Networks and Mis-allocation: Insurance, Migration, and the Rural-Urban Wage Gap

Kaivan Munshi

University of Cambridge

Mark R. Rosenzweig

Yale University

Removing obstacles to resource mis-allocation can be very productive

Requires identifying the source of the mis-allocation - “bottleneck”

In India, observe large and persistent real wage gap between rural and urban areas and across villages in rural areas (not due to selection)

Larger rural-urban wage gap than most other countries of the world.

Also, among the lowest rural-urban and rural-rural male migration rates in the world.

[High migration rates for women, but for marriage.]

What is the cause? A cause unique to India.

What is the remedy to the mis-allocation?

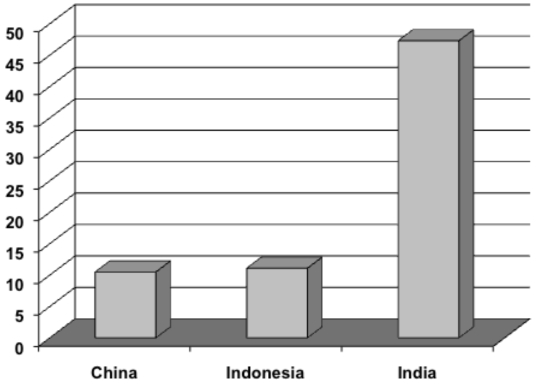
How do we know what it is, and whether it will work?

Table 1
Rural-Urban Wage and Expected Wage Gaps in India in 2004
(Daily Wages, Rupees)

Sector	Nominal	PPP-adjusted (rural consumption)	PPP & unemployment-adjusted
Urban	62.7	54.1	51.2
Rural	42.5	42.5	38.8
% Gain	47.3	27.1	31.9

Source: National Sample Survey (NSS)

Figure 1: Rural-Urban Wage Gap, by Country



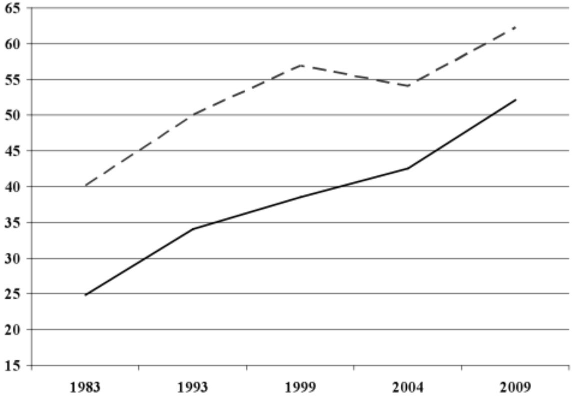
Source: 2006 Chinese mini-census, 2007 IFLS, 2004 NSS

Table 1A
Changes in the Nominal Rural-Urban Wage Gap, 1993-2007/9
in India and Indonesia

Country/Sector	India		Indonesia	
Year	1993	2009	1993	2007
Urban	52.8	80.3	22.9	37.3
Rural	84.2	104.7	39.3	41.5
% Gain	59.5	30.3	71.8	11.1

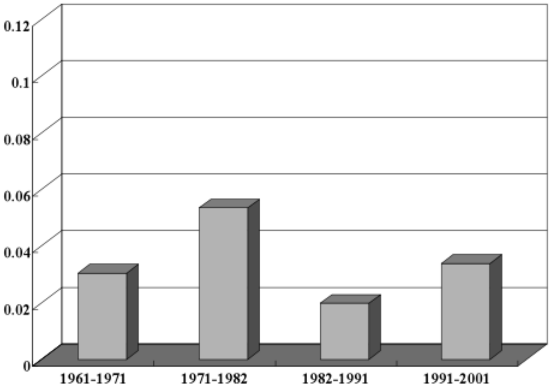
Sources: National Sample Surveys (NSS) and Indonesia Family Life Surveys (IFLS)

