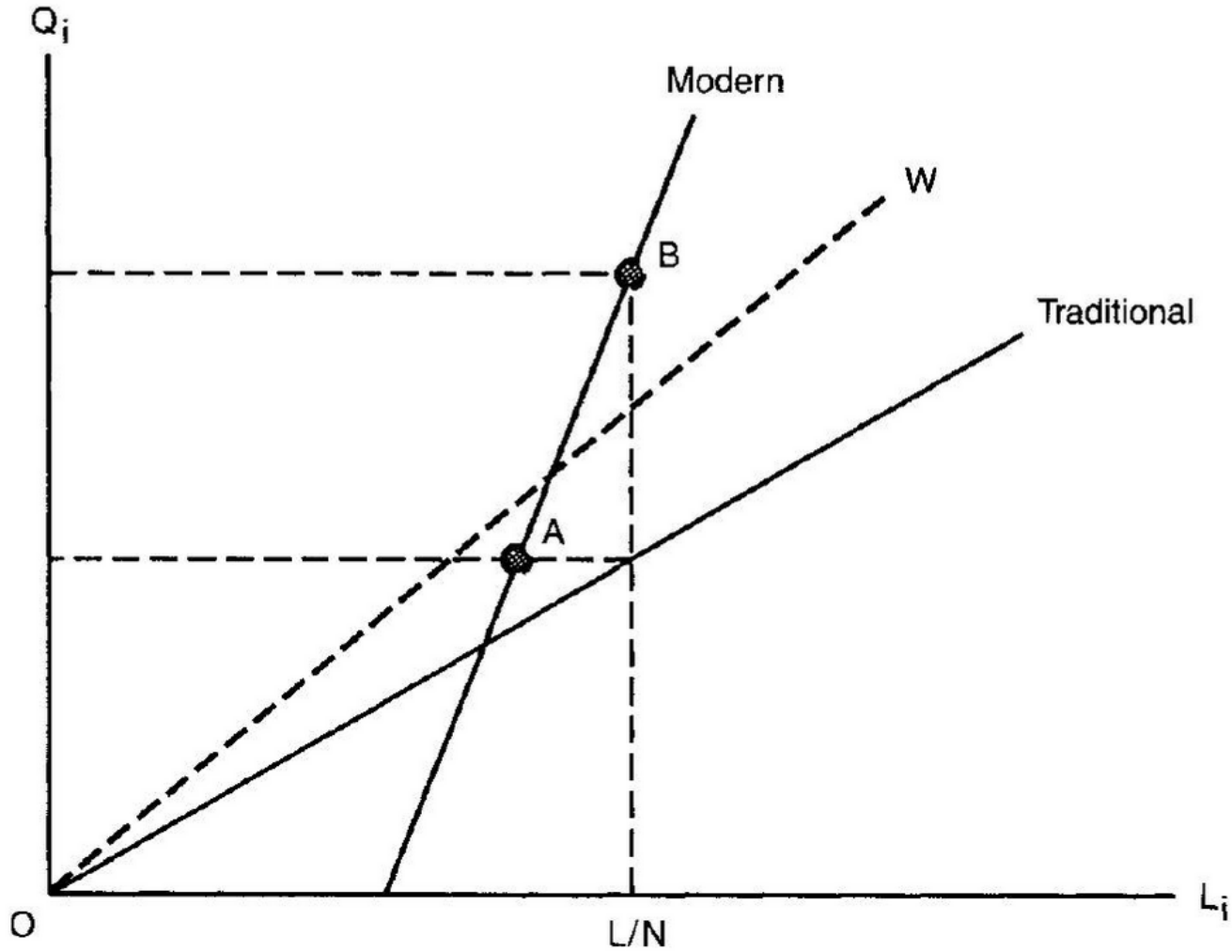
$$L\left(1 - \frac{1+v}{\alpha}\right) - F(1+v) \leq 0$$

$\rightarrow$ industrialisation equilibrium: $y(1) = \alpha(L - F)$, if

$$L\left(1 - \frac{1+v}{\alpha}\right) - F \geq 0$$

$\rightarrow$ multiple equilibria possible!

MSV (1989) with wage premium (Krugman (1995))



Source: Krugman, P. (1995). *Development, geography, and economic theory*.

MSV (1989) with wage premium (Krugman (1995))

- suppose all labour L/N either producing traditional or modern
- if all produce traditional, gain output Q_1
- if all produce modern, gain output Q_2, but $Q_2 > Q_1$
- first entrepreneur to produce modern (while everybody else produces traditional) will produce at point A

↯ however, in this example, point A lies below $0W$ line, so not profitable for modern entrepreneur (wage premium higher than benefit)

→ multiple equilibria for intermediate parameter values of wage premium or fixed cost

MSV (1989) with industry wage premium: conclusion

- complementarity that leads to multiple equilibria
- link between a firm's profit and its contribution to demand for products of other sectors is now broken!
- if firm industrialises and sets up factory, it pays a wage premium, thereby increasing the size of the market for all other producers $\rightarrow$ even if its own investment is losing money
- firm's profit not an adequate measure of its contribution to aggregate demand anymore
- extra wage paid by industrialising firm **not** captured in profits

MSV (1989) with industry wage premium: conclusion

- complementarity that leads to multiple equilibria
- link between a firm's profit and its contribution to demand for products of other sectors is now broken!
- if firm industrialises and sets up factory, it pays a wage premium, thereby increasing the size of the market for all other producers $\rightarrow$ even if its own investment is losing money
- firm's profit not an adequate measure of its contribution to aggregate demand anymore
- extra wage paid by industrialising firm **not** captured in profits

$\rightarrow$ 'Big Push' à la Rosenstein-Rodan (1943): firms underinvest in no-industrialisation equilibrium, resulting in inefficiency

$\rightarrow$ industrialisation equilibrium Pareto superior, since workers as well off as before as wage earners, but now also receive share of profits, ceteris paribus

MSV (1989) essentials: IRS plus elastic factor supply

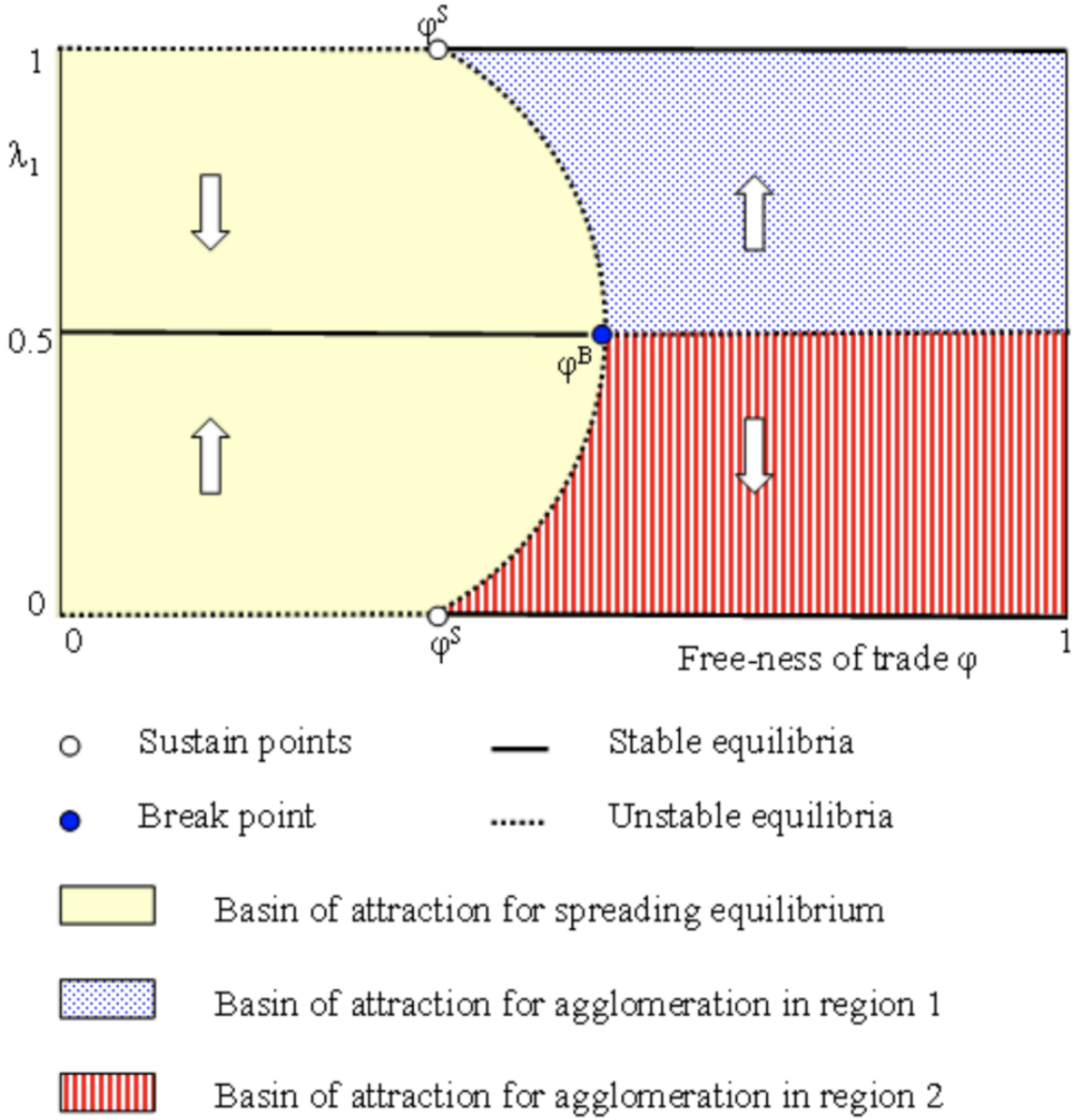
Two crucial assumptions to arrive at strategic complementarity:

- internal economies of scale in industry
  (here: IRS from fixed cost to set up)
- elastic factor supply to industry
  (here: underemployed factors from elsewhere, Lewis (1954))

$\rightarrow$ necessary conditions: interaction of IRS with elastic factor supply gives rise to complementarity ('external economies')

$\rightarrow$ still require assumption on market structure that ensures model will be tractable (here: monopolist deciding to enter)

Development as function of trade cost: multiple equilibria



Source: Fujita, M., Krugman, P., & Venables, A. (1999). *The spatial economy: Cities, regions and international trade*. MIT Press.

Do transport cost lock in 'bad' equilibria?

1. why do so many people work in such an unproductive sector?
   - large numbers of people in today's LICs work in agriculture (e.g. 60-80%)
   - employment share of agriculture $>$ GDP share of agriculture
2. within agriculture, why do so many people live in quasi-subsistence (as opposed to selling cash crops)?
   - majority in quasi-subsistence (e.g. 58% of HHs in Uganda)
   - produce identical staple crops on very small plots
   - sell small amount on local market, consume most themselves
3. what prevents the emergence of an alternative equilibrium?
   - a few farmers produce food on large farms
   - farmers sell food to cities
   - most people live in cities
   - city population produces non-agricultural goods

*** Asher, S., & Novosad, P. (2019). Rural roads and local economic development. *American Economic Review*, *110*(3), 797–823

Asher & Novosad (2019): rural roads & local development

Research question:

Does poor rural transport infrastructure prevent structural transformation, growth and poverty alleviation?

Setting:

- over a billion people around the world lack paved road access
- however, existing findings in literature inconclusive
- empirical test: Asher and Novosad (2019) estimate causal effect of India's national rural road construction program
- theoretically ambiguous: can decreasing transport costs **alone** cause local economic development?
  - road access enables trade in factors (intermediate inputs) and goods (import manufactured goods, export agricultural)
  - road access increases labour mobility (migration, commuting)
  - potentially missing complementary interventions, e.g. to increase agricultural productivity (Gollin & Rogerson, 2014)

Empirical challenges to analyse road construction

1. **road placement endogenous:** follows political and economic considerations
   → exploit discontinuities in treatment probability (fuzzy RDD) due to program rules

2. **data availability:** lack of data at required spatial resolution, i.e. require visibility at the village- or even household-level
   → new geo-identified microdata: 825m individuals in 600k villages

India's rural road expansion scheme: launched in 2000

- Pradhan Mantri Gram Sadak Yojana (PMGSY) scheme to connect all villages to road network – progress by 2015:
  - 113,000 roads constructed (approx. 400,000km)
  - 107,000 previously unconnected villages now connected
  - $37bn funds disbursed
- scheme funded by central government, but construction administered by states
- transparent, systematic electronic records, including details of every habitation and road built
- objective eligibility rules: villages over population of 500 or 1,000 (depending on state) prioritised

$\rightarrow$ fuzzy RDD: probability to receive road should increase disproportionately at population threshold(s)

RDD first stage: 22pp↑ in road treatment at threshold

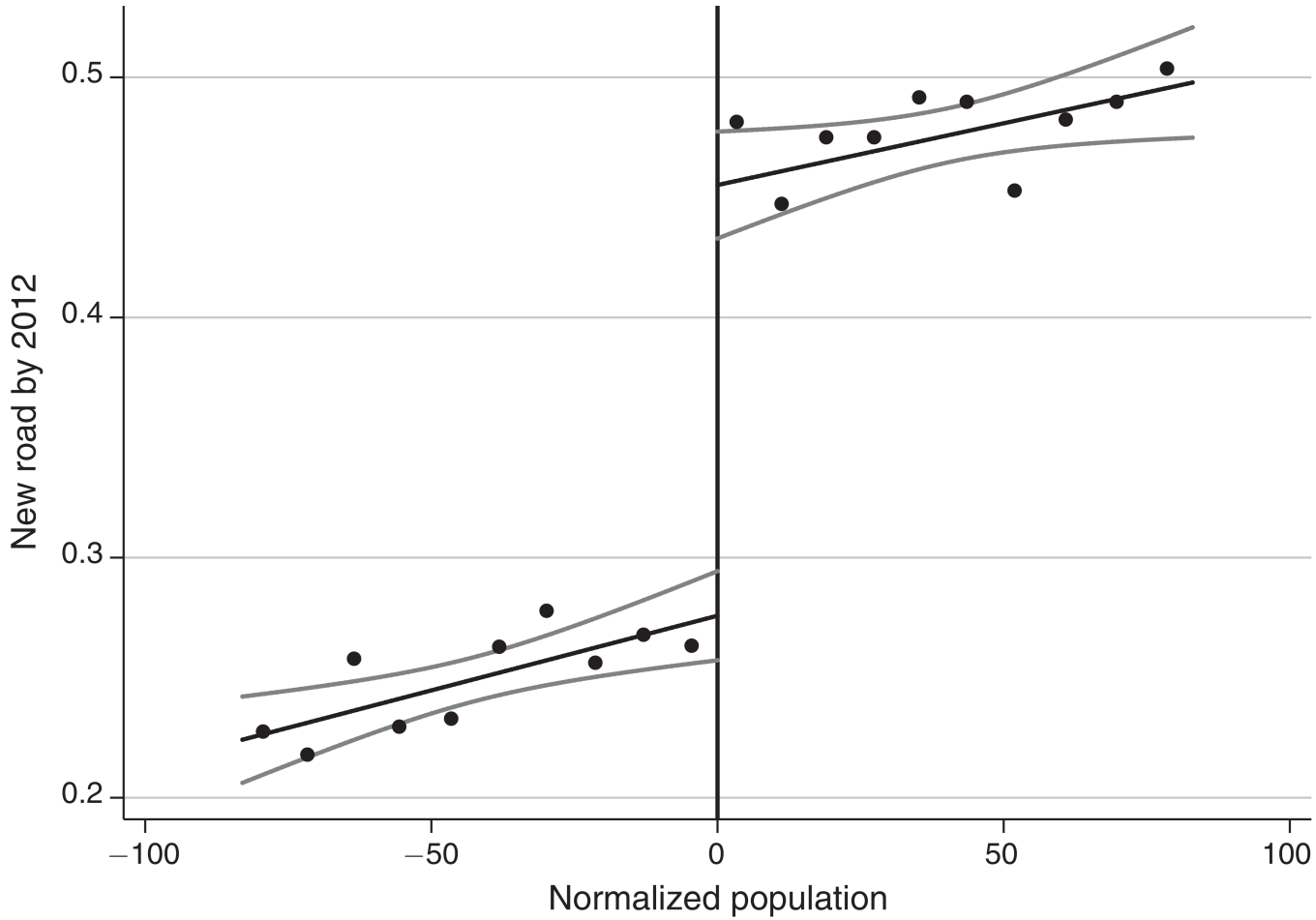


FIGURE 4. FIRST STAGE: EFFECT OF ROAD PRIORITIZATION ON PROBABILITY OF NEW ROAD BY 2012

Source: Asher & Novosad (2019), Rural roads and local economic development, *American Economic Review, 110(3), 797–823*.

Empirical strategy and specification: fuzzy RDD

- exploit discontinuities in villages' treatment probability
- sample restricted to villages without paved road in 2001
- estimate LATE for complier villages: villages that would not have obtained a road without meeting eligibility threshold

Empirical strategy and specification: fuzzy RDD

- exploit discontinuities in villages' treatment probability
- sample restricted to villages without paved road in 2001
- estimate LATE for complier villages: villages that would not have obtained a road without meeting eligibility threshold

$$Road_{v,j} = \gamma_0 + \gamma_1 \mathbb{1}\{pop_{v,j} \geq T\} + \gamma_2 (pop_{v,j} - T) + \gamma_3 (pop_{v,j} - T) \times \mathbb{1}\{pop_{v,j} \geq T\} + \nu X_{v,j} + \mu_j + \upsilon_{v,j} \quad (1)$$

$$Y_{v,j} = \beta_0 + \beta_1 Road_{v,j} + \beta_2 (pop_{v,j} - T) + \beta_3 (pop_{v,j} - T) \times \mathbb{1}\{pop_{v,j} \geq T\} + \zeta X_{v,j} + \eta_j + \epsilon_{v,j} \quad (2)$$

Empirical strategy and specification: fuzzy RDD

- exploit discontinuities in villages' treatment probability
- sample restricted to villages without paved road in 2001
- estimate LATE for complier villages: villages that would not have obtained a road without meeting eligibility threshold

$$Road_{v,j} = \gamma_0 + \gamma_1 \mathbb{1}\{pop_{v,j} \geq T\} + \gamma_2 (pop_{v,j} - T) + \gamma_3 (pop_{v,j} - T) \times \mathbb{1}\{pop_{v,j} \geq T\} + \nu X_{v,j} + \mu_j + \upsilon_{v,j} \quad (1)$$

$$Y_{v,j} = \beta_0 + \beta_1 Road_{v,j} + \beta_2 (pop_{v,j} - T) + \beta_3 (pop_{v,j} - T) \times \mathbb{1}\{pop_{v,j} \geq T\} + \zeta X_{v,j} + \eta_j + \epsilon_{v,j} \quad (2)$$

→ two-stage least squares, using local linear regression for villages near cutoff (Gelman & Imbens, 2014) in both stages

→ $X_{v,j}$ baseline controls, e.g. school, health centre, electricity, distance to town, irrigation share, illiteracy share, etc.

→ η_j district × threshold fixed effects

Rich high-resolution, geo-identified, cross-referenced data

- PMGSY administrative data (2000 – 2018): road connection status at village-level (and time of connection), including road characteristics
- GIS data: geo-coordinates of villages
- remote sensing: nightlights, NDVI, EVI
- population census (2001, 2011):
  - baseline village characteristics: controls, heterogeneity
  - outcome variables: amenities, population
- economic census (1998, 2005, 2013)
- socioeconomic census microdata (2012): 825m individuals, 166m households, 600k villages

RDD results: five different outcome classes

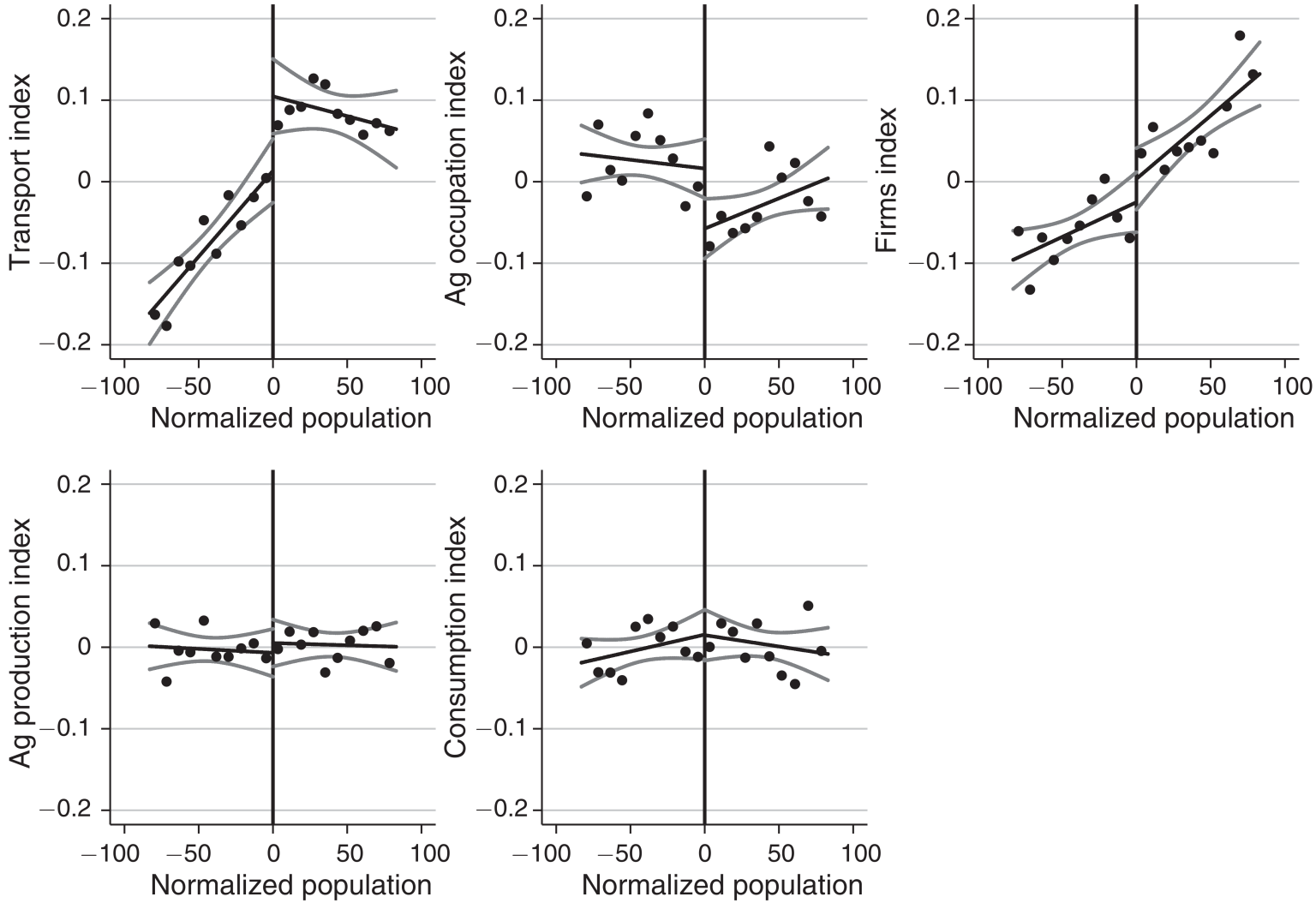


FIGURE 5. REDUCED FORM: EFFECT OF ROAD PRIORITIZATION ON INDICES OF MAJOR OUTCOMES

Source: Asher & Novosad (2019), Rural roads and local economic development, *American Economic Review, 110(3), 797–823*.

