Wealth Heterogeneity, Income Shocks, and International Migration: Theory and Evidence from Indonesia

Samuel Bazzi

UC San Diego

October 6, 2012

Motivation: Migration and Development

- $>$1 billion aspire to migrate for temporary work
  (Gallup, 2012)
- Global welfare gains to greater international labor mobility
  (Benhabib & Jovanovic, 2012; Klein & Ventura, 2009)
- Substantial income gains to migrants from poor countries
  (see Clemens, 2011)
- Countervailing effects of rising home income on migration

Motivation: Migration and Development (II)

- "Emigration life cycle" debate using historical cross-country data
  (see Hatton and Williamson, 2011)

- Standard models presume ready financing of migration costs
  (e.g., Borjas, 1987)

- Relaxing this assumption $\Longrightarrow$ new implications for self-selection
  (McKenzie & Rapoport, 2007; Orrenius & Zavodny, 2005)

Motivation: Migration and Development (II)

- "Emigration life cycle" debate using historical cross-country data
  (see Hatton and Williamson, 2011)

- Standard models presume ready financing of migration costs
  (e.g., Borjas, 1987)

- Relaxing this assumption $\Longrightarrow$ new implications for self-selection
  (McKenzie & Rapoport, 2007; Orrenius & Zavodny, 2005)

This Paper

To what extent do financial constraints limit international migration flows from low-income settings?

This Paper

- *Theory*: *microfounded model of migration flows*
    1. fixed migration costs and imperfect credit markets
    2. transitory and permanent income shocks
    3. idiosyncratic wealth heterogeneity

This Paper

- *Theory*: *microfounded model of migration flows*
  1. fixed migration costs and imperfect credit markets
  2. transitory and permanent income shocks
  3. idiosyncratic wealth heterogeneity

- *Testing*: *Indonesian villages*
  1. new administrative *panel* data on international labor migration from $> 65,000$ villages in 2005/8
  2. rainfall and rice price shocks
  3. land-holdings data from universal Agricultural Census 2003

This Paper

- *Theory*: *microfounded model of migration flows*
  1. fixed migration costs and imperfect credit markets
  2. transitory and permanent income shocks
  3. idiosyncratic wealth heterogeneity

- *Testing*: *Indonesian villages*
  1. new administrative *panel* data on international labor migration from $> 65,000$ villages in 2005/8
  2. rainfall and rice price shocks
  3. land-holdings data from universal Agricultural Census 2003

- *Quantitative*
  1. mapping from household to village migration elasticities
  2. structural estimates of village-specific migration costs

Preview of Results

- Strong evidence of liquidity constraints consistent with theory
  1. positive rainfall and price shocks $\Longrightarrow$ $\Uparrow$ flow migration rates
  2. larger increases in villages with less dispersion in land-holdings

Preview of Results

- Strong evidence of liquidity constraints consistent with theory
  1. positive rainfall and price shocks $\Longrightarrow$ $\Uparrow$ flow migration rates
  2. larger increases in villages with less dispersion in land-holdings

- Other results support financial constraints interpretation
  1. rainfall shocks have larger effects in villages with no banks, lower mean household expenditures, lower agg. productivity of land
  2. auxiliary micro data: consistent with aggregate village elasticities

Preview of Results

- Strong evidence of liquidity constraints consistent with theory
  1. positive rainfall and price shocks $\Longrightarrow$ $\Uparrow$ flow migration rates
  2. larger increases in villages with less dispersion in land-holdings

- Other results support financial constraints interpretation
  1. rainfall shocks have larger effects in villages with no banks, lower mean household expenditures, lower agg. productivity of land
  2. auxiliary micro data: consistent with aggregate village elasticities

- Large estimated migration costs, implied net income differentials

Contributions

1. Tractable way to disentangle offsetting effects of income shocks
   - extends insights from RCTs showing liquidity constraints bind for poorest potential migrants (Angelucci, 2012; Bryan et al, 2012)
2. Zero migrant stocks are a possible equilibrium outcome
   - important given network externalities in costs (Carrington et al, 1996)
3. Destination policy, informational barriers relatively constant
   - new insights on migration using admin data (e.g., McKenzie et al, 2012)
4. Novel application of aggregation methods from trade (Melitz, 2003)
   - relevant to broader literature on heterogeneity and aggregation (see Blundell and Stoker, 2005/7)

Outline of Today's Talk

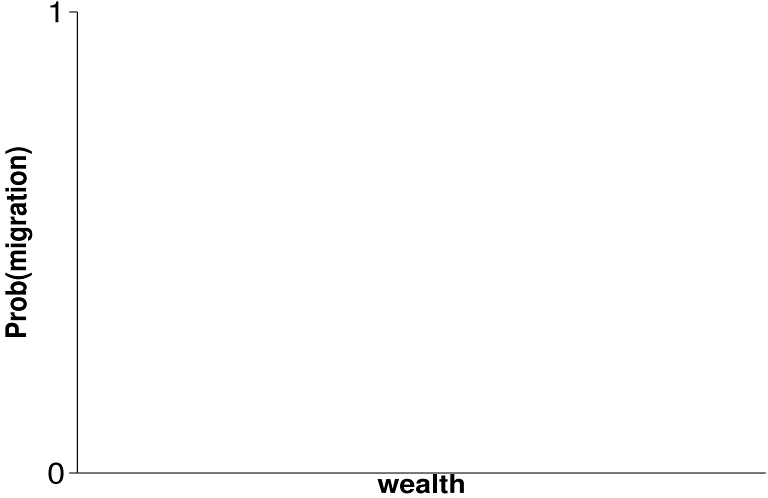
Theory

Data and Context

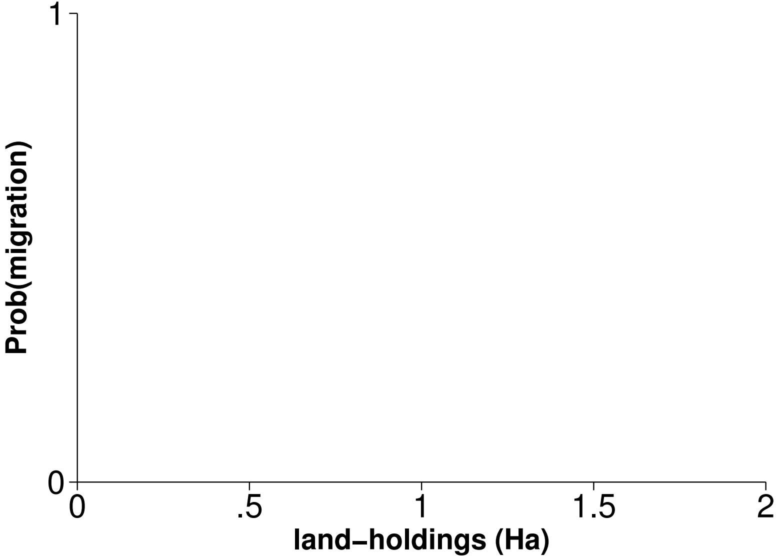
Empirical Results

Conclusion

Migration Choice and Wealth



Migration Choice and Wealth



Model Background

- Agricultural income at home:

$$\Pi_{ivt} = p_{vt}\sigma_{vt}K_v^{\theta}R_{iv}^{\beta}$$

  - land-holding R, public capital K, farmgate price p, rainfall σ

- Net income abroad (j):

$$W_{ivj,t+1} - C_{vj,t+1}$$

- Fraction τ_{vj} of fixed migration costs C must be paid upfront

Model Background

- Agricultural income at home:

$$\Pi_{ivt} = p_{vt}\sigma_{vt}K_v^{\theta}R_{iv}^{\beta}$$

  - ▷ land-holding R, public capital K, farmgate price p, rainfall σ

- Net income abroad (j):

$$W_{ivj,t+1} - C_{vj,t+1}$$

- Fraction τ_{vj} of fixed migration costs C must be paid upfront

$\Longrightarrow$ **Migrate if:**

$$\begin{aligned} W_{ivj,t+1} - C_{vj,t+1} &\geq \mathbb{E}_t[\Pi_{iv,t+1}] \\ \tau_{vj}C_{vj,t+1} &\leq \Pi_{ivt} \end{aligned}$$

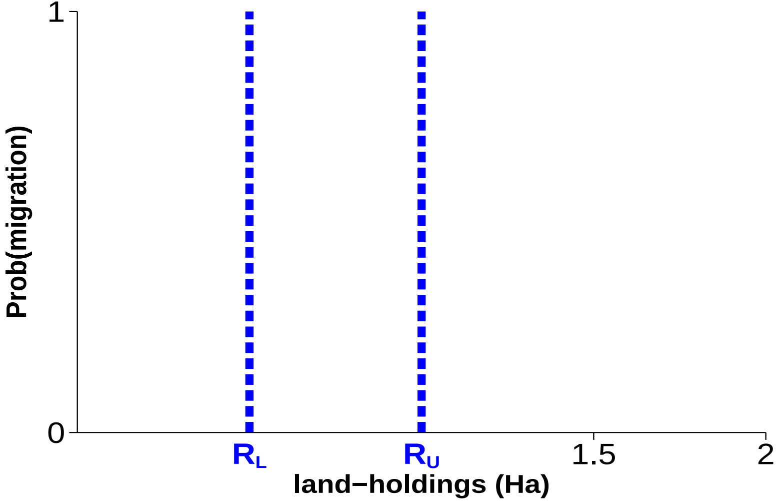
Land-holdings Thresholds in Migration Choice

Individuals abroad in $t+1$ characterized by

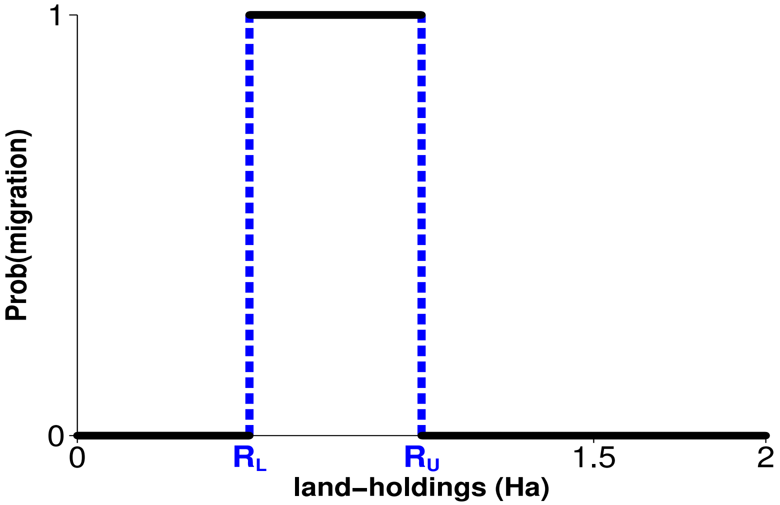
$$\underbrace{\left(\frac{\tau_{vj} C_{vj,t+1}}{p_{vt}\sigma_{vt} K_v^\theta}\right)^{\frac{1}{\beta}}}_{R_L} \leq R_{iv} \leq \underbrace{\left(\frac{W_{ivj,t+1} - C_{vj,t+1}}{\mathbb{E}_t[p_{v,t+1}\sigma_{v,t+1}] K_v^\theta}\right)^{\frac{1}{\beta}}}_{R_U}$$

- R_L is the **liquidity threshold**
- R_U is the **profitability threshold**

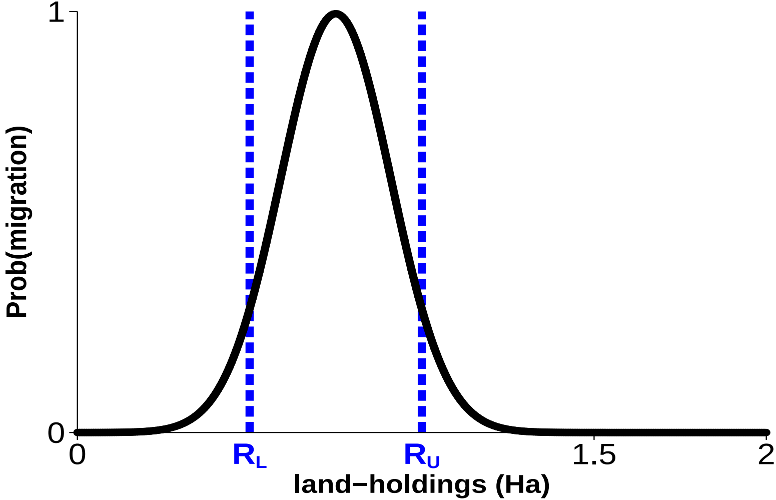
Land-holdings Thresholds in Migration Choice



Land-holdings Thresholds in Migration Choice



Land-holdings Thresholds in Migration Choice



Assumptions

- **Land-holdings**: distributed Pareto ▸ pdf

$$g_v(R_{iv}) = \lambda_v \underline{R}^{\lambda_v} R_{iv}^{-\lambda_v - 1}$$

  - *Foundations*: random population growth on fixed land mass (à la Gabaix)
    - inheritance pressures counteracted by urbanization, demographic transition, and structural transformation $\Longrightarrow$ *Pareto steady-state*
  - *Empirical relevance*: good fit for $R_{iv} > \underline{R} = 0.1$ Ha ▸ ccdf plot
  - *Tractability*: (i) closed form CDF; (ii) shape invariant to truncation

