The Curse of Connectivity:

Evidence from Indonesia's Village Resettlement Program

Harrison Mitchell
hamitchell@ucsd.edu

January 9, 2026

Motivation

- Infrastructure and market access are central to (spatial) development policy
  - Dominant thinking: $\uparrow$ market access $\rightarrow$ $\uparrow$ productivity and welfare
  - But, mixed evidence in developing contexts, difficulty gauging long-run impacts
- Another core policy focus centers around land ownership and use in both developing and developed countries
  - Frictions around ownership an impediment to agricultural advancement
  - Concerns around land supply primarily an urban focus

Question 1: What is the effect of market access on rural development?
Question 2: What role does land and land development play in explaining the effects of market access?

Talk Preview

- Indonesia's Transmigration Program as a unique setting to answer question
- Increased market access leads to *worse* outcomes in transmigration villages
- Not driven by sectoral stagnation, migration, increased external competition, or compensating differentials
- Despite increased firm entry, no increase in land devoted to firm activities
- An adapted firm dynamics model can explain these results

Talk Outline

- Background on study design and context
  - Village location determination
- Constructing market access and data used
  - Data digitization process
- Empirical strategy
  - Identification and robustness measures
- Empirical results
  - Headline results and mechanisms
- Quantitative model outline

Related Literature

- The positive effects of market access:
  - Donaldson and Hornbeck (2016), Hornbeck and Rotemberg (2022), Redding and Strum (2008): Long term variation, granular data
- Mixed evidence in developing countries:
  - Asher and Novosad (2020), Gertler et al. (2024), Morten and Oliveira (2024): Find *negative* effects
  - Faber (2014): Rule out "trade channel" in this context
- STEG 2026:
  - Fan, Peters, Zilibotti: Increase in services associated with negative effects
  - Brooks, Donovan, Thomas: Different land constraints, long run outcomes

Roadmap

- Context and Data
- Empirical Strategy
- Empirical Results
- Ruling Out Mechanisms
- Land Channel
- Quantitative Model

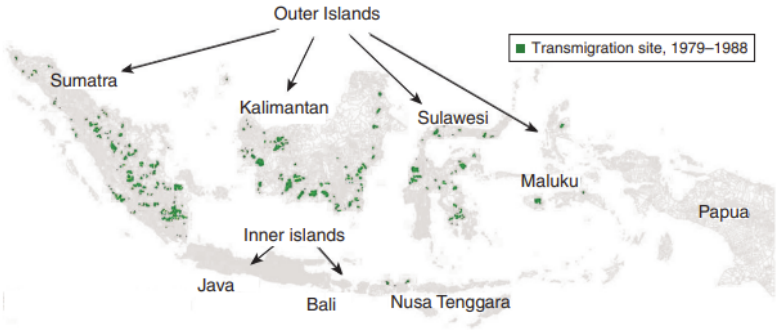
Context

- The government of Indonesia relocated over 2 million individuals into *newly created* villages during the 1950s-1990s
  - Households received the same sized plots of land via lottery upon arrival - Evidence
  - Migrants were uniformly without assets prior to migration - Migrant Survey
  - Newly created villages are balanced on observable characteristics throughout the resettlement period - Balance
  - Program intensity determined by external factor - Oil Prices
- At the time of creation, the traits of villagers, village infrastructure, and sectoral composition of the village were *orthogonal* to all other factors

Exogenous Assignment of Migrants

- No incentive to match productive migrants with productive locations (Dove, 1985)
- Migrants almost uniform in prior experience and income
- Explicit nation-building aim led to intentional mixing (Levang, 1995)
- Small number of ports meant no/low wait times (Hardjono, 1988)
- Migrants had almost no knowledge of destination prior to migration - Migrant Survey

Geographic Distribution of Villages - Outer Islands



Notes: Each colored location on the map corresponds to a Transmigration village settled between 1979 and 1988. The white areas outlined in gray are other villages.

Figure 1: Transmigration Village Locations in Outer Islands, from Bazzi et al., 2016

Constructing Initial Connectivity

- Maps from circa 1950 via Perry-Castañeda Library Map Collection and US Army
- Hand digitized and stitched together Digitization Process
- Estimate cost between each transmigration origin o and all other Desa's using FMM and vareity of cost parameters, baseline population information
- Initial market access approximated following Donaldson and Hornbeck (2016) as:

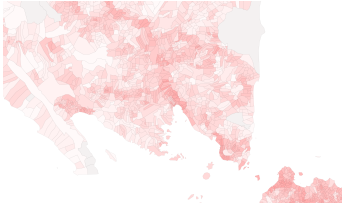
$$MA_o \approx \sum_d \tau_{od}^{-\theta} N_d \tag{1}$$

- Results robust to other measures of connectivity
- Relationship between θ and estimated elasticities

Map Digitization And Market Access Construction

(a) Original Image

(b) Digitized Road Network



(c) Constructed Market Access

- Across island variation in MA in levels
- Higher MA correlated with other measures of connectivity, modern measures

Village Location Assignment

- Unrealistic relocation targets, administrative difficulties, and poor land-use plans meant that villages were often simply placed wherever land was available
- Provincial government often had final say in location decision
- Find no evidence that observable characteristics of villages vary with market access

"As a consequence of the focus on numbers, the land-use plans developed during the 1970s were totally abandoned. Transmigrants were placed on whatever land was submitted by provincial governments for settlement purposes" (Hardjono, 1988)

Balance

Dependent variable (std): ln(Market Access) (std)	
Wetland Rice Potential Yield	0.023 (0.033)
Palm Oil Potential Yield	0.016 (0.046)
Coffee Potential Yield	-0.014 (0.070)
Cassava Potential Yield	-0.010 (0.034)
Vector Ruggedness Measure	-0.019 (0.038)
Individuals Placed	-0.096*** (0.035)
Agrosimilarity Score	0.053 (0.047)
Share of Land with Poor Drainage	0.012 (0.060)
Share of Land with Good Drainage	0.005 (0.059)
Organic Carbon (%)	0.016 (0.053)
Topsoil Salinity (Elco) (dS/m)	0.014 (0.018)
Topsoil Base Saturation (%)	-0.013 (0.043)
Avg. rainfall, 1948–1978	0.043 (0.039)
Avg. temp (Celsius), 1948–1978	0.071* (0.039)
ln(District Population, 1980)	-0.058 (0.039)
District Own Electricity, 1980	0.053 (0.038)
District Own Radio, 1980	0.042 (0.041)
District Own TV, 1980	0.059 (0.038)