Figure 2: Real Rural and Urban Wages in India



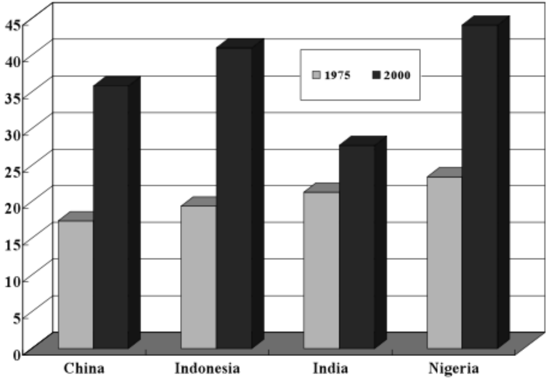
Source: 1983-2009 NSS

Figure 3: Change in Rural-Urban Migration Rates in India, 1961- 2001



Source: 1961-2001 Indian Population Census

Figure 4: Change in Percent Urbanized, by Country, 1975-2000



Source: UNDP 2002

An Explanation for Low Mobility

- Why have rural Indian workers not taken advantage of the economic opportunities associated with spatial wage differentials?
  - Combination of well-functioning rural insurance networks and the absence of formal insurance (Banerjee and Newman 1998)

An Explanation for Low Mobility

- In rural India, insurance networks are organized along caste lines
- Commitment and information problems are greater for households with male migrants
- If the resulting loss in network insurance is sufficiently large, and alternative sources of insurance are unavailable, then large wage gaps could persist without generating a flow of workers to higher wage areas

Strategies to Increase Mobility

- Move as a group (Munshi and Rosenzweig 2006, Munshi 2011)
    - Only available to members of select castes
- Temporary/seasonal migration (Morten 2012)
    - Cannot be used for permanent jobs

Distribution and Clustering of Castes (2006 REDS Census)

Mean number of castes per state:	66
Mean number of castes per village:	12
Mean number of hh's per village:	326
Mean number of hh's per caste in a village:	27

Also observe within-village spatial clustering of castes
(Based on street-level location information)

Table 2: Participation in the Caste-Based Insurance Arrangement

Survey year:	1982	1999
	(1)	(2)
Households participating (%)	25.44	19.62
Percent of income sent	5.28	8.74
Percent of income received	19.06	40.26
Number of observations	4981	7405

Source: Rural Economic Development Survey (REDS) 1982 and 1999

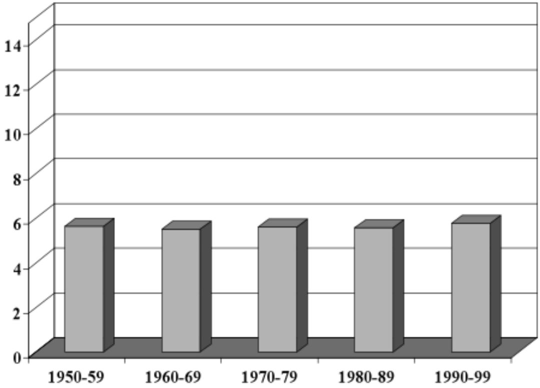
Table 3: Percent of Loans by Purpose and Source

Data source:	1982 REDS				
Purpose:	investment	operating expenses	contingencies	consumption expenses	all
	(1)	(2)	(3)	(4)	(5)
Sources:					
Bank	64.11	80.80	27.58	25.12	64.61
Caste	16.97	6.07	42.65	23.12	13.87
Friends	2.11	11.29	2.31	4.33	7.84
Employer	5.08	0.49	21.15	15.22	5.62
Moneylender	11.64	1.27	5.05	31.85	7.85
Other	0.02	0.07	1.27	0.37	0.22
Total	100.00	100.00	100.00	100.00	100.00