RDD results: structural transformation out of agriculture

TABLE 5—IMPACT OF NEW ROAD ON OCCUPATION AND INCOME SOURCE

	Occupation		Household income source	
	Agriculture	Manual labor	Agriculture	Manual labor
New road	−0.092	0.072	−0.030	−0.011
	(0.043)	(0.043)	(0.044)	(0.044)
Control group mean	0.476	0.448	0.418	0.507
Observations	11,432	11,432	11,432	11,432
R^2	0.28	0.26	0.31	0.28

Notes: This table presents regression discontinuity estimates from the main estimating equation of the effect of new road construction on occupational choice and household source of income. Column 1 estimates the impact on the share of workers in agriculture. Column 2 estimates the effect on the share of workers in manual labor (excluding agriculture). Columns 3 and 4 provide estimates of the impact of a new road on the share of households reporting cultivation and manual labor as the primary source of income. For each regression, the outcome mean for the control group (villages with population below the threshold) is also shown. The specification includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects (see Section IV for details). Heteroskedasticity robust standard errors are reported below point estimates.

Source: Asher & Novosad (2019), Rural roads and local economic development, *American Economic Review, 110(3), 797–823*.

RDD results: mostly landless young men leave agriculture

Table A5: Impact of new road on agricultural labor share by land, age, and gender

Panel A. Impact by household landholding

	Landless	0-2 Acres	2-4 Acres	4+ Acres
New road	-0.117	-0.100	-0.075	-0.063
	(0.047)	(0.052)	(0.054)	(0.053)
Control group mean	0.352	0.514	0.590	0.653
N	11101	10698	10380	9945
R2	0.22	0.18	0.19	0.22

Panel B. Impact by age and gender

	All		Male		Female	
	21-40	41-60	21-40	41-60	21-40	41-60
New road	-0.085	-0.093	-0.085	-0.094	-0.020	-0.044
	(0.045)	(0.045)	(0.045)	(0.044)	(0.056)	(0.061)
Control group mean	0.430	0.578	0.450	0.612	0.268	0.330
N	11421	11379	11410	11369	10781	10184
R2	0.28	0.29	0.28	0.29	0.21	0.24

Notes: This table presents regression discontinuity estimates from the main estimating equation of the effect of new road construction on occupational choice. The dependent variable in each regression is the share of workers in agriculture, for that specific category. Panel A examines whether treatment effects vary by the size of the household landholding. Column 1 estimates the impact for workers in households without agricultural land, Column 2 for workers in households with greater than 0 acres but but weakly less than two acres, Column 3 for workers in households with more than 2 acres but weakly less than 4 acres, and Column 4 for households with 4 or more acres of land. Panel B examines whether treatment effects vary by age and gender. The first two columns present results for workers aged 21-40 and 41-60. The next two present the same results for males workers only, while the final two present the same results for female workers. For each regression, the outcome mean for the control group (villages with population below the threshold) is also shown. The specification includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects (see Section V for details). Heteroskedasticity robust standard errors are reported below point estimates.

Source: Asher & Novosad (2019), Rural roads and local economic development, *American Economic Review, 110(3), 797–823.*

RDD results: firm outcomes imprecise/inconclusive

TABLE 6—IMPACT OF NEW ROAD ON FIRMS

	Total	Livestock	Manufacturing	Education	Retail	Forestry
Panel A. Log employment growth, by sector						
New road	0.273	0.252	0.260	0.198	0.333	−0.107
	(0.159)	(0.188)	(0.193)	(0.143)	(0.154)	(0.107)
Observations	10,678	10,678	10,678	10,678	10,678	10,678
R^2	0.30	0.42	0.23	0.18	0.23	0.35
Panel B. Level employment growth, by sector						
New road	4.219	−1.962	2.802	0.686	1.831	2.381
	(7.596)	(3.364)	(3.794)	(0.973)	(1.534)	(4.002)
Mean employment (level)	32.1	6.9	5.8	5.1	4.5	2.8
Observations	10,678	10,678	10,678	10,678	10,678	10,678
R^2	0.30	0.46	0.18	0.13	0.17	0.36

Notes: This table presents regression discontinuity estimates from the main estimating equation of the effect of new road construction on employment in in-village nonfarm firms. Panel A examines the impact on log employment in all nonfarm firms (column 1) and in the five largest sectors in our sample: livestock, manufacturing, education, retail, and forestry. Panel B presents estimates for the same regressions, instead specifying the level of employment as the dependent variable. The specification includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects (see Section IV for details). Heteroskedasticity robust standard errors are reported below point estimates.

Source: Asher & Novosad (2019), Rural roads and local economic development, *American Economic Review, 110(3), 797–823*.

RDD results: consumption, earnings & assets unchanged

TABLE 8—IMPACT OF NEW ROAD ON PREDICTED CONSUMPTION, EARNINGS, AND ASSETS

	Consumption per capita (log)	Poverty rate	Night lights (log)	Share of HH earning ≥ INR 5k	
Panel A. Consumption and earnings					
New road	0.022	−0.010	0.033	−0.001	
	(0.038)	(0.042)	(0.165)	(0.032)	
Control group mean	9.571	0.282	1.444	0.147	
Observations	11,432	11,432	11,102	11,432	
R^2	0.41	0.30	0.66	0.25	
	Asset index	Solid house	Refrigerator	Vehicle	Phone
Panel B. Asset ownership					
New road	0.107	0.033	0.005	−0.001	0.033
	(0.132)	(0.029)	(0.013)	(0.023)	(0.041)
Control group mean	−0.015	0.222	0.036	0.140	0.443
Observations	11,432	11,432	11,432	11,432	11,432
R^2	0.52	0.67	0.26	0.38	0.48

Notes: This table presents regression discontinuity estimates from the main estimating equation of the effect of new road construction on various measures of welfare. Panel A examines the impact on measures of predicted consumption and earnings. We use imputed log consumption per capita (outcome for column 1, see online Appendix for details of variable construction) and share of the population below the poverty line (column 2). The dependent variable for column 3 is the log of mean total night light luminosity in 2011–2013, with an extra control for log light at baseline in 2001. The dependent variable for column 4 is the share of households whose highest earning member earns more than INR 5,000 per month. Panel B examines the impact on asset ownership as measured in the 2012 SECC. The dependent variable for column 1 is the village-level average of the primary component of indicator variables for all household assets measured in the SECC. The remaining four columns present estimates for the impact on the share of households in the village that own each of these assets. The specification includes baseline village-level controls for amenities and economic indicators, as well as district-cutoff fixed effects (see Section IV for details). Heteroskedasticity robust standard errors are reported below point estimates for all estimates except for consumption and poverty, which report bootstrapped standard errors as described in the online Appendix.

Source: Asher & Novosad (2019), Rural roads and local economic development, *American Economic Review, 110(3), 797–823.*

Asher & Novosad (2019): rural roads expensive & useless?

Takeaways:

- estimate the local economic impact of the largest rural road scheme in history
- transport cost significantly constrain participation in non-agricultural labour markets
- however, overall effects disappointing: small, noisy effects on consumption/assets, agriculture, firms, source of income
- even with better infrastructure, villages may not be in a position to grow out of poverty

Asher & Novosad (2019): what are we missing?

Takeaways, continued:

→ unclear what we are missing, if anything
- transport cost reduced?
  (e.g. intermediary market power)
- lifting one binding constraint, but tightening another?
  (e.g. credit, education)
- multiple, interacted market failures?
  (e.g. nutrition, storage, savings, land)
  → complementary investments?
  (e.g. agricultural productivity: fertilizer, tractors, cash crops)
- time horizon?
  (e.g. long time for effects to unfold, especially firm growth)
- spatial general equilibrium forces?
  (e.g. some places lose, others gain, aggregate welfare unclear)

→ what was our prior regarding the effects of constructing a road to remote, small, poor rural places?

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

Donaldson (2018): railroads of the Raj

Research questions:

1. How large are the economic benefits of transportation infrastructure expansions (which aim to reduce trade costs)?
2. What economic mechanism explains these benefits?

Setting:

- more World Bank funding for transport infrastructure projects than for education, health and social services combined
- here: study large-scale expansion of transport infrastructure, construction of railroad network in colonial India ("the Raj")
- construct new dataset on Indian economy during this period:
  - outstanding effort: manual archival data mining
  - output, prices, internal and external trade
  - annual (1861–1930) at district-level (N = 239)
- employ GE spatial/trade model to guide empirical approach
  - Ricardian model of trade, i.e. driven by comparative advantage
  - derive **four testable predictions** from model

Donaldson (2018): model-guided empirical analysis

Donaldson proceeds carefully:

1. validate workings of model empirically, railroads supposed to:
   - reduce transport cost [Step 1]
   - increase trade flows [Step 2]
   - improve real incomes [Step 3]

Donaldson (2018): model-guided empirical analysis

Donaldson proceeds carefully:

1. validate workings of model empirically, railroads supposed to:
   - reduce transport cost [Step 1]
   - increase trade flows [Step 2]
   - improve real incomes [Step 3]
2. then ask how much of railroad–income effects can be plausibly explained as having arisen through model mechanism [Step 4]
   - model predicts gains from trade to drive railroad–income effect
   - but railroads can affect many dimensions that increase incomes (e.g. labour mobility, technology/ideas, capital, demand shock)

Donaldson (2018): model-guided empirical analysis

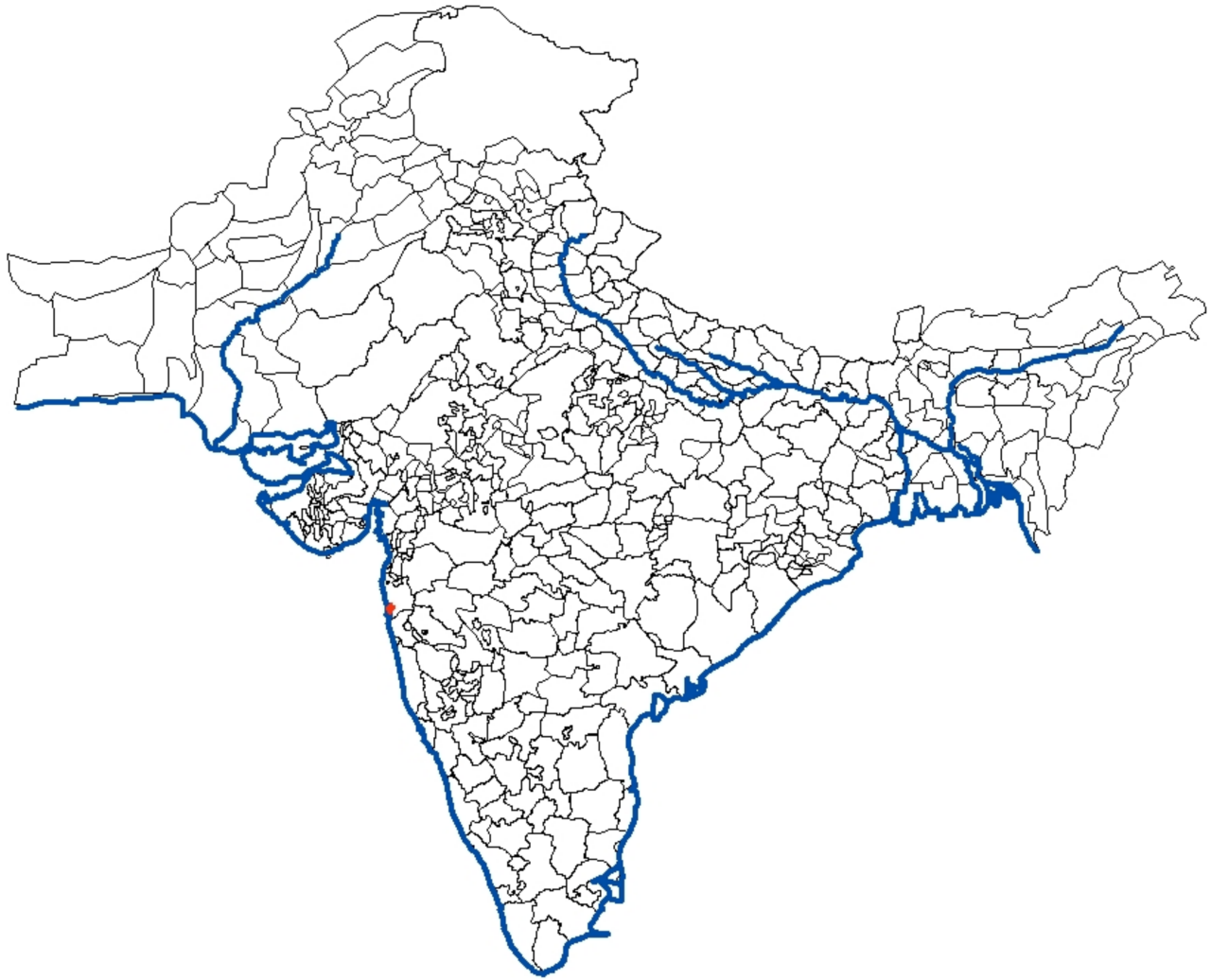
Donaldson proceeds carefully:

1. validate workings of model empirically, railroads supposed to:
   - reduce transport cost [Step 1]
   - increase trade flows [Step 2]
   - improve real incomes [Step 3]
2. then ask how much of railroad–income effects can be plausibly explained as having arisen through model mechanism [Step 4]
   - model predicts gains from trade to drive railroad–income effect
   - but railroads can affect many dimensions that increase incomes (e.g. labour mobility, technology/ideas, capital, demand shock)

→ gain new insights into **how exactly** transport infrastructure may affect growth and development

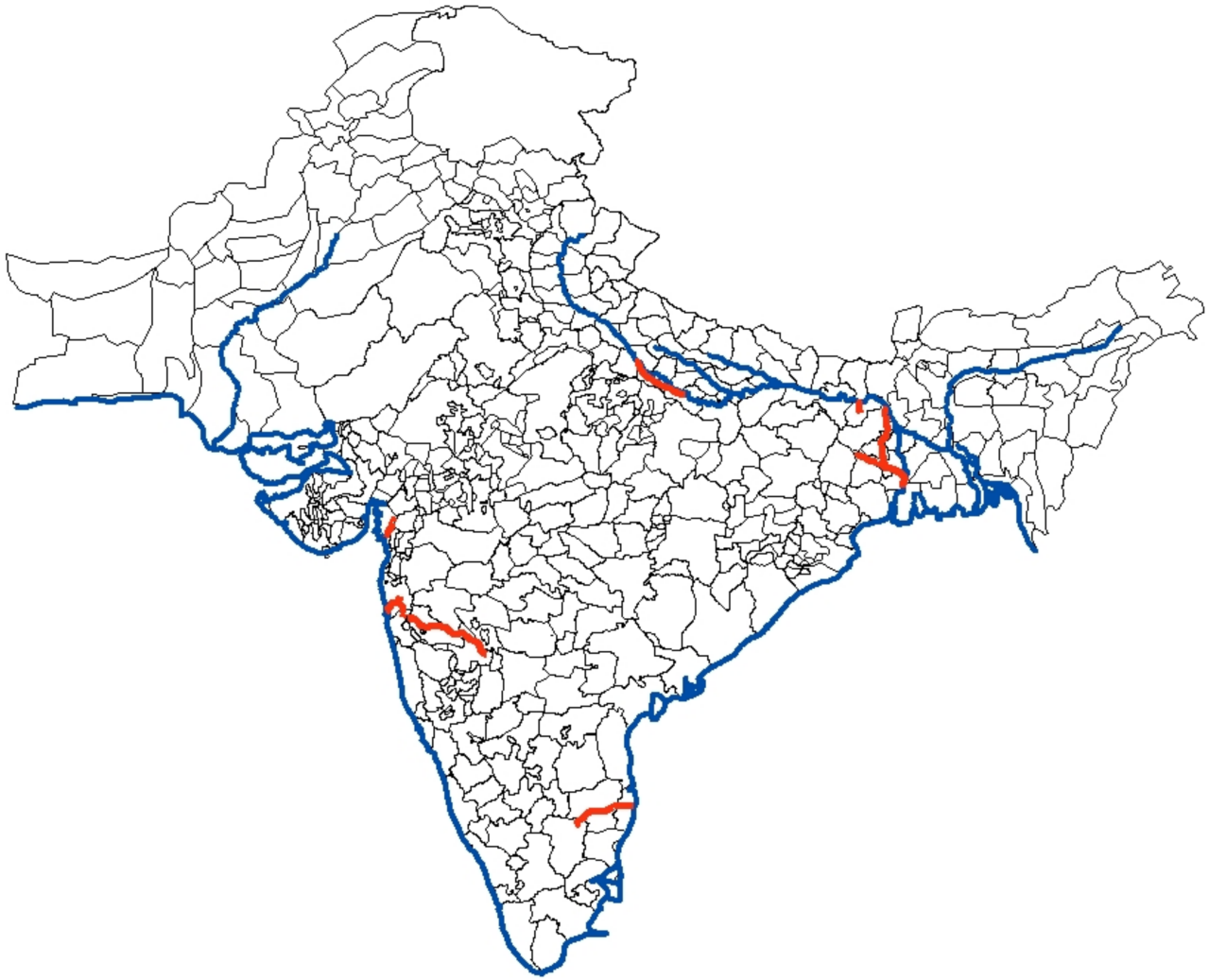
→ precisely the kind of structure previously missing in literature

Donaldson (2018): Indian railroad network in 1853



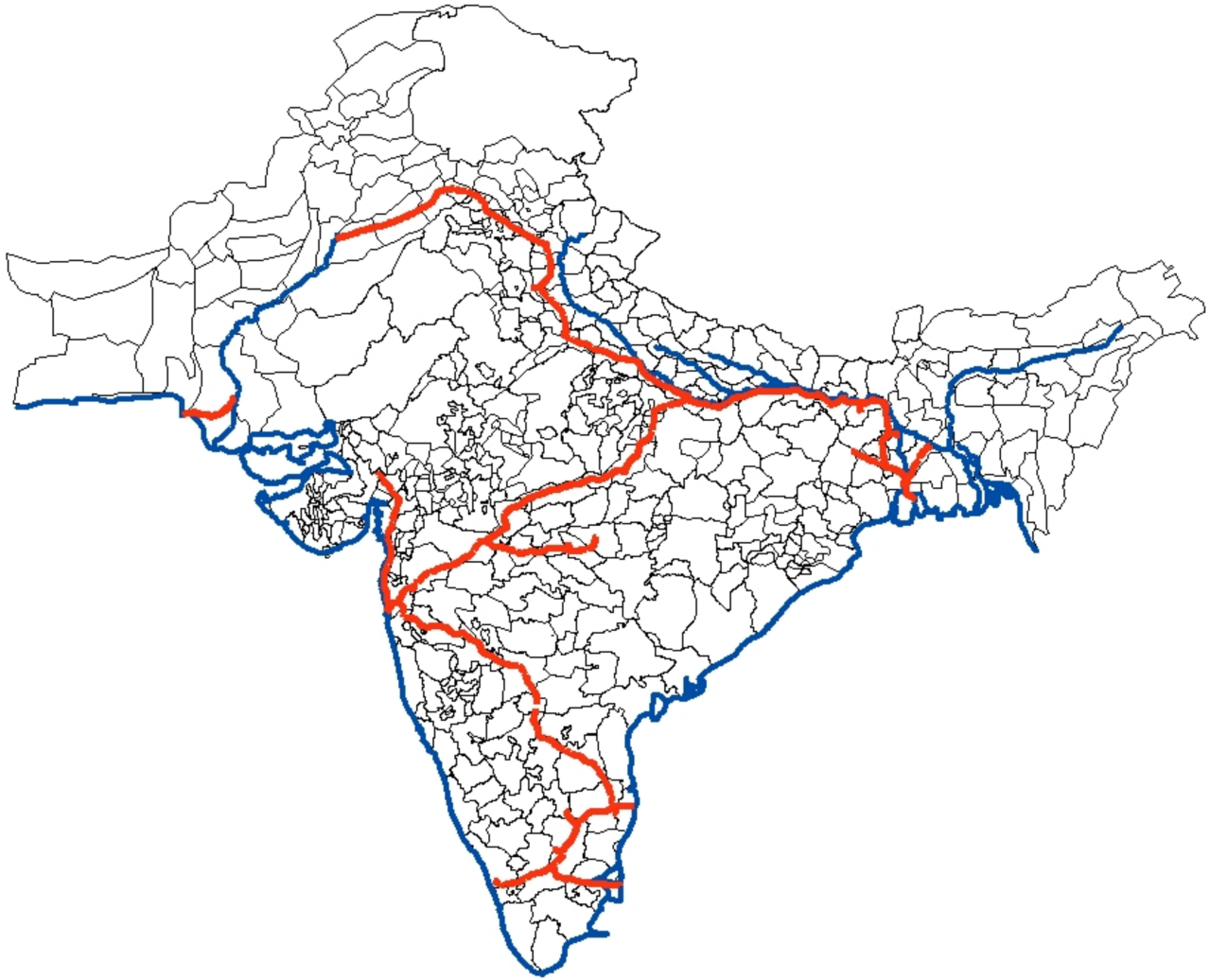
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1860