Dependent variable (std): ln(Market Access) (std)	
District Literacy, 1980	-0.034 (0.039)
District Avg Years Schooling, 1980	-0.019 (0.038)
District Agricultural Employment Share, 1980	-0.043 (0.042)
District Wetland Rice Potential Yield, 1980	0.001 (0.033)
District Palmoil Potential Yield, 1980	0.026 (0.036)
District Coffee Potential Yield, 1980	-0.044 (0.042)
District Cassava Potential Yield, 1980	-0.003 (0.035)
District Mining Employment Share, 1980	0.022 (0.039)
District Manufacturing Employment Share, 1980	-0.020 (0.039)
District Trade Employment Share, 1980	0.037 (0.045)
District Service Employment Share, 1980	0.053 (0.038)
District Wage Worker Share, 1980	-0.005 (0.038)
District Share HHs with Piped Water, 1980	0.092** (0.040)
District Share HHs with Sewage, 1980	0.061 (0.039)
District Share HHs with Modern Fuel, 1980	0.072* (0.039)
District Share HHs with Modern Floor, 1980	-0.037 (0.033)
District Share HHs with Modern Roof, 1980	-0.016 (0.038)

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. F-Stat (omnibus) $= 1.594$. Observations $= 1214$.

Specification Details

Additional Data Used

- Transmigration Census
  - Location and settlement details
- Local GDP per Capita
  - From Rossi-Hansberg and Zhang (2025), validated with nightlights
- Population Census
  - Sector of employment, employment type, migrant status, education
- Economic Census
  - Medium and large firm location and employment
- PODES
  - Ag productivity, amenities, migration patterns, small firm counts, land use
- SUSENAS
  - Food prices, expenditure for non-food, housing characteristics, wages
- Firm surveys
  - Indonesian manufacturing firm panel (SI), Informal Sector Enterprise Survey (8 countries, between 2010-2014)

Roadmap

- Context and Data
- Empirical Strategy
- Empirical Results
- Ruling Out Mechanisms
- Land Channel
- Quantitative Model

Empirical Strategy

- Estimate effect of initial market access for village v, in region r, settled in year c, on outcomes in year t:

$$y_{vrct} = \beta_1 ln(MA)_{vrc} + \beta_2 E(ln(MA))_{vrc} + \gamma_c + \eta_r + \mathbf{X}'_{vrc}\boldsymbol{\delta} + \alpha_t + \epsilon_{vrct}$$

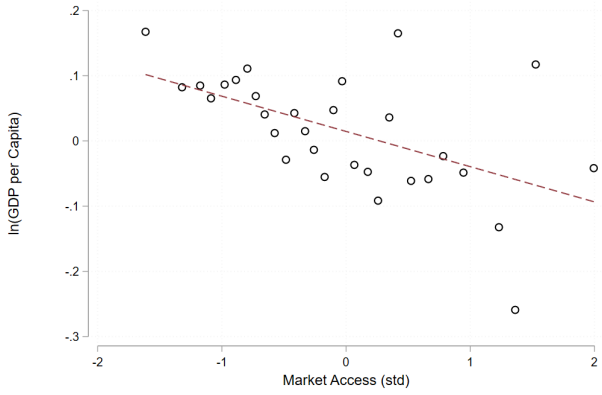
- Cohort FE motivated by temporal variation in program intensity
- Include "expected market access" as suggested by Borusyak and Hull (2023)
- Controls include rich set of info on: soil quality, geography, agricultural productivity of settlers, and initial conditions of district
- Time FE are included if multiple waves of data exist
- Additional work done for the identification argument and robustness Identification

Roadmap

- Context and Data
- Empirical Strategy
- Empirical Results
- Ruling Out Mechanisms
- Land Channel
- Quantitative Model

GDP per Capita and Market Access

Figure 3: Binscatter of GDP per Capita and Market Access



- 1 std $\uparrow$ MA $\rightarrow$ $\sim 5\% \downarrow$ GDP per capita. Includes controls, island FE

Market Access and GDP per Capita

Table 1: Effects of MA on Aggregate Output

	ln(GDP per Cap)	ln(GDP per Cap)	ln(GDP per Cap)	ln(GDP per Cap)
ln(MA)	-0.0060***	-0.0059***	-0.0057***	-0.0058***
	(0.0022)	(0.0019)	(0.0019)	(0.0020)
Expected MA Control	No	No	No	Yes
Local Controls	Yes	Yes	Yes	Yes
District Controls	No	No	Yes	Yes
Cohort FE	No	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Unique Villages	1152	1152	1152	1152
DMP Fraction	0.714	0.643	0.643	0.857

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- More connected villages are worse off today than less connected villages
- Effect of moving from 25th to 75th percentile of MA: $\downarrow$ GDP per Capita of 10.6%

Market Access and Productivity

Table 2: Effects of MA on Productivity

	Employees per Employer	Tractors per 1000 Individuals	Tractors per 1000 Ag Workers	ln(Productivity of Top 5 Crops)	ln(Rice Productivity)	ln(Palm Oil Productivity)
ln(MA)	-0.6994**	-0.0501**	-0.0555**	-0.0095*	-0.0097*	-0.0073
	(0.2858)	(0.0203)	(0.0247)	(0.0053)	(0.0050)	(0.0217)
Expected MA Control	Yes	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	No	No	No	No
Unique Villages	1105	1030	1030	974	583	296

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Proxies for productivity in non-ag (employees per employer) and ag all lower - Wages

Roadmap

- Context and Data
- Empirical Strategy
- Empirical Results
- Ruling Out Mechanisms
- Land Channel
- Quantitative Model

Ruling Out Potential Channels

- Nearby demand forces specialization in agriculture (Matsuyama, 1992)
  - More connected locations specialize in agriculture to serve urban markets
- Increased connectivity implies improved migration opportunities Results
  - Lower population density, lower positive agglomeration
  - Selective out-migration, "brain drain"
  - Potential for commuting and improved work opportunities outside village
- Amenities as a compensating differential Results
  - Lower output made up for by better amenities in high MA locations
- Competition effects force local firms to exit
  - Fewer firms (but larger), lower prices from imports

Agricultural Employment

Table 3: Effects of MA on Structural Transformation

	Agriculture Share	Change in Ag Share	Processing Share	Services Share	Trade Share
ln(Market Access)	-0.0038***	-0.0027***	0.0013***	0.0017***	0.0009***
	(0.0008)	(0.0007)	(0.0003)	(0.0004)	(0.0002)
Expected MA Control	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	Yes	Yes
Unique Villages	1152	1147	1152	1152	1152