Assumptions

- **Land-holdings**: distributed Pareto ▸ pdf

$$g_v(R_{iv}) = \lambda_v \underline{R}^{\lambda_v} R_{iv}^{-\lambda_v - 1}$$

  - *Foundations*: random population growth on fixed land mass (à la Gabaix)
    - inheritance pressures counteracted by urbanization, demographic transition, and structural transformation $\Longrightarrow$ *Pareto steady-state*
  - *Empirical relevance*: good fit for $R_{iv} > \underline{R} = 0.1$ Ha ▸ ccdf plot
  - *Tractability*: (i) closed form CDF; (ii) shape invariant to truncation

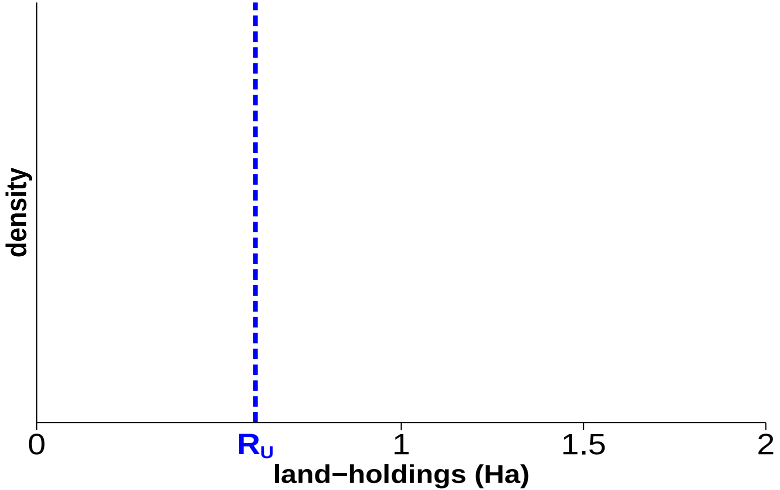
- **Farmgate price**: ARMA(1,Q) with possible unit root,

$$p_{vt} = \alpha_v p_{v,t-1} + \sum_{q=0}^{Q} \theta_q e_{v,t-q}$$

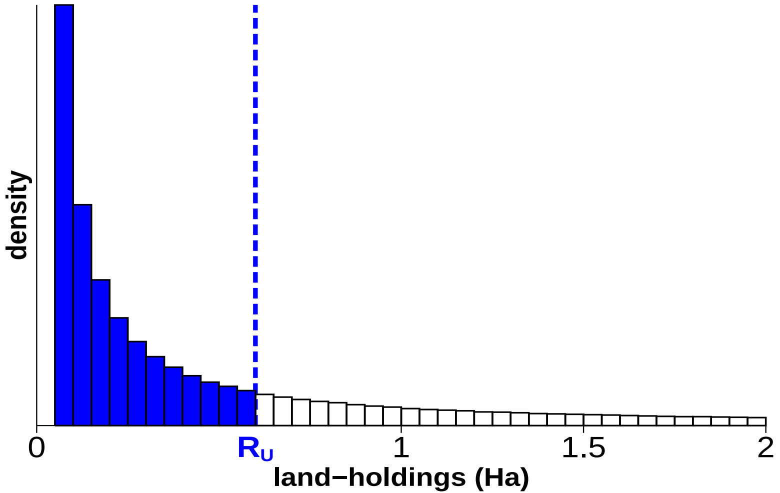
- **Rainfall**: mean-reverting with serially uncorrelated shocks a_{vt},

$$\sigma_{vt} = \overline{\sigma}_v + a_{vt}$$

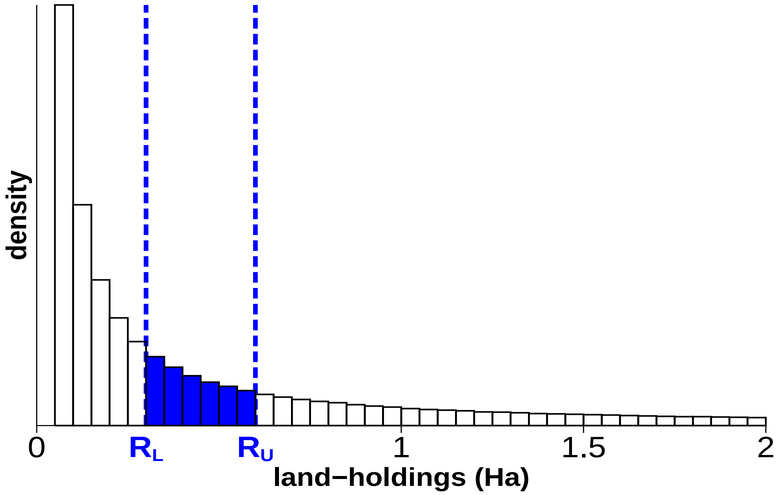
Unconstrained *Stock* Migration Rates



Unconstrained *Stock* Migration Rates



Constrained *Stock* Migration Rates



Closed-Form Flow Migration Rates

If liquidity constraints are not binding ($\tau_{vj} = 0$ or $R_L < \underline{R}$):

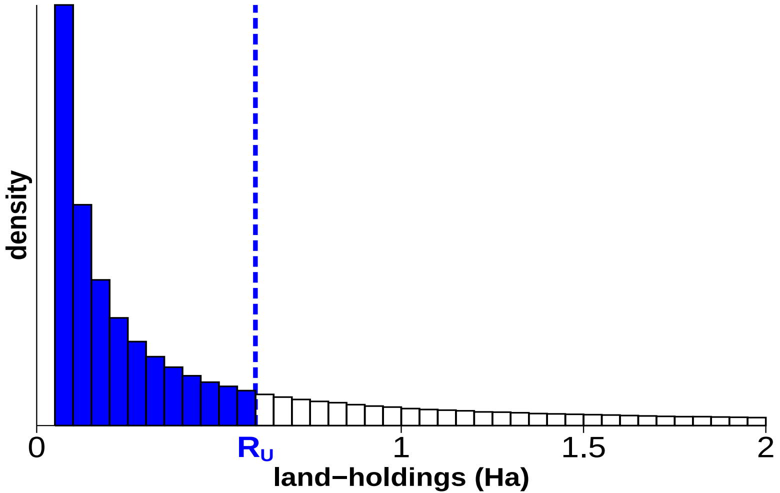
$$\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) = \ln \left[1 - \left(\frac{\alpha_v p_{vt} \overline{\sigma}_v K_v^\theta}{\omega_{vj,t+1} - C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} \right] - \ln \left[1 - \left(\frac{\alpha_v p_{v,t-1} \overline{\sigma}_v K_v^\theta}{\omega_{vjt} - C_{vjt}} \right)^{\frac{\lambda_v}{\beta}} \right]$$

If liquidity constraints are binding ($\tau_{vj} > 0$ and $R_L \geq \underline{R}$):

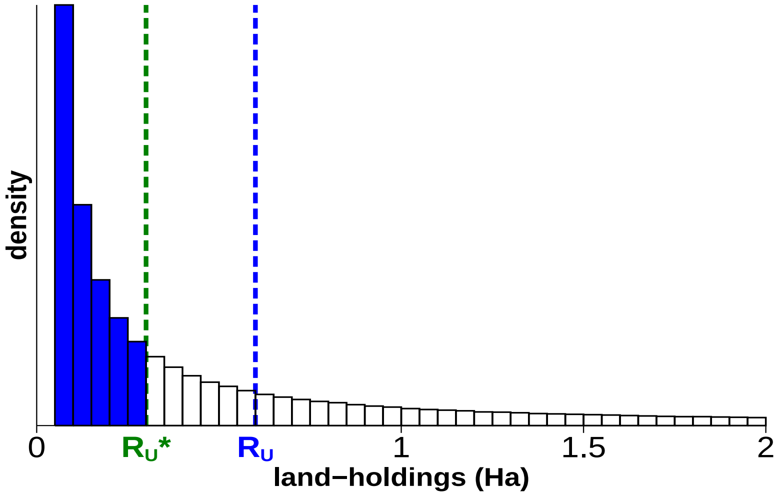
$$\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) = \frac{\lambda_v}{\beta} \Delta \ln p_{vt} + \Delta \ln \left[\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} - \left(\frac{\overline{\sigma}_v \alpha_v}{\omega_{vj,t+1} - C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} \right]$$

- M_{vs}: number of village residents working abroad in s
- N_{vs}: village population in period s (including emigrants)

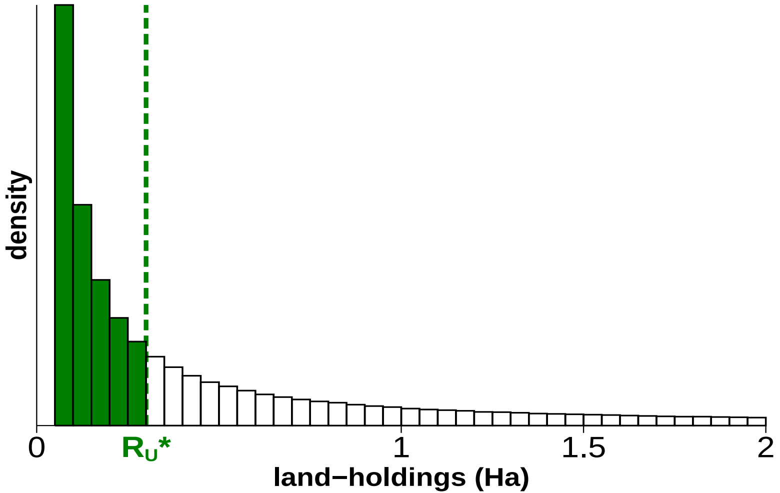
Unconstrained *Flow* Migration Rates and Income Shocks



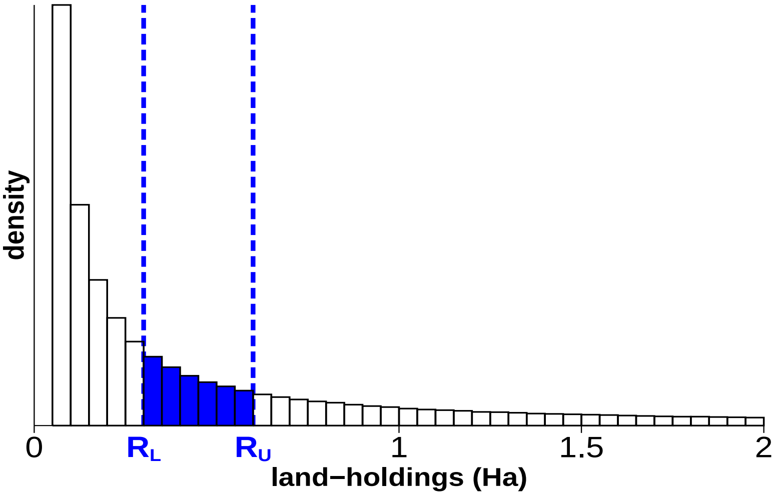
Unconstrained *Flow* Migration Rates and Income Shocks



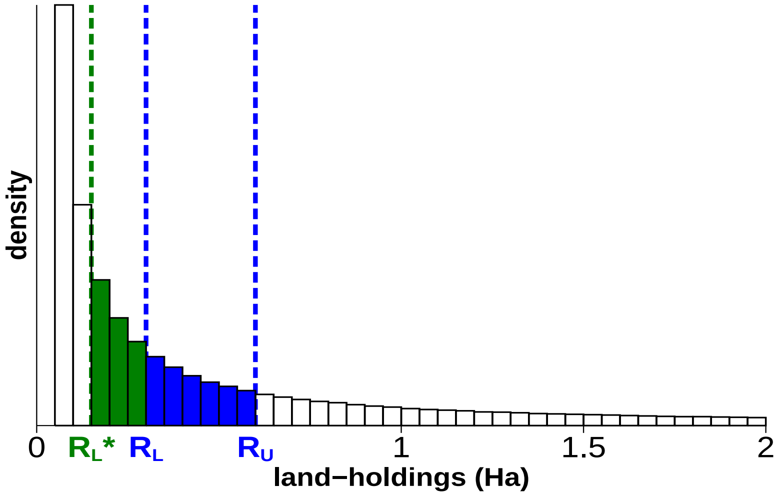
Unconstrained *Flow* Migration Rates and Income Shocks



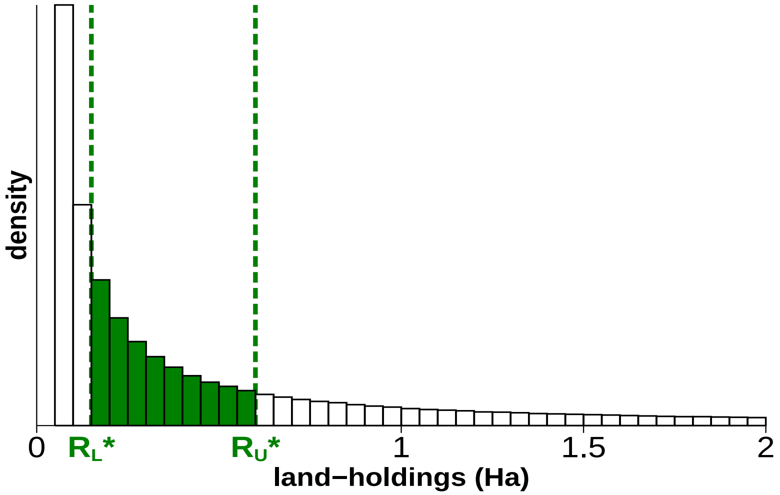
Constrained *Flow* Migration Rates and Income Shocks