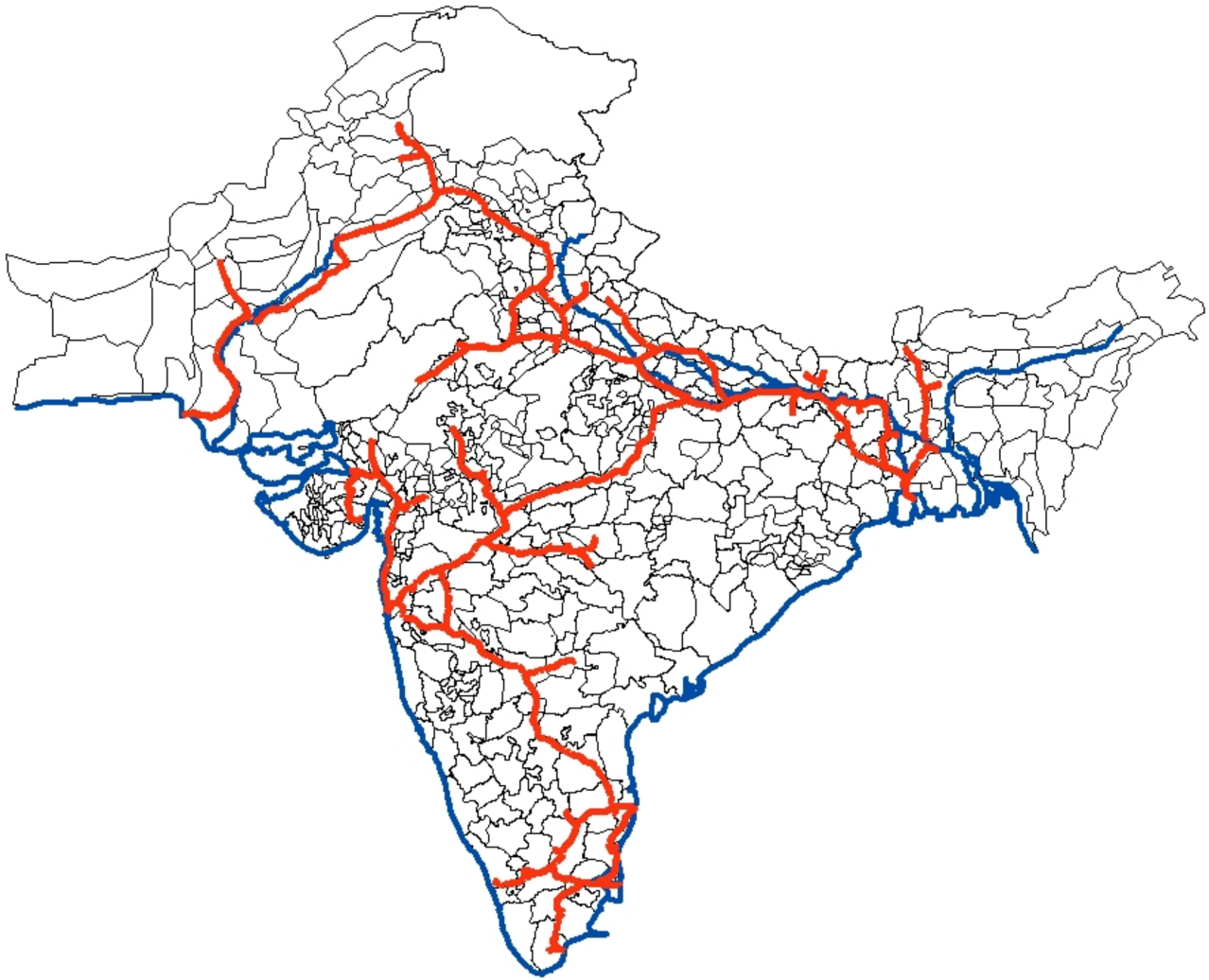
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1870



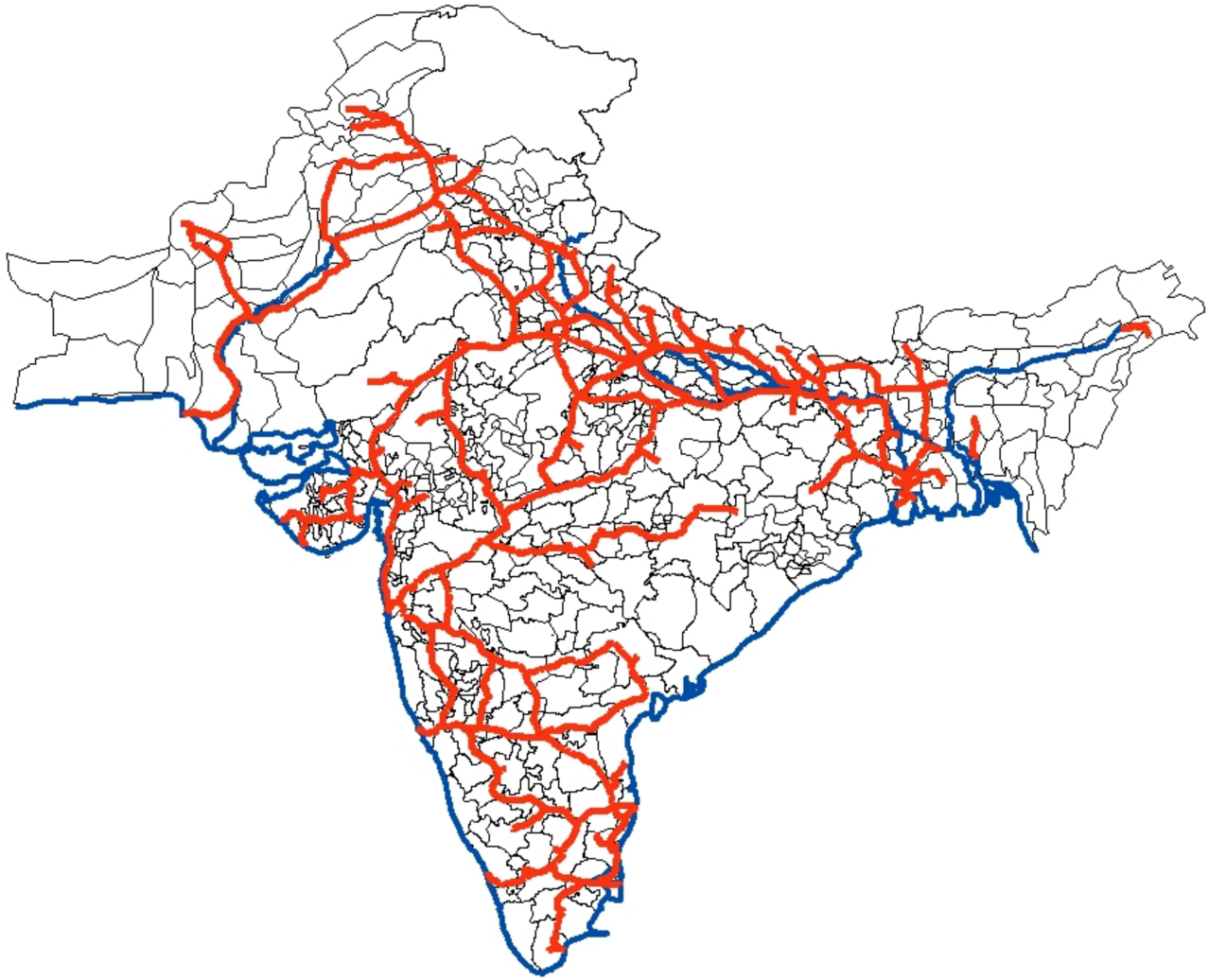
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1880



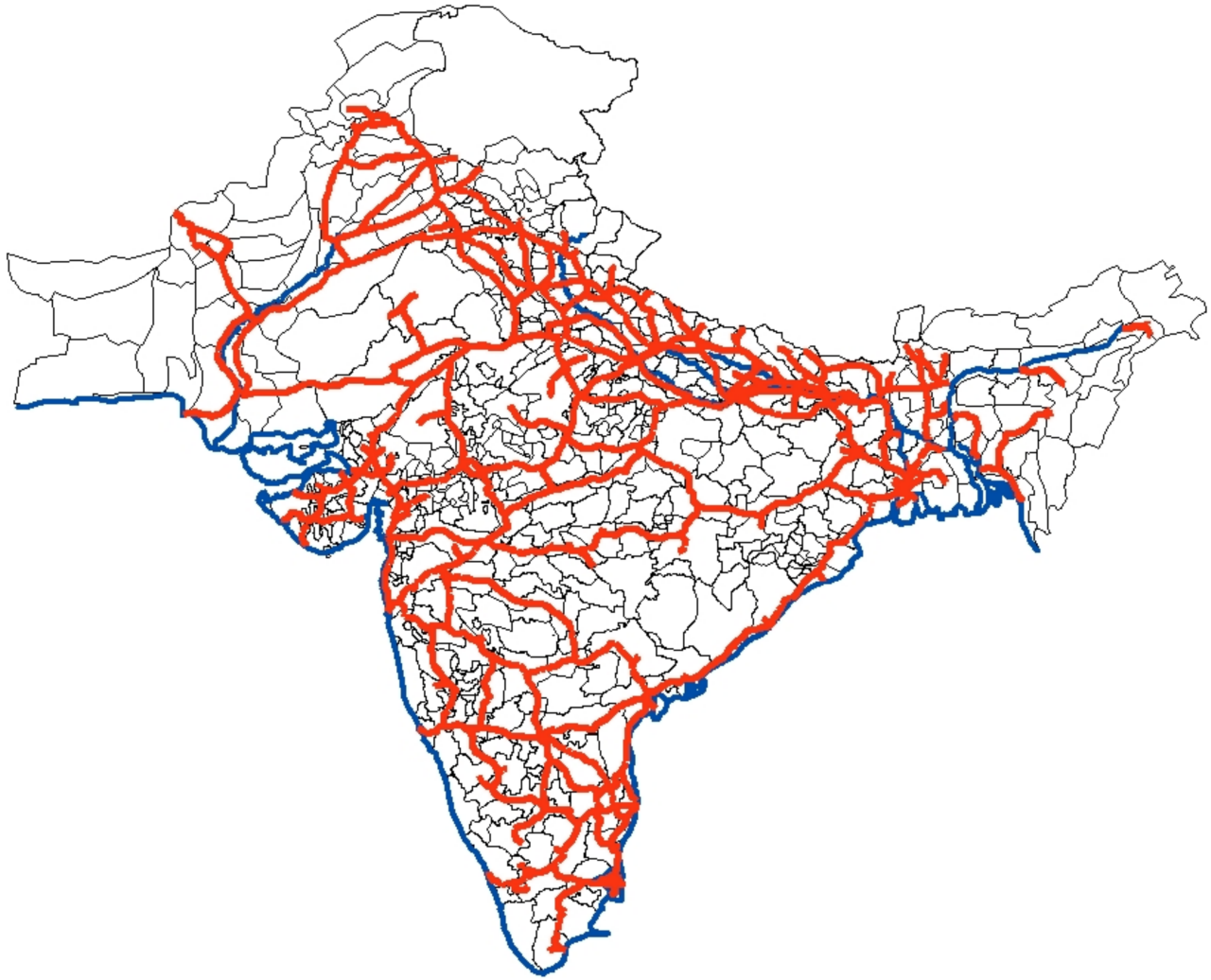
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1890



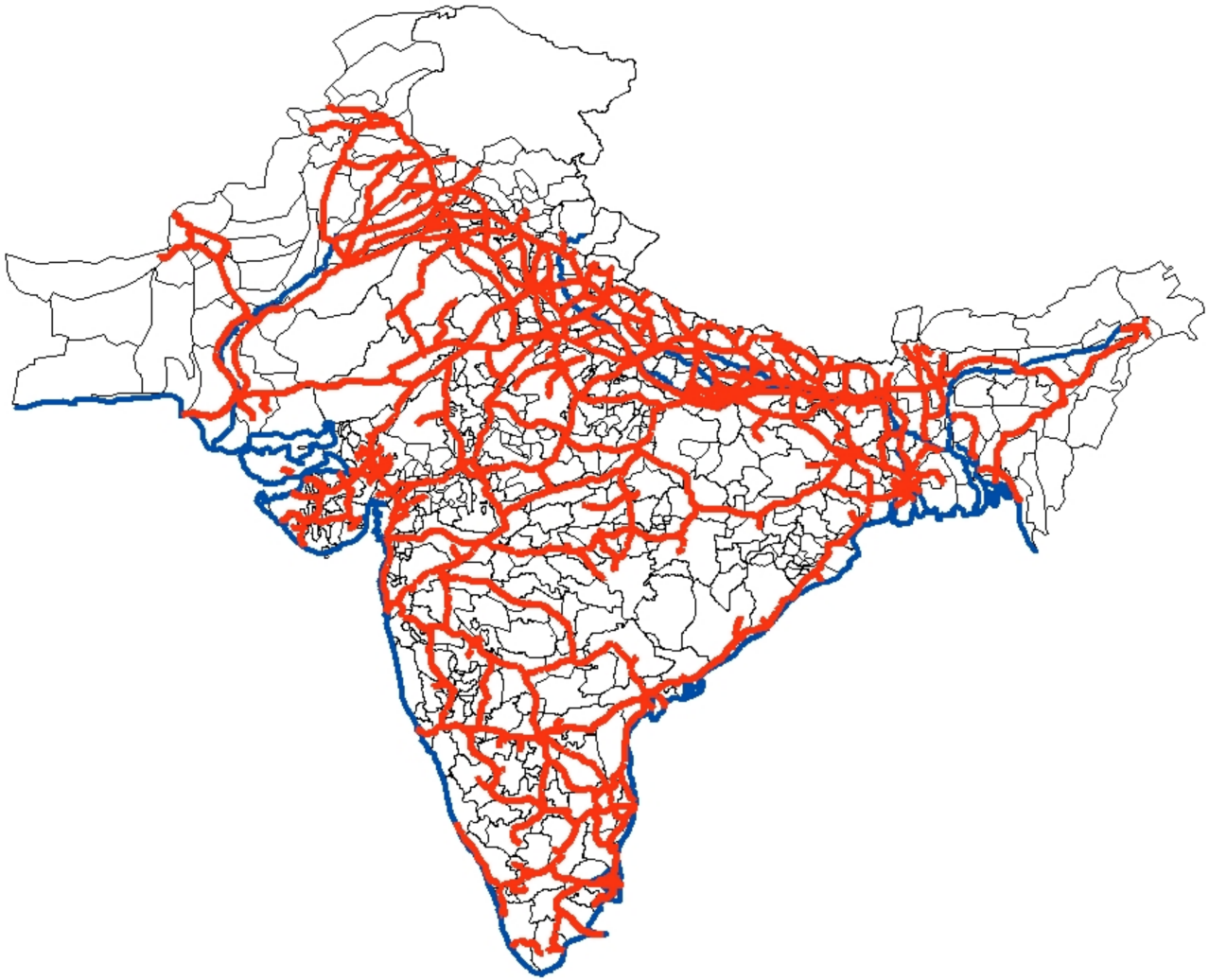
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1900



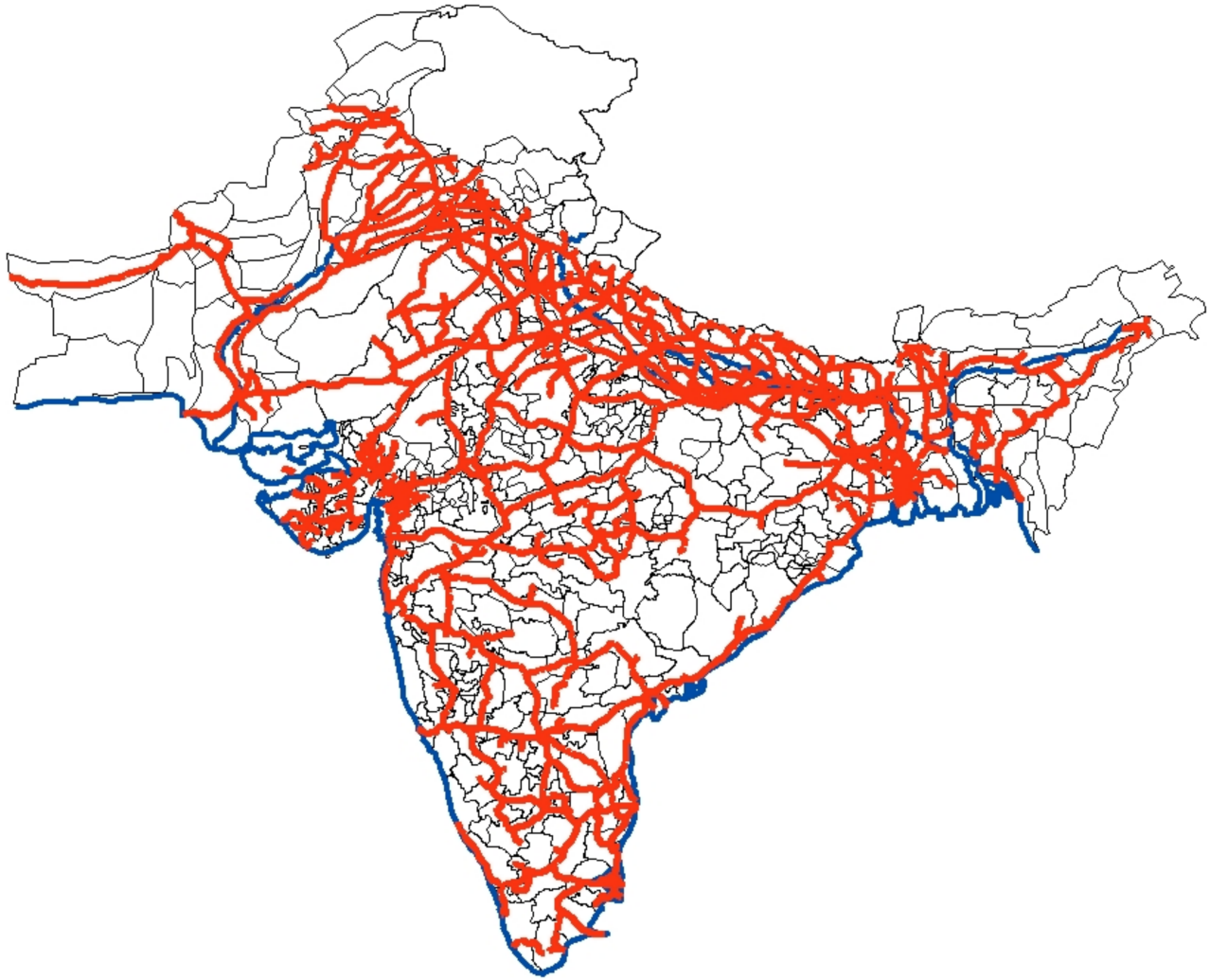
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1910



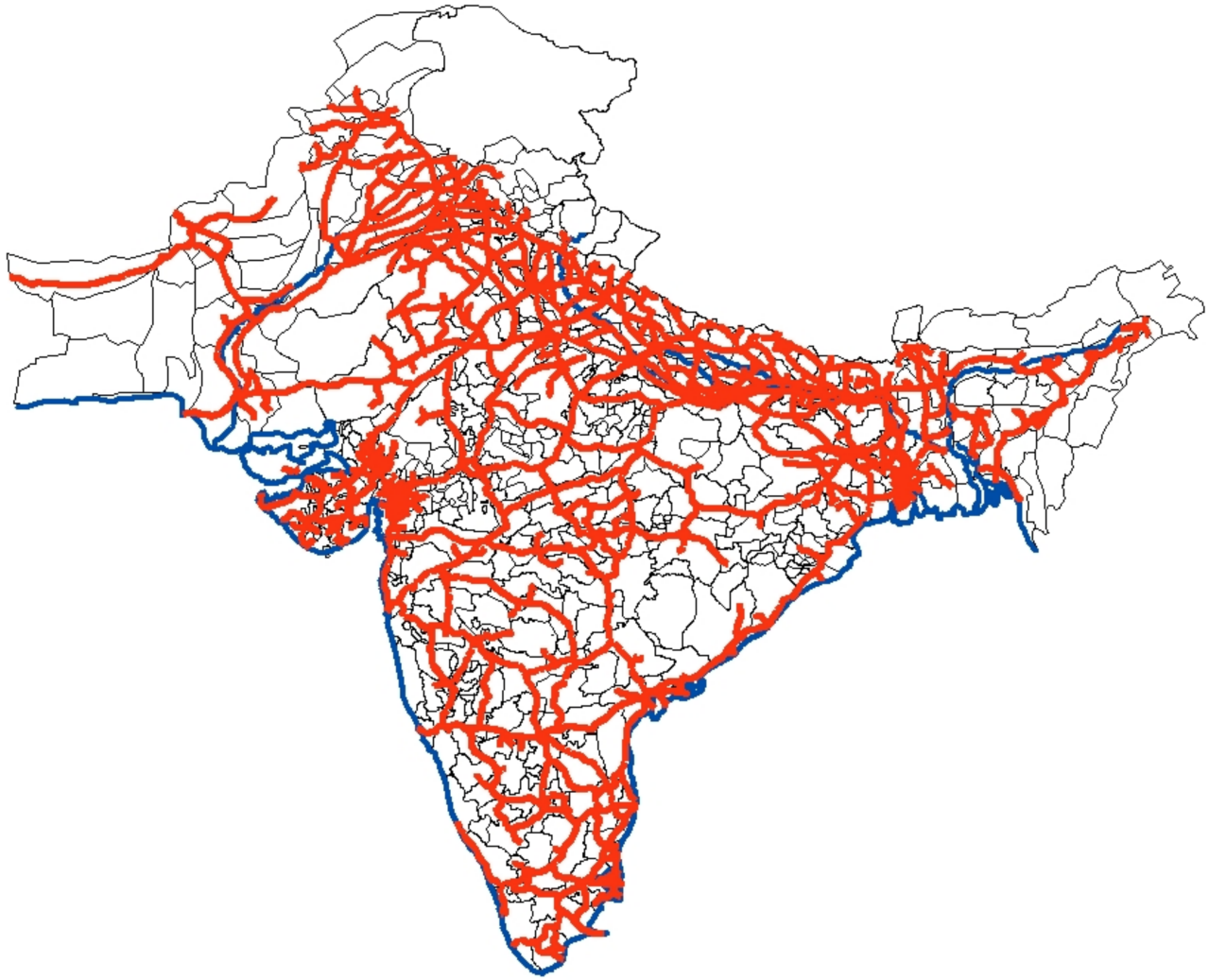
Source: Donaldson, D. (2018), Railroads of the Raj: Estimating the impact of transportation infrastructure, *American Economic Review, 108(4-5), 899–934*.