Standard errors in parentheses

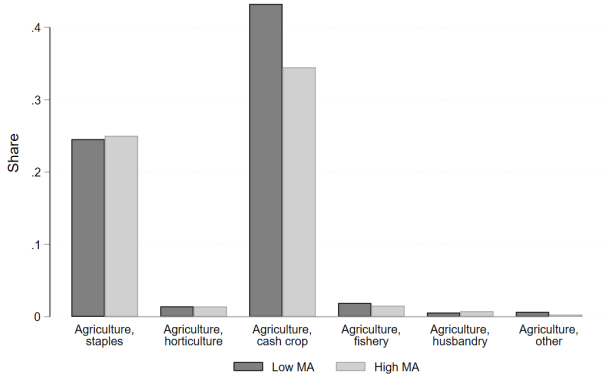
Standard errors clustered at the village level

$^{*}\ p < .1,\ ^{**}\ p < .05,\ ^{***}\ p < .01$

- See *increased* structural transformation in more connected villages
- Evidence for divergence over time Evidence

Type of Agricultural Employment

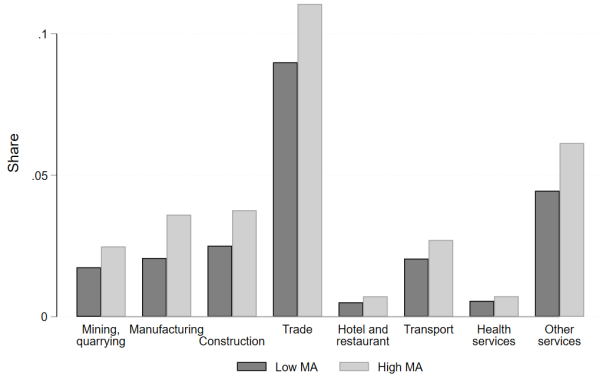
Figure 4: Share of Workers by Agricultural Employment Type



- Predominately decline in *cash crop* employment

Other Employment Shares

Figure 5: Share of Workers by Non-Agricultural Employment Type



- Large increase in retail services, transport

Firm Dynamics

	Employers per 1000	Total Number of Firms	ln(Employment)	ln(Employment)
ln(MA)	1.7717***	0.1331*		
	(0.5354)	(0.0785)		
ln(Market Access)			-0.0289**	-0.0244**
			(0.0130)	(0.0102)
Expected MA Control	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes
Year FE	Yes	No	No	No
Unique Villages	1152	944	128	128

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- No evidence that competition lead more exposed villages to have fewer, more productive firms

Prices and Household Expenditure

	Price Index^{Ind}	Price Index^{Ref}	Quantity Index^{Ind}	Quantity Index^{Ref}	ln(Food Expenditure)
ln(MA)	0.0135***	0.0106***	0.0060	0.0133***	0.0284***
	(0.0034)	(0.0031)	(0.0048)	(0.0044)	(0.0088)
Expected MA Control	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Unique Villages	84	84	84	84	84
Unique Households	1234	1234	1234	1234	1234

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Ind uses household-specific expenditure weights, Ref uses avg expenditure share for rural non-TM households on same island
- Lower output not compensated for by decreased price level

Prices and Household Expenditure

	ln(Housing Expenditure)	ln(Housing Expenditure)	ln(House Area)	ln(Price Per Sq Meter)
ln(MA)	0.0242***	0.0208***	-0.0006	0.0215***
	(0.0078)	(0.0069)	(0.0046)	(0.0072)
Household Size		0.0540***	0.0771***	0.0141
		(0.0090)	(0.0091)	(0.0111)
House Area		0.0101***		
		(0.0021)		
House Area, Sq		-0.0000		
		(0.0000)		
Expected MA Control	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes
District Controls	No	No	No	No
Cohort FE	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes
Unique Villages	70	70	70	70
Unique Households	1086	1086	1086	1086

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Largest non-food expenditure category, only where I can measure price
- Price of land higher in high MA villages, quantity unchanged

Overview of Alternative Channel Results

- More connected villages do not specialize in agriculture
- See increased number of firms in higher MA locations, higher prices
- No evidence of migration and commuting playing a role
- See opposite of brain drain results
- Amenities not a factor
- See evidence of increased price of land

Roadmap

- Context and Data
- Empirical Strategy
- Empirical Results
- Ruling Out Mechanisms
- Land Channel
- Quantitative Model

Land Use and Connectivity

- Given structural transformation, increased firm presence, and more population, expect to see changes in land use within a village
- Find evidence that land devoted to ag shrinks, but is *not* replaced by land use by firms and households
- Suggests high land conversion costs could impede growth

Empirical Evidence of Land Changes

Table 4: Effects of MA on Land Use Dynamics

	ln(Ag Land)	ln(Housing + Firm Land)	Change in Ag Land	Change in Housing + Firm Land	New Housing Development
ln(Market Access)	-0.0237***	-0.0155**	-2.7412**	-1.7831*	-0.0000
	(0.0056)	(0.0071)	(1.1285)	(1.0409)	(0.0009)
Expected MA Control	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	No
Unique Villages	940	935	944	944	1147

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- No new housing, ag land no replaced by housing and firm land
- Results qualitatively similar if land use instead measured in shares

Empirical Results Recap

- Increased initial market access leads to lower GDP per capita and productivity in villages
- Higher rates of firm entry, but smaller firm size
- Not driven by a suite of potential explanations previously studied
- Potential mechanism: relatively fixed supply of land for firms and housing

Roadmap

- Context and Data
- Empirical Strategy
- Empirical Results
- Ruling Out Mechanisms
- Land Channel
- Quantitative Model

Model Overview

- Extension of Hopenhayn (1992)
- Key addition: Land as an additional factor of production
  - (Relatively) fixed supply
  - Lumpy, likely non-homothetic production function
- Contribution 1: Direct evidence of non-homotheticity wrt land in firm production function
  - Evidence using a panel of Indonesian manufacturing firms, and support for relevance of this function from global surveys of informal firms Estimation Section
- Contribution 2: Increased exposure to "market access" can lead to declines in wages, productivity, even with increased firm entry Main Result

Model (No Land) - Incumbent Firm Problem

- Have productivity z_t, choose employment n_t, and stopping time (when to exit) τ. Must pay fixed per period cost $\bar{c}_f$

$$v(z) = \max_{\{n_t\}_{t\geq 0},\tau} \left\{ \mathbb{E}_0 \int_0^\tau e^{-\rho t}(pf(z_t, n_t) - wn_t - w\bar{c_f})dt \right\}$$
$$dz_t = \mu(z_t)dt + \sigma(z_t)dW_t, \quad z_0 = z$$

- let: $\pi(z) = \max_n\{pzn^\alpha - wn\} - w\bar{c_f}$, so

$$v(z) = \max_\tau \left\{ \mathbb{E}_0 \int_0^\tau e^{-\rho t}\pi(z_t)dt \right\}$$