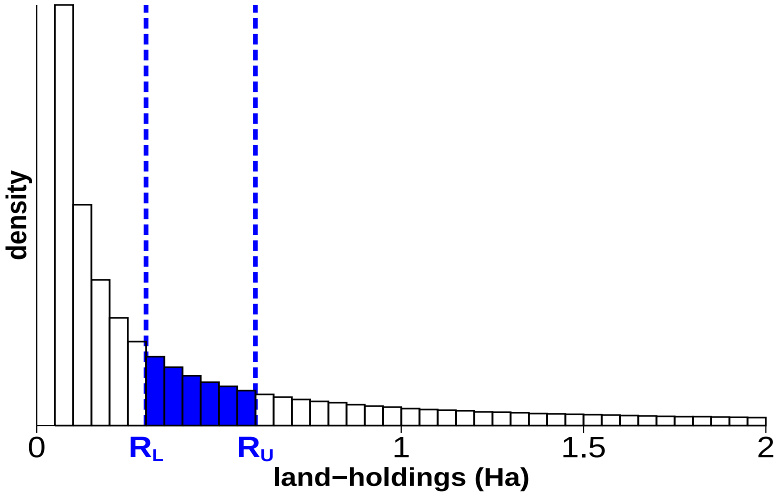
Constrained *Flow* Migration Rates and Income Shocks



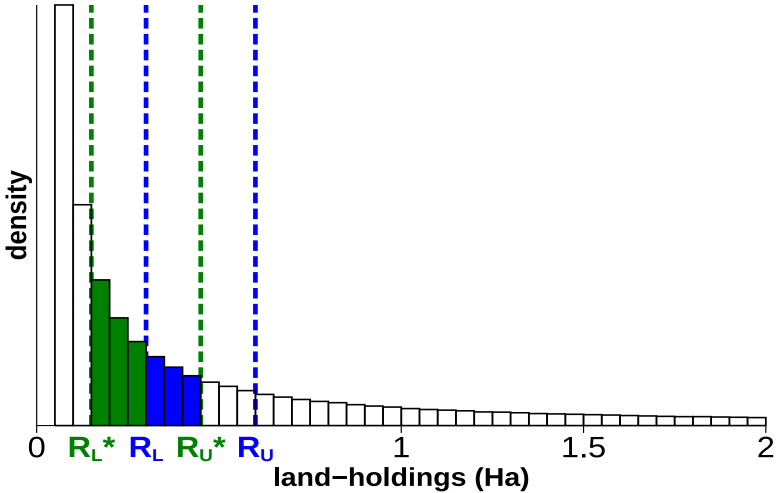
Constrained *Flow* Migration Rates and Income Shocks



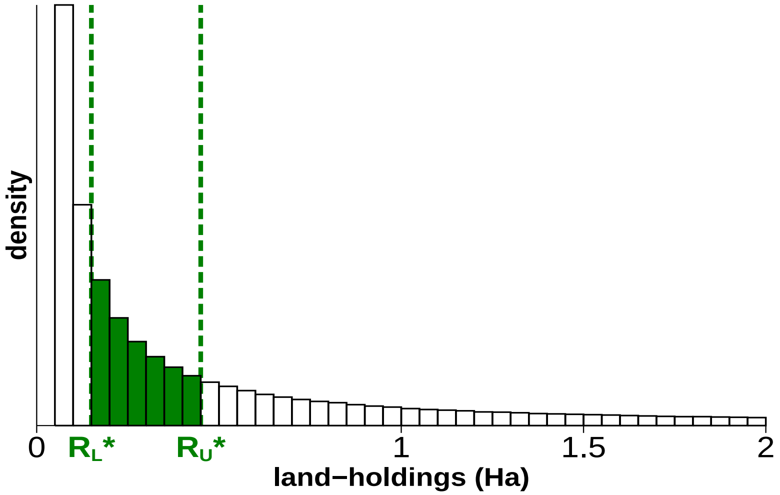
Constrained *Flow* Migration Rates and Income Shocks



Constrained *Flow* Migration Rates and Income Shocks



Constrained *Flow* Migration Rates and Income Shocks



Key Implications for the Intensive Margin

Propositions 1 and 2

If liquidity constraints are not binding, then flow migration rate is

1. *uncorrelated* with rainfall shocks
2. *decreasing* in price shocks

$\implies$ larger declines in villages with greater dispersion in land-holdings (lower λ_v)

Key Implications for the Intensive Margin

Propositions 1 and 2

If liquidity constraints are not binding, then flow migration rate is

1. *uncorrelated* with rainfall shocks
2. *decreasing* in price shocks

$\Longrightarrow$ larger declines in villages with greater dispersion in land-holdings (lower λ_v)

If liquidity constraints are binding, then flow migration rate is

1. *increasing* in rainfall shocks
2. *increasing* in price shocks

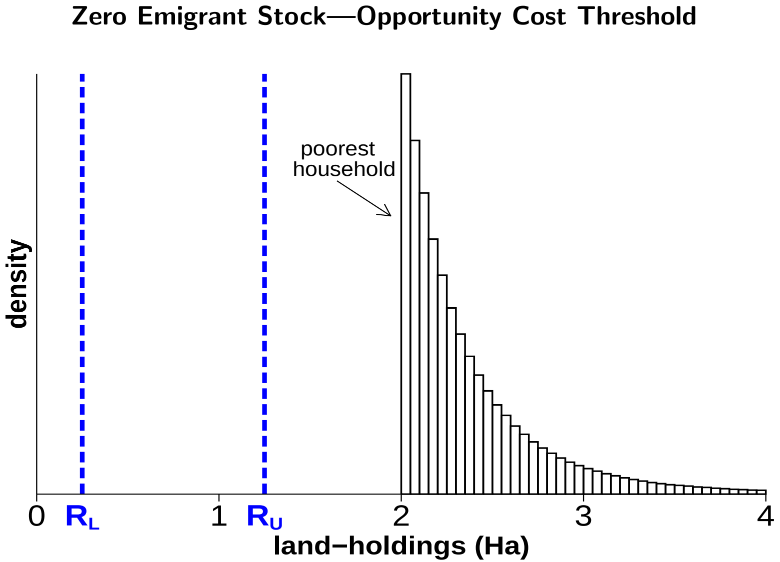
$\Longrightarrow$ larger increases in villages with less dispersion in land-holdings (higher λ_v)

Intuition

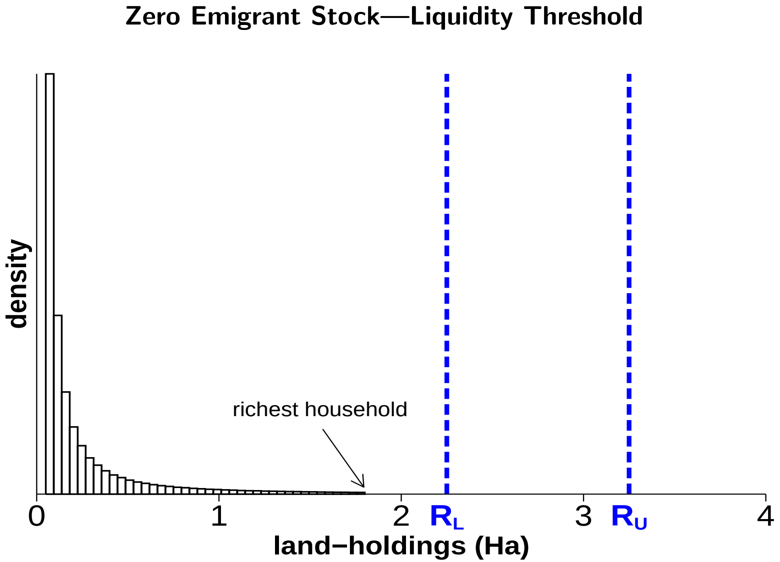
1. Transitory shocks only matter if liquidity constraints are binding
2. Limited income smoothing under serial correlation (Deaton, 1991)
3. In villages with less dispersion in land-holdings (higher λ_v)
   - mean household savings/liquidity lower
   - low inequality $\Longrightarrow$ thin (potential) informal credit market

Implications for the Extensive Margin

Implications for the Extensive Margin



Implications for the Extensive Margin



Implications for the Extensive Margin

- Identically zero emigration from village v if
  - poorest household $\underset{\sim}{R}_v \equiv \min_\ell R_{\ell v}$ deems migration unprofitable

    or
  - wealthiest household $\widetilde{R}_v \equiv \max_k R_{kv}$ cannot afford to migrate

- Probability of *any* emigrants from v is
  - increasing in population size N_v
  - increasing (decreasing) in λ_v if liquidity constraints (do not) bind

Outline of Today's Talk

Theory

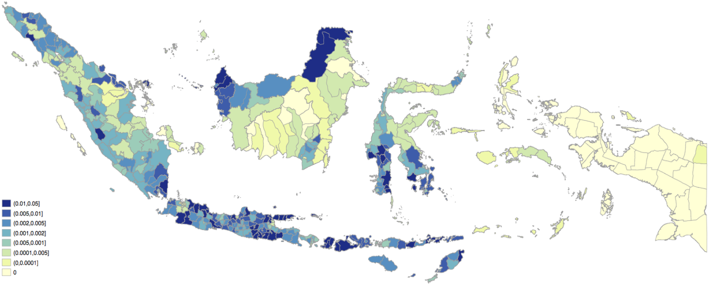
Data and Context

Empirical Results

Conclusion

International Migration from Indonesia

Stock Migration Rates, Districts of Indonesia



- $\sim$ 700,000 annual legal departures
- $>$ 85% of emigrants are from rural areas
- *duration*: majority for 2-3 years
- *destinations*: MYS, SGP, Gulf, HKG, TWN, KOR, JPN
- *occupations*: agriculture, construction, housemaids

New Village-Level Panel Data on International Migration

Village Potential (*Podes*) administrative census, April 2005 & 2008

International (Labor) Migrants:

"those working abroad for a fixed wage and time period"

New Village-Level Panel Data on International Migration

Village Potential (*Podes*) administrative census, April 2005 & 2008

International (Labor) Migrants:

"those working abroad for a fixed wage and time period"

	Summary Stats, 2008			
	mean	*median*	*std. dev*	*max*
population	3,377	2,187	4,330	82,215
1(migrants>0)	0.59	—	—	—
migrants/population $\mid > 0$	0.012	0.004	0.026	0.759
Δ log (migrants/population)	0.106	0.062	1.012	5.669

Notes: Includes all urban and rural villages. "$\mid > 0$" denotes the sample of villages with at least one migrant. The full, balanced village-level panel comprises 65,966 villages excluding the provinces of Papua. A number of villages located in urban areas are dropped from the main analysis. In total, there are around 1 million migrants in the roughly 54,000 (semi-)rural villages. A probabilistic balls-and-bins test suggests that the incidence of zeros not a statistical artifact. ▸ Distribution ▸ Balls & Bins

New Village-Level Panel Data on International Migration

Village Potential (*Podes*) administrative census, April 2005 & 2008

International (Labor) Migrants:

"those working abroad for a fixed wage and time period"

	Summary Stats, 2008			
	mean	*median*	*std. dev*	*max*
population	3,377	2,187	4,330	82,215
1(migrants>0)	0.59	—	—	—
migrants/population $\mid > 0$	0.012	0.004	0.026	0.759
Δ log (migrants/population)	0.106	0.062	1.012	5.669

Notes: Includes all urban and rural villages. "$\mid > 0$" denotes the sample of villages with at least one migrant. The full, balanced village-level panel comprises 65,966 villages excluding the provinces of Papua. A number of villages located in urban areas are dropped from the main analysis. In total, there are around 1 million migrants in the roughly 54,000 (semi-)rural villages. A probabilistic balls-and-bins test suggests that the incidence of zeros not a statistical artifact. ▸ Distribution ▸ Balls & Bins

New Village-Level Panel Data on International Migration

Village Potential (*Podes*) administrative census, April 2005 & 2008

International (Labor) Migrants:

"those working abroad for a fixed wage and time period"

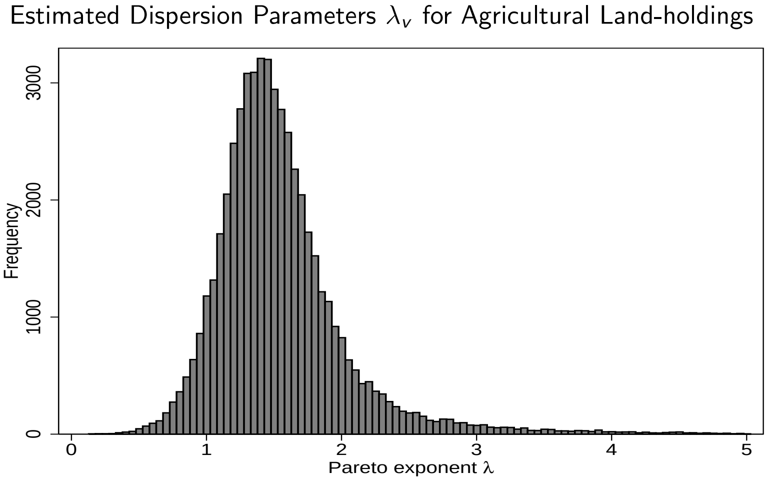
	Summary Stats, 2008			
	mean	*median*	*std. dev*	*max*
population	3,377	2,187	4,330	82,215
1(migrants>0)	0.59	—	—	—
migrants/population $\mid > 0$	0.012	0.004	0.026	0.759
Δ log (migrants/population)	0.106	0.062	1.012	5.669