Table 4: Percent of Loans by Type and Source

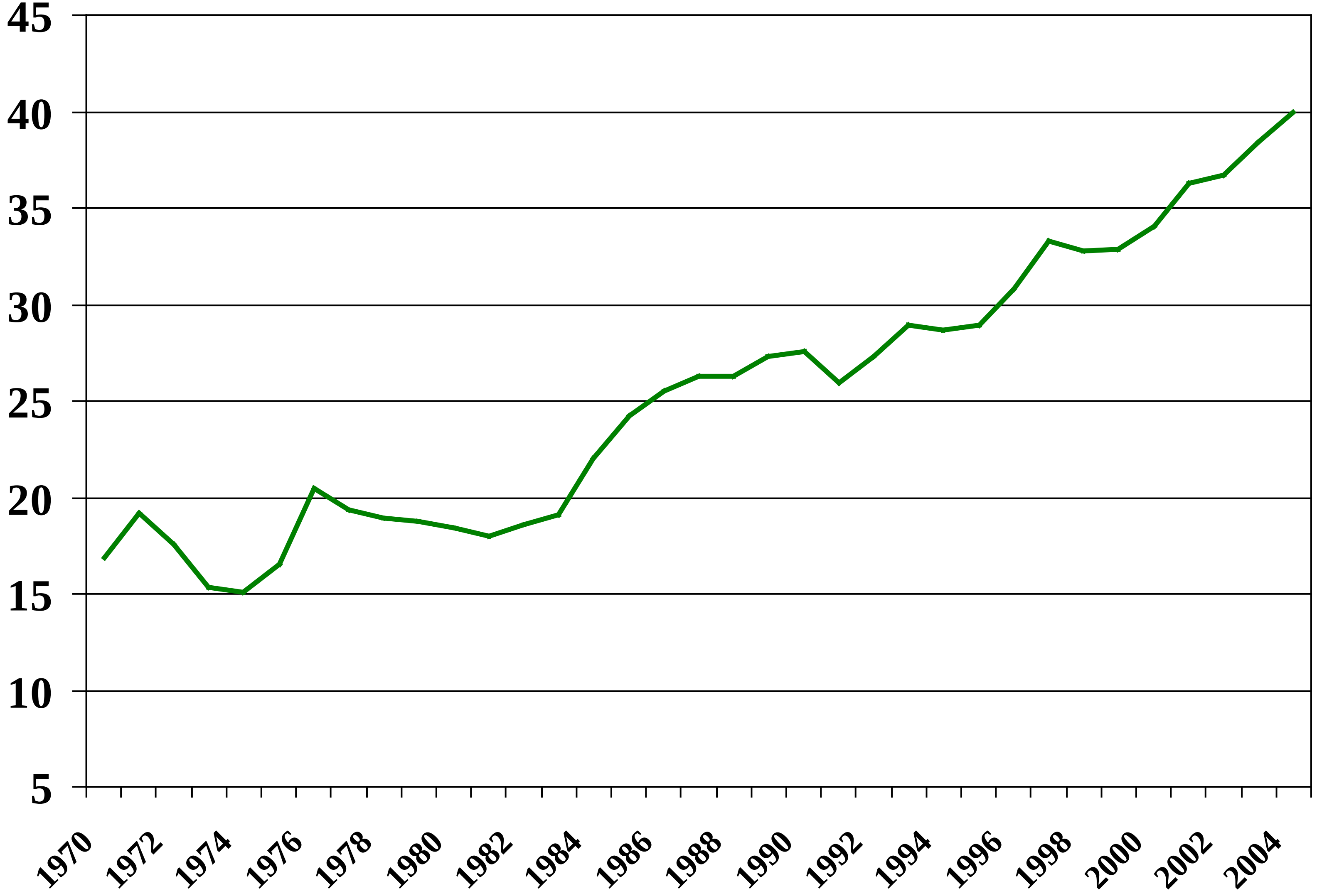
Data source:	1982 REDS			2005 IHDS
Loan type:	without interest	without collateral	without collateral or interest	without interest
	(1)	(2)	(3)	(4)
Sources:				
Bank	0.57	23.43	0.38	0.00
Caste	28.99	60.27	20.38	44.62
Friends	9.35	91.72	3.89	21.5
Employer	0.44	65.69	0.44	10.75
Moneylender	0.00	98.71	0.00	0.27

Figure 5: Change in Out-Marriage Percent in Rural India, 1950- 1999



Source: 1999 REDS

Figure 5A. Real Daily Agricultural Wages in India, 1970-2004 (Source: Bhalla and Das, 2006)