Donaldson (2018): Indian railroad network in 1920



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

Donaldson (2018): Indian railroad network in 1930



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

Results for step 1: railroads ↑ → trade costs ↓

TABLE 2—RAILROADS AND TRADE COSTS: STEP 1

Dependent variable: log salt price at destination	(1)	(2)
log effective distance to source, along lowest-cost route (at historical freight rates)	0.088 (0.028)	
log effective distance to source, along lowest-cost route (at estimated mode costs)		0.169 [0.062, 0.296]
Estimated mode costs per unit distance: Railroad (normalized to 1)		1 N/A
Road		2.375 [1.750, 10.000]
River		2.250 [1.500, 6.250]
Coast		6.188 [5.875, 10.000]
Observations	7,345	7,345
R^2	0.946	0.946

Notes: Regressions estimating equation (12) using data on 6 types of salt (listed in online Appendix A), from 133 districts in Northern India, annually from 1861 to 1930. Column 1 and column 2 estimated by OLS and NLS respectively; both include salt type × year and salt type × destination fixed effects. "Effective distance to source, along lowest-cost route" measures the railroad-equivalent kilometers (because railroad freight rate is normalized to 1) between the salt source and the destination district, along the lowest-cost route given relative mode costs per unit distance. "Historical freight rates" used are 4.5, 3.0, and 2.25 respectively for road, river, and coastal mode costs per unit distance, all relative to rail transport. Standard errors corrected for clustering at the destination district level are reported in parentheses of column 1, and bootstrapped 95 percent confidence intervals are reported in column 2.

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

Results for step 2: railroads ↑ → trade flows ↑

TABLE 3—RAILROADS AND TRADE FLOWS: STEP 2

Dependent variable: log value of exports	(1)	(2)
log effective distance beween origin and destination along lowest-cost route	−1.603	−1.701
	(0.533)	(1.141)
(log effective distance beween origin and destination along lowest-cost route) × (weight per unit value of commodity in 1890)		−0.946
		(3.634)
(log effective distance beween origin and destination along lowest-cost route) × (high-value railroad freight class of commodity in 1859)		1.286
		(1.243)
Observations	142,541	142,541
R^2	0.901	0.901

Notes: Regressions estimating equation (14) using data on 15 commodities and 47 trade blocks annually from 1882 to 1920. Regressions include origin and destination fixed effects, separately for each commodity and year. "Effective distance between origin and destination along lowest-cost route" measures the railroad-equivalent kilometers (due to the normalization of railroad distance cost to 1) between the centroid of the origin and destination trade blocks in question, along the lowest-cost route given relative freight rates for each mode of transport (as estimated in Table 2). "Weight per unit value in 1890" is the weight (in maunds) per rupee, as measured by 1890 prices. "Railroad freight class in 1859" is an indicator variable for all commodities that were classified in the higher (more expensive) freight class in 1859; salt is in the omitted category (low-value commodities). Heteroskedasticity robust standard errors adjusted for clustering at the exporter-importer block level are reported in parentheses for columns 1 and 2 respectively.

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

Results for step 3: railroads ↑ → real incomes ↑

TABLE 4—RAILROADS AND REAL INCOME LEVELS: STEP 3

Dependent variable: log real agricultural income	(1)	(2)	(3)	(4)
Railroad in district	0.164	0.158	0.160	0.167
	(0.049)	(0.048)	(0.050)	(0.050)
Unbuilt railroad in district, abandoned after proposal stage		0.057		
		(0.058)		
Unbuilt railroad in district, abandoned after reconnaissance stage		0.013		
		(0.099)		
Unbuilt railroad in district, abandoned after survey stage		−0.069		
		(0.038)		
(Unbuilt railroad in district, included in Lawrence Plan 1869–1873) × (post-1871 indicator)			0.067	
			(0.104)	
(Unbuilt railroad in district, included in Lawrence Plan 1874–1878) × (post-1874 indicator)			−0.019	
			(0.092)	
(Unbuilt railroad in district, included in Lawrence Plan 1879–1883) × (post-1879 indicator)			0.095	
			(0.084)	
(Unbuilt railroad in district, included in Lawrence Plan 1884–1888) × (post-1884 indicator)			−0.072	
			(0.075)	
(Unbuilt railroad in district, included in Lawrence Plan 1889–1893) × (post-1889 indicator)			0.047	
			(0.049)	
(Unbuilt railroad in district, included in Lawrence Plan 1894–1898) × (post-1894 indicator)			−0.088	
			(0.086)	
(Unbuilt railroad in district, included in Kennedy plan, high-priority) × (year-1848)				−0.0001
				(0.002)
(Unbuilt railroad in district, included in Kennedy plan, low-priority) × (year-1848)				0.001
				(0.003)
Observations	7,086	7,086	7,086	7,086
R^2	0.848	0.848	0.848	0.848

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

Results for step 4: (railroads $\uparrow$ | τ^k_{oot}) $\rightarrow$ real incomes $\nearrow$

TABLE 5—A SUFFICIENT STATISTIC FOR RAILROAD IMPACT: STEP 4

log real ag. income, corrected for rainfall:	(1)	(2)
Railroad in district	0.258	0.124
	(0.050)	(0.050)
"Trade share," as computed in model		−1.587
		(0.177)
Observations	7,086	7,086
R^2	0.835	0.844

Notes: OLS Regressions estimating equation (18) using real income constructed from crop-level data on 17 principal agricultural crops (listed in online Appendix A), from 192 districts in India, annually from 1870 to 1930. Dependent variable is log real income, corrected for crop-specific rainfall of each of 17 crops, weighted across crops as in equation (18). Regressions include district fixed effects and year fixed effects. "Railroad in district" is a dummy variable whose value is one if any part of the district in question is penetrated by a railroad line. "Trade share" is the share of a district's expenditure that it buys from itself; this variable is computed in the equilibrium of the model, where the model parameters are set to those estimated in Steps 1 and 2, and the exogenous variables (the transportation network, rainfall, and district land sizes) are as observed. Heteroskedasticity-robust standard errors corrected for clustering at the district level are reported in parentheses.

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

Donaldson (2018): findings and takeaways

- railroads substantially improved the trading environment in colonial India
  - trade costs (and price gaps) decreased
  - trade flows increased
  - price responsiveness decreased
- railroads raised real incomes
  - reduced real income volatility
  - volatility $\downarrow$ had important implications for incidence of famine (cf. Burgess & Donaldson, 2012)
- welfare gains from railroads well accounted for by Ricardian model of trade

$\rightarrow$ static gains from trade an important mechanism behind the economic benefits of transport infrastructure

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway system. *The Review of Economic Studies*, *81*(3)

Faber (2014): core–periphery highways & industrialisation

Research questions:

Do large-scale transport infrastructure investments that connect core regions to the periphery ..

a .. diffuse the existing asymmetric spatial distribution?

b .. reinforce/amplify the existing concentration pattern?

Setting:

- substantial welfare improvements of transport cost reductions motivate policymakers to integrate domestic markets
- by construction, mostly connecting pre-existing metropolitan areas (e.g. industrial core) to (possibly autarkic) periphery
- core-periphery inequality extremely stark in LICs
- here: study large-scale expansion of transport infrastructure to integrate markets, construction of Chinese highway system
- quasi-experimental approach: inconsequential units IV
  - exploit accidentally connected peripheral locations on the way from one core metropolitan area to another

The completed highway system in 2007

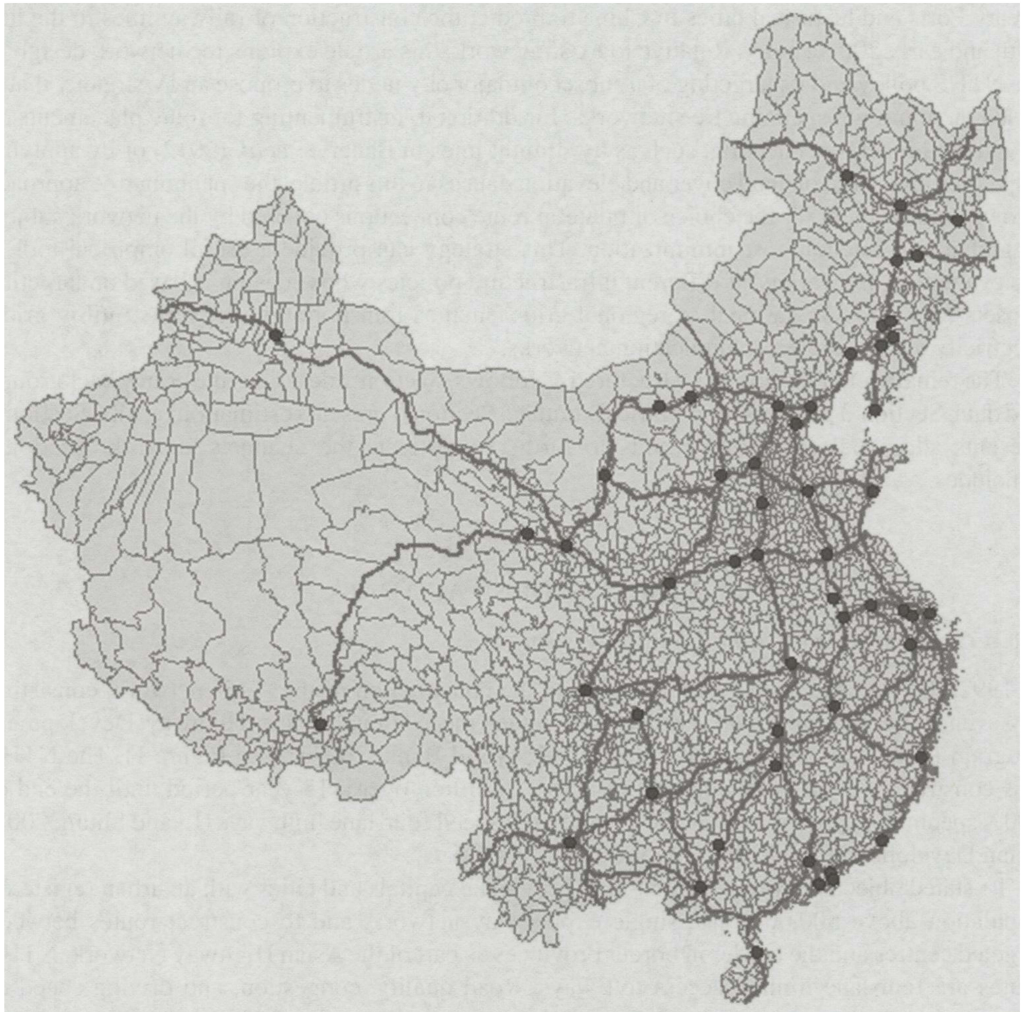


FIGURE 1

Source: Faber, B. (2014), Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway system, *The Review of Economic Studies, 81(3)*.

IV: least-cost spanning tree (dark = pred., bright = actual)

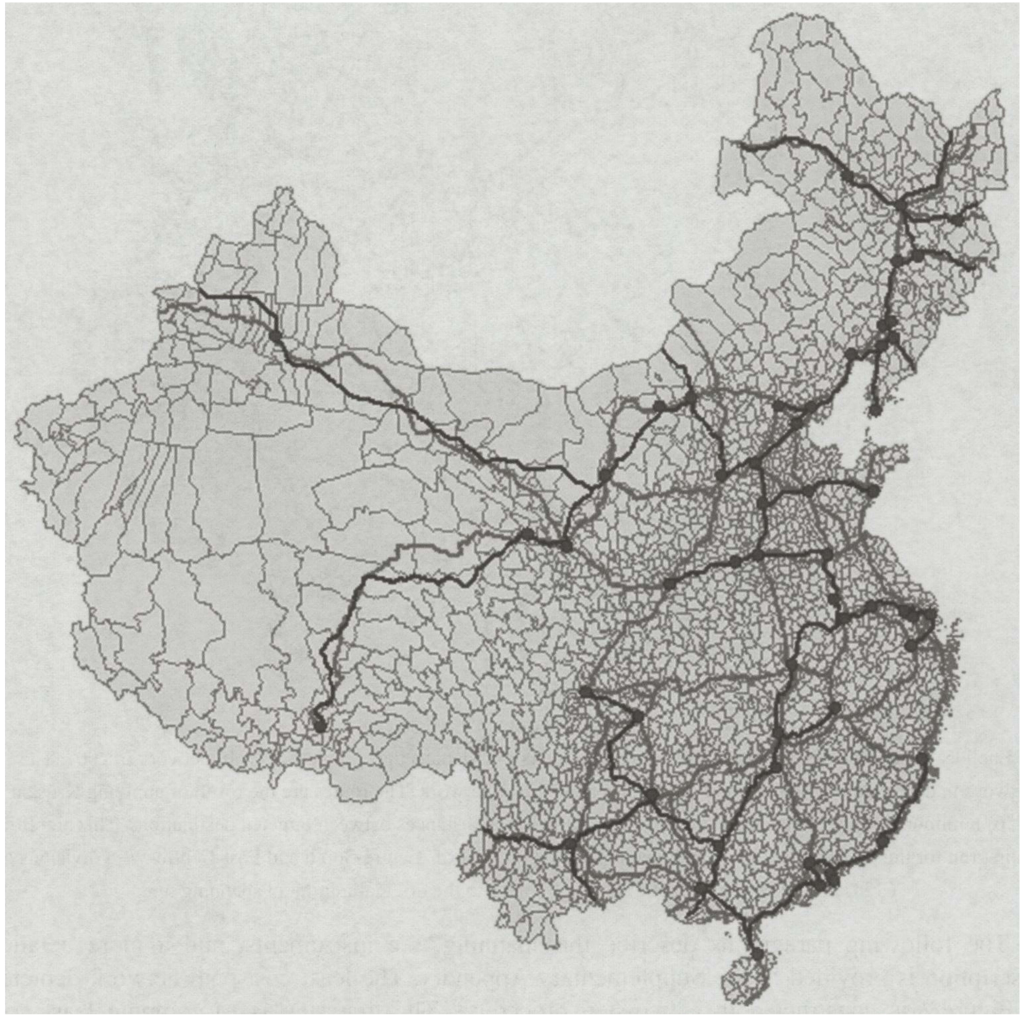


FIGURE 2

Source: Faber, B. (2014), Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway system, *The Review of Economic Studies, 81(3)*.

Faber (2014): findings and takeaways

- transportation infrastructure can display efficiency–equity trade-off:
  - GDP ↑, hence national allocative efficiency improves, cf. Banerjee et al. (2010), Donaldson (2018)
  - regional equity, however, may decrease
- Chinese national highways:
  - decreased industrial and total output growth in connected, previously peripheral locations (relative to non-connected peripheral locations)
  - (probably) increased aggregate welfare due to gains from trade
- key mechanism: pre-existing asymmetries in spatial allocation exacerbated by falling trade costs
- unclear to what extent limited to highways vs roads/railroads

→ potentially unintended GE consequences of large-scale transport infrastructure investments

Can Big Push infrastructure unlock development?

- Economic development strongly associated with structural transformation out of agriculture
- Long literature studying effects of single infrastructure expansion on structural transformation
- In reality, infrastructure expansions commonly sequenced or bundled (e.g. New Deal, TVA, GOELRO, Marshall Plan, BRI)
- However, how different forms of infrastructure investments interact is not well understood

Moneke, N. (2020). Can Big Push infrastructure unlock development? Evidence from Ethiopia. *University of Oxford mimeo*

Moneke (2020): Big push infrastructure and development

In this paper, I

i collect new, geo-identified data on big push infrastructure in Ethiopia

$\Rightarrow$ track roads and electrification expansions across space, over time

ii test for reduced-form causal effects of big push infrastructure

$\Rightarrow$ estimate sectoral employment changes of infrastructure investments

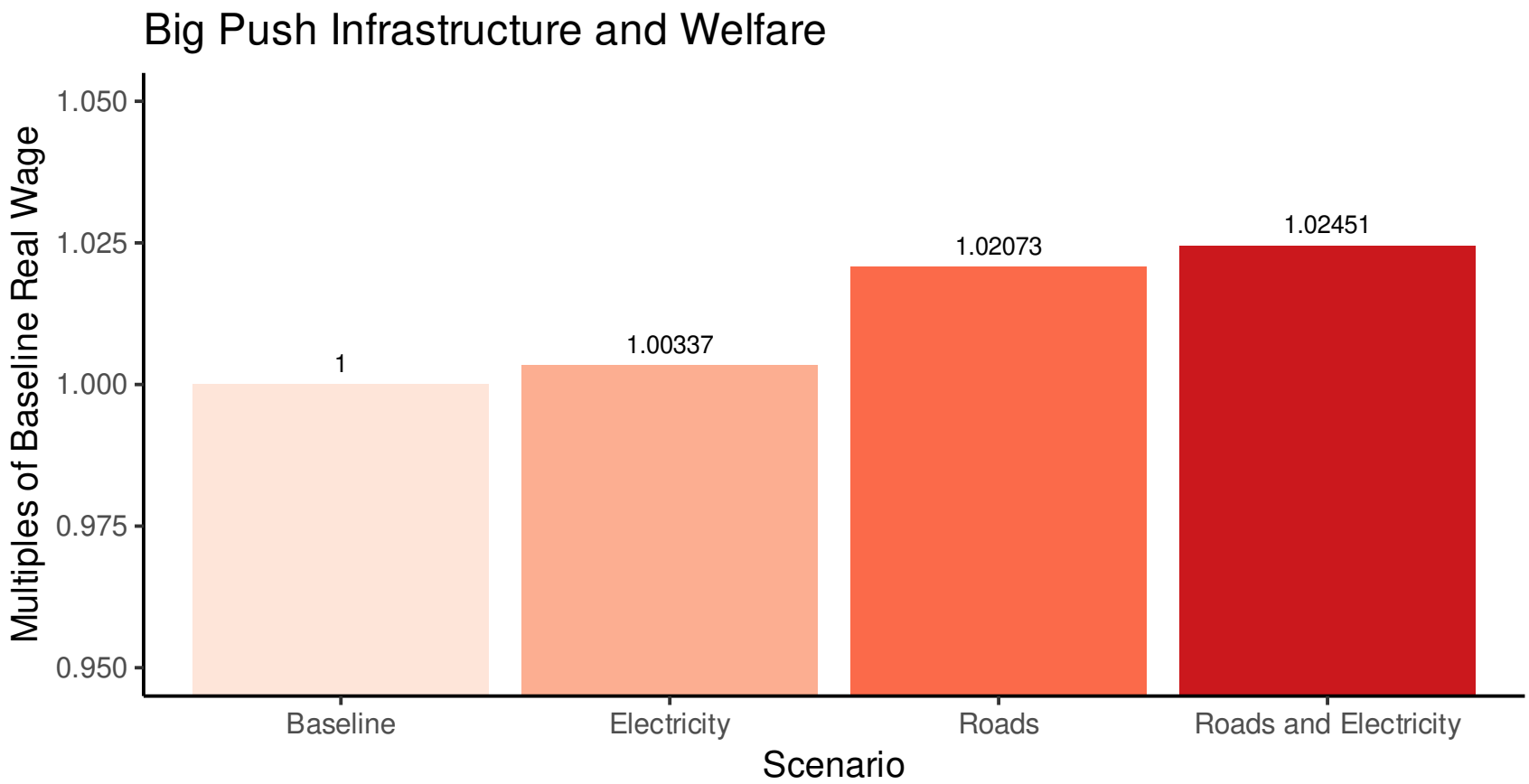
iii develop a spatial GE model of big push infrastructure

$\Rightarrow$ show how diverging sectoral employment patterns can arise

iv structurally estimate productivity and welfare implications

$\Rightarrow$ disentangle welfare effects of big push v isolated infrastructure

Moneke (2020): welfare estimate of big push infrastructure



Moneke (2020): takeaways

1. Causal evidence on a big push infrastructure interaction effect for manufacturing employment
2. Road access alone (via improved market access) causes retail services employment to emerge, at the expense of traditional manufacturing
3. Adverse effect on manufacturing reverses, however, with additional access to electricity (via improved productivity)
4. Structural estimates imply that big push infrastructure investments are material to welfare

Why public, not private education?

- demand side complications:
  - choices made by parents on behalf of children
  - externalities in human capital production
  - credit constraints
  - incorrect or inaccurate beliefs about returns to education
  - difficulty of assessing educational quality
  - social norms associated with female educational attainment
- supply side complications:
  - fixed cost of setting up school
  - externalities in educational provision on labour market
  - natural local monopoly in supply of education
  - quality of education function of teachers' expected utility
  - complementarities with other infrastructure investments

$\rightarrow$ focus on public provision of education infrastructure

Returns to education infrastructure: estimation issues

Estimate causal (aggregate) effect of school expansion on wages:

- ability θ unobserved, suppose of two types $\theta \in \{L, H\}$
- estimate returns to education: $lnw_i = \alpha + \beta S_i + \theta_i + \epsilon_i$

Returns to education infrastructure: estimation issues

Estimate causal (aggregate) effect of school expansion on wages:

- ability θ unobserved, suppose of two types $\theta \in \{L, H\}$
- estimate returns to education: $lnw_i = \alpha + \beta S_i + \theta_i + \epsilon_i$

↯ reduced-form complications (Duflo, 2001):
  - OVB – cost of schooling decreasing in ability, e.g. $C = \frac{c}{\theta}$
  - reverse causality – require S as signal to achieve high w
  - measurement error – wage $\neq$ income, attendance $\neq$ learning
  - selection – attend school if return $>$ cost, e.g. $\beta = \frac{c}{\theta}$

Returns to education infrastructure: estimation issues

Estimate causal (aggregate) effect of school expansion on wages:

- ability θ unobserved, suppose of two types $\theta \in \{L, H\}$
- estimate returns to education: $lnw_i = \alpha + \beta S_i + \theta_i + \epsilon_i$

↯ reduced-form complications (Duflo, 2001):
  - OVB – cost of schooling decreasing in ability, e.g. $C = \frac{c}{\theta}$
  - reverse causality – require S as signal to achieve high w
  - measurement error – wage $\neq$ income, attendance $\neq$ learning
  - selection – attend school if return $>$ cost, e.g. $\beta = \frac{c}{\theta}$

↯ GE complications:
  - returns to skill function of demand and supply (Khanna, 2022)
  - young vs old educational attainment
  - returns to education arise where jobs are, not where schools are
  - complementary decisions: learn and migrate (Hsiao, 2023)

Khanna, G. (2022). Large-scale education reform in general equilibrium: Regression discontinuity evidence from India. *The Journal of political economy*

Large-Scale Education Reform in General Equilibrium:
Regression Discontinuity Evidence from India

Gaurav Khanna

UCSD

7.24.19

Benefits of Nation-Wide Education Schemes

- How do **large-scale** investments in education benefit households?

Benefits of Nation-Wide Education Schemes

- How do **large-scale** investments in education benefit households?
  - Schooling expansion policy in India

Benefits of Nation-Wide Education Schemes

- How do **large-scale** investments in education benefit households?
  - Schooling expansion policy in India
- Economic benefits:
  1. **Education Sector:** In short-run
  2. **Labor Market:** In long-run

Benefits of Nation-Wide Education Schemes

- How do **large-scale** investments in education benefit households?
  - Schooling expansion policy in India
- Economic benefits:
  1. **Education Sector:** In short-run
  2. **Labor Market:** In long-run
- General Equilibrium Effects:

Benefits of Nation-Wide Education Schemes

- How do **large-scale** investments in education benefit households?
  - Schooling expansion policy in India
- Economic benefits:
  1. **Education Sector:** In short-run
  2. **Labor Market:** In long-run
- General Equilibrium Effects:
  1. **Education Sector:** Crowd-out or crowd-in of private Schools?