Notes: Includes all urban and rural villages. "$\mid > 0$" denotes the sample of villages with at least one migrant. The full, balanced village-level panel comprises 65,966 villages excluding the provinces of Papua. A number of villages located in urban areas are dropped from the main analysis. In total, there are around 1 million migrants in the roughly 54,000 (semi-)rural villages. A probabilistic balls-and-bins test suggests that the incidence of zeros not a statistical artifact. ▸ Distribution ▸ Balls & Bins

Data and Measurement at the Village Level

- *Migration costs*
  - *distance*: to nearest emigration center, (sub)district capitals
  - *information*: population share of Chinese, Arabs, and Muslims
- *Destination shocks*
  - fixed effects for plurality destination by village, 2005
- *Land-holdings distribution*
  - estimates of λ_v for total agri. land, wetland, paddy area planted

Data and Measurement at the Village Level

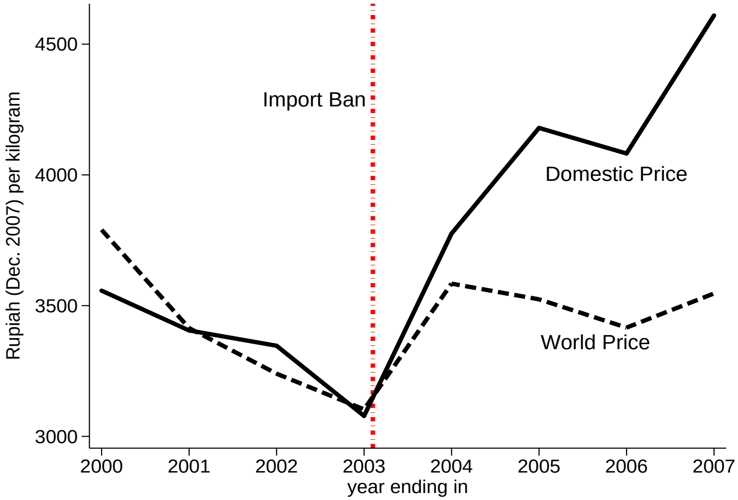


Notes: The Pareto distribution is given by $\lambda_v \underline{R}^{\lambda_v} R_{iv}^{-\lambda_v - 1}$. The figure shows the kernel density of Gabaix & Ibragimov (2011) log rank(-1/2) - log size OLS estimates of λ_v using the average log rank for a given log total agricultural land-holding size and imposing $\underline{R} = 0.1$ hectares. The estimates were calculated independently across all villages with at least 3 distinct land-holding sizes of the given type recorded in the Agricultural Census 2003. In the figure, the top 2 % of estimates are trimmed.

Data and Measurement at the Village Level

- *Migration costs*
  - *distance*: to nearest emigration center, (sub)district capitals
  - *information*: population share of Chinese, Arabs, and Muslims
- *Destination shocks*
  - fixed effects for plurality destination by village, 2005
- *Land-holdings distribution*
  - estimates of λ_v for total agri. land, wetland, paddy area planted
- *Rainfall shocks*
  - log deviations from long-run district-level mean Distribution

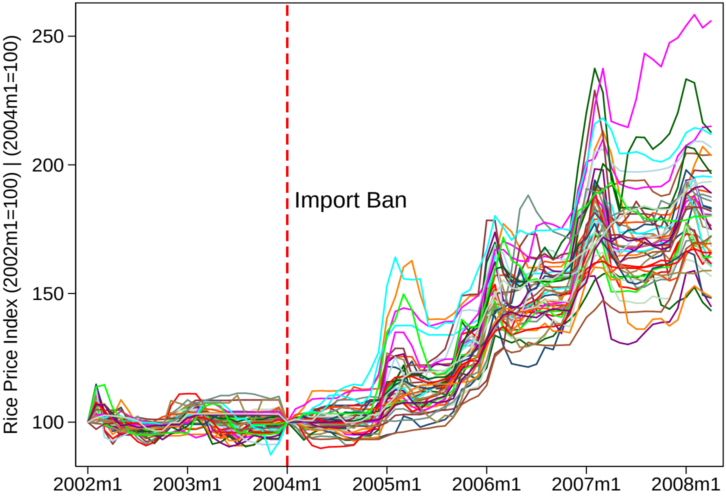
Rice Price Shock



Notes: Year-end average farmgate/producer prices from 2000 to 2007 across Indonesia reported by the Food and Agriculture Organization (FAO). Nominal prices are deflated by the national CPI reported by Bank of Indonesia. Exchange rate and world price data are obtained from the IMF. Further adjustments are made as suggested in Dawe (2008): Thai 100B f.o.b. adjusted to retail level by USD 20 per ton and 10% markup from wholesale to retail, adjusted downward for quality by 20% from 1991-2000 and by 10% from 2001-2007 based on trends in quality preferences in the world market.

Spatial Heterogeneity in Price Shocks

Evolution of Rice Prices across Indonesian Cities: 2002-2008



Notes: Each line represents average of prices in major markets located in cities throughout Indonesia. The index is normalized to 100 in January 2002, and for comparison purposes, I re-initialize and renormalize the index to 100 at the time of the import ban in January 2004.

Outline of Today's Talk

Theory

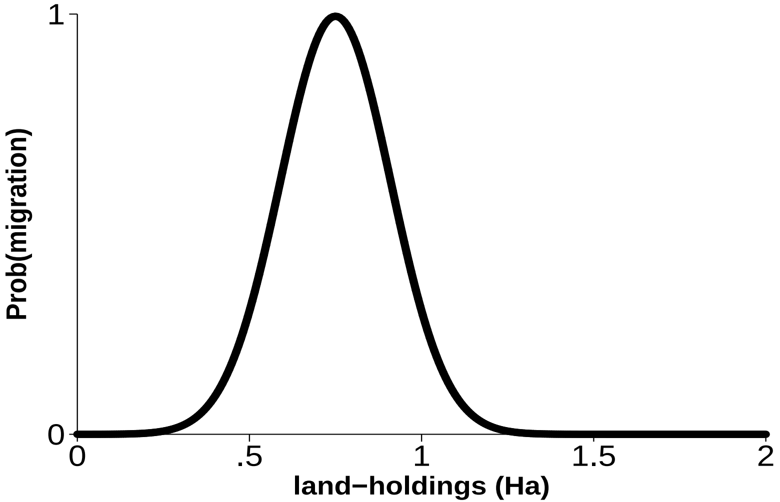
Data and Context

Empirical Results

Conclusion

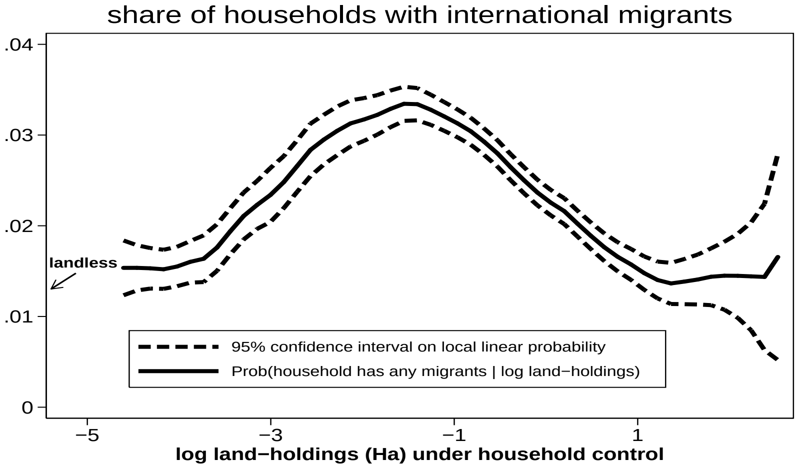
Inverted U in Land-holdings and Migration Choice

What does the theory say?



Inverted U in Land-holdings and Migration Choice

What do the micro data say?



Notes: Calculations based on nationally representative household survey (*Susenas*) data collected in July 2005. The nonparametric regression curve and analytic confidence band is based on a local linear probability regression of an indicator for whether a household member worked abroad from 2002-2005 on log land-holdings under household control. There are a total of 257,906 households in the data and 124,472 report controlling any land-holdings at the time of enumeration. All estimates employ sampling weights. Top percentile of land-holdings are trimmed.

Village-Level Regressions: Reduced Form

$$M_{v,t+1}/N_{v,t+1} = \theta_{\mathbf{a}} \textit{rainfall shock}_{vt} + \theta_{\mathbf{p}} \textit{price shock}_{vt} + \xi_t + \xi_v + \varepsilon_{v,t+1}$$

- *rainfall shock*$_{vt(t+1)}$:= cumulative seasonal shocks 2003-5 (2006-8)
- *price shock*$_{vt(t+1)}$:= annualized log difference 2002m1-5m3 (2005m4-8m3)

Village-Level Regressions: Reduced Form

	FE-OLS		Semipar. FE-Tobit	
	(1)	(2)	(3)	(4)
rainfall shock	**0.0011**		**0.0083**	
	[0.0002]***		[0.0011]***	
	(0.0011)		(0.0047)*	
rice price shock	**0.0086**		**0.0254**	
	[0.0014]***		[0.0046]***	
	(0.0058)		(0.0172)	
Village Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	103,196	103,196	103,196	103,196

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% . Standard errors are clustered by village in brackets and district in parentheses. Semiparametric Tobit is the trimmed LAD estimator of Honore (1992). *rainfall shock* is the cumulative log deviation from long-run mean rainfall in the growing seasons ending in 2006-2008 or 2002-2005. *rice price shock* is the annualized log growth rate in the nearest rice price index between 2005m4-2008m3 or 2002m1-2005m3. The estimated Pareto exponent $\widehat{\lambda}_V$ is for total agricultural land-holdings. Results are qualitatively similar using a conditional fixed effects Poisson specification.

▸ Table

Village-Level Regressions: Reduced Form

	FE-OLS		Semipar.	FE-Tobit
	(1)	(2)	(3)	(4)
rainfall shock	**0.0011**	**0.0014**	**0.0083**	**0.0033**
	[0.0002]***	[0.0006]**	[0.0011]***	[0.0047]
	(0.0011)	(0.0012)	(0.0047)*	(0.0051)
rice price shock	**0.0086**	**0.0023**	**0.0254**	**-0.0185**
	[0.0014]***	[0.0023]	[0.0046]***	[0.0135]
	(0.0058)	(0.0064)	(0.0172)	(0.0176)
$\widehat{\lambda}_v$ × **rainfall shock**		**-0.0002**		**0.0032**
		[0.0004]		[0.0032]
		(0.0006)		(0.0033)
$\widehat{\lambda}_v$ × **price shock**		**0.0040**		**0.0293**
		[0.0013]***		[0.0093]***
		(0.0026)		(0.0125)**
Village Fixed Effects	Yes	Yes	Yes	Yes
Number of Observations	103,196	103,196	103,196	103,196

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% . Standard errors are clustered by village in brackets and district in parentheses. Semiparametric Tobit is the trimmed LAD estimator of Honore (1992). *rainfall shock* is the cumulative log deviation from long-run mean rainfall in the growing seasons ending in 2006-2008 or 2002-2005. *rice price shock* is the annualized log growth rate in the nearest rice price index between 2005m4-2008m3 or 2002m1-2005m3. The estimated Pareto exponent $\widehat{\lambda}_v$ is for total agricultural land-holdings. Results are qualitatively similar using a conditional fixed effects Poisson specification.

▸ Table

Two-Step Model

- Theory suggests two-period latent variable framework

$$\begin{aligned} m^*_{vt} &= \eta_t' \mathbf{Z}_{v,t-1} + u_{vt} \\ m^*_{v,t+1} &= \eta_{t+1}' \mathbf{Z}_{vt} + u_{v,t+1} \\ \Delta \ln(M_{v,t+1}/N_{v,t+1}) &= \Theta' \Delta \mathbf{X}_{vt} + \Delta \varepsilon_{v,t+1} \text{ iff } \quad m^*_{v,t+1} > 0, \; m^*_{vt} > 0 \end{aligned}$$

  - *parametric* correction (Poirier, 1980)
  - *semiparametric* correction (Das et al, 2003)

Two-Step Model

- Theory suggests two-period latent variable framework

$$\begin{aligned} m^*_{vt} &= \eta'_t \mathbf{Z}_{v,t-1} + u_{vt} \\ m^*_{v,t+1} &= \eta'_{t+1} \mathbf{Z}_{vt} + u_{v,t+1} \\ \Delta \ln(M_{v,t+1}/N_{v,t+1}) &= \Theta' \Delta \mathbf{X}_{vt} + \Delta \varepsilon_{v,t+1} \;\; \text{iff} \quad m^*_{v,t+1} > 0, \;\; m^*_{vt} > 0 \end{aligned}$$

  – *parametric* correction (Poirier, 1980)
  – *semiparametric* correction (Das et al, 2003)

- Candidate exclusion restrictions

  1. actual max- $\widetilde{R}_v$ and min- $\underset{\sim}{R}_v$ land-holdings $\perp$ intensive margin
     - in theory, location of $(\underset{\sim}{R}_v, \widetilde{R}_v)$ determined by N_v and λ_v
  2. appeal of neighboring villages to recruiters ("traveling salesman")
     - # villages in district, inter-village travel distance, district pop.