Testing our Explanation

- The simplest test of the hypothesis that the potential loss in network services restricts mobility in India would be to compare migration-rates in populations with and without caste-based insurance
    - This exercise is infeasible, given the pervasiveness of caste networks
- What we do is to look *within* the caste and theoretically identify which households benefit less (more) from caste-based insurance
    - We then proceed to test whether those households are more (less) likely to have migrant members

The Model

- The literature on mutual insurance is concerned with *ex post* risk-sharing, taking the size of the network and the income sharing rule as given
- To derive the connection between networks and permanent migration, it is necessary to derive *ex ante* participation and the sharing rule (which determines which households choose to stay)

Income

- The decision-making unit is the household, which consists of multiple earners
- Each household derives income from its local activities
- Income varies independently across households in the community and over time
- In addition, one or more members of the household receive a job opportunity in the city
  - The key decision is whether or not to send them to the city

Preferences

- We assume that the household has logarithmic preferences
- This allows us to express the expected utility from consumption, C, as an additively separable function of mean consumption, M, and normalized risk, $R \equiv \frac{V}{M^2}$, where V is the variance of consumption

$$EU(C) = \log(M) - \frac{1}{2}\frac{V}{M^2}.$$

Risk-Sharing

- Rural incomes vary over time and so risk-averse households benefit from a community-based insurance network to smooth their consumption
- Because our interest is in the *ex ante* decision to participate in the rural insurance network, we assume that complete risk-sharing can be maintained *ex post*
  - Consistent with high levels of risk-sharing documented in India and other developing countries (Townsend 1994, Grimard 1997, Ligon 1998, Fafchamps and Lund 2003, Mazzocco and Saini 2012, Angelucci, de Giorgi, and Rasul)

Risk-Sharing

- *Ex post* commitment is supported by social sanctions
- These sanctions are less effective when someone from the household has migrated to the city
    - With full risk-sharing, each household is either in the network or out of the network
    - We assume that households with migrants cannot commit to reciprocating at the level needed for full risk-sharing and so will be excluded from the network
- If the migrant's income cannot be observed by the rural community, his household has an incentive to over-report this income *ex ante* and under-report this income *ex post*
    - This information problem is another reason why households with migrants will be excluded from the network

Risk-Sharing

- Each household thus has two options:
  1. It can remain in the village and participate in the insurance network, benefiting from the accompanying reduction in the variance of its consumption
  2. It can send one or more of its members to the city and add to its income but forego the services of the rural network

The Participation Decision

- The household will choose to participate in the network and remain in the village if

$$\log(M_I) - \frac{1}{2}\frac{V_I}{M_I^2} \geq \log(M_A) - \frac{1}{2}\beta\frac{V_A}{M_A^2} + \epsilon \qquad (1)$$

  - M_A, V_A are the mean and variance of the household's income when all its members remain in the village
  - M_I, V_I are the corresponding mean and variance of consumption
  - $M_A(1+\tilde{\epsilon})$ is the household's mean income when one or more members move to the city, $\epsilon \equiv \log(1+\tilde{\epsilon})$
  - β reflects both the change (decline) in income-risk due to migration and the availability of alternative insurance

The Participation Decision

- With full risk-sharing and log preferences, each household's consumption is a fixed fraction of total income in each state of nature
  - Mean rural income, M_A, is the same for all households
  - ϵ, which is uncorrelated with M_A, is private information
  - We will thus have an equal sharing rule

The Participation Decision

- The equal sharing rule implies that

$$M_I = E\left(\frac{1}{N}\sum_i y_{is}\right) = \frac{1}{N}(NM_A) = M_A$$

$$V_I = V\left(\frac{1}{N}\sum_i y_{is}\right) = \frac{1}{N^2}(NV_A) = \frac{V_A}{N}$$

- Assume that migration increases the risk that the household faces, $R_I < \beta R_A$, even if $\beta < 1$
  - where $R_I \equiv \frac{V_I}{M_I^2}, R_A \equiv \frac{V_A}{M_A^2}$
- Participation will thus depend on the gain from insurance, $\beta R_A - R_I$, versus the income-gain from migration, ϵ, since $M_I = M_A$