Benefits of Nation-Wide Education Schemes

- How do **large-scale** investments in education benefit households?
  - Schooling expansion policy in India
- Economic benefits:
  1. **Education Sector:** In short-run
  2. **Labor Market:** In long-run
- General Equilibrium Effects:
  1. **Education Sector:** Crowd-out or crowd-in of private Schools?
  2. **Labor Market:** More skilled workers:
     - Depress skill-premium; workers migrate
     - Adopt skill-biased technology; Increase productivity

Labor Market Benefits

Labor Market Benefits

- **Returns to schooling**:
  - Individual returns – **Becker 1967, Mincer 1974**

Labor Market Benefits

- **Returns to schooling**:
  - Individual returns – **Becker 1967, Mincer 1974**
  - Empirically: use macro-policies – **Tuition subsidies, Compulsory schooling** or **School-building**

Labor Market Benefits

- **Returns to schooling**:
  - Individual returns – **Becker 1967, Mincer 1974**
  - Empirically: use macro-policies – **Tuition subsidies, Compulsory schooling** or **School-building**
- GE Effects **Theoretically** Important :
  - Large GE effects in the US. Depress college skill-premium – **(Heckman et al 1998, Abbot et al 2013)**
  - GE effects when scaling up experiments – **(Muralidharan et al 2018, Acemoglu 2010, Deaton 2012)**

This Paper

- Model in GE - **Firms, Households, Schools**
- I combine:
  1. **Mincer 1958:** education choice for given wage distribution
  2. **Card & Lemieux 2001:** wage distribution for given education choices
- Causally estimate sufficient statistics using **large-scale school-expansion policy** in India
- Variation in:
  1. Districts - **Regression Discontinuity**
  2. Age cohort exposure
  3. Skill level

I find....

1. **Education Sector:**
   - More upper primary schools; more education
   - Private schools – **No crowd out**

I find....

1. **Education Sector:**
   - More upper primary schools; more education
   - Private schools – **No crowd out**

2. **Labor Market**
   - Returns to Education: **With GE: 13.4%**
   - GE effects: earnings - **lowers returns by 1/3**
   - GE effects: welfare - **Skilled 'lose' & unskilled 'gain'**
   - Adoption of **skill-biased technology**

District Primary Education Program (DPEP)

- 1994-2007 in 271/600 districts
- Built 160,000 schools, hired & trained 1.1 million teachers, upgrade infrastructure, 'affected' 51.3 million children
- Largest donor-funded program. Increased funds for primary education by 17-20% - **(Jalan and Glinskaya 2013, Azam and Chan 2016)**
- Reduced household costs of schooling by 20-40%

DPEP: Selection Procedure

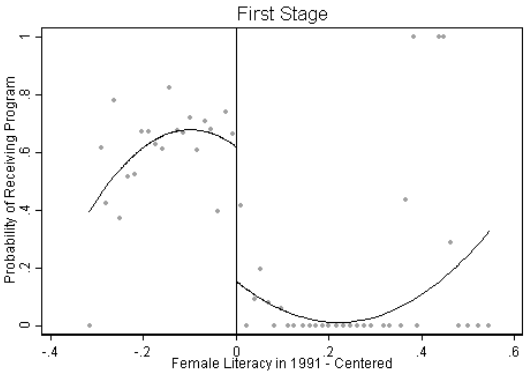
Districts: female literacy rate $<$ national average were eligible

Fuzzy Regression Discontinuity

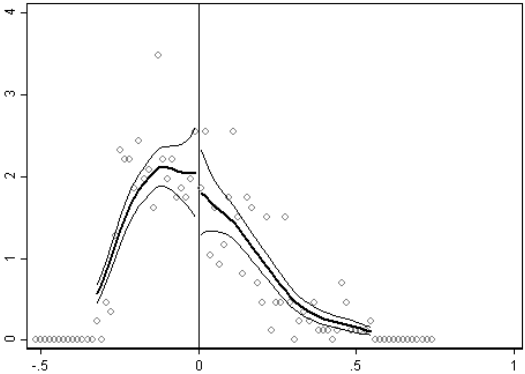
DPEP: Selection Procedure

Districts: female literacy rate $<$ national average were eligible

Fuzzy Regression Discontinuity



First Stage



McCrary Density Test

Pieces of the Model

1. District-Public Education Model
2. Private schools Model
3. Household optimization Model
4. **Economic Production and the Labor Market**
5. Market Clearing & Economic Benefits Model

Pieces of the Model

1. District-Public Education Model
   - Maximize access to education
   - Funds –> build more schools
2. Private schools Model
3. Household optimization Model
4. **Economic Production and the Labor Market**
5. Market Clearing & Economic Benefits Model

Pieces of the Model

1. District-Public Education Model

2. Private schools Model
   - Maximize Profits. Entry or Exit?
   - Crowd-out or crowd-in

3. Household optimization Model

4. **Economic Production and the Labor Market**

5. Market Clearing & Economic Benefits Model

Pieces of the Model

1. District-Public Education Model
2. Private schools Model
3. Household optimization Model
   - Choose education given ability, costs, returns
   - Fall in MC of schooling –> more education
4. **Economic Production and the Labor Market**
5. Market Clearing & Economic Benefits Model

Pieces of the Model

1. District-Public Education Model
2. Private schools Model
3. Household optimization Model
4. **Economic Production and the Labor Market**
   - Returns to education
   - Change in skill distribution & skill biased technology
5. Market Clearing & Economic Benefits Model

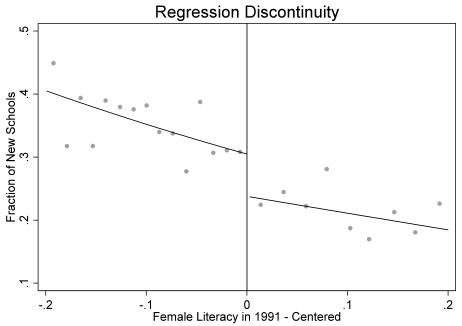
Fraction of New Schools in 2005

12 years after policy started – more schools in DPEP regions

(a) Fraction of New Gov Schools

ITT: 0.049

(b) Fraction of New Schools



ITT: 0.068

Full Bandwidth | Old Schools | Total | Per Capita

Pieces of the Model

- District-Public Education: Choose inputs to maximize access to education – reduce household MC of schooling

2. **Private schools:**
   - **Crowd-out?** Eq. fees falls p_d^*– lower profits **(Filmer et al 1998)**
   - **Crowd-in?**
     1. **Costs fall**: teachers/infrastructure **(Andrabi et al 2013, Jagnani and Khanna 2019)**
     2. Demand could rise if peer effects **(Bobonis and Finan 2009)** – higher fees p_d^*.

Crowd in of Private Schools & Lower Expenditure

More private schools in DPEP regions; lower tuition expenditure

Private schools built post 1993

Crowd in of Private Schools & Lower Expenditure

More private schools in DPEP regions; lower tuition expenditure

Private schools built post 1993

RD – Household Expenditure	
Log(Total Edu Exp)	-0.0857
	(0.058)
Log(Fees and Tuition)	-0.205
	(0.0806)**
Log(Other edu exp)	0.0675
	(0.055)

Full Bandwidth and Totals | Non Teachings Assignments | Crowd Out of Other Funds

Pieces of the Model

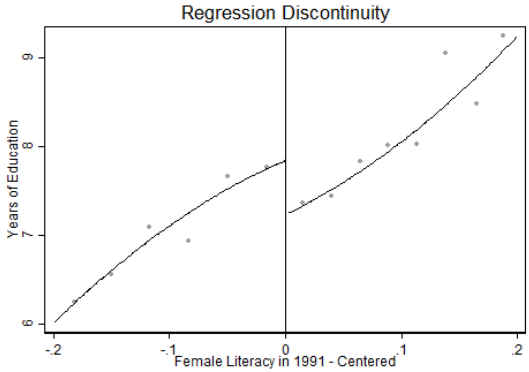
- District-Public Education: Choose inputs to maximize access to education (reduces marginal costs of schooling)
- Private schools: Maximize profits - heterogeneous productivity - entry and exit

3. **Household optimization:** Choose years of schooling given:
   1. Returns to education β_d
   2. Marginal costs of schooling $r_{id} = r(A_d, p_d, \eta_i)$
   3. Ability ϵ_i ; where $Corr(\epsilon_i, \eta_i) \neq 0$ Schooling Equilibrium

   - **60% finish primary school**
   - **5.7 years of school on average**

Years of Education

Years of Education - Old



Years of Education - Young

No Migration Schooling Equilibrium

RD - Returns to Education

	Years of Education	Log(Earnings)	2SLS Returns
RD	0.720	0.112	0.155
	(0.199)***	(0.0312)***	(0.0427)***
Observations	10,175	10,175	10,175
Bandwidth	CCT	CCT	CCT

- Same 2SLS-returns if use Diff-in-Diff: RD = Diff-in-Diff

Full Sample | Parametric RD | By Age Cohort | Reporting Earnings | By Gender | Cells | DISE only | Test Scores

Pieces of the Model

1. District-Public Education
2. Private schools
3. Household optimization
4. **Economic Production and the Labor Market**
5. Market Clearing & Economic Benefits Schooling Equilibrium

Economic Production & The Labor Market

- Nested CES:

$$Y_d = L_d^{\varrho} K_d^{(1-\varrho)}$$

$$L_d = \left(\sum_s \theta_{sd} L_{sd}^{\frac{\sigma_E - 1}{\sigma_E}} \right)^{\frac{\sigma_E}{\sigma_E - 1}}$$

$$L_{sd} = \left(\sum_a \psi_a \ell_{asd}^{\frac{\sigma_A - 1}{\sigma_A}} \right)^{\frac{\sigma_A}{\sigma_A - 1}}$$

- **s - education/skill; a - age; d - district**
- **K - capital; ϱ - labor share**
- **σ_E - elasticity of substitution across education**
- **σ_A - elasticity of substitution across ages**
- **θ_{sd} - productivity technology (skill biased capital)**
- **ψ_a - age effect (e.g. experience)**

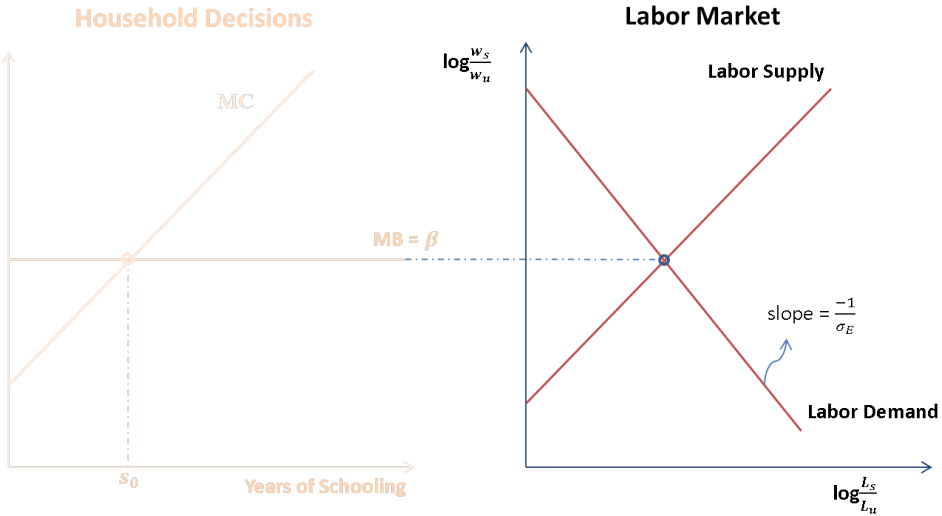
Earnings Equation

- Workers paid marginal product: $w_{asd} = \frac{\partial Y}{\partial \ell_{asd}}$
- Comparing two people's earnings:

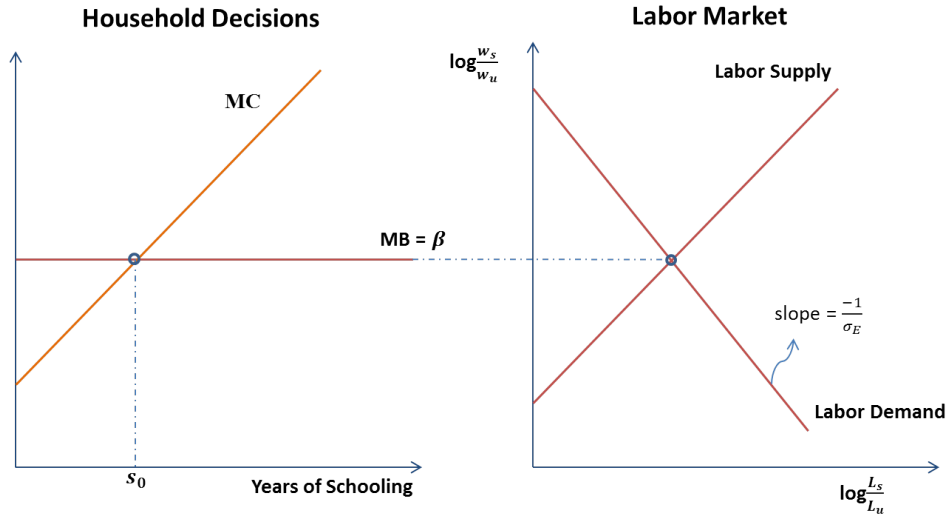
$$\log\left(\frac{w_{asd}}{w_{a's'd'}}\right) = \underbrace{\log\left(\frac{\theta_{sd}}{\theta_{s'd'}}\right)}_{\text{productivity}} + \underbrace{\log\left(\frac{\psi_a}{\psi_{a'}}\right)}_{\text{age}} + \underbrace{\frac{1}{\sigma_E}\log\left(\frac{Y_d}{Y_{d'}}\right)}_{\text{output}} + \underbrace{\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right)\log\frac{L_{sd}}{L_{s'd'}}}_{\text{Agg skill distribution}} \underbrace{- \frac{1}{\sigma_A}\log\frac{\ell_{asd}}{\ell_{a's'd'}}}_{\text{skill-cohort distribution}}$$

Full version: Explicit Skill Biased Capital

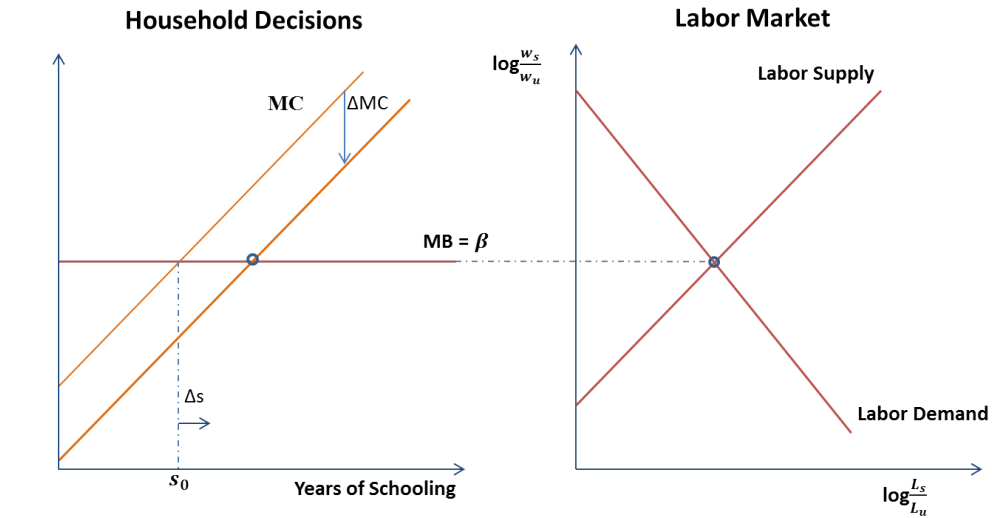
The Labor Market



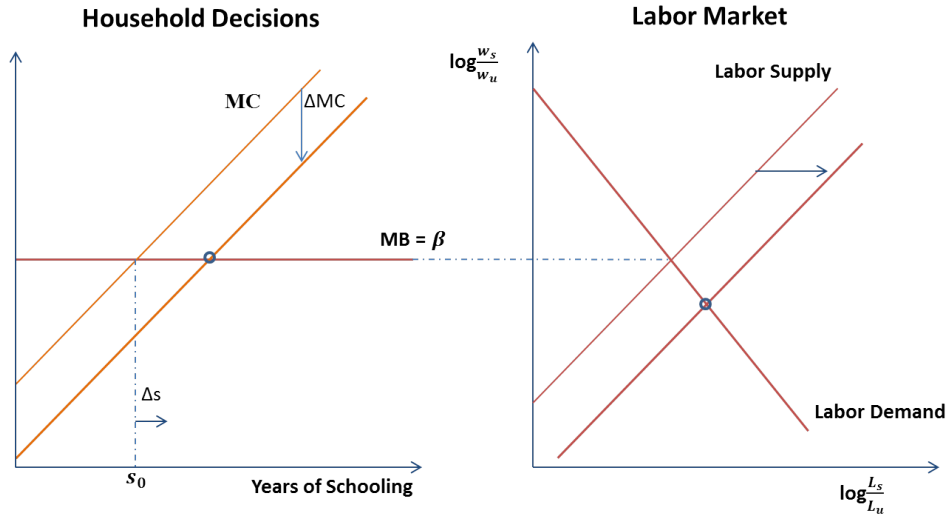
Household Decisions



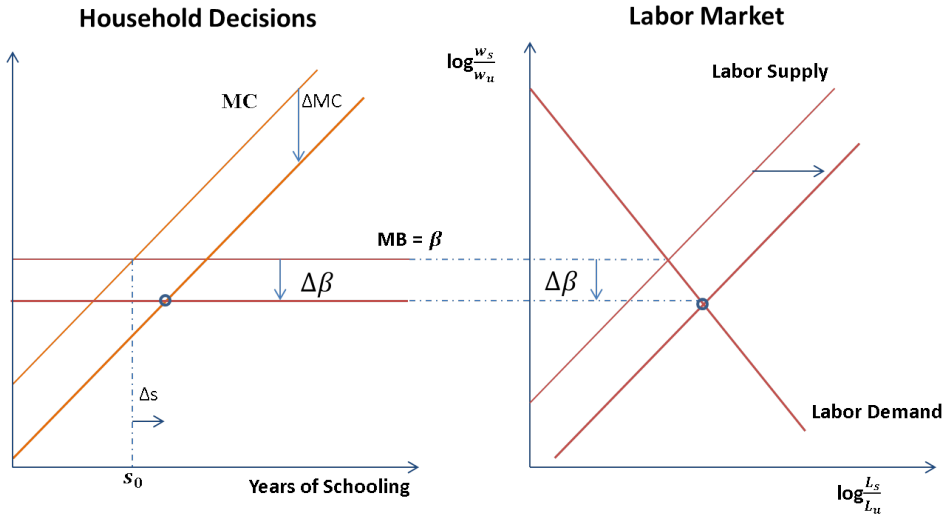
New Schools Lower the Marginal Cost of Schooling



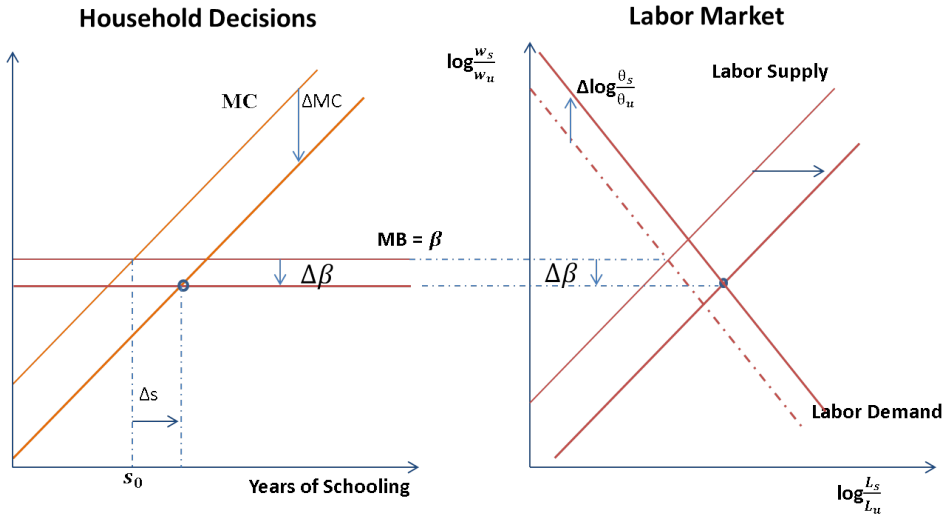
And Shift Out the Relative Labor Supply



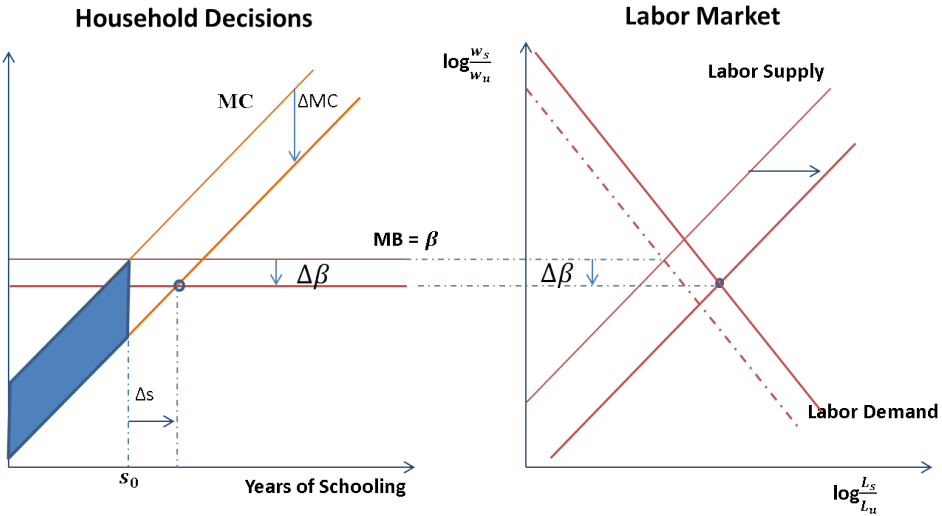
Depress the Skill Premium



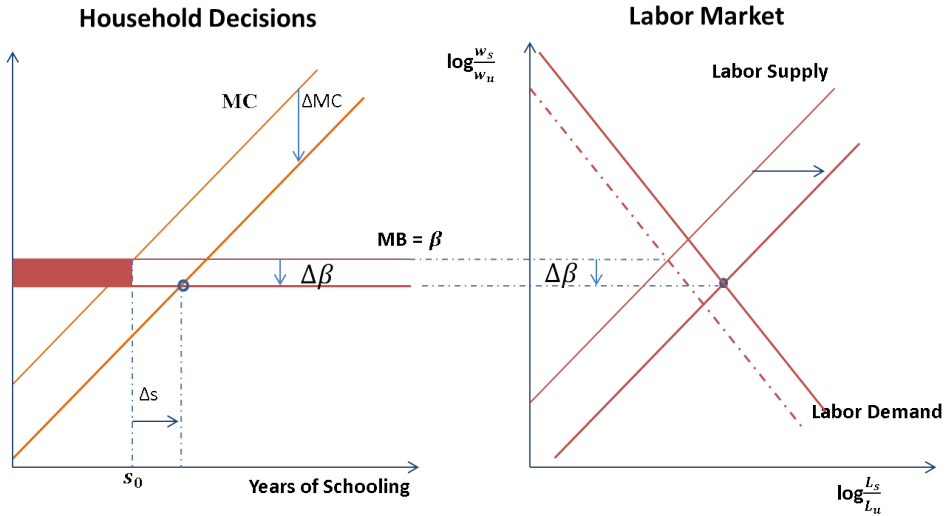
Skill Biased Capital May Move In



Welfare of Students



Distributional Effects



Returns to Skill

- In Untreated district $D = 0$:

$$log\frac{w_{s,D=0}}{w_{u,D=0}} = \underbrace{log\frac{\theta_{s,D=0}}{\theta_{u,D=0}}}_{\text{Productivity}} + \underbrace{\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=0}}{L_{u,D=0}}}_{\text{Agg skill}} - \underbrace{\frac{1}{\sigma_A} log\frac{\ell_{as,D=0}}{\ell_{au,D=0}}}_{\text{age-skill dist}} \equiv \beta_{as,D=0}$$