Two-Step Estimates: Extensive Margin

$$
\begin{aligned}
\mathbb{P}(M_{vt} > 0) &= \eta_t' \mathbf{Z}_{v,t-1} + u_{vt} \\
\mathbb{P}(M_{v,t+1} > 0) &= \eta_{t+1}' \mathbf{Z}_{vt} + u_{v,t+1} \\
\text{corr}(u_{vt}, u_{v,t+1}) &\neq 0
\end{aligned}
$$

Estimator	SU-LPM		Bivariate Probit	
	2008	2005	2008	2005
	(1)		(2)	
log maximum landholdings in v	0.023	0.031	0.088	0.109
	(0.005)***	(0.006)***	(0.020)***	(0.020)***
log minimum landholdings in v	-0.064	-0.057	-0.242	-0.222
	(0.012)***	(0.011)***	(0.043)***	(0.044)***
Number of Villages	51,592	51,592	51,592	51,592

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% . Standard errors clustered by district. SU-LPM refers to seemingly unrelated linear probability models (Zellner, 1965). The minimum and maximum landholdings are calculated over all agricultural landholdings above $\underline{R} = 0.1$ Ha. The specification is suggested by the latent variable model prior to integrating over observable land-holdings extrema. See paper for additional covariates and discussion.

Two-Step Estimates: Extensive Margin

Estimator	SU-LPM		Bivariate Probit	
	2008	2005	2008	2005
	(3)		(4)	
Pareto exponent $\widehat{\lambda}_v$	-0.011	-0.016	-0.049	-0.069
	(0.005)**	(0.006)***	(0.021)**	(0.023)***
log village population	0.081	0.074	0.304	0.277
	(0.006)***	(0.006)***	(0.020)***	(0.021)***
rice price shock	0.041	0.139	0.193	0.499
	(0.396)	(0.421)	(1.364)	(1.217)
rainfall shock	0.027	0.034	0.078	0.108
	(0.027)	(0.025)	(0.095)	(0.084)
log district population less v	0.095	0.091	0.316	0.303
	(0.034)***	(0.031)***	(0.109)***	(0.101)***
log district area less v	-0.047	-0.053	-0.156	-0.178
	(0.018)**	(0.017)***	(0.056)***	(0.051)***
log distance to subdistrict capital	-0.022	-0.023	-0.087	-0.087
	(0.005)***	(0.005)***	(0.016)***	(0.016)***
log distance to nearest emigration ctr.	-0.026	-0.021	-0.021	-0.016
	(0.029)	(0.029)	(0.087)	(0.080)
Number of Villages	51,592	51,592	51,592	51,592

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% . Standard errors clustered by district. SU-LPM refers to seemingly unrelated linear probability models (Zellner, 1965). This specification is suggested by the latent variable model after integrating over land-holdings extrema. See paper for additional covariates and discussion.

Two-Step Estimates: Extensive Margin

Estimator	SU-LPM		Bivariate Probit	
	2008	2005	2008	2005
	(3)		(4)	
Pareto exponent $\widehat{\lambda}_v$	-0.011	-0.016	-0.049	-0.069
	(0.005)**	(0.006)***	(0.021)**	(0.023)***
log village population	0.081	0.074	0.304	0.277
	(0.006)***	(0.006)***	(0.020)***	(0.021)***
rice price shock	0.041	0.139	0.193	0.499
	(0.396)	(0.421)	(1.364)	(1.217)
rainfall shock	0.027	0.034	0.078	0.108
	(0.027)	(0.025)	(0.095)	(0.084)
log district population less v	0.095	0.091	0.316	0.303
	(0.034)***	(0.031)***	(0.109)***	(0.101)***
log district area less v	-0.047	-0.053	-0.156	-0.178
	(0.018)**	(0.017)***	(0.056)***	(0.051)***
log distance to subdistrict capital	-0.022	-0.023	-0.087	-0.087
	(0.005)***	(0.005)***	(0.016)***	(0.016)***
log distance to nearest emigration ctr.	-0.026	-0.021	-0.021	-0.016
	(0.029)	(0.029)	(0.087)	(0.080)
Number of Villages	51,592	51,592	51,592	51,592

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% . Standard errors clustered by district. SU-LPM refers to seemingly unrelated linear probability models (Zellner, 1965). This specification is suggested by the latent variable model after integrating over land-holdings extrema. See paper for additional covariates and discussion.

Two-Step Estimates: Intensive Margin

$$\Delta \ln \left(M_{v,t+1}/N_{v,t+1} \right) = \theta_{\mathbf{a}} \Delta \textit{rainfall shock}_{vt} + \theta_{\mathbf{p}} \Delta \textit{price shock}_{vt} + \alpha \widehat{\lambda}_v + \boldsymbol{\zeta}' \mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta \varepsilon_{v,t+1}$$

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Landholdings type	**Agricultural**		**Wetland**		**Paddy Planted**	
	(1)	(2)	(3)	(4)	(5)	(6)
Pareto exponent $\widehat{\lambda}_v$	0.039	0.038	0.072	0.050	0.043	0.048
	(0.017)*	(0.018)*	(0.018)***	(0.017)***	(0.016)**	(0.017)**
Δ rainfall shock	**0.415**	**0.296**	**0.305**	**0.212**	**0.390**	**0.253**
	(0.133)* **	**(0.128)** **	**(0.134)* **	**(0.130)* **	**(0.139)** **	**(0.136)** **
Δ price shock	**0.409**	**0.283**	**1.287**	**0.625**	**0.919**	**0.457**
	(0.448)	**(0.426)**	**(0.502)* **	**(0.451)**	**(0.487)**	**(0.447)**
Number of villages	26,527	26,527	24,537	24,537	24,855	24,855

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% based on a block bootstrap$-t$ procedure with standard errors clustered at the district level. SU-LPM is seemingly unrelated linear probability model (Zellner, 1965). Semiparametric refers to Das, Newey, and Vella (2003) and includes a 3rd degree polynomial in the propensity scores for 2005/8. Parametric includes bivariate Mills ratio terms. Additional controls: log distance to (sub)district capital and nearest emigration center; ethnic Arab/Chinese share, Muslim share, and schooling distribution from 2000; share of households with land-holdings below 0.1 Ha. See paper for additional covariates and exclusion restrictions.

Landholdings Heterogeneity, Income Shocks, and Migration Flows

$$\Delta \ln\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) = \theta_a \Delta \textit{rainfall shock}_{vt} + \theta_{a\lambda}(\widehat{\lambda}_v \times \Delta \textit{rainfall shock}_{vt}) + \theta_p \Delta \textit{price shock}_{vt} + \theta_{p\lambda}(\widehat{\lambda}_v \times \Delta \textit{price shock}_{vt}) + \alpha \widehat{\lambda}_v + \boldsymbol{\zeta}' \mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta \varepsilon_{v,t+1}$$

Landholdings Heterogeneity, Income Shocks, and Migration Flows

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Landholdings type	**Agricultural**		**Wetland**		**Paddy Planted**	
	(1)	(2)	(3)	(4)	(5)	(6)
Pareto exponent $\widehat{\lambda}_v$	-0.007	-0.019	-0.113	-0.070	-0.083	-0.084
	(0.035)	(0.034)	(0.030)***	(0.030)***	(0.046)*	(0.042)***
Δ rainfall shock	0.225	0.188	0.262	0.135	0.167	0.110
	(0.169)	(0.162)	(0.168)	(0.161)	(0.176)	(0.174)
Δ price shock	-0.016	-0.616	-2.234	-1.503	-1.031	-1.750
	(0.688)	(0.686)	(0.709)**	(0.688)*	(0.822)	(0.776)**
$\widehat{\lambda}_v \times \Delta$ rainfall shock	**0.119**	**0.073**	**0.028**	**0.048**	**0.140**	**0.087**
	(0.074)*	**(0.070)**	**(0.052)**	**(0.051)**	**(0.065)****	**(0.061)**
$\widehat{\lambda}_v \times \Delta$ price shock	**0.267**	**0.586**	**1.913**	**1.155**	**1.116**	**1.314**
	(0.329)	**(0.335)***	**(0.335)*****	**(0.327)*****	**(0.423)*****	**(0.400)*****
Number of villages	26,527	26,527	24,537	24,537	24,855	24,855

Notes: Significance levels: * 10% ** 5% *** 1% based on a block bootstrap$-t$ procedure with standard errors clustered at the district level. SU-LPM is seemingly unrelated linear probability model (Zellner, 1965). Semiparametric refers to Das, Newey, and Vella (2003) and includes a 3rd degree polynomial in the propensity scores for 2005/8. Parametric includes bivariate Mills ratio terms. Additional controls: log distance to (sub)district capital and nearest emigration center; ethnic Arab/Chinese share, Muslim share, and schooling distribution from 2000; share of households with land-holdings below 0.1 Ha. See paper for additional covariates and exclusion restrictions.

Interpreting Elasticities

- $\widehat{\theta}_{\mathbf{p}\lambda}$: moving from 25th to 75th percentile of $\widehat{\lambda}_v$
  - ▷ $\lambda_{v,[0.25]}$: $mean_v(R) = 0.50$ Ha, $Gini_v(R) = 0.66$
  - ▷ $\lambda_{v,[0.75]}$: $mean_v(R) = 0.24$ Ha, $Gini_v(R) = 0.40$

  $\Longrightarrow$ tripling elasticity with respect to price shocks from $\sim 0.3 - 0.9$

- $\widehat{\theta}_{\mathbf{p}\lambda}$: largest and most precise for λ_v specific to rice production

- Estimates robust to allowing shocks to vary with other features of the land-holdings distribution and agricultural labor markets ▸ Tables

- $(\widehat{\theta}_{\mathbf{p}\lambda}, \widehat{\theta}_{\mathbf{a}\lambda})$: lower bounds for true effects if λ_v estimated with any error

- $\widehat{\theta}_{\mathbf{p}\lambda} \approx 1.9 \implies \widehat{\beta} \approx 0.55$
  - ▷ statistically indistinguishable from available production function estimates (Bazzi, 2012; Fuglie, 2010; Mundlak et al, 2004)

Other Evidence of Liquidity Constraints

$$\Delta \ln \left(M_{v,t+1} / N_{v,t+1} \right) = \theta_z z_{vt} + \theta_a \Delta \textit{rain shock}_{vt} + \theta_{az} (\Delta \textit{rain shock}_{vt} \times z_{vt}) + \zeta' \mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta \varepsilon_{v,t+1}$$

Correction Procedure *1st Stage Estimator*	Semipar. SU-LPM	Param. BiProbit	Semipar. SU-LPM	Param. BiProbit	Semipar. SU-LPM	Param. BiProbit
	z := bank presence in subdistrict		**z := log mean HH exp./capita**		**z := tech. irrigation in village**	
	(1)	(2)	(3)	(4)	(5)	(6)
z × Δ rainfall shock	**-0.226** **(0.087)****	**-0.162** **(0.085)****	**-0.909** **(0.163)*****	**-0.841** **(0.150)*****	**-0.256** **(0.068)*****	**-0.187** **(0.068)*****
Δ rainfall shock	0.572 (0.155)***	0.419 (0.147)***	10.897 (1.912)***	10.037 (1.742)***	0.527 (0.137)***	0.396 (0.132)***
z	-0.114 (0.026)***	-0.082 (0.025)***	-0.014 (0.059)	-0.062 (0.057)	-0.014 (0.024)	0.007 (0.022)
Number of Villages	26,527	26,527	26,127	26,127	26,527	26,527

Notes: Significance levels: * 10% ** 5% * * * 1% based on a block bootstrap−t procedure with standard errors clustered at the district level. Bank presence equals one if any banks located in village's subdistrict and zero otherwise; log mean household expenditures/capita obtained from Poverty Map estimates (SMERU, 2006) without any information on household land-holdings; technical irrigation equals one if village has any land irrigated by technical system not reliant on rainfall. See paper for additional covariates and exclusion restrictions.

Results Are Robust to...

1. Including any two of the four instruments in the 2nd stage
2. Controlling for natural disasters
3. Controlling for other commodity price and agri. GDP shocks
4. Alternative specifications for rainfall and price shocks
   - rainfall shock in levels not diffs, fully elaborated annual shocks
   - alternative proxies for local rice price changes
5. Controls for demog./age structure, avg. HH size, $F(\textit{schooling})$
6. Alternative choices of $\underline{R}$ in estimating λ_v
7. Outliers in migrant stocks and $\widehat{\lambda}_v$; quality of pop. registers
8. Addition of 3rd year to panel in 2011 (preliminary)

Income Shocks, Land-holdings and Migration *Choice*

What do the micro data say?

$$migrate_{iv,t+1} = \alpha + \beta rainfall\ shock_{vt} + \gamma price\ shock_{vt} + \eta_i + \eta_t + e_{ivt}$$

- $migrate_{iv,t+1} = 1$ if household i in village v had migrant(s) depart in year $t+1$
- recall data from nationally representative household survey (*Susenas*)
- conditional fixed effects logit with panel of *ever*-migrants, 2000-6

Income Shocks, Land-holdings and Migration *Choice*

What do the micro data say?

$$migrate_{iv,t+1} = \alpha + \beta rainfall\ shock_{vt} + \gamma price\ shock_{vt} + \eta_i + \eta_t + e_{ivt}$$

	(1)	(2)
rainfall shock, t	0.200	0.212
	(0.115)*	(0.112)*
price shock, t		0.762
		(0.340)**
Average Marginal Effect (AME)	Yes	Yes
Observations	1,902	1,380
Years	6	5

Notes: Significance levels: ∗ 10% ∗∗ 5% ∗ ∗ ∗ 1% . Average marginal effects (AMEs) based on conditional fixed effects logit estimates. Price shock is log growth from $tm1$-$tm12$; rainfall shock is annual log deviation from long-run mean. By necessity, computation of the AMEs requires imposing that the fixed effects equal zero. Both columns include year fixed effects. Standard errors clustered at the district level. The sample size drops when including price shocks because that data is only available from 2002 forward.