Equilibrium Participation

- There is a strategic element to the participation decision because the gain from insurance depends on the number of participants
- To solve this fixed-point problem,
  - We first derive the threshold ϵ_I at which the participation condition holds with equality
  - Let the ϵ distribution be characterized by the function $F(\epsilon)$
  - Then set $F(\epsilon_I)$ to be equal to $\frac{N}{P}$
- $\frac{N}{P} = F(\Delta M + \Delta R)$
  - where $\Delta M \equiv \log(M_I) - \log(M_A)$ equals zero
  - $\Delta R \equiv \frac{1}{2}\beta R_A - \frac{1}{2}R_I$ is a function of N

Equilibrium Participation

- We make the following assumptions about the distribution of ϵ

A1. The left support is equal to zero
A2. The right support is unbounded
A3. The density, f, is decreasing in ϵ

- Given these distributional assumptions:

Lemma 1. *Equilibrium participation is characterized by a unique fixed point,* $N^* \in (0, P)$.

Participation and Income-Sharing with Inequality

- Divide the community into K income classes of equal size, P_k
- With log preferences and full risk-sharing, $C_{ks}/C_{Ks} = \lambda_k$

$$M_{Ik} = \left(\frac{\lambda_k}{\sum_k \lambda_k N_k}\right) \sum_k N_k M_{Ak} \qquad V_{Ik} = \left(\frac{\lambda_k}{\sum_k \lambda_k N_k}\right)^2 \sum_k N_k V_{Ak}$$

$$R_I = \frac{\sum_k N_k V_{Ak}}{\left(\sum_k N_k M_{Ak}\right)^2}$$

Participation and Income-Sharing with Inequality

- Fixed-point condition in each income class:

$$\frac{N_k}{P_k} = F(\Delta M_k + \Delta R_k)$$

  - $\Delta M_k \equiv \log(M_{Ik}) - \log(M_{Ak})$, $\Delta R_k \equiv \frac{1}{2}\beta R_{Ak} - \frac{1}{2}R_I$
  - If we knew λ_k, then we could solve for N_k

Participation and Income-Sharing with Inequality

- To derive λ_k, maximize social surplus W, subject to the fixed point conditions
- For $\beta < 1, W =$

$$\sum_k P_k \int_0^{\epsilon_{Ik}} \left\{ \left[\log(M_{Ik}) - \frac{1}{2} R_I \right] - \left[\log(M_{Ak}) - \frac{1}{2} \beta R_{Ak} + \epsilon \right] \right\} f(\epsilon) d\epsilon$$

- $W = \sum_k N_k \epsilon_{Ik} - P_k \int_0^{\epsilon_{Ik}} \epsilon f(\epsilon) d\epsilon$
  - Where $\epsilon_{Ik} = \Delta M_k + \Delta R_k$

Relative Wealth, Rural Risk, and Migration

- If participation in the network were fixed, the community could increase surplus (given diminishing marginal utility) by redistributing income
- But the sharing-rule must be attentive to increased exit by wealthier households, which makes it smaller and reduces its ability to smooth consumption

Proposition 1. Some redistribution is socially optimal, which implies that (relatively) wealthy households in the community should *ceteris paribus* be more likely to have migrant members

Relative Wealth, Rural Risk, and Migration

- A household that faces greater rural income-risk benefits more from the insurance network and is less likely to have migrant members
- Must account for redistribution favoring safe households

Proposition 2. Households that face greater rural income-risk are *ceteris paribus* less likely to have migrant members

Testing the Theory

- The theory generates three testable predictions:
  1. Income is redistributed in favor of poor households within the caste
  2. Relatively wealthy households, who benefit less from the network, should be more likely to have migrant members
  3. Households facing greater rural income-risk, who benefit more from the network, should be less likely to have migrant members
- Additional tests validate the key assumption that permanent male migration is associated with a loss in network services

Testing the Theory

- Urban caste networks can also explain low migration and large wage gaps
- Alternative explanations are available for redistribution and increased exit by relatively wealthy households
  - No alternative can deliver all three predictions (especially the third)

Evidence on Redistribution within Castes

- 2005-2011 Indian ICRISAT panel survey
  - household income over 7 years
  - consistent consumption data for 4 years
- 2006 REDS Census
  - 119,000 households in 242 villages in 17 major states
  - permanent migration information is collected but income is only available in the year prior to the survey
  - impute average income and average consumption using ICRISAT data