- In Treated district $D = 1$:

$$log\frac{w_{s,D=1}}{w_{u,D=1}} = log\frac{\theta_{s,D=1}}{\theta_{u,D=1}} + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=1}}{L_{u,D=1}} - \frac{1}{\sigma_A} log\frac{\ell_{as,D=1}}{\ell_{au,D=1}} \equiv \beta_{as,D=1}$$

$$\Delta\beta_{as} = \underbrace{\Delta\text{productivity} + \Delta\text{Agg-skill}}_{\text{GE on all}} + \underbrace{\Delta\text{age-skill}}_{\text{Additional GE on young}}$$

Returns to Skill

- In Untreated district $D = 0$:

$$log\frac{w_{s,D=0}}{w_{u,D=0}} = \underbrace{log\frac{\theta_{s,D=0}}{\theta_{u,D=0}}}_{\text{Productivity}} + \underbrace{\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=0}}{L_{u,D=0}}}_{\text{Agg skill}} - \underbrace{\frac{1}{\sigma_A} log\frac{\ell_{as,D=0}}{\ell_{au,D=0}}}_{\text{age-skill dist}} \equiv \beta_{as,D=0}$$

- In Treated district $D = 1$:

$$log\frac{w_{s,D=1}}{w_{u,D=1}} = log\frac{\theta_{s,D=1}}{\theta_{u,D=1}} + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=1}}{L_{u,D=1}} - \frac{1}{\sigma_A} log\frac{\ell_{as,D=1}}{\ell_{au,D=1}} \equiv \beta_{as,D=1}$$

$$\Delta\beta_{as} = \underbrace{\Delta\text{productivity} + \Delta\text{Agg-skill}}_{\text{GE on all}} + \underbrace{\Delta\text{age-skill}}_{\text{Additional GE on young}}$$

Returns to Skill Differ Across Districts

- In Untreated district $D = 0$:

$$log\frac{w_{s,D=0}}{w_{u,D=0}} = \underbrace{log\frac{\theta_{s,D=0}}{\theta_{u,D=0}}}_{\text{Productivity}} + \underbrace{\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=0}}{L_{u,D=0}}}_{\text{Agg skill}} - \underbrace{\frac{1}{\sigma_A} log\frac{\ell_{as,D=0}}{\ell_{au,D=0}}}_{\text{age-skill dist}} \equiv \beta_{as,D=0}$$

- In Treated district $D = 1$:

$$log\frac{w_{s,D=1}}{w_{u,D=1}} = log\frac{\theta_{s,D=1}}{\theta_{u,D=1}} + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=1}}{L_{u,D=1}} - \frac{1}{\sigma_A} log\frac{\ell_{as,D=1}}{\ell_{au,D=1}} \equiv \beta_{as,D=1}$$

$$\Delta\beta_{as} = \underbrace{\Delta\text{productivity} + \Delta\text{Agg-skill}}_{\text{GE on all}} + \underbrace{\Delta\text{age-skill}}_{\text{Additional GE on young}}$$

Returns to Skill Differ Across Districts

- In Untreated district $D = 0$:

$$log\frac{w_{s,D=0}}{w_{u,D=0}} = \underbrace{log\frac{\theta_{s,D=0}}{\theta_{u,D=0}}}_{\text{Productivity}} + \underbrace{\left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=0}}{L_{u,D=0}}}_{\text{Agg skill}} - \underbrace{\frac{1}{\sigma_A} log\frac{\ell_{as,D=0}}{\ell_{au,D=0}}}_{\text{age-skill dist}} \equiv \beta_{as,D=0}$$

- In Treated district $D = 1$:

$$log\frac{w_{s,D=1}}{w_{u,D=1}} = log\frac{\theta_{s,D=1}}{\theta_{u,D=1}} + \left(\frac{1}{\sigma_A} - \frac{1}{\sigma_E}\right) log\frac{L_{s,D=1}}{L_{u,D=1}} - \frac{1}{\sigma_A} log\frac{\ell_{as,D=1}}{\ell_{au,D=1}} \equiv \beta_{as,D=1}$$

$$\Delta\beta_{as} = \underbrace{\Delta\text{productivity} + \Delta\text{Agg-skill}}_{\text{GE on all cohorts}} + \underbrace{\Delta\text{age-skill}}_{\text{Additional GE on young}}$$

Returns: Estimation

- Diff in young workers' skill-premium

$$\underbrace{\left[log\frac{w_{sy,D=1}}{w_{sy,D=0}} - log\frac{w_{uy,D=1}}{w_{uy,D=0}}\right]}_{\Delta \text{ young skill premium}} = \Delta\beta_{ys} = \text{GE on all} + \text{Addl. on young}$$

Returns: Estimation

- Diff in young workers' skill-premium

$$\underbrace{\left[log\frac{w_{sy,D=1}}{w_{sy,D=0}} - log\frac{w_{uy,D=1}}{w_{uy,D=0}}\right]}_{\Delta \text{ young skill premium}} = \Delta\beta_{ys} = \text{GE on all} + \text{Addl. on young}$$

- Older population - GE affecting all cohorts

$$\underbrace{log\frac{w_{so,D=1}}{w_{so,D=0}} - log\frac{w_{uo,D=1}}{w_{uo,D=0}}}_{\Delta \text{ old skill premium}} = \text{GE on all}$$

Identification and Error Structure

Estimating β

Returns in untreated district $\beta_{as,D=0}$

$$log\frac{w_{y,D=1}}{w_{y,D=0}} = \Delta\ell_{sy}\underbrace{log\frac{w_{sy,D=0}}{w_{uy,D=0}}}_{\beta_{as,D=0}}$$

Estimate b1

Estimating β

Returns in untreated district $\beta_{as,D=0}$

$$log\frac{w_{y,D=1}}{w_{y,D=0}} = \ell_{sy,D=1} log\frac{w_{sy,D=1}}{w_{sy,D=0}} + \ell_{uy,D=1} log\frac{w_{uy,D=1}}{w_{uy,D=0}} + \Delta\ell_{sy} \underbrace{log\frac{w_{sy,D=0}}{w_{uy,D=0}}}_{\beta_{as,D=0}}$$

Estimate b1

Estimating β

Returns in untreated district $\beta_{as,D=0}$

$$log\frac{w_{y,D=1}}{w_{y,D=0}} = \ell_{sy,D=1}\underbrace{log\frac{w_{sy,D=1}}{w_{sy,D=0}}}_{\neq 0} + \ell_{uy,D=1}\underbrace{log\frac{w_{uy,D=1}}{w_{uy,D=0}}}_{\neq 0} + \Delta\ell_{sy}\underbrace{log\frac{w_{sy,D=0}}{w_{uy,D=0}}}_{\beta_{as,D=0}}$$

Estimate b1

Labor Market GE effects

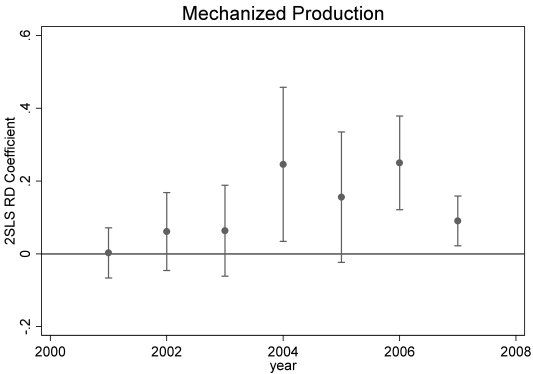
	Fraction Switched	Change in Returns $\Delta\beta = \Delta$young skill premium
Estimate	0.173	-0.065
SE	(0.059)	(0.030)

	Returns without GE $\beta_{D=0}$	Returns with GE $\beta_{D=1}$	% change in β due to GE
Estimate	0.199	0.134	-32.5%
p-value	[0.047]	[0.084]	

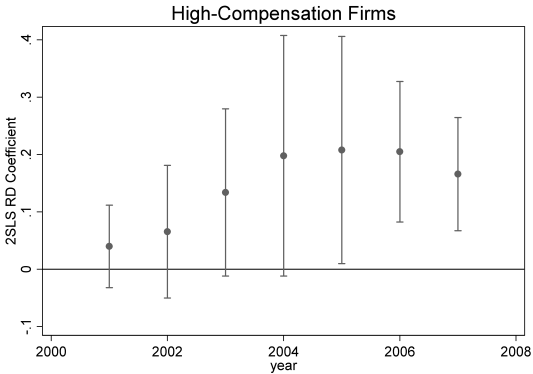
Δ all cohorts	Additional on Young	% Δ on young
-0.0053	-0.0594	91.87%

Estimating β Estimating $log\frac{\theta_{ss}}{\theta_{ff}}$ $\Delta\beta$ details Finer Skill Groups

Capital Adoption – But Workers Stay



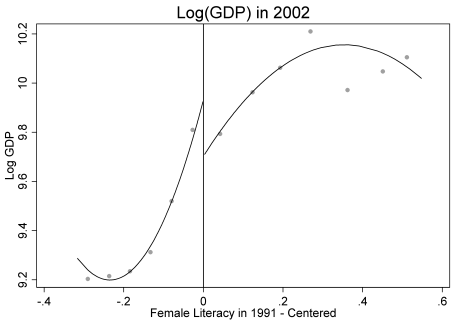
Mechanized Production



High-Wage Paying Firms

(Lack of) Migration Unemployment & Frequency Paid Density of Firms Test Scores

Total Output



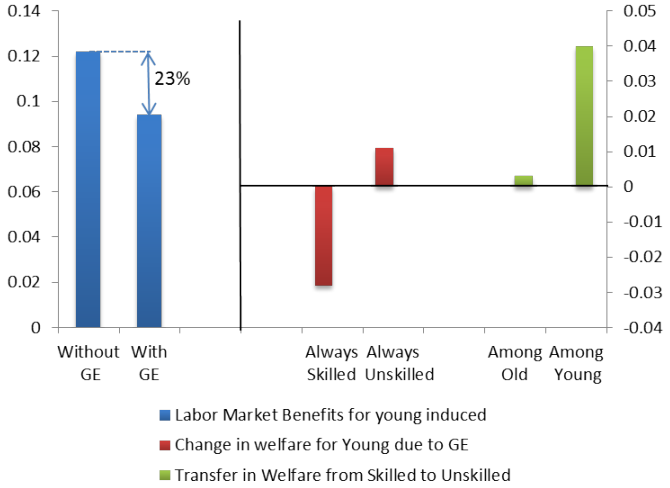
District GDP increase 6.8%

	Log(Consumption Expenditure) 2009-10	
RD Estimate	0.0589	0.0575
	(0.0177)***	(0.0171)***
Observations	12,563	26,420
Bandwidth	CCT	I and K

Lifetime Welfare Changes Including Costs

Labor Market Benefits

Calculating Welfare



Benefits of Educational Expansion

1. **Education Sector:** Effectiveness of policy

Benefits of Educational Expansion

1. **Education Sector:** Effectiveness of policy
   - GE on private schools? **crowd-in**

2. **Labor Market** GE effects? **Depress returns by 33%**
   - Returns to skill? **13.4% (with GE)**
   - Distributional consequences

Benefits of Educational Expansion

1. **Education Sector:** Effectiveness of policy
   - GE on private schools? **crowd-in**

2. **Labor Market** GE effects? **Depress returns by 33%**
   - Returns to skill? **13.4% (with GE)**
   - Distributional consequences

- **Research implications** -
  1. Estimating β using macro-policies
  2. Scaled up experiment may be more/less effective

Benefits of Educational Expansion

1. **Education Sector:** Effectiveness of policy
   - GE on private schools? **crowd-in**

2. **Labor Market** GE effects? **Depress returns by 33%**
   - Returns to skill? **13.4% (with GE)**
   - Distributional consequences

- **Research implications** -
  1. Estimating β using macro-policies
  2. Scaled up experiment may be more/less effective

- **Policy implications** - Estimate GE effects when evaluating cost-effectiveness

Benefits of Educational Expansion

1. **Education Sector:** Effectiveness of policy
   - GE on private schools? **crowd-in**

2. **Labor Market** GE effects? **Depress returns by 33%**
   - Returns to skill? **13.4% (with GE)**
   - Distributional consequences

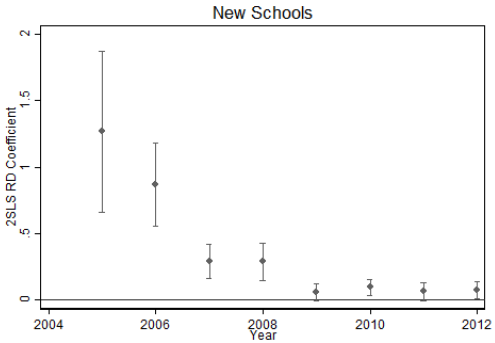
- **Research implications** -
  1. Estimating β using macro-policies
  2. Scaled up experiment may be more/less effective

- **Policy implications** - Estimate GE effects when evaluating cost-effectiveness

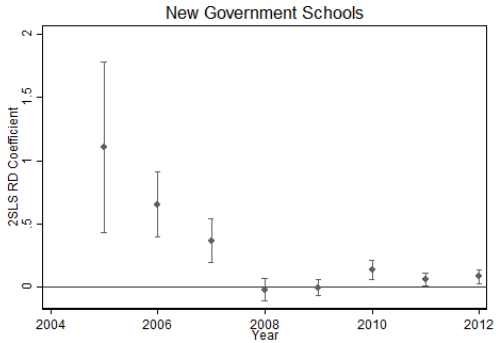
Thank you - gakhanna@ucsd.edu

Over Time: Fraction of New Schools

Over time, as the funding is cut, difference in schools is lower....



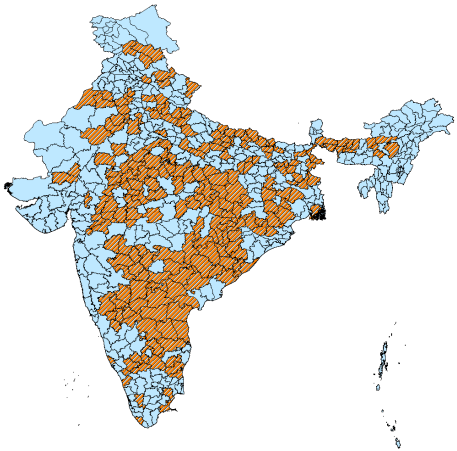
All Schools post 1993



Government Schools post 1993

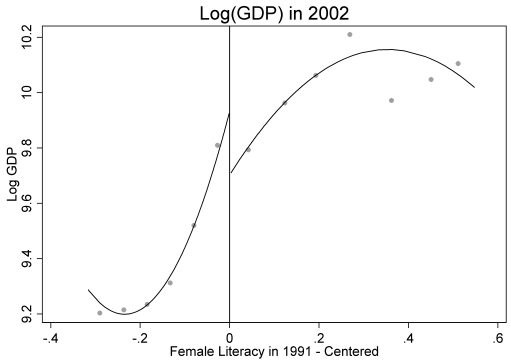
No discontinuity for Old Schools

Map of DPEP districts



Orange and shaded districts received DPEP, whereas blue-unshaded districts did not.

Total Output and Consumption



District GDP increase 6.8%

	Log(Consumption)	
RD	0.179	0.172
	(0.0372)***	(0.0334)***
Obs	27,372	33,758
BW	CCT	I and K

Lifetime Welfare Changes Including Costs

*** Hsiao, A. (2023). Educational investment in spatial equilibrium: Evidence from Indonesia. *Princeton mimeo*

Educational Investment in Spatial Equilibrium: Evidence from Indonesia

Allan Hsiao
Princeton University

January 7, 2023

Motivation

- 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
- **How does migration shape educational investment?**

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)

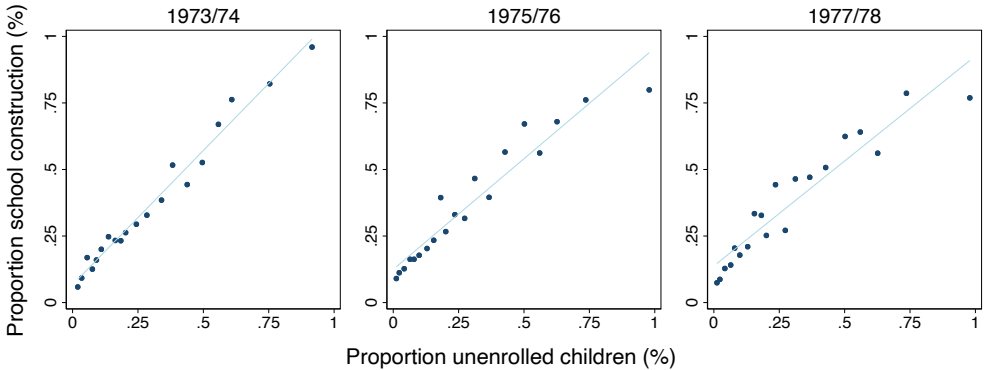
- **INPRES program** evaluation with aggregate effects and counterfactuals
  - Duflo 2001/2004, Martinez-Bravo 2017, Ashraf et al. 2020, Bazzi et al. 2021

- **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


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

INPRES built 62,000 new primary schools



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 (Duflo 2001)

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

1. **Young vs. old** students by age cohort k
2. **More vs. less** school construction by origin district j

Long-run education and wage effects

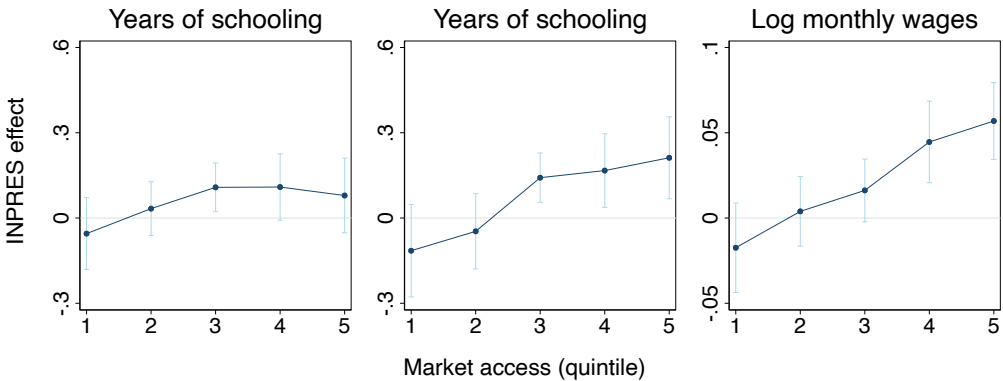
Outcomes	Estimate	SE	Obs
Years of schooling	0.103**	(0.0424)	233,517
— For wage earners	0.121**	(0.0495)	89,404
Log monthly wages	0.020**	(0.0092)	89,404

Driven by labor market access

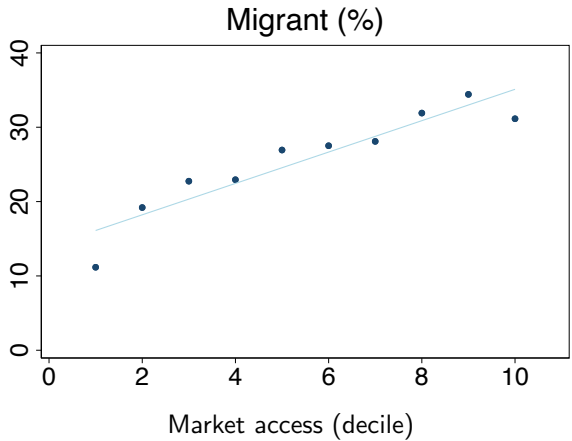
$$\text{MA}_d = \sum_{d'} \text{w}_{d'} \text{popden}_{d'} \quad \text{for} \quad \text{w}_{d'} \propto (1 + \text{dist}_{dd'})^{-2}$$

- Captures access to high urban wages
  - Population density in 1971 + Euclidean distances

Driven by labor market access



Migration rates are high



People benefit, but not places

	People		Places	
Outcomes	Estimate	SE	Estimate	SE
Years of schooling	0.103**	(0.0424)	0.052	(0.0452)
— For wage earners	0.121**	(0.0495)	0.026	(0.0506)
Log monthly wages	0.020**	(0.0092)	0.011	(0.0076)

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 → better access → 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)

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)

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

- Individual i, origin $j(i)$, age cohort $k(i)$, destinations ℓ

1. Choose education for future labor

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

2. Choose labor migration given education and skill draws (Fréchet)

$$U_{i\ell}(e, \epsilon_{i\ell}) = \left(\frac{a_\ell w_\ell s_{jk\ell}}{\tau^m_{jk\ell}} \right) e^\eta \epsilon_{i\ell}$$