Model (With Land) - Incumbent Firm Problem

- Also choose land use l_t, fixed per period cost $\bar{c}_f$ paid in land p_l and inside production function

$$v(z) = \max_{\{n_t\}_{t\geq 0}, \{l_t\}_{t\geq 0}, \tau} \left\{ \mathbb{E}_0 \int_0^\tau e^{-\rho t} (pf(z_t, n_t, l_t, \bar{c}_f) - wn_t - p_l l_t dt \right\}$$
$$dz_t = \mu(z_t)dt + \sigma(z_t)dW_t, \quad z_0 = z$$

- Now: $\pi(z) = \max_{n,l}\{pzn^\alpha(l - \bar{c}_f)^\beta - wn - p_l l\}$, so

$$v(z) = \max_\tau \left\{ \mathbb{E}_0 \int_0^\tau e^{-\rho t} \pi(z_t) dt \right\}$$

- Importance of functional form

Model - Entrant Problem

- Free entry condition determines number of firms
- Must pay entry cost, in labor, c_e, draws initial productivity from $\psi(z)$

$$wc_e = \int_0^1 v(z)\psi(z)dz$$

- Numerically

$$m = \bar{m}\text{exp}(\eta(\int_0^1 v(z)\psi(z)dz - wc_e))$$

Model - Closing Model

- Standard: Exogenous product demand and labor supply

$$p = D(Q), \quad w = W(N),$$
$$Q = \int_Z q(z)g(z)dz, \quad N = \int_Z n(z)g(z)dz$$

- Addition: Add land supply

$$p_l = Q_L(L), \quad L = \int_Z l(z)g(z)dz$$

- Parametrize as

$$L^S(p_l) = L_0 p_l^{\lambda}, \quad p_l = \left(\frac{L}{L_0}\right)^{\frac{1}{\lambda}}$$

Model - Equilibrium Solution

- Stationary equilibrium given by value function v, firm distribution g, prices (w, p, p_l), entry rate m, and exit region Z that solves:

$$0 = \min\Big\{\rho v(z) - v'(z)\mu(z) - \frac{1}{2}v''(z)\sigma^2(z) - \pi(z), v(z)\Big\} \quad (HJBVI)$$

$$0 = -\left(\mu(z)g(z)\right)' + \frac{1}{2}(\sigma^2(z)g(z))'' + m\psi(z), \forall z \in Z \qquad (KF)$$

$$m = \bar{m}\exp(\eta(\int_0^1 v(z)\psi(z)dz - wc_e)),$$

$$p_l = Q_L(L), \quad p = D(Q), \quad w = W(N),$$

$$L = \int_Z l(z)g(z)dz, \quad Q = \int_Z q(z)g(z)dz, \quad N = \int_Z n(z)g(z)dz$$

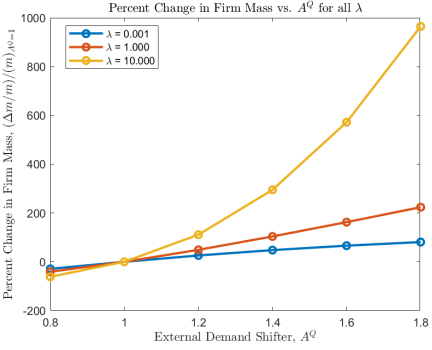
Model - The Role of Connectivity

- Currently modeling connectivity as an exposure to aggregate demand
- Increased access $\rightarrow$ higher exogenous demand for output
  - Intuition: being closer to other locations/consumers leads to higher demand
- Example of MA in the model

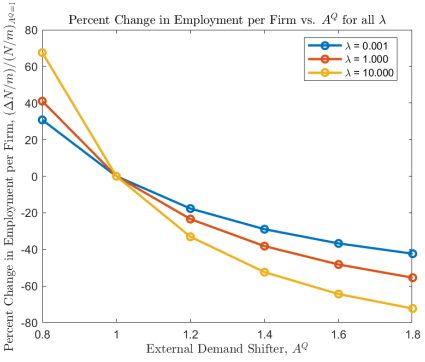
$$p = D(Q) \equiv A_Q Q^{-\epsilon}$$

- where $A_Q \equiv A\tau^{-\theta}$ is effective demand shifter
  - Higher MA from Donaldson and Hornbeck (2016) $\rightarrow A \uparrow$ or $\tau \downarrow$
- **Importantly**, this shuts down competition and "import" price effects

Effects of Increased Connectivity

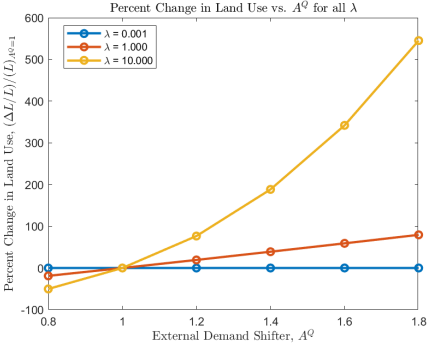


(a) Number of Firms

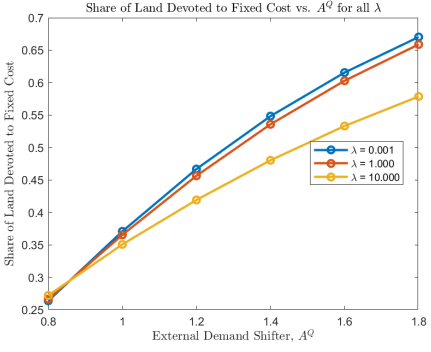


(b) Employment per Firm

Effects of Increased Connectivity

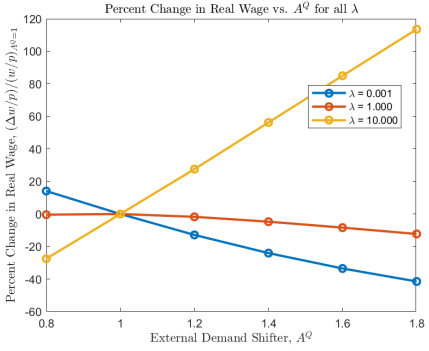


(a) Land Use

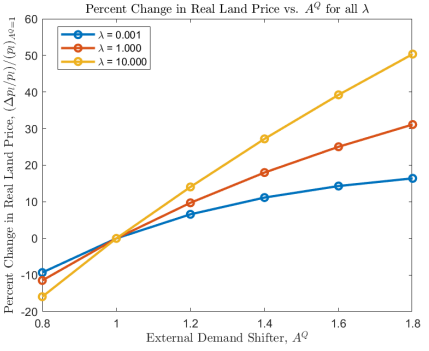


(b) Share of Land to $\bar{c}_f$

Effects of Increased Connectivity



(a) Wage



(b) Rent

Next Steps

- Conclusion:
  - Find evidence that higher initial market leads to welfare losses in my setting
  - Effects can be partially explained by inclusion of land into a firm dynamics model
  - Model results rely on shutting down two previously emphasized trade channels
  - Why sectoral transformation?
    - The wrong type of services, the palm oil boom, learning by doing?
- Thank you!