Table 5: Income and Consumption within the Caste

Data Source:	ICRISAT		
	relative income	relative consumption	consumption-income ratio
	(1)	(2)	(3)
Income class:			
1	0.119	0.460	3.871
2	0.281	0.625	2.224
3	0.373	0.626	1.680
4	0.510	0.673	1.319
5	1.000	1.000	1.000

Disadvantages of the REDS listing data:

A. Income is only for the prior year; want more permanent measure as in the ICRISAT data.

B. No consumption data.

Estimate relationships between average income net of transfers and average total consumption (7 years) and household and village characteristics (common) in ICRISAT data (2005-2011):

Landholdings, irrigation, number of adult earners, soil color and depth, mean and variance of village rainfall. (R^2=.3.).

Use the estimated coefficients to impute average income and consumption for REDS hh's *in the same states* (ICRISAT = 2005-2011) using the same variables.

To test proposition 2 we also need a measure of income variability.

> Impute from the association between the log income variance and hh characteristics and hh characteristics*the variance of rainfall in the ICRISAT data (R^2=.3).

Standard errors are obtained from bootstrapping, with clustering at two levels: caste and village

Table 5: Income and Consumption within the Caste

Data Source:	REDS 2006			
	relative income	relative consumption	consumption-income ratio	migration
	(4)	(5)	(6)	(7)
Income class:				
1	0.316	0.843	2.665	0.032
2	0.416	0.854	2.052	0.034
3	0.513	0.871	1.697	0.051
4	0.627	0.887	1.413	0.046
5	1.000	1.000	1.000	0.051

Reduced-Form Estimates

- Proposition 1 indicates that *relatively* wealthy households are more likely to have migrant members

$$M_i = \pi_0 + \pi_1 y_i + \pi_2 \overline{y}_i + \epsilon_i$$

  - $\pi_1 > 0, \pi_2 < 0$
  - cannot interpret π_1 once we allow household income to have a direct effect on migration

- Proposition 2 indicates that households facing greater rural income-risk should be less likely to have migrant members

Table 6: Relative Wealth, Rural Income-Risk, and Migration

Dependent variable:	migration			
	(1)	(2)	(3)	(4)
Household Income	0.0059	0.0051	0.0026	0.0020
	(0.0024)	(0.0024)	(0.0045)	(0.0032)
Caste Income	-0.016	-0.018	-0.022	-0.028
	(0.0043)	(0.0055)	(0.010)	(0.0090)
Income Risk	–	-0.00038	-0.00037	-0.00056
		(0.00015)	(0.00013)	(0.00015)
Village Income			0.007	–
			(0.013)	–
Village/ Caste Income				0.0076
				(0.012)
Village Fixed Effects	No	No	No	No
Infrastructure Variables	No	No	No	Yes
Joint sig. of infrastructure variables:				
χ^2	–	–	–	16.59
	–	–	–	[0.00090]
Number of observations	19,362	19,362	19,362	19,362

Source: 2006 REDS Census

Structural Estimates

- The structural estimates are used to
  - (i) provide independent support for the redistribution within castes predicted by the theory (external validation)
  - (ii) carry out counter-factual simulations
- There are two exogenous variables in the model: $M_{Ak}, R_{Ak} \equiv V_{Ak}/M_{Ak}^2$
  - Although there is a single community in the theoretical analysis, there are 100 castes in the 2006 REDS census
  - Within each caste, j, we thus construct M_{Akj}, R_{Akj}

Structural Estimates

- Suppose, to begin with, that the β parameter and the F function are known
  - For a given λ_{kj} vector, we can then solve for N_{kj}/P_j from the fixed-point condition
  - Total surplus can then be computed for each caste, j
  - If the model is correctly specified, predicted migration at the surplus-maximizing λ_{kj} should match actual migration

Structural Estimates

- Now suppose that β is unknown
- For an arbitrary β, we can go through the same steps
  - But predicted migration will not match actual migration
- As β increases, migration will decline in each income-class in each caste
  - Thus there exists a unique β for which (overall) predicted and actual migration match