Migration, education, and wages

- Closed form for $m_{jk\ell}$, e_{jk}, and $w_{jk\ell}$ (data)
- Labor market access ↑ $\Rightarrow$ migration ↑, education ↑, wages ↑

$$MA_{jk} \equiv \sum_{\ell} \left(\frac{a_{\ell} w_{\ell} s_{jk\ell}}{\tau^{m}_{jk\ell}} \right)^{\theta}$$

Equilibrium and output

$$H_\ell^{\text{supply}}(w_\ell) = H_\ell^{\text{demand}}(w_\ell) \quad \forall\ \ell$$

- Supply: individuals choose e_{jk} and $m_{jk\ell}$ in response to w_ℓ
- Demand: firms set w_ℓ to marginal product

- Production uses human capital $(0 < \kappa < 1)$

$$Y_\ell = A_\ell H_\ell^\kappa, \quad w_\ell = \kappa A_\ell H_\ell^{\kappa-1}, \quad Y = \frac{1}{\kappa}\sum_{j,k,\ell} N_{jk} m_{jk\ell} w_{jk\ell} \quad (Y^U, Y^R)$$

Estimation

1. Education and migration costs (INPRES as DD)

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

$$\Downarrow$$

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

$$\Downarrow$$

$$\widehat{\sigma} = 0.110^{**}(0.047)\,, \quad \widehat{\delta} = 0.042^{***}(0.004)$$

1. Education and migration costs (INPRES as DD)

$$\frac{w_{jk\ell}}{e_{jk}} = \frac{\tau^{e}_{jk}\tau^{m}_{jk\ell}}{\eta a_{\ell}}$$

$$\Downarrow$$

$$\tilde{w}_{jk\ell} - \tilde{e}_{jk} = -\tilde{\eta} - \sigma\tilde{S}_{jk} + \delta\tilde{D}_{j\ell} - \tilde{a}_{\ell} + \tilde{\tau}_{j} + \tilde{\tau}_{k} + \tilde{\varepsilon}^{\tau}_{jk\ell}$$

$$\Downarrow$$

$$\widehat{\sigma} = 0.110^{**}(0.047)\,, \quad \widehat{\delta} = 0.042^{***}(0.004)$$

1. Education and migration costs (INPRES as DD)

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

$$\Downarrow$$

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

$$\Downarrow$$

$$\widehat{\sigma} = 0.110^{**}(0.047)\,, \quad \widehat{\delta} = 0.042^{***}(0.004)$$

2. Human capital and skill (INPRES as IV)

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

$$\Downarrow$$

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

$$\Downarrow$$

$$\widehat{\eta} = 0.688^{**}(0.311)\,, \quad \widehat{\theta} = 21.31^{***}(10.52)$$

2. Human capital and skill (INPRES as IV)

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

$$\Downarrow$$

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

$$\Downarrow$$

$$\widehat{\eta} = 0.688^{**}(0.311)\,, \quad \widehat{\theta} = 21.31^{***}(10.52)$$

2. Human capital and skill (INPRES as IV)

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

$$\Downarrow$$

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

$$\Downarrow$$

$$\widehat{\eta} = 0.688^{**}(0.311)\,, \quad \widehat{\theta} = 21.31^{***}(10.52)$$

3. Production (INPRES as IV)

$$Y_\ell = A_\ell H_\ell^\kappa$$

$$\Downarrow$$

$$\tilde{Y}'_\ell = \tilde{\kappa} + \kappa \tilde{H}_\ell + \tilde{A}_\ell$$

$$\Downarrow$$

$$\widehat{\kappa} = 0.767^{***} (0.101)$$

3. Production (INPRES as IV)

$$Y_\ell = A_\ell H_\ell^\kappa$$

$$\Downarrow$$

$$\tilde{Y}'_\ell = \tilde{\kappa} + \kappa \tilde{H}_\ell + \tilde{A}_\ell$$

$$\Downarrow$$

$$\widehat{\kappa} = 0.767^{***} (0.101)$$

3. Production (INPRES as IV)

$$Y_\ell = A_\ell H_\ell^\kappa$$

$$\Downarrow$$

$$\tilde{Y}'_\ell = \tilde{\kappa} + \kappa \tilde{H}_\ell + \tilde{A}_\ell$$

$$\Downarrow$$

$$\widehat{\kappa} = 0.767^{***}(0.101)$$

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

The program increased aggregate output by 8%

	Aggregate output
Zero construction	**1.00**
+ Direct effect of construction	1.03
+ Migration	1.05
+ Migration-induced schooling	1.09
+ 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**
+ INPRES construction	**0.95**

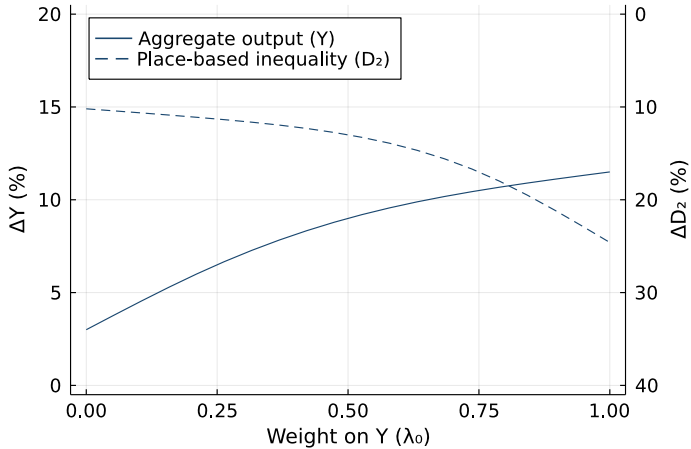
- Expanded opportunity for rural students with high marginal returns
  - Decreased inequality between rural and urban students by 5%

But also increased inequality across places by 12%

	Inequality (places)
Zero construction	**1.00**
+ INPRES construction	**1.12**

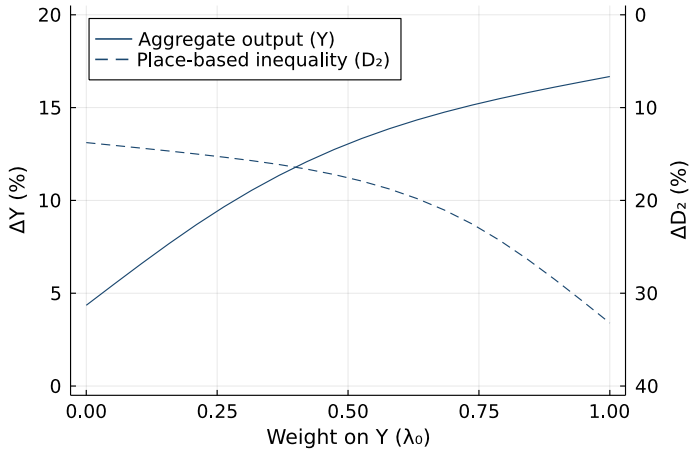
- The program explicitly aimed to encourage regional convergence
  - But mobility places convergence in tension with output gains
  - Rural regions still enjoy net gains, but urban regions gain more

Equity-efficiency tradeoff for policymaker



- Targeting rural areas: output ↑, but rural-urban gap ↑ (implied 50-50 weight)

Equity-efficiency tradeoff for policymaker



- Especially with schools + roads, but roads drain rural areas

Conclusion

Summary

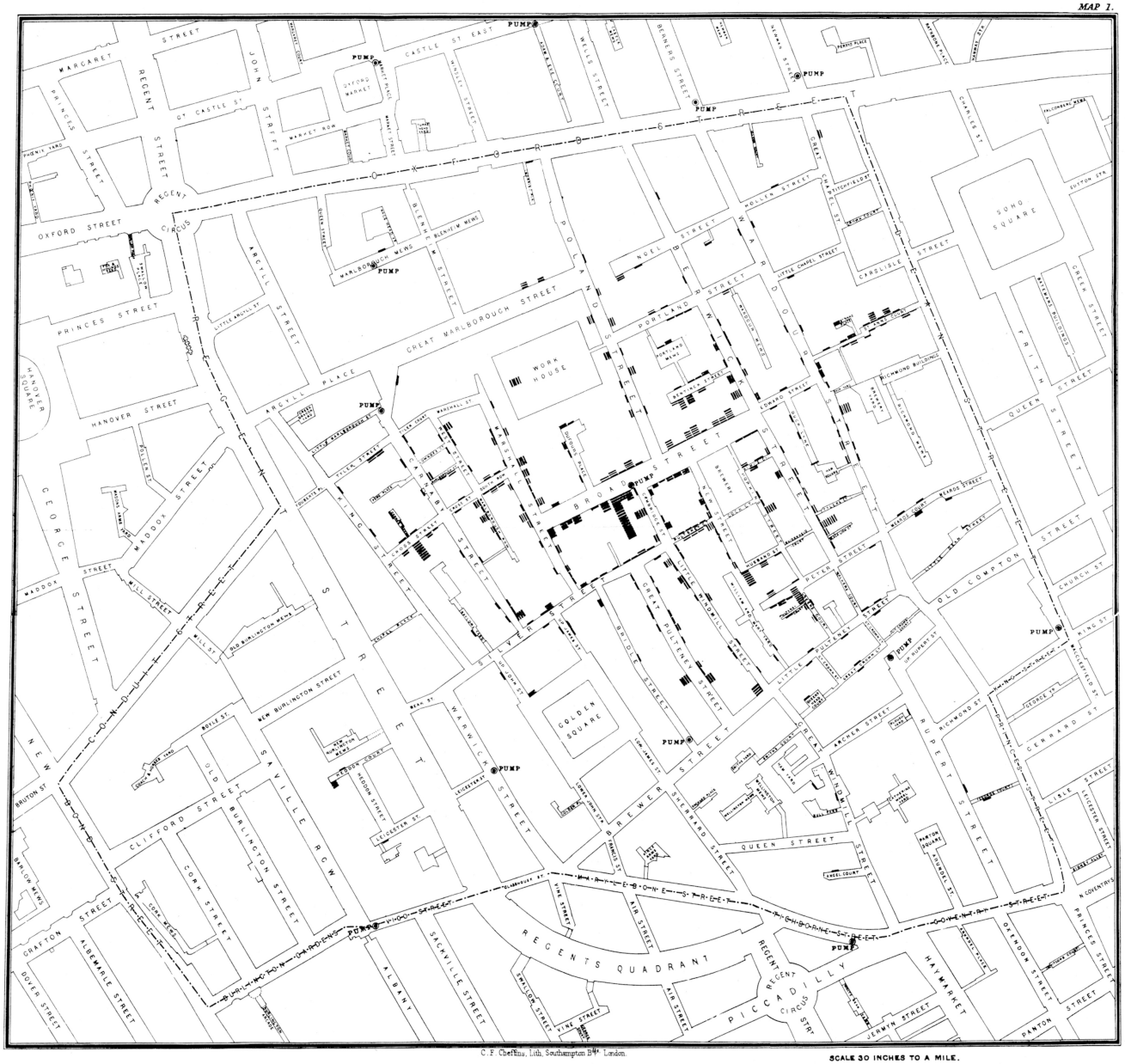
- Evaluating large-scale educational investment in spatial equilibrium
  - Indonesia's INPRES program built 62,000 primary schools in 1970s
  - Aggregate output $\uparrow$ (8%), person-based inequality $\downarrow$ (5%), place-based $\uparrow$ (12%)

- Education and migration are **complementary**
  - Big gains for rural students who leave rural regions behind

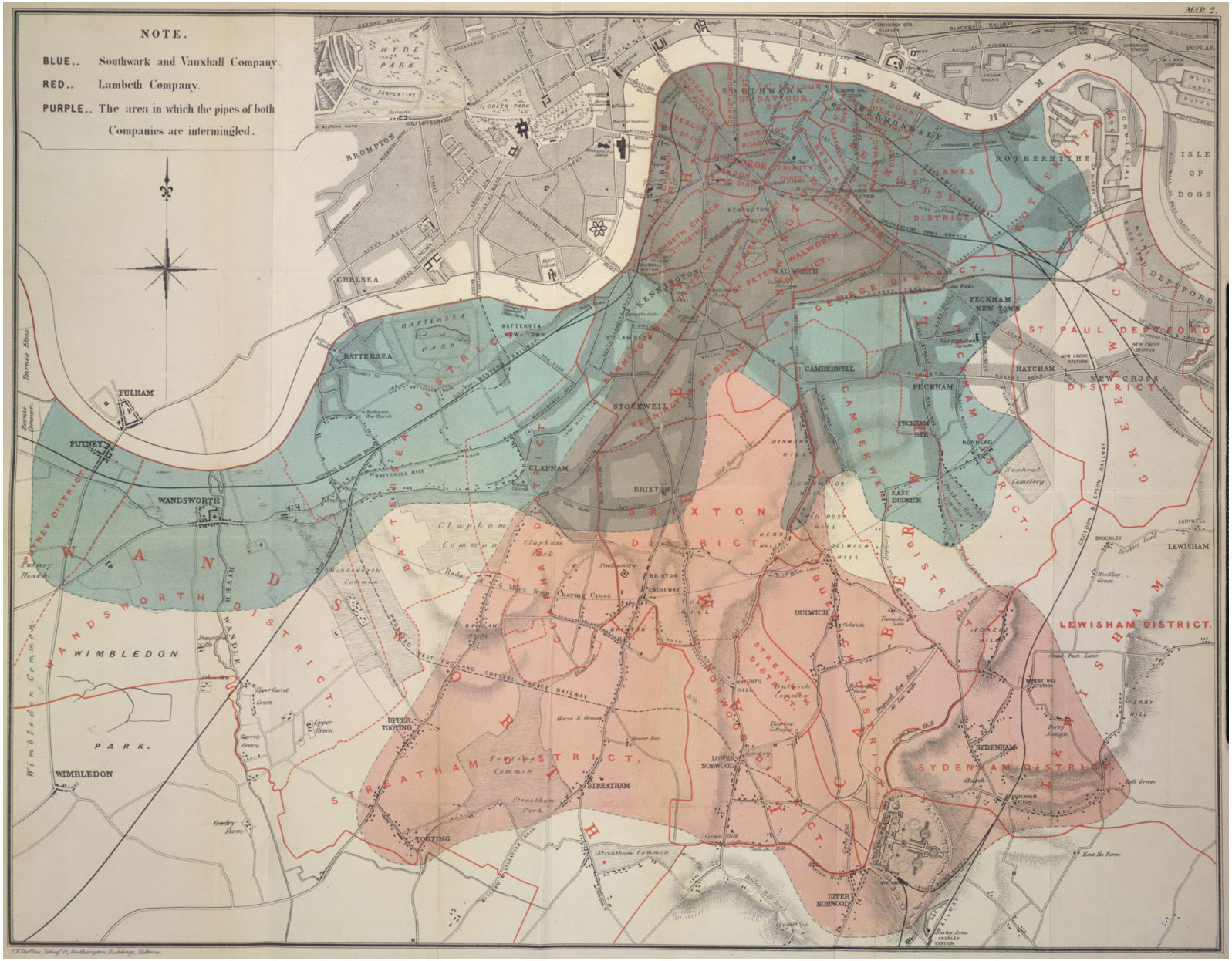
Water infrastructure crucial for public health

- sewerage key to public health, large mortality implications
- famous example: John Snow and the 1854 cholera outbreak
  - → first documented analysis of a natural experiment:
    - overlaid water pumps and deaths by house around Broad St
    - studied two water companies in London
    - both supplied adjacent neighbourhoods with water
  - → during 1849 cholera epidemic:
    - both drew from contaminated region of Thames
    - . . . and consumer death rates were similar
  - → cholera reappeared in 1853-54:
    - but one company had moved its intake upstream
    - . . . and only consumers with polluted water died
- later credited as 'founding event' of epidemiology
- clean water hugely improved public health in Victorian Britain

Snow's analysis of 1854 cholera outbreak: Broad St pump



Snow's analysis of 1854 cholera outbreak: water intake



How to provide clean water and sewerage infrastructure?

- SDGs: reduce child mortality, clean water $\rightarrow$ actually linked
- clean water critical to contain spread of infectious diseases
- $>$3m children die from preventable water-related diseases
- most low income countries aim to increase clean water access

How to provide clean water and sewerage infrastructure?

- SDGs: reduce child mortality, clean water $\rightarrow$ actually linked
- clean water critical to contain spread of infectious diseases
- $>$3m children die from preventable water-related diseases
- most low income countries aim to increase clean water access

↯ unclear how to actually achieve access to clean water:

  - public provision prohibitively expensive without extra funding
  - private provision may provide own funding and efficiency gains
  - unclear if and how efficiency gains can be translated into expanded access, enhanced quality and better health outcomes

How to provide clean water and sewerage infrastructure?

- SDGs: reduce child mortality, clean water $\rightarrow$ actually linked
- clean water critical to contain spread of infectious diseases
- $>$3m children die from preventable water-related diseases
- most low income countries aim to increase clean water access

↯ unclear how to actually achieve access to clean water:
  - public provision prohibitively expensive without extra funding
  - private provision may provide own funding and efficiency gains
  - unclear if and how efficiency gains can be translated into expanded access, enhanced quality and better health outcomes
  - ↯ private water companies may provide suboptimal service levels since they fail to internalise large health externalities
  - ↯ private water provision potentially regressive due to higher fees
  - ↯ exploitation of natural monopoly, threat of regulatory capture
  - ↯ potential efficiency–equity trade-off (and unclear pass-through of efficiency gains)

*** Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120

Galiani et al. (2005): 'water for life'

- examine water privatisation on child mortality in Argentina:
  - focus on children, highly vulnerable to water-related disease
  - weak body defenses, higher susceptibility, and greater exposure
- clean water could easily prevent several deadly diseases:
  - e.g. diarrhea, septicemia, and gastrointestinal infections are three of top ten causes of death for children under five
- two main disease transmission mechanisms:
  1. waterborne diseases that occur by drinking contaminated water
  2. water-washed diseases due to lack of water and sanitation

Galiani et al. (2005): Argentinian partial water privatisation

- exploit that local governments responsible for water services
- only selection of municipalities privatised water services:
  - in 1990s, about 30% of Argentinian municipalities privatised
  - approx. 60% of population covered by private water services
- variation in ownership across time and space
- provides a potential instrument to identify the causal effect of privatisation on child mortality

Galiani et al. (2005): privatisation progress

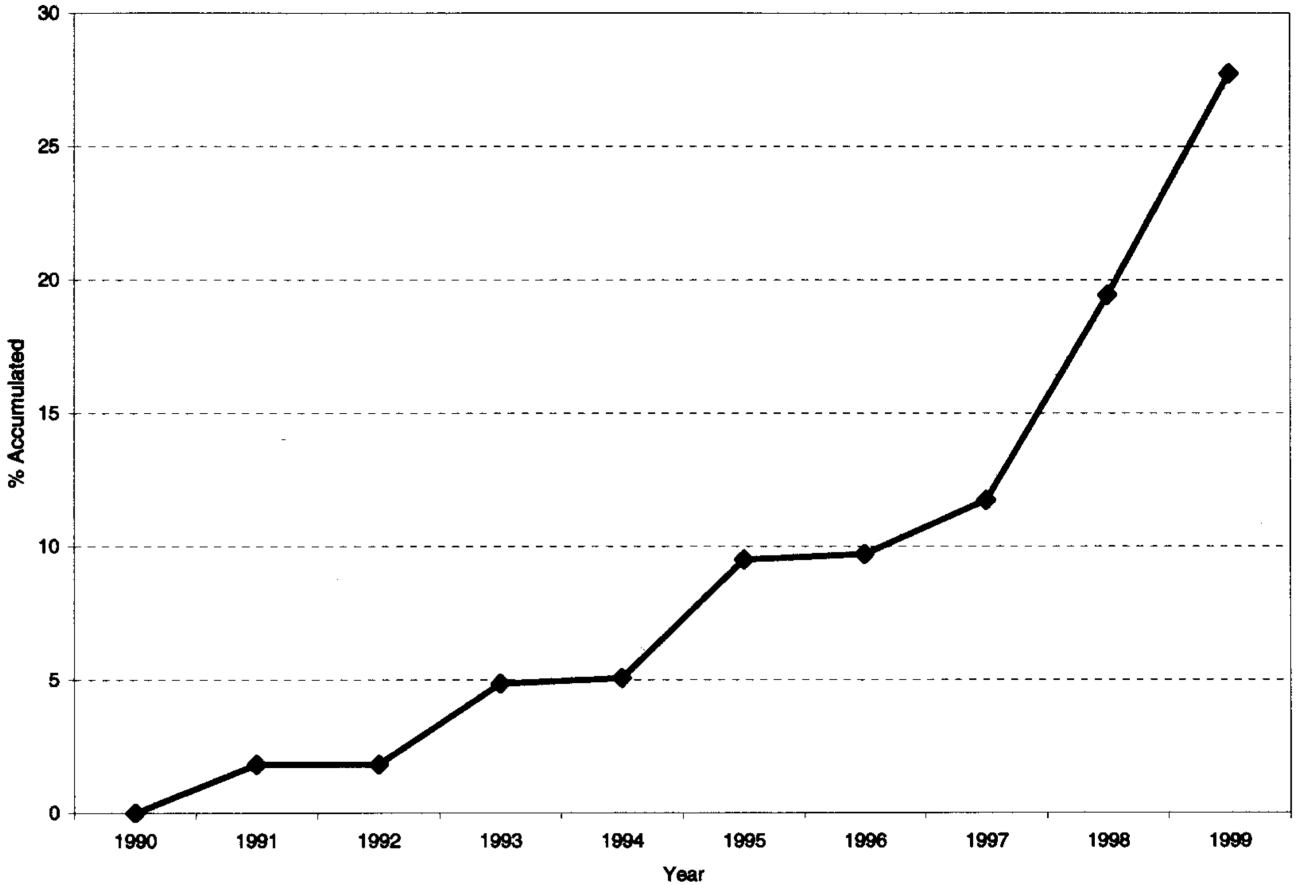


FIG. 2.—Percentage of municipalities with privatized water systems

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): water company ownership changes

TABLE 1
CHANGE IN OWNERSHIP OF WATER SYSTEMS, 1990–99

Ownership	Number of Municipalities	Percentage
Always public	196	39.7
Always private not-for-profit cooperative	143	28.9
Transferred from public to private for-profit	137	27.7
Always private for-profit	1	.2
No service or missing information	17	3.4
Total	494	100.0

NOTE.—In municipalities in which more than one company provides water services, we defined the ownership status of the municipality as the ownership of the company supplying the largest fraction of the population. Source: SPIDES, ENOHSA.