Village Lottery

Figure 9: From Journey to a New Life. The Human Face of Indonesia

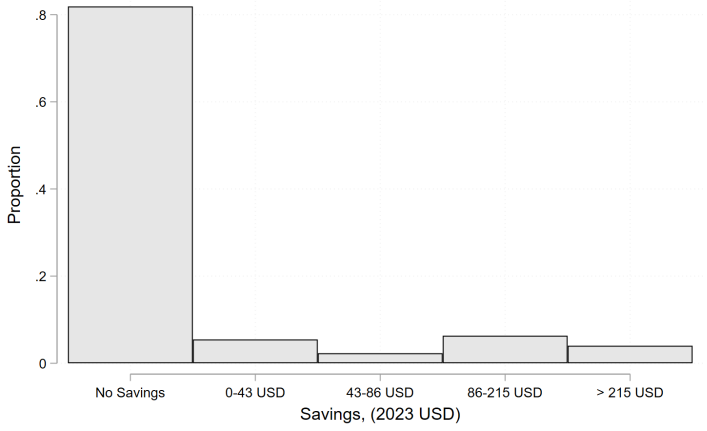
Example of New Village

Figure 10: From Transmigration Article, 2022

Back

Prior to Migration

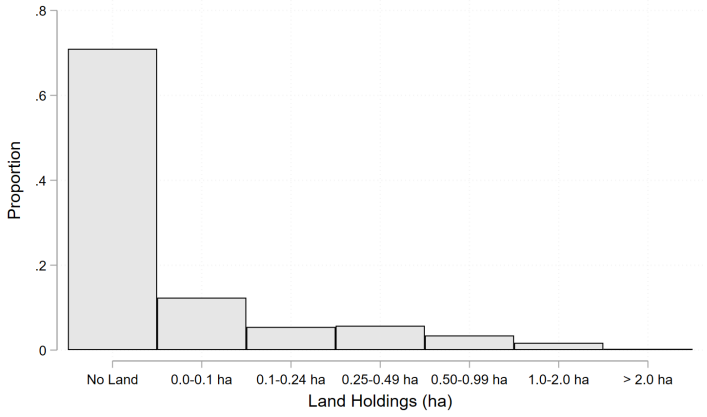
Figure 11: Asset Holdings of Sample of Migrants



Author's calculations, data from Kebschull (1986)

Prior to Migration

Figure 12: Land Holdings of Sample of Migrants



1 ha = 2.5 acres
Author's calculations, data from Kebschull (1986)

Back

Balance Across Time

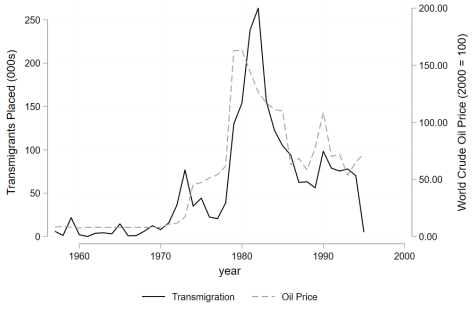
Table 5: Village Characteristics Across Creation Time

Dependent Variable (std)	Year Created
Wetland Rice Potential Yield	-0.011**
	(0.005)
Palm Oil Potential Yield	-0.006
	(0.004)
Coffee Potential Yield	-0.004
	(0.004)
Cassava Potential Yield	-0.011**
	(0.005)
Vector Ruggedness Measure	0.010
	(0.006)
Individuals Placed	-0.014*
	(0.008)
Agroclimlarity Score	-0.008
	(0.007)
Share of Land with Poor Drainage	-0.001
	(0.008)
Share of Land with Good Drainage	0.007
	(0.007)
Organic Carbon (%)	0.007
	(0.008)
Topsoil Salinity (Elco) (dS/m)	-0.004
	(0.003)
Topsoil Base Saturation (%)	-0.007
	(0.005)
Avg. rainfall, 1948-1978	0.008
	(0.005)
Avg. temp (Celcius), 1948-1978	0.007
	(0.005)
ln(District Population, 1960)	0.001
	(0.005)
District Own Electricity, 1960	0.003
	(0.005)
District Own Radio, 1960	-0.005
	(0.005)
District Own TV, 1960	0.004
	(0.005)
District Literacy, 1960	-0.002
	(0.005)
District Avg Years Schooling, 1960	0.001
	(0.005)
District Agricultural Employment Share, 1960	-0.005
	(0.005)
District Wetland Rice Potential Yield, 1960	-0.006
	(0.005)
District Palmoil Potential Yield, 1960	0.005
	(0.005)
District Coffee Potential Yield, 1960	0.002
	(0.005)
District Cassava Potential Yield, 1960	-0.005
	(0.005)
District Mining Employment Share, 1960	-0.002
	(0.004)
District Manufacturing Employment Share, 1960	0.005
	(0.004)
District Trade Employment Share, 1960	0.003
	(0.005)
District Service Employment Share, 1960	0.004
	(0.005)
District Wage Worker Share, 1960	0.006*
	(0.005)
District Share HHs with Piped Water, 1960	0.004
	(0.004)
District Share HHs with Sewage, 1960	0.001
	(0.005)
District Share HHs with Modern Fuel, 1960	0.004
	(0.005)
District Share HHs with Modern Floor, 1960	0.002
	(0.005)
District Share HHs with Modern Roof, 1960	-0.006
	(0.005)
F-Stat of Omnibus Test	2.339
Observations	1219

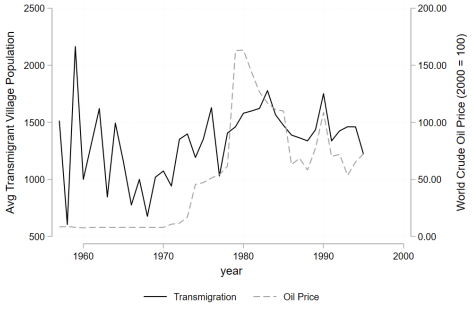
Back

Transmigration and Oil

Figure 13: Transmigration and Oil Prices



(a) Total Transmigration by Year



(b) Average TM Village Size by Year

Back

Migrant Knowledge of Destination

Table 6: Share of Migrants with Knowledge of Village Characteristics

Percent with Knowledge of	All Sending Locations	Excluding Central Java
Village Name	70.98	58.95
Nearest Town	31.61	24.21
Climate	18.10	10.53
Ethnic Group	15.52	7.89
Observations	348	190