Income Shocks, Land-holdings and Migration *Choice*

What do the micro data say?

$$migrate_{iv,t+1} = \alpha + \beta rainfall\ shock_{vt} + \gamma price\ shock_{vt} + \eta_i + \eta_t + e_{ivt}$$

	(1)	(2)	(3)	(4)
rainfall shock, t	0.200	0.212	1.109	1.712
	(0.115)*	(0.112)*	(0.578)*	(0.605)***
price shock, t		0.762	3.752	3.892
		(0.340)**	(1.687)**	(1.729)**
rainfall shock, t × land-holdings (Ha)			-0.267	-3.844
			(0.580)	(1.416)***
rainfall shock, t × land-holdings (Ha) squared				1.583
				(0.647)**
price shock, t × land-holdings (Ha)			-0.581	-3.128
			(1.176)	(1.875)*
price shock, t × land-holdings (Ha) squared				0.802
				(0.513)
Average Marginal Effect (AME)	Yes	Yes	No	No
Observations	1,902	1,380	1,380	1,380
Years	6	5	5	5

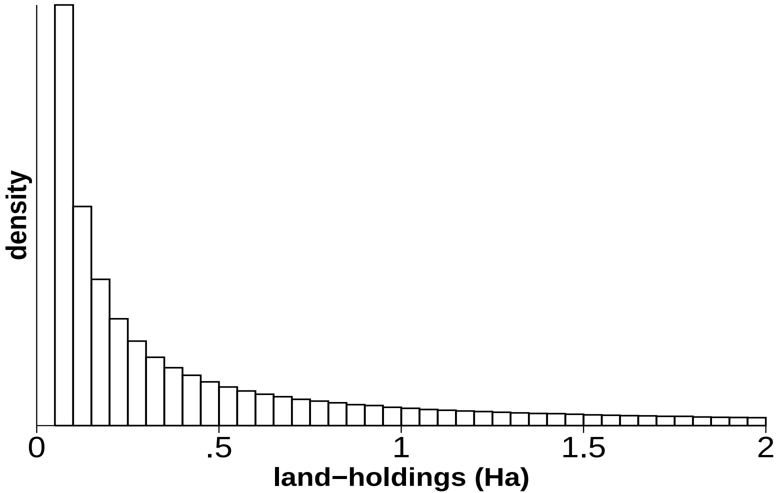
Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% . Average marginal effects (AMEs) based on conditional fixed effects logit estimates. Price shock is log growth from $tm1$-$tm12$; rainfall shock is annual log deviation from long-run mean. By necessity, computation of the AMEs requires imposing that the fixed effects equal zero. Both columns include year fixed effects. Standard errors clustered at the district level. The sample size drops when including price shocks because that data is only available from 2002 forward.

Validating the Mapping from Migration Choices to Flows

1. Average marginal effects (AMEs) of income shocks at each land-holding size R_{iv} in micro-data

2. Calculate village-specific elasticity by reweighting AMEs using $\widehat{\lambda}_v$

Validating the Mapping from Migration Choices to Flows



Comparing Elasticities

Elasticity Summary Statistic→	**mean**	std. dev.	**median**	correl.
Δprice shock				
Aggregating Micro-Data AMEs (linear R_{iv})	**0.771**	0.018	**0.775**	0.898
Aggregating Micro-Data AMEs (quadratic R_{iv})	**0.698**	0.030	**0.702**	0.757
Village-Level Regression	**0.747**	0.473	**0.680**	—
Δrainfall shock				
Aggregating Micro-Data AMEs (linear R_{iv})	**0.223**	0.006	**0.225**	0.794
Aggregating Micro-Data AMEs (quadratic R_{iv})	**0.226**	0.018	**0.226**	0.977
Village-Level Regression	**0.389**	0.059	**0.381**	—

Notes: *Aggregating Micro Data AMEs* elasticities are computed based on columns 3-4 in previous table as the sum of nationally representative average marginal effects (AMEs) of shocks at land-holding sizes $\in \{0.1, 0.2, \ldots, 2.5\}$ weighted by the share of households in the village falling within each of the given size ranges as implied by the estimated Pareto dispersion parameter $\widehat{\lambda}_v$. *Village-Level Regression* elasticities are based on the estimates of θ parameters using village-level data in main baseline specification with interactions of $\widehat{\lambda}_v$ and income shocks. All elasticities are based on λ_v estimated for total wetland holdings and are restricted to villages in the second-step sample.

Comparing Elasticities

Elasticity Summary Statistic$\rightarrow$	mean	std. dev.	median	**correl.**
Δprice shock				
Aggregating Micro-Data AMEs (linear R_{iv})	0.771	0.018	0.775	**0.898**
Aggregating Micro-Data AMEs (quadratic R_{iv})	0.698	0.030	0.702	**0.757**
Village-Level Regression	0.747	0.473	0.680	—
Δrainfall shock				
Aggregating Micro-Data AMEs (linear R_{iv})	0.223	0.006	0.225	**0.794**
Aggregating Micro-Data AMEs (quadratic R_{iv})	0.226	0.018	0.226	**0.977**
Village-Level Regression	0.389	0.059	0.381	—

Notes: *Aggregating Micro Data AMEs* elasticities are computed based on columns 3-4 in previous table as the sum of nationally representative average marginal effects (AMEs) of shocks at land-holding sizes $\in \{0.1, 0.2, \ldots, 2.5\}$ weighted by the share of households in the village falling within each of the given size ranges as implied by the estimated Pareto dispersion parameter $\widehat{\lambda}_v$. *Village-Level Regression* elasticities are based on the estimates of θ parameters using village-level data in main baseline specification with interactions of $\widehat{\lambda}_v$ and income shocks. All elasticities are based on λ_v estimated for total wetland holdings and are restricted to villages in the second-step sample.

Comparing Elasticities

Elasticity Summary Statistic→	mean	**std. dev.**	median	correl.
Δprice shock				
Aggregating Micro-Data AMEs (linear R_{iv})	0.771	**0.018**	0.775	0.898
Aggregating Micro-Data AMEs (quadratic R_{iv})	0.698	**0.030**	0.702	0.757
Village-Level Regression	0.747	**0.473**	0.680	—
Δrainfall shock				
Aggregating Micro-Data AMEs (linear R_{iv})	0.223	**0.006**	0.225	0.794
Aggregating Micro-Data AMEs (quadratic R_{iv})	0.226	**0.018**	0.226	0.977
Village-Level Regression	0.389	**0.059**	0.381	—

Notes: *Aggregating Micro Data AMEs* elasticities are computed based on columns 3-4 in previous table as the sum of nationally representative average marginal effects (AMEs) of shocks at land-holding sizes $\in \{0.1, 0.2, \ldots, 2.5\}$ weighted by the share of households in the village falling within each of the given size ranges as implied by the estimated Pareto dispersion parameter $\widehat{\lambda}_V$. *Village-Level Regression* elasticities are based on the estimates of θ parameters using village-level data in main baseline specification with interactions of $\widehat{\lambda}_V$ and income shocks. All elasticities are based on λ_V estimated for total wetland holdings and are restricted to villages in the second-step sample.

Recovering Village-Specific Migration Costs

- Empirical tests $\Longrightarrow$ liquidity-constrained regime

$$\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) = \frac{\lambda_v}{\beta} \Delta \ln p_{vt} + \Delta \ln \left[\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} - \left(\frac{\overline{\sigma}_v \alpha_v}{\omega_{vj,t+1} - C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} \right]$$

Recovering Village-Specific Migration Costs

- Empirical tests $\Longrightarrow$ liquidity-constrained regime

$$\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) = \frac{\lambda_v}{\beta} \Delta \ln p_{vt} + \Delta \ln \left[\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} - \left(\frac{\overline{\sigma}_v \alpha_v}{\omega_{vj,t+1} - C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} \right]$$

- Estimate C_{vj}'s for *all* villages using two-step model
  1. predicted $\Delta \ln \left(M_{v,t+1} / N_{v,t+1} \right)$
  2. estimated β

- Auxiliary inputs
  1. α_v estimable using monthly local rice price indices
  2. ω_{vj}'s from published contracts by destination

Systematic Variation in Implied Costs

Estimated Village-Specific Migration Costs (USD)

	mean	std. dev.	median
all villages	1,485	430	1,373
villages with migrants in...			
2005 & 2008	1,474	481	1,334
neither year	1,509	299	1,431
only in 2008	1,491	320	1,430
only in 2005	1,487	479	1,358

Notes: The village-specific costs (in 2006 USD, roughly) are recovered from the structural equation and two-step model for a total of 42,063 villages, 24,540 of which have migrants in both 2005 and 2008, 9,504 in neither year, 4,875 only in 2008, and 3,144 only in 2005.

- placement+pre-departure fees: ∼800-1200 USD (quoted by recruiters/government)

Systematic Variation in Implied Costs

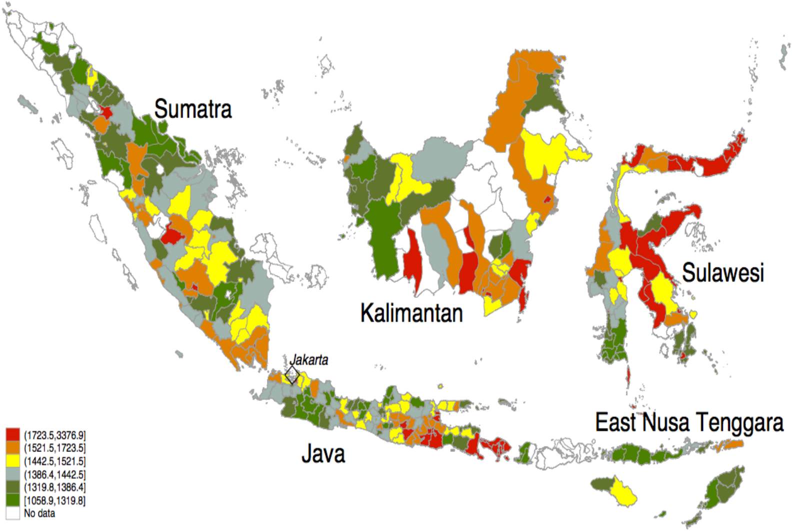
Estimated Village-Specific Migration Costs (USD)

	mean	std. dev.	median
all villages	1,485	430	1,373
villages with migrants in...			
2005 & 2008	1,474	481	1,334
neither year	1,509	299	1,431
only in 2008	1,491	320	1,430
only in 2005	1,487	479	1,358

Notes: The village-specific costs (in 2006 USD, roughly) are recovered from the structural equation and two-step model for a total of 42,063 villages, 24,540 of which have migrants in both 2005 and 2008, 9,504 in neither year, 4,875 only in 2008, and 3,144 only in 2005.

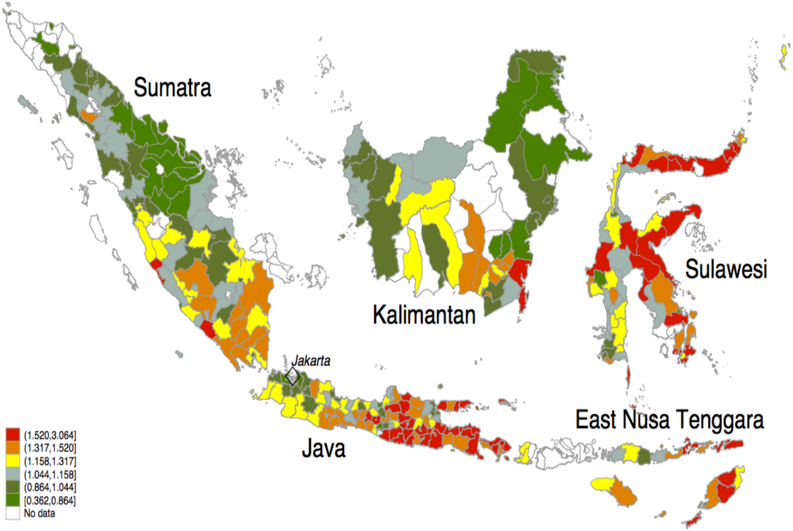
▷ placement+pre-departure fees: ∼800-1200 USD (quoted by recruiters/government)

Average Village-Specific Costs for Two-Year Contracts, by District



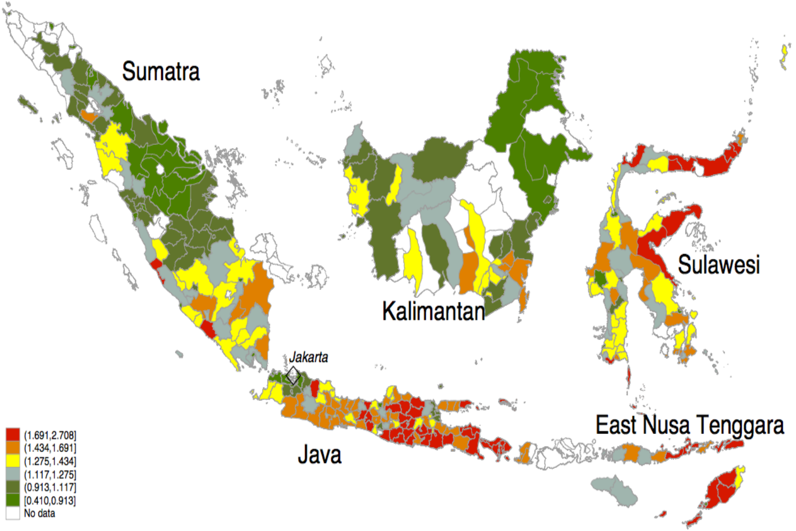
Notes: The village-specific costs (in 2006 USD, roughly) are recovered from the structural equation and two-step model.