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): privatisation determinants

TABLE 2
DISCRETE-TIME HAZARD ESTIMATE OF THE PROBABILITY OF BEING PRIVATIZED

	Mean (Standard Deviation) (1)	Model 1 (2)	Model 2 (3)
Time-varying covariates:			
Federal government operates services (=1)	.018 (.134)	15.975*** (2.719)	16.035*** (2.727)
Local government by Radical party (=1)	.139 (.346)	−3.198*** (1.067)	−3.204*** (1.067)
Local government by Peronist party (=1)	.719 (.449)	−.042 (.401)	−.054 (.402)
Δ log GDP per capita$_{t-1}$	.047 (.135)	4.294 (3.567)	4.259 (3.561)
Δ unemployment rate$_{t-1}$	.006 (.029)	−6.692 (5.696)	−6.805 (5.711)
Δ income inequality$_{t-1}$	.005 (.014)	.483 (7.483)	.139 (7.503)
Δ child mortality rate$_{t-1}$	−.266 (2.994)		.034 (.043)
Fixed pretreatment characteristics as of 1991:			
GDP per capita	60.601 (30.388)	−.022*** (.007)	−.022*** (.008)
Unemployment rate	.045 (.023)	12.871** (5.384)	12.790** (5.383)
Income inequality	.452 (.021)	−3.591 (5.820)	−3.469 (5.805)
Child mortality rate	6.208 (3.683)		−.009 (.036)
Population is 5,000–25,000 (=1)	.419 (.493)	.227 (.471)	.225 (.480)
Population is 25,000–50,000 (=1)	.202 (.402)	.106 (.535)	.110 (.540)
Population is 50,000–100,000 (=1)	.114 (.318)	−.261 (.605)	−.256 (.610)
Population is 100,000–250,000 (=1)	.079 (.269)	.663 (.612)	.668 (.615)
Population is more than 250,000 (=1)	.066 (.249)	1.159* (.631)	1.151* (.640)
Proportion of families with UBN	.246 (.151)	−13.660** (6.067)	−13.328** (6.226)
Proportion of families living in overcrowded housing	.097 (.059)	13.560* (7.150)	13.444* (7.200)
Proportion of families living in poor housing	.060 (.049)	6.980** (3.472)	6.987** (3.451)
Proportion of families living below subsistence	.036 (.022)	5.221 (7.418)	4.917 (7.449)
Proportion of houses with no toilet	.095 (.117)	10.143** (4.429)	9.798** (4.563)

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): expansion in connections

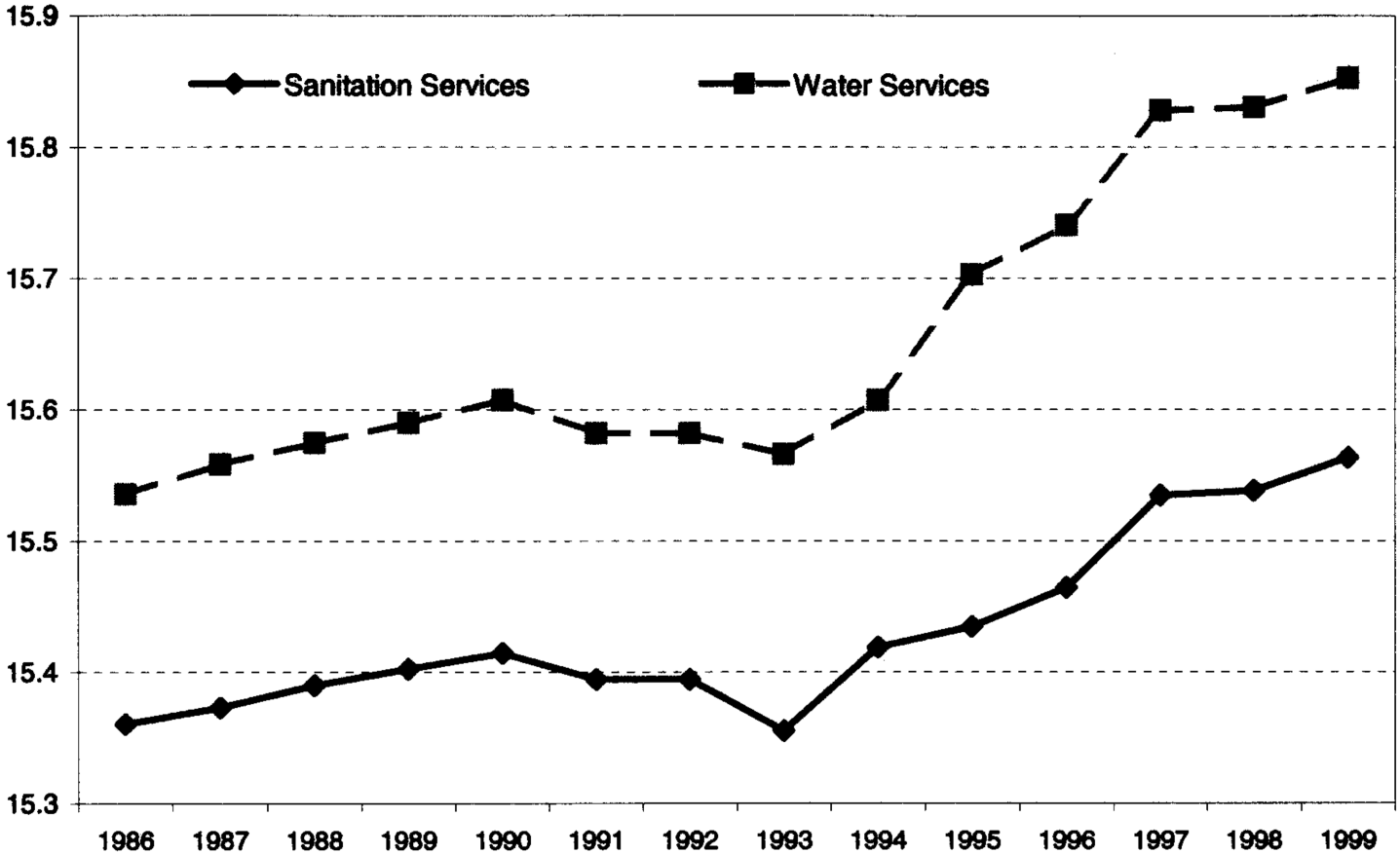


FIG. 3.—Logarithm of population connected to OSN–Aguas Argentinas network

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): mortality rates by ownership

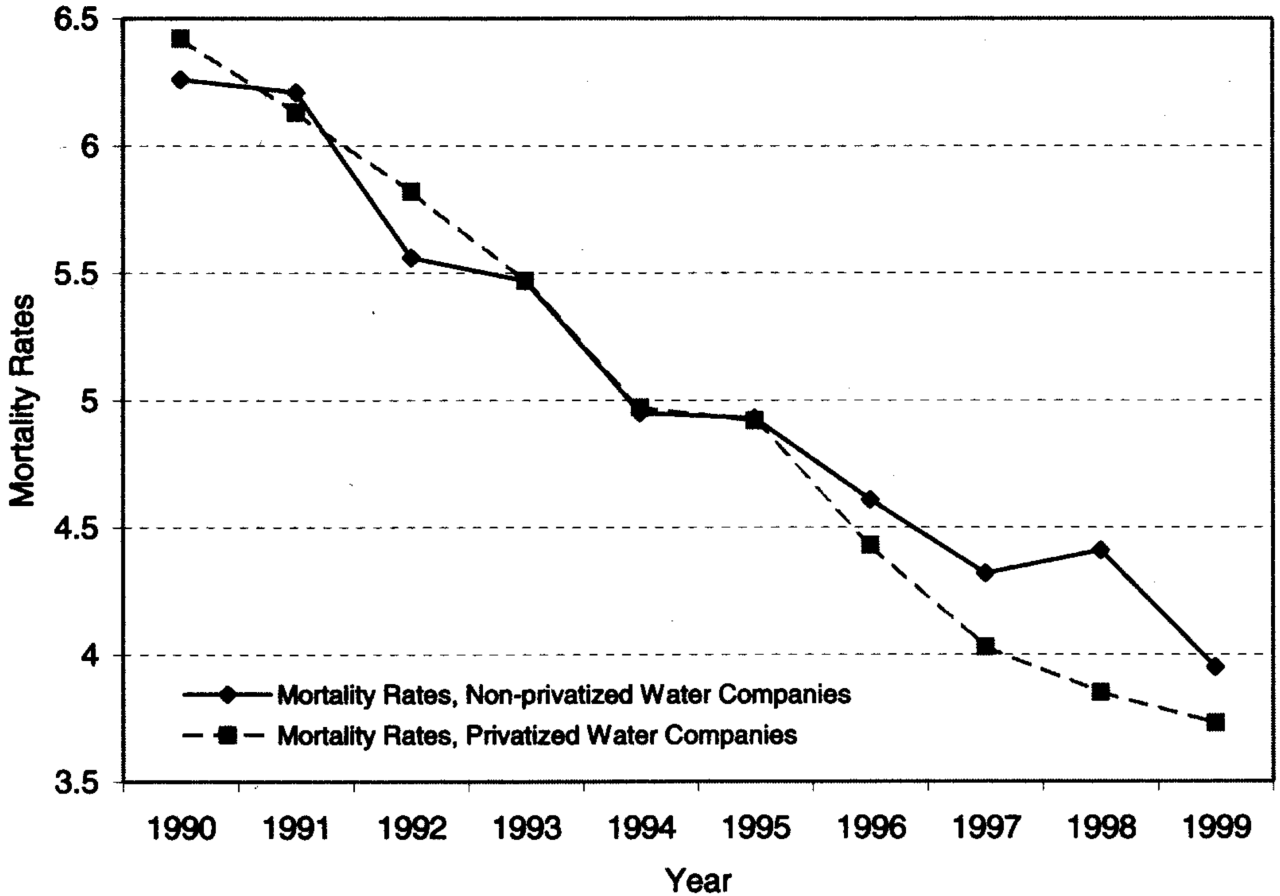


FIG. 1.—Evolution of mortality rates for municipalities with privatized vs. nonprivatized water services

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): child mortality on privatisation

TABLE 3
IMPACT OF PRIVATIZATION OF WATER SERVICES ON CHILD MORTALITY

	Full Sample			Using Observations on Common Support			Kernel Matching on Common Support[a]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Private water services (=1)	−.334	−.320	−.283	−.540	−.541	−.525	− .604
	(.169)**	(.170)*	(.170)*	(.177)***	(.178)***	(.178)***	(.168)***
	[.157]**	[.163]**	[.162]*	[.191]***	[.198]***	[.195]***	
	{.195}*	{.203}	{.194}	{.261}**	{.274}**	{.266}**	
%Δ in mortality rate	−5.3	−5.1	−4.5	−8.6	−8.6	−8.4	−9.7
Other covariates:							
Real GDP per capita		.007	.009		.005	.006	
		(.005)	(.006)		(.006)	(.006)	
		[.006]	[.006]		[.007]	[.007]	
		{.007}	{.007}		{.007}	{.008}	
Unemployment rate		−.555	−.636		−.778	−.836	
		(1.757)	(1.758)		(1.797)	(1.802)	
		[2.161]	[2.166]		[2.249]	[2.263]	
		{2.862}	{2.846}		{2.635}	{2.635}	

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): child mortality by cause of death

TABLE 4
IMPACT OF PRIVATIZATION ON CHILD MORTALITY BY CAUSE OF DEATH

	1990 Mean Mortality Rate	Estimated Impact Coefficients	%Δ in Mortality Rate
Infectious and parasitic diseases	.565	−.103 (.048)** [.055]* {.068}	−18.2
Perinatal deaths	2.316	−.266 (.105)** [.107]** {.123}**	−11.5
All other causes in aggregate	2.565	−.082 (.114) [.101] {.109}	−3.2
All other causes disaggregated:			
Accidents	.399	−.004 (.057)	…
Congenital anomalies	.711	−.022 (.056)	…
Skin and soft-tissue diseases	.000	.000 (.001)	…
Blood and hematologic diseases	.024	−.002 (.008)	…
Nervous system disorders	.163	.025 (.026)	…
Cardiovascular diseases	.236	.006 (.030)	…
Gastrointestinal tract disorders	.051	−.001 (.010)	…
Genital and urinary diseases	.020	−.006 (.007)	…
Osteoarticular and connective tissue diseases	.003	−.001 (.001)	…
Respiratory diseases	.511	−.038 (.051)	…
Immunodeficiencies, endocrine, and nutrition system diseases	.376	−.035 (.033)	…
Mental disorders	.002	.001 (.001)	…
Tumors	.068	−.006 (.015)	…

NOTE.—Each cell reports the estimated coefficient on the private water services dummy from a different difference-in-differences regression. Standard errors are in parentheses. Standard errors clustered at the municipality level are in brackets. Standard errors clustered at the province-year level are in braces. All the regressions include year and municipality fixed effects. All the regressions use the 3,870 observations on the common support and the socioeconomic

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): type of newly connected households

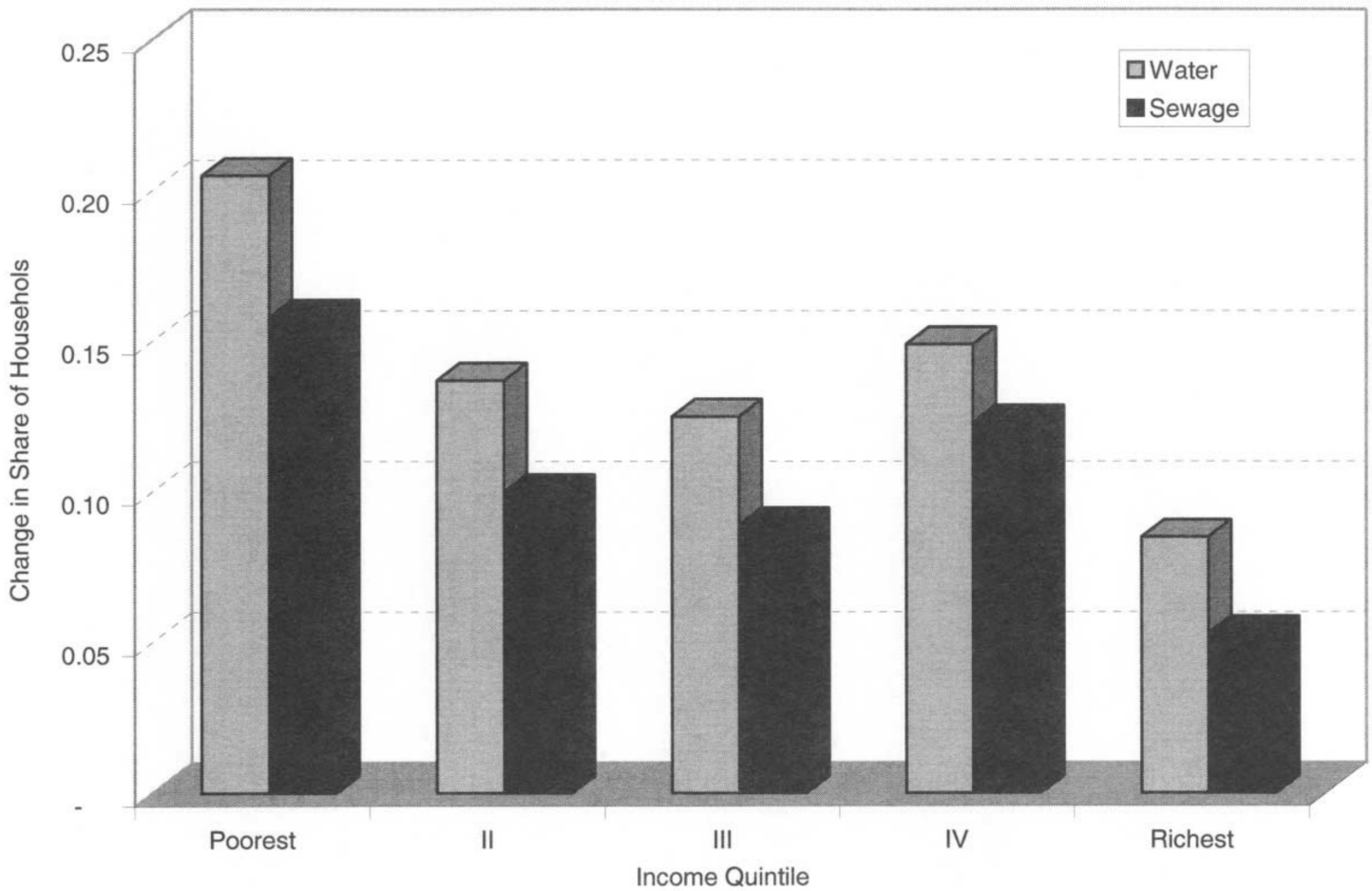


FIG. 4.—Change in share of households connected to water and sewerage, 1992–2002

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): expansion by income group

TABLE 7
NETWORK EXPANSION BY INCOME GROUP IN GREATER BUENOS AIRES, 1993–2000

Income Level	New Connections	Percentage
High and upper-middle income	90,200	15.4
Lower-middle income	282,250	48.3
Low income	211,800	36.3
Total	584,250	100.0

SOURCE.—Subsecretaría de Recursos Hídricos, from Abdala and Spiller (1999).

Source: Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.

Galiani et al. (2005): conclusion

- child mortality fell by approximately 8 percent in the areas in which water systems were privatized

- several pieces of evidence to support causal chain:
  1. privatisation across municipalities and time does not depend on time-varying variables that may also affect mortality rates
  2. similar trends in pre-intervention period
  3. water privatisation affected child mortality from water-related diseases, but it showed no effect on deaths from other causes
  4. the impact of privatisation was largest in poorest areas

Galiani et al. (2005): policy implications

Policy implications from Argentinian experience:

- private operation reduces costs, raises productivity and profits

↯ however, unclear if privatisation actually increases welfare

→ privatisation reduces child mortality, a good welfare indicator

→ no evidence for underinvestment compared to highly inefficient public water provision

→ efficiency gains so large that it generated private profits, improved access, expanded service, reduced child mortality

↯ growing public perception that privatisation hurts the poor

→ poorest population experienced the largest gains

→ privatisation appears progressive in reducing health inequality

- external validity?

Testing spatial equilibrium in LICs

Spatial GE models have a simple, but powerful prediction:

→ in equilibrium, real wages (or: expected utility) across all locations equalised
  - in other words, in equilibrium, no worker has an incentive to move anywhere else
  - intuition follows from classic Rosen-Roback model of compensating differentials (Roback, 1982)
  - e.g. require higher wage to counter lower amenities; or high rents in exchange for high wages in a given location, etc.

Does spatial equilibrium approximately hold in today's low income countries?

Testing spatial equilibrium in LICs (Gollin et al., 2017)

- examine differences in living standards across space within countries
- compare households on continuum of population density (instead of potentially ambiguous urban vs rural)
- use consistent real measures of living standards derived from internationally comparable DHS household surveys

Guiding questions

- do observed rural-urban differences show up in spatially disaggregated data?
- what kinds of variation do we see *within* rural areas?
- what does this new evidence allow us to infer about the reasons why geographic disparities exist and persist?

Data

- data on living standards from DHS data
  - nationally representative surveys covering large numbers of households
  - consistent methodology and definitions across countries
  - use all DHS surveys from 2005 with available GPS coordinates for survey clusters (293,517 households, 25 countries)
- data on population density measures from WorldPop
  - provides population density measures at a resolution of about 100m at the equator
  - combines multiple data sources and uses various interpolation measures
  - e.g., for Tanzania, correlation between WorldPop density estimates and census estimates of population density is 0.93 with a p-value of 0.000

Key findings

- living standards appear to vary within countries in consistent ways:
  - strong positive relationship between measured living standards and population density
  - relationship is nearly monotonic and cuts across a range of different indicators
- relationship holds even after we take into account differences in education
- migrants to cities do not appear to face very serious risk of bad outcomes, at least in the medium term

Implications

- evidence seems clear that living standards vary across the density distribution
- consistent with rural-urban differences and sectoral differences, but more puzzling in some ways:
  - many barriers or frictions could discourage people from moving from rural areas to big cities
  - but fewer obvious barriers to stop people from moving from rural areas to slightly more densely populated areas

↯ limited data on additional sources of utility and disutility
  - crime
  - direct measures of life satisfaction

→ ambiguous, but suggest some relatively modest decreases in subjective well-being

Conclusions on spatial equilibrium in LICs

- living standards differ strikingly within countries
  - not just individual heterogeneity and inequality
  - systematic differences across space and sector
- differences remain even when we use real measures and unambiguous spatial data
- differences across space persist *within* populations at similar educational levels
- migrants do not appear to face significant risk of bad material outcomes – but subjective experiences may be important
- need to think further about the kinds of costs, barriers, frictions, and inefficiencies that drive behaviour
- but our observations do not necessarily imply misallocation
  - possible to have 'gaps' without 'wedges'

References I

- Asher, S., & Novosad, P. (2019). Rural roads and local economic development. *American Economic Review*, *110*(3), 797–823.
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2010). The miracle of microfinance? Evidence from a randomized evaluation. *MIT Working Paper*.
- Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, *108*(4-5), 899–934.
- Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, *91*(4), 795–813.
- Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway system. *The Review of Economic Studies*, *81*(3).
- Fleming, M. (1955). External economies and the doctrine of balanced growth. *The Economic journal (London)*, *65*(258), 241–256.

References II

- Fujita, M., Krugman, P., & Venables, A. (1999). *The spatial economy: Cities, regions and international trade*. MIT Press.
- Galiani, S., Gertler, P., & Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *The Journal of political economy*, *113*(1), 83–120.
- Gollin, D., Kirchberger, M., & Lagakos, D. (2017). In search of a spatial equilibrium in the developing world. *NBER Working Paper Series*, (23916).
- Gollin, D., & Rogerson, R. (2014). Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, *107*.
- Hirschman, A. (1958). *The strategy of economic development*. Yale University Press.
- Hsiao, A. (2023). Educational investment in spatial equilibrium: Evidence from Indonesia. *Princeton mimeo*.
- Khanna, G. (2022). Large-scale education reform in general equilibrium: Regression discontinuity evidence from India. *The Journal of political economy*.

References III

- Krugman, P. (1995). *Development, geography, and economic theory*.
- Lewis, W. A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, *22*(2), 139–191.
- Moneke, N. (2020). Can Big Push infrastructure unlock development? Evidence from Ethiopia. *University of Oxford mimeo*.
- Murphy, K., Shleifer, A., & Vishny, R. (1989). Industrialization and the big push. *The Journal of Political Economy*, *97*(5), 1003–1026.
- Nurkse, R. (1953). *Problems of capital formation in underdeveloped countries.*. Blackwell.
- Roback, J. (1982). Wages, rents, and the quality of life. *The Journal of political economy*, *90*(6), 1257–1278.
- Rosenstein-Rodan, P. (1943). Problems of industrialization of Eastern and Southeastern Europe. *Economic Journal*, *53*, 202–211.