Author's calculations, data from Kebschull (1986)

Back

Balance Tests

- For first test for balance across market access, run the following:

$$X_{vdic} = \beta_1 ln(MA)_{vdic} + \beta_2 ln(E(MA))_d + \eta_i + \gamma_c + \epsilon_{vdic} \quad (2)$$

- and record the β_1 for each control variable, includes 1980 district controls too
- standard errors clustered at Level 2 for district controls, Level 3 for village controls, eventually add Conley SE
- Additionally perform omnibus test of difference and record F-stat
- Test of balance across time villages created the same, but now record $\beta_1 Year\_Placed_c$ and drop cohort FE

Back

Wages and Market Access

Table 7: Effects of MA on Wages

	ln(Wage)	ln(Hourly Wage)	ln(Hours Worked)	ln(Wage)	ln(Wage)
(max) ln_MA_full_noself	-0.0027	-0.0027	-0.0003	-0.0022	-0.0038*
	(0.0023)	(0.0027)	(0.0017)	(0.0023)	(0.0023)
Borusyak and Hull	Yes	Yes	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes	Yes	Yes
District Controls	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Sample	Full	Full	Full	Limited	Limited
Sector FE	No	No	No	No	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes
Island FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Unique Villages	502	501	503	501	501
Unique Households	5781	5706	5989	5462	5462

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Negative impact of MA appears in wages too

Back

Agricultural Employment, by Median MA

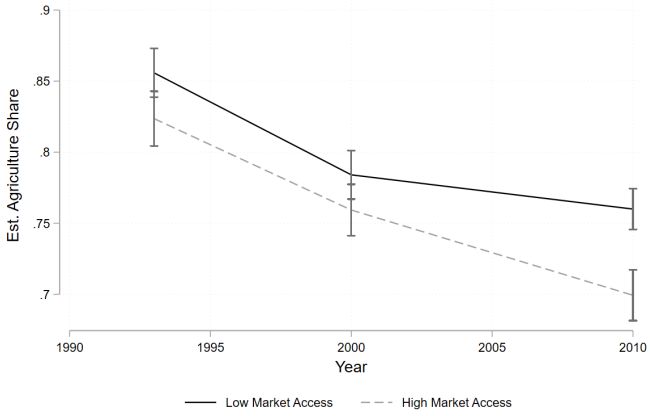
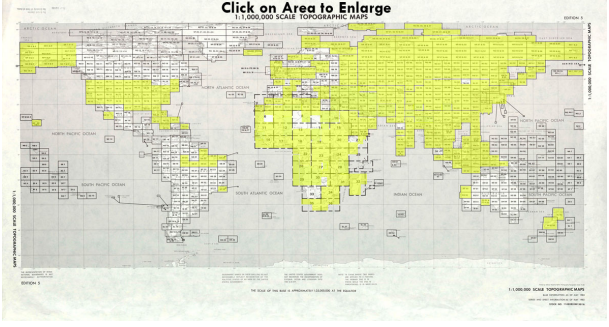
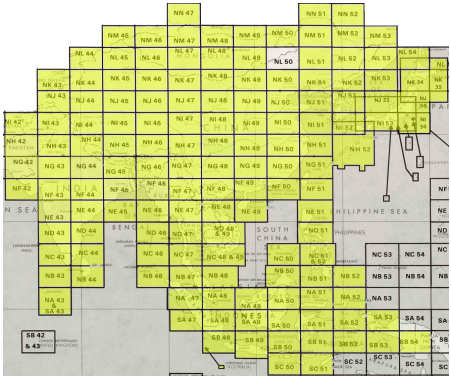


Figure 14: Agricultural Employment Share, by Initial Market Access

PCL Maps



(a) Global Coverage



(b) Asia

PCL Maps

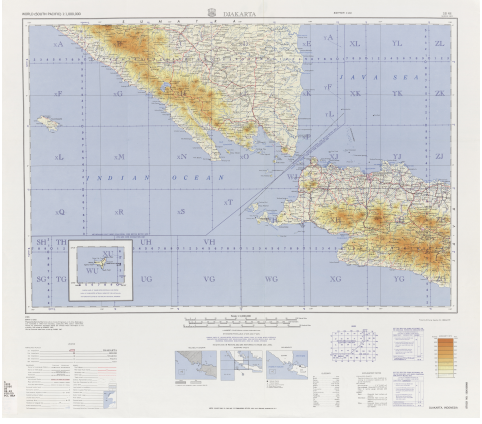


Figure 16: South Sumatra

Back

Identification

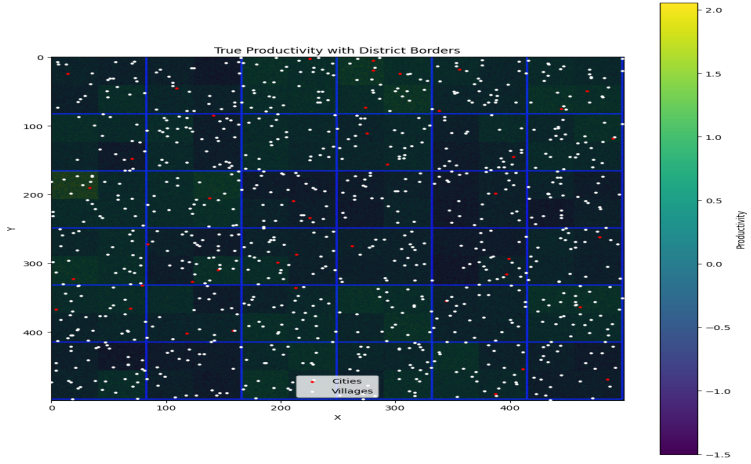
- So far: convincing evidence of no selection on observables, evidence site selection did not seem to consider market access
- Even with *random* placement of new villages, no guarantee market access will be uncorrelated with unobserved drivers of growth
- With positive selection of preexisting settlements, introduces positive correlation between randomly assigned MA and productivity
- Importantly, potential bias is *positive*, and controls designed to address issue

Simulation Exercise

- Space is a grid, equally sized provinces, (sub)districts
- Cells assigned partially unobserved productivity, correlation across spatial units
- 50 initial "cities" placed on the grid, more likely to be placed in high productivity cells
- New villages then randomly placed within provinces, mimics TM program
- Can then construct expected MA and local and district controls from observed component of productivity
- Predict relationship between true productivity and initial market access