Average Village-Specific Costs / Annual Household Expenditure, by District



Notes: The village-specific costs (in 2006 USD, roughly) are recovered from the structural equation and two-step model. Household expenditure estimates are based on survey data from 2006 representative at the district level.

Average Village-Specific Net Income Abroad / Cumulative Household Expenditure, by District



Notes: The wages are village-specific based on plurality destination (of neighbors), and the village-specific costs (in 2006 USD, roughly) are recovered from the structural equation and two-step model. Household expenditure estimates are based on survey data from 2006 representative at the district level.

Ruling Out Other Potential Explanations

1. destination demand shocks
2. idiosyncratic preferences or costs
3. skill heterogeneity
4. internal migration flows (changes in N_v)

Conclusion

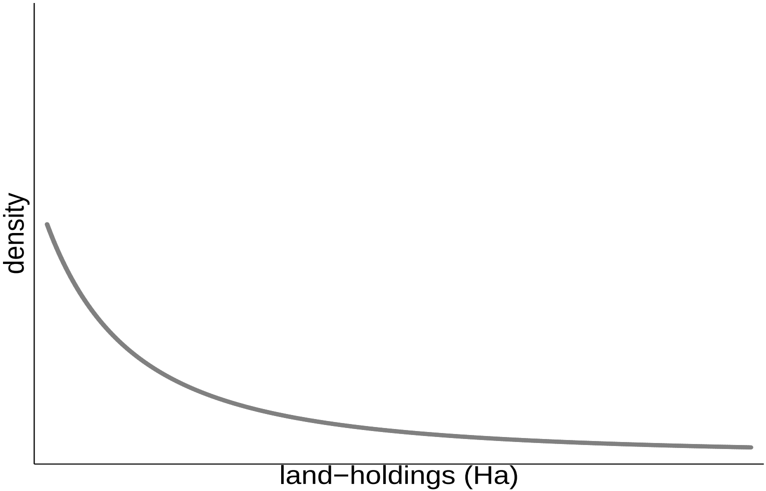
- *Theory*: tests for prevalence of financial barriers to migration
- *Empirics*: evidence of binding financial constraints in Indonesia
- Implications for other developing countries (e.g., MEX, PHL)?
- Future theoretical extensions
  - multiple destinations (including internal) with different fixed costs
  - alternative investment option (in agriculture)

Thank You.

sbazzi@ucsd.edu

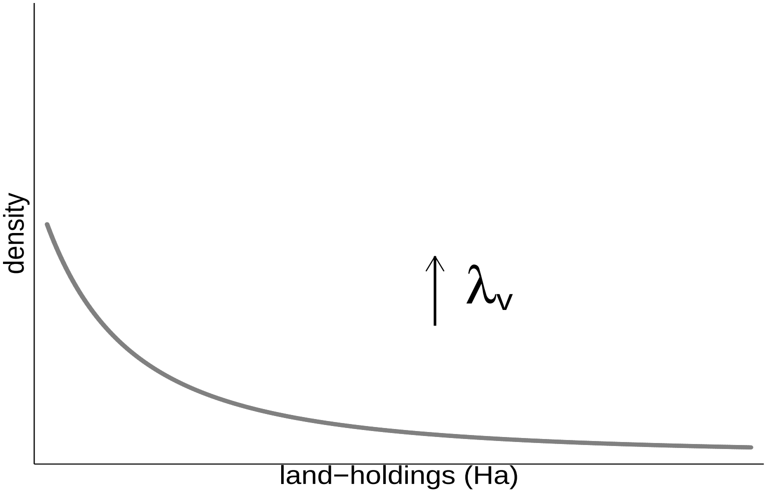
APPENDIX

Pareto Distribution: Example



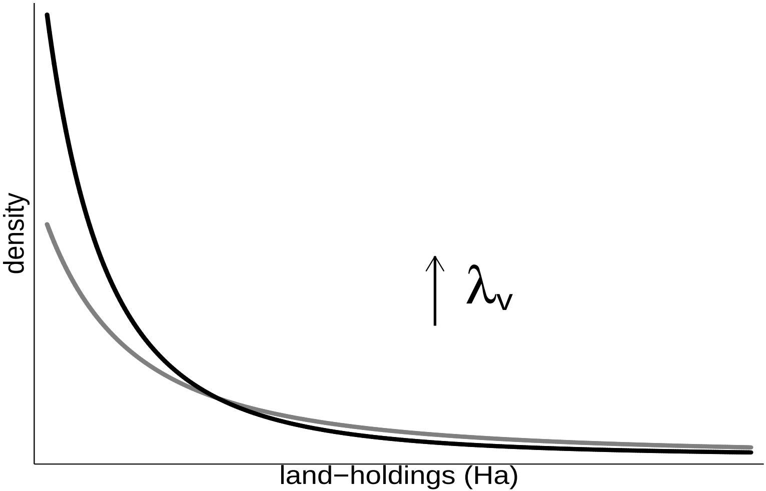
back

Pareto Distribution: Example



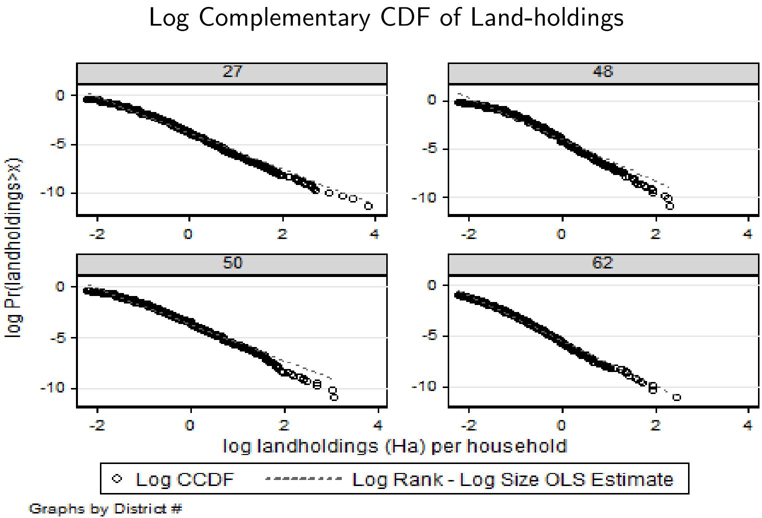
back

Pareto Distribution: Example



back

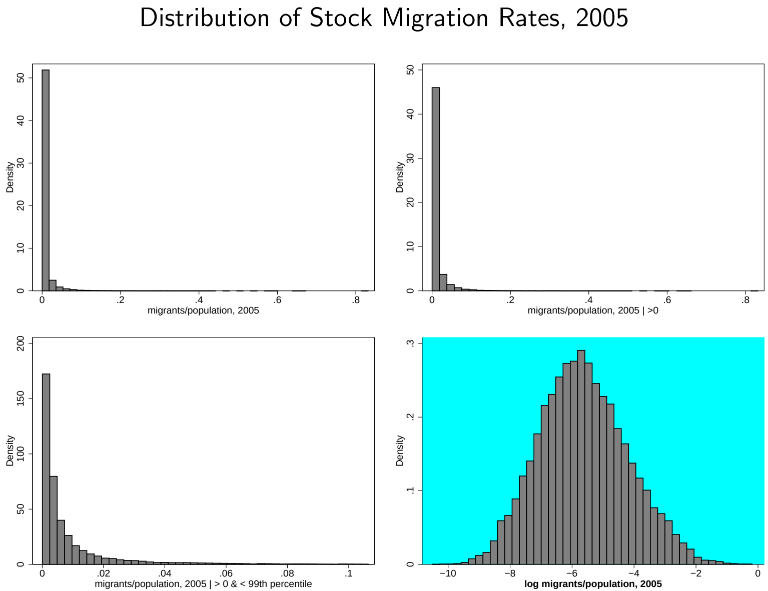
Pareto Power Law Properties



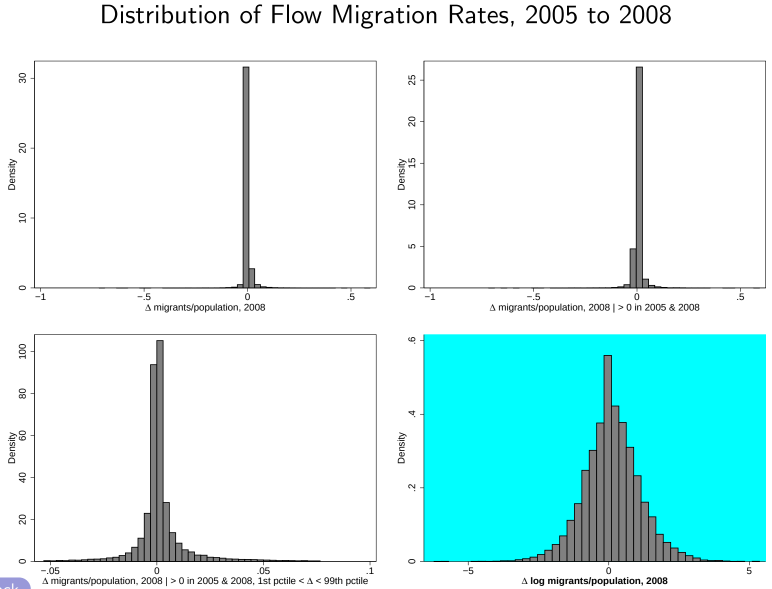
Notes: The figures report, for four randomly chosen districts, the log CCDF – log size observations for total wetland holdings in the 2002-3 growing season for all Indonesian households recorded in the Agricultural Census of 2003. The graphs impose lower thresholds of $\underline{R} = 0.1$ Ha in estimating the CCDF. The dashed lines are the linear OLS estimates of λ_V.

back

Distribution of Migration Rates



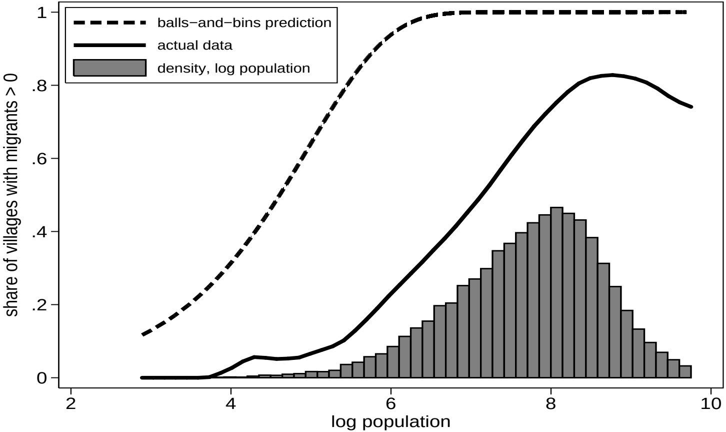
Distribution of Migration Rates



back

Balls-and-Bins Test

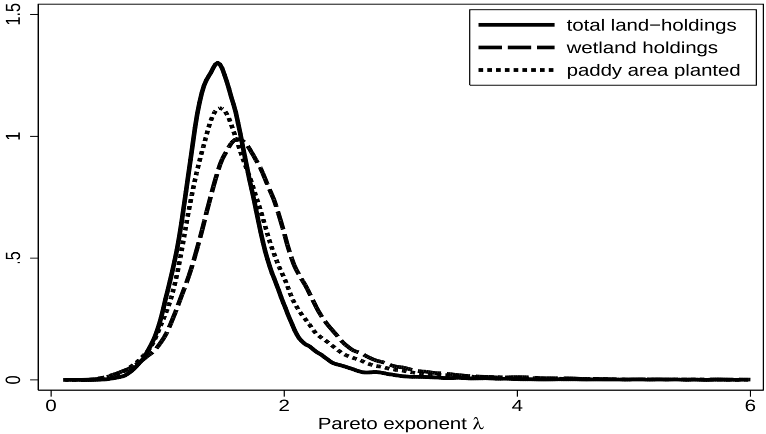
$\Longrightarrow$ $<$ 5.5 percent of the 27,297 zeros in 2005 can be deemed an atheoretical regularity in sparse data



Notes: The estimates are based on local linear regressions using an Epanechnikov kernel, bandwidth of 400, and trimming the top 4 percent of villages for presentational purposes. The confidence bands are computed using analytic standard error formulas. The test is adapted from the Armenter and Koren (2008) model for trade data.