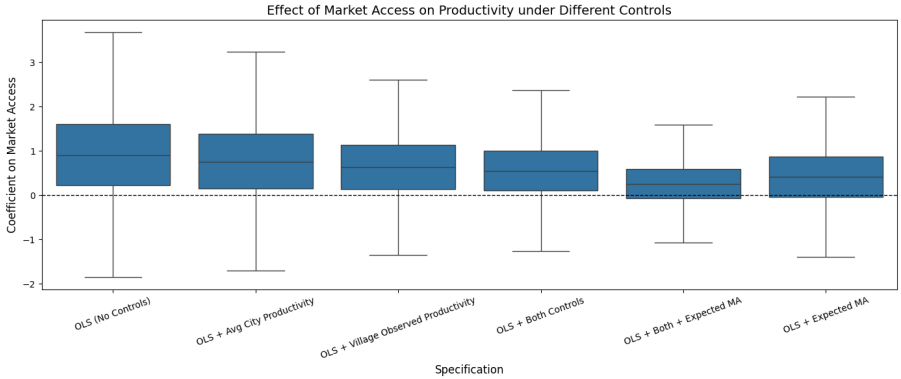
Simulation: Map

Figure 17: Map of Simulated Space



Simulation: Results

Figure 18: Simulation Results



Robustness

- Given potential for OVB, formally assess potential for unobservables to explain estimated effects
- Implement test proposed by Diegert et al. (2025), an extension of Oster (2019)
- Compare breakdown point $\bar{r}_X^{bp}$ to ratio of fitted variation in treatment explained by local controls k relative to all the other observed controls combined: ρ_k
- Robustness summarized as fraction of times that $\bar{r}_X^{bp} > \rho_k$
- $\bar{r}_X^{bp}$ most conservative measure of breakdown point, less conservative assumptions yield more robust results

Back

Agricultural Employment, by Median MA

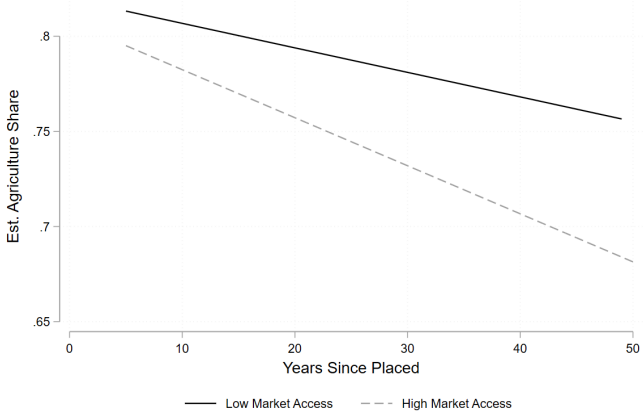


Figure 19: Agricultural Employment Share, by Initial Market Access

Back

Migration and Commuting

Table 8: Effects of MA on Migration and Commuting

	ln(Inner Islanders)	Share of Inner Islanders	Commuting in Village (0/1)
ln(MA)	-0.0113*	-0.0006	0.0006
	(0.0058)	(0.0011)	(0.0013)
Expected MA Control	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes
District Controls	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Island FE	Yes	Yes	Yes
Year FE	Yes	Yes	No
Unique Villages	1136	1147	944

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- No differences in commuting patterns, small, mostly insignificant differences in share from inner islands

Brain Drain

Table 9: Effects of MA on Human Capital

	Years School: Migrants	Years School: Next Gen
ln(MA)	0.0222**	0.0160**
	(0.0103)	(0.0077)
Expected MA Control	Yes	Yes
Local Controls	Yes	Yes
District Controls	Yes	Yes
Cohort FE	Yes	Yes
Island FE	Yes	Yes
Year FE	Yes	Yes
Individuals	742896	1099700
Unique Villages	739	740

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Evidence of brain *gain*, negative GDP/productivity results more surprising

Back

Amenities

	Public Goods Index - Full	Public Goods Index - Excl Roads	Public Goods Index - Roads Only
ln(Market Access)	0.0023	0.0005	0.0106***
	(0.0020)	(0.0022)	(0.0035)
Expected MA Control	Yes	Yes	Yes
Local Controls	Yes	Yes	Yes
District Controls	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes
Island FE	Yes	Yes	Yes
Year FE	No	No	No
Unique Villages	1147	1147	1147

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- No evidence of improved amenities *except* better road quality

Back

Data Used

- Going to see results from two different sources
- SI: Indonesian formal manufacturing firms, long panel, large sample
  - Potential Issue: observe rl, not separately (report value of land)
- Global IFS: Pooled sample of 8 developing/middle income countries from early wave of IFS, see l across wider range of firm types
  - Potential Issues: relatively small sample, not a panel
  - Countries (between 2010-2014): Argentina, DRC, Ghana, Guatemala, Kenya, Myanmar, Peru, Rwanda

Non-homothetic CES Function

- For clarity, going to stick to the "iso-MRS" special case of NH-CES (find strong support for this assumption later)
- From Sato (1977), can write two factor example:

$$log\left(\frac{L}{N}\right) = \alpha - \sigma log\left(\frac{w_L}{w_N}\right) - \sigma\delta log(Y) \tag{3}$$

- Rearranging terms to instead express as expenditure shares:

$$log\left(\frac{w_L L}{w_N N}\right) = \alpha + (1-\sigma) log\left(\frac{w_L}{w_N}\right) - \sigma\delta log(Y) \tag{4}$$

- where δ is the non-homotheticity parameter

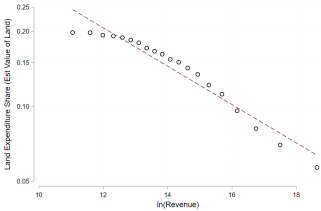
Estimated Value of Property

- Assume that for some subsample of data we see instead the estimated value of the property, q
- Relationship assumed to be $q = \kappa * (w_L L)$
- Plugging in to Eq 5:

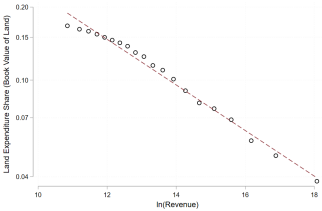
$$log\left(\frac{q}{w_N N}\right) = \alpha + (1-\sigma) log\left(\frac{w_L}{w_N}\right) - \sigma\delta log(Y) + log(\kappa) \quad (5)$$

- If no covariance (after FE) between κ and Y then not an issue
- For limited sample of firms, observe both rent and estimated value, can calculate κ
- Median value of κ is high (30), but not unreasonable (and likely upper bound)