back

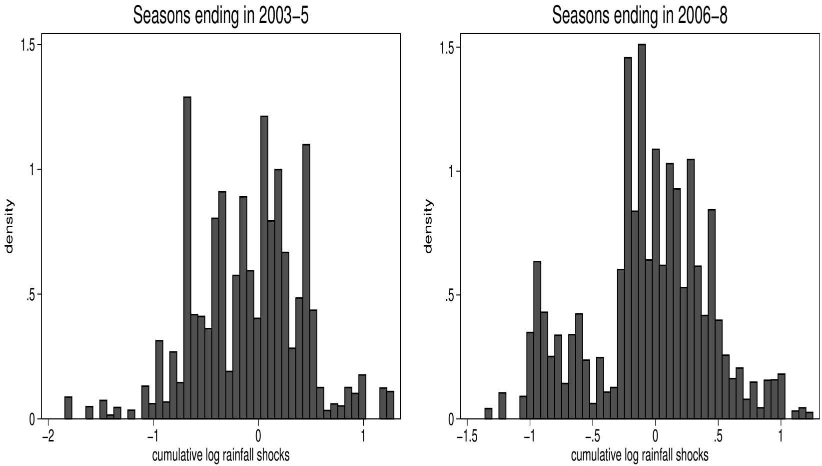
Estimated Pareto Exponents λ_v



Notes: The Pareto distribution is given by $\lambda_{i}v\underline{R}^{\lambda_v} R_{iv}^{-\lambda_v - 1}$. The figure shows the kernel density of Gabaix & Ibragimov (2011) log rank(-1/2) - log size OLS estimates of λ_v using the average log rank for a given log land-holding size and imposing $\underline{R} = 0.1$ hectares. The estimates were calculated independently across all villages with at least 3 distinct land-holding sizes of the given type recorded in the Agricultural Census 2003. In the figure, the top 2 % of estimates are trimmed.

back

Distribution of Rainfall Shocks



Notes: The figure shows the cumulative log deviation of province-specific seasonal rainfall levels from their long-run mean (1952-2008) by district.

back

Reduced Form

$$\begin{aligned} M_{v,t+1}/N_{v,t+1} &= \theta_{\mathbf{a}} \mathit{rainfall\ shock}_{vt} + \theta_{\mathbf{a}\lambda}(\widehat{\lambda}_v \times \mathit{rainfall\ shock}_{vt}) \\ &+ \theta_{\mathbf{p}} \mathit{price\ shock}_{vt} + \theta_{\mathbf{p}\lambda}(\widehat{\lambda}_v \times \mathit{price\ shock}_{vt}) \\ &+ \xi_t + \xi_v + \varepsilon_{v,t+1} \end{aligned}$$

$$\begin{aligned} M_{v,t+1} &= \theta_{\mathbf{a}} \mathit{rainfall\ shock}_{vt} + \theta_{\mathbf{a}\lambda}(\widehat{\lambda}_v \times \mathit{rainfall\ shock}_{vt}) + \phi(=1) \ln N_{v,t+1} \\ &+ \theta_{\mathbf{p}} \mathit{price\ shock}_{vt} + \theta_{\mathbf{p}\lambda}(\widehat{\lambda}_v \times \mathit{price\ shock}_{vt}) \\ &+ \xi_t + \xi_v + \varepsilon_{v,t+1} \end{aligned}$$

Reduced Form

	FE-OLS		Semipar.	FE-Tobit	CFE-Poisson	
	(1)	(2)	(3)	(4)	(5)	(6)
rainfall shock	**0.0011**	**0.0014**	**0.0083**	**0.0033**	**0.3141**	**0.3325**
	(0.0002)***	(0.0006)**	(0.0011)***	(0.0047)	(0.0485)***	(0.1619)**
	[0.0011]	[0.0012]	[0.0047]*	[0.0051]	[0.0950]**	[0.2465]
rice price shock	**0.0086**	**0.0023**	**0.0254**	**-0.0185**	**1.0627**	**-0.2001**
	(0.0014)***	(0.0023)	(0.0046)***	(0.0135)	(0.2371)***	(0.4258)
	[0.0058]	[0.0064]	[0.0172]	[0.0176]	[0.5805]*	[0.9002]
$\widehat{\lambda}_v \times$ **rainfall shock**		**-0.0002**		**0.0032**		**0.0075**
		(0.0004)		(0.0032)		(0.0932)
		[0.0006]		[0.0033]		[0.1360]
$\widehat{\lambda}_v \times$ **price shock**		**0.0040**		**0.0293**		**0.8630**
		(0.0013)***		(0.0093)***		(0.2486)***
		[0.0026]		[0.0125]**		[0.4634]*
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% Standard errors are clustered by village in parentheses and district in brackets. Sample size: 103,196 (51,598 villages) in columns 1-4; 72,650 (36,325 villages) in columns 5-6 (since the CFE estimators retains villages with any migrants in 2005 or 2008). Semiparametric Tobit is the trimmed LAD estimator of Honore (1992). *rainfall shock* is the cumulative log deviation from long-run mean rainfall in the growing seasons ending in 2006-2008 or 2002-2005. *rice price shock* is the annualized log growth rate in the nearest rice price index between 2005m4-2008m3 or 2002m1-2005m3. The estimated Pareto exponent $\widehat{\lambda}_v$ is for total agricultural land-holdings.

back

θ_a, θ_p, and Other Features of Land Distribution and Labor Market

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Landholdings type	**Agricultural**		**Wetland**		**Paddy Planted**	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ rainfall shock	0.740	0.668	0.561	0.450	0.518	0.445
	(0.179)***	(0.164)***	(0.217)***	(0.450)***	(0.210)***	(0.205)***
Pareto exponent $\widehat{\lambda}_v$	0.015	0.034	0.046	0.037	0.021	0.037
	(0.018)	(0.018)*	(0.018)**	(0.017)**	(0.017)	(0.017)**
$\times$ Δ rainfall shock	0.062	0.030	0.042	0.031	0.110	0.071
	(0.068)	(0.064)	(0.055)	(0.057)	(0.058)*	(0.060)
share households $<$ 0.1 Ha	-0.026	-0.019	0.039	0.022	-0.048	-0.029
	(0.047)	(0.047)	(0.047)	(0.047)	(0.045)	(0.045)
$\times$ Δ rainfall shock	-0.789	-0.682	-0.486	-0.330	-0.487	-0.366
	(0.133)***	(0.138)***	(0.151)***	(0.141)**	(0.148)***	(0.142)**
share pop. paid agricultural labor	-0.451	-0.316	-0.337	-0.250	-0.389	-0.291
	(0.087)***	(0.075)***	(0.090)***	(0.081)***	(0.090)***	(0.081)***
$\times$ Δ rainfall shock	-1.124	-0.922	-0.792	-0.709	-0.997	-0.849
	(0.334)***	(0.306)***	(0.347)**	(0.316)**	(0.342)***	(0.320)***
Number of villages	26,527	26,527	24,537	24,537	24,855	24,855

Notes: Significance levels: * 10% ** 5% *** 1% Standard errors are clustered by district. See paper for additional covariates and exclusion restrictions.

θ_a, θ_p, and Other Features of Land Distribution and Labor Market

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Landholdings type	**Agricultural**		**Wetland**		**Paddy Planted**	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ price shock	0.493	0.165	0.048	-0.025	1.047	-0.452
	(0.840)	(0.838)	(0.998)	(0.908)	(1.025)	(0.926)
Pareto exponent $\widehat{\lambda}_v$	0.027	0.011	-0.070	-0.037	-0.043	-0.048
	(0.036)	(0.036)	(0.029)**	(0.029)	(0.039)	(0.038)
$\times$ Δ price shock	0.069	0.245	1.416	0.761	0.770	0.931
	(0.326)	(0.339)	(0.324)***	(0.314)**	(0.352)**	(0.360)**
share households $<$ 0.1 Ha	0.147	0.131	0.280	0.173	0.222	0.124
	(0.081)*	(0.080)	(0.088)***	(0.073)**	(0.083)***	(0.073)*
$\times$ Δ price shock	-1.979	-1.478	-3.124	-1.853	-3.621	-1.781
	(0.794)**	(0.768)*	(0.873)***	(0.736)**	(0.887)***	(0.765)**
share pop. paid agricultural labor	-0.875	-0.515	-0.857	-0.572	-0.920	-0.533
	(0.158)***	(0.150)***	(0.172)***	(0.158)***	(0.168)***	(0.150)***
$\times$ Δ price shock	5.199	2.912	5.823	3.866	5.769	3.253
	(1.351)***	(1.255)**	(1.498)***	(1.319)***	(1.429)***	(1.239)***
Number of villages	26,527	26,527	24,537	24,537	24,855	24,855

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% Standard errors are clustered by district. See paper for additional covariates and exclusion restrictions.

back

Income Shocks, Banks, and Migration Choice

$$migrate_{iv,t+1} = \alpha + \beta rainfall\ shock_{vt} + \zeta' \mathbf{shock}_{vt} \times bank_v + \eta_i + \eta_t + e_{ivt}$$

- $migrate_{iv,t+1} = 1$ if household i in village v had any migrants in year $t+1$
- recall data from nationally representative household survey (*Susenas*)
- conditional fixed effects logit with panel of ever-migrants, 2000-6

	(1)	(2)	(3)	(4)
rainfall shock, t	0.891	0.979	1.510	1.574
	(0.514)*	(0.528)*	(0.696)**	(0.676)**
rainfall shock, $t-1$		0.163		0.100
		(0.494)		(0.518)
rainfall shock, t × bank in subdistrict			-3.796	-3.877
			(1.381)***	(1.428)***
rainfall shock, $t-1$ × bank in subdistrict				0.049
				(0.075)
Observations	1,902	1,902	1,902	1,902

Significance levels: * 10% ** 5% *** 1% Standard errors clustered at the district level. Estimates based on recall data from a nationally representative household survey (*Susenas*) conducted in mid-2006.

back

Robustness of Θ to Invalid Instruments

	(1)	(2)	(3)	(4)	(5)	(6)
Pareto exponent $\hat{\lambda}_v$	0.073	-0.113	0.082	-0.111	0.091	-0.109
	(0.018)***	(0.030)***	(0.018)***	(0.030)**	(0.019)***	(0.031)**
Δ rainfall shock	1.304	-2.234	1.317	-2.343	1.397	-2.381
	(0.503)*	(0.709)*	(0.515)	(0.724)	(0.509)	(0.734)
Δ price shock	0.162	0.262	0.150	0.260	0.165	0.277
	(0.167)	(0.168)	(0.165)	(0.166)	(0.166)	(0.167)
$\hat{\lambda}_v \times \Delta$ rainfall shock	0.082	0.028	0.103	0.040	0.105	0.040
	(0.049)	(0.052)	(0.048)*	(0.052)	(0.048)*	(0.053)
$\hat{\lambda}_v \times \Delta$ price shock		1.913		1.973		2.031
		(0.335)***		(0.334)***		(0.344)***
log # villages in district			-0.076	-0.065	-0.131	-0.114
			(0.025)**	(0.025)*	(0.034)*	(0.033)
log district area less v					0.058	0.052
					(0.025)	(0.024)
Number of Villages	24,537	24,537	24,537	24,537	24,537	24,537

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% The table reports estimates of the key elasticity parameters sequentially relaxing one or two of the four baseline exclusion restrictions. All estimates are based on the Das et al (2003) semiparametric correction procedure and the measure of $\widehat{\lambda}_v$ for wetland holdings. Standard errors are clustered at the district level.

Robustness of Θ to Invalid Instruments

	(7)	(8)	(9)	(10)	(11)	(12)
Pareto exponent $\hat{\lambda}_v$	0.065	-0.108	0.065	-0.108	0.069	-0.108
	(0.019)**	(0.030)***	(0.019)**	(0.030)**	(0.019)**	(0.030)**
Δ rainfall shock	1.183	-2.131	1.173	-2.134	1.178	-2.168
	(0.519)	(0.726)	(0.522)	(0.726)	(0.523)	(0.724)*
Δ price shock	0.123	0.217	0.120	0.215	0.116	0.216
	(0.168)	(0.169)	(0.168)	(0.169)	(0.168)	(0.169)
$\hat{\lambda}_v \times \Delta$ rainfall shock	0.086	0.034	0.087	0.035	0.093	0.036
	(0.050)*	(0.052)	(0.049)	(0.052)	(0.049)	(0.052)
$\hat{\lambda}_v \times \Delta$ price shock		1.787		1.785		1.805
		(0.337)***		(0.338)***		(0.337)***
log village pop., t	0.089	0.086	0.090	0.087	0.073	0.078
	(0.028)*	(0.027)	(0.027)	(0.027)	(0.036)	(0.034)
log district pop. less v, t					-0.030	-0.016
					(0.038)	(0.038)
log district area less v			-0.006	-0.005		
			(0.019)	(0.018)		
Number Of Villages	24,537	24,537	24,537	24,537	24,537	24,537

Notes: Significance levels: $*$ 10% $**$ 5% $***$ 1% The table reports estimates of the key elasticity parameters sequentially relaxing one or two of the four baseline exclusion restrictions. All estimates are based on the Das et al (2003) semiparametric correction procedure and the measure of $\widehat{\lambda}_v$ for wetland holdings. Standard errors are clustered at the district level.

back