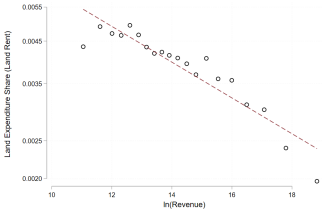
Descriptive Evidence



(a) Land Value Est



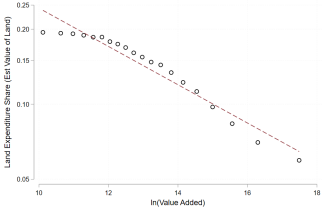
(b) Land Book Value



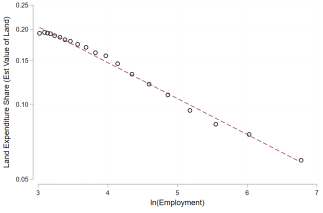
(c) Rent Payments

- Slopes extremely similar, fewer observations for firms that rent
- Denominator includes capital est. value, slopes are similar if excluded

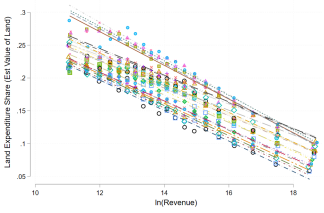
Descriptive Evidence



(a) Value Added



(b) Employment



(c) Land Value Est, by Year

- Slopes extremely similar, holds across years

Regression Specification

- If input prices (and technology) vary by Region X Product X Year, then cleanly estimate non-homotheticity elasticity ϵ_{LN}
- If firms also pay different prices within those cells, can absorb with Firm FE
- Let $E_{LN} \equiv \left(\frac{w_L L}{w_N N}\right)$
- For firm i in region r primarily producing product p in year t:

$$log\left(E_{LN}\right)_{irpt} = \beta_1 ln(Size)_{irpt} + \alpha_i + \gamma_{rpt} + \epsilon_{irpt} \tag{6}$$

- Where size is measured by revenue, value added, or employment count

Results - SI

	ln(E_{LN} (Land Est))	ln(E_{LN} (Rent))	ln(E_{LN} (Land Est))	ln(E_{LN} (Rent))	ln(E_{LN} (Land Est))	ln(E_{LN} (Rent))
ln(Revenue)	-0.2751***	-0.2649***	-0.2321***	-0.1863***	-0.1893***	-0.1846***
	(0.0069)	(0.0290)	(0.0054)	(0.0229)	(0.0051)	(0.0170)
FE Type	Region X Ind X Year + Firm	Region X Ind X Year + Firm	Region X Ind X Year	Region X Ind X Year	Firm	Firm
N	193487	7870	195619	9524	232076	16504
Unique Firms	23602	1826	25541	3301	27131	3614

Standard errors in parentheses

Standard errors clustered at the firm level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Rent seems equivalent to land value estimates, results relatively consistent across FE's
- Can also control (flexibly) for firm age, doesn't change results

Results - SI

	ln(E_{LN} (Land Est))	ln(E_{LN} (Rent))	ln(E_{LN} (Land Est))	ln(E_{LN} (Rent))	ln(E_{LN} (Land Est))	ln(E_{LN} (Rent))
ln(Revenue)	-0.2751***	-0.2649***				
	(0.0069)	(0.0290)				
ln(Value Added)			-0.2806***	-0.2086***		
			(0.0054)	(0.0231)		
ln(Employment)					-0.6731***	-0.5401***
					(0.0135)	(0.0596)
FE Type	Region X Ind X Year + Firm	Region X Ind X Year + Firm	Region X Ind X Year + Firm	Region X Ind X Year + Firm	Region X Ind X Year + Firm	Region X Ind X Year + Firm
N	193487	7870	207090	8652	207117	8654
Unique Firms	23602	1826	24846	2006	24842	2004

Standard errors in parentheses

Standard errors clustered at the firm level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Non-homotheticity very apparent regardless of measure for output

Results - SI

	ln(E_{LN} (Land Est))	ln(E_{LN} (Land Est))
ln(Revenue)	-0.2751***	-0.1037***
	(0.0069)	(0.0057)
ln(Avg Wage)		-0.8704***
		(0.0055)
FE Type	Region X Ind X Year + Firm	Region X Ind X Year + Firm
N	193487	186826
Unique Firms	23602	23395

Standard errors in parentheses

Standard errors clustered at the firm level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Bigger firms pay lower wages, would bias estimate downwards, see evidence of this

Results - IFS

	ln(Land-Labor Input Ratio)	ln(Land-Labor Input Ratio)	ln(Land-Labor Input Ratio)	ln(Land-Labor Input Ratio)	ln(Land-Labor Input Ratio)	ln(Land-Labor Input Ratio)
ln(Revenue)	-0.0514*			-0.1389***		
	(0.0280)			(0.0362)		
ln(Employment)		-0.5164***			-0.6135***	
		(0.0492)			(0.0573)	
ln(Value Added)			-0.0891***			-0.1322***
			(0.0345)			(0.0443)
FE Type	Ctry X Region X Sector	Ctry X Region X Sector	Ctry X Region X Sector	Ctry X Region X Sector	Ctry X Region X Sector	Ctry X Region X Sector
Controls	No	No	No	Yes	Yes	Yes
N	1638	1808	1057	1378	1453	931

Standard errors in parentheses
Robust standard errors implemented
* $p < .1$, ** $p < .05$, *** $p < .01$

- Very similar estimates for non-homotheticity, especially when controlling for wages

Controlling for (average) wage and bounds

- Sticking with simple price-taking example for example, so ignoring $\omega_i(l, n)$

$$\log \frac{m_l}{m_n} = \alpha_i + (1-\sigma)ln(r_i) - (1-\sigma)ln(w_i) + \beta logY \tag{7}$$

- No wage control, $\beta = -0.2$, including wage control causes $\beta = -0.1$
- Cov(Y, w) $> 0 \rightarrow \sigma < 1$
- When $\sigma < 1$, would likely expect that including rent (if observed) would push coefficient to be more negative

$$\text{Bias} = (1-\sigma)\frac{Cov(log(Y), log(r))}{Var(log(Y))} \tag{8}$$

- If land behaves in a similar direction as wages, then bias is positive

Back

Wages and Market Access

Table 10: Effects of MA on Wages

	ln(GDP per Capita)
ln(MA)	-0.0049***
	(0.0017)
Rice Potential Yield, std	0.0321
	(0.0384)
ln(MA) × Rice Potential Yield, std	0.0037***
	(0.0011)
Expected MA Control	Yes
Local Controls	Yes
District Controls	Yes
Cohort FE	Yes
Island FE	Yes
Year FE	Yes
Unique Villages	1152

Standard errors in parentheses

Standard errors clustered at the village level

* $p < .1$, ** $p < .05$, *** $p < .01$

- Most productive locations see wage *increase* in response to increased MA

Back