Structural Estimates

- Finally describe how the $F(\epsilon)$ function is derived
- Let ϵ be characterized by the exponential distribution
  - $F(\epsilon) = 1 - e^{\nu\epsilon}$, $E(\epsilon) = 1/\nu$
  - Satisfies A1-A3

Structural Estimates

- ν is estimated in two steps
  1. Use REDS and NSS data to compute the average income-gain from migration for households with migrants, $\tilde{\epsilon}$, and its utility-equivalent $\hat{\epsilon} = \log(1 + \tilde{\epsilon})$
  2. Use the percent of households with migrants, x, together with the properties of the exponential distribution, to derive ν

$$\nu = \frac{-\log(x/200)}{\hat{\epsilon}}$$

- As a robustness test, estimate ν within absolute income classes and within castes

Table 7: Structural Estimates

	measured		single ν	
	relative consumption	migration	relative consumption	migration
	(1)	(2)	(3)	(4)
Relative Income Class:				
1	0.843	0.032	0.801	0.000
			(0.071)	(0.00020)
2	0.854	0.034	0.817	0.014
			(0.070)	(0.0073)
3	0.871	0.051	0.834	0.039
			(0.063)	(0.0083)
4	0.887	0.046	0.868	0.060
			(0.044)	(0.0089)
5	1.000	0.051	1.000	0.100
				(0.014)
overall		0.043		0.043
β			1.410	
			(0.91)	
α			–	
γ			–	

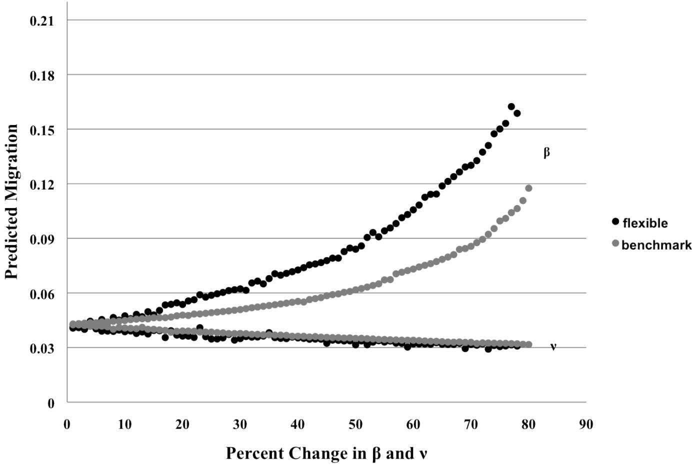
Source: 2006 REDS Census

Table 7: Structural Estimates

	estimating ν by caste			
	relative consumption	migration	relative consumption	migration
	(7)	(8)	(9)	(10)
Relative Income Class:				
1	0.751	0.000	0.730	0.032
	(0.097)	(0.000081)	(0.083)	(0.0095)
2	0.767	0.011	0.744	0.032
	(0.092)	(0.010)	(0.064)	(0.052)
3	0.792	0.029	0.765	0.046
	(0.070)	(0.025)	(0.055)	(0.027)
4	0.842	0.055	0.825	0.044
	(0.044)	(0.033)	(0.037)	(0.013)
5	1.000	0.119	1.000	0.051
		(0.062)		(0.0074)
overall		0.043		0.041
β	0.845		0.991	
	(0.92)		(0.18)	
α	–		0.012	
			(0.050)	
γ	–		4.45	
			(0.91)	

Source: 2006 REDS Census

Figure 6: Counter-Factual Simulation



Testing the Mechanism

- Key assumption is that permanent male migration is associated with a loss in network services
- Test this assumption by examining how a household's relative wealth affects: out-migration, network participation, and out-marriage
  - Use household sample from the 1982 and 1999 REDS rounds

$$X_{it} = \pi_1 y_{it} + \pi_2 \overline{y}_{it} + f_i + \epsilon_{it}$$

$$\Delta X_{it} = \pi_1 \Delta y_{it} + \pi_2 \Delta \overline{y}_{it} + \Delta \epsilon_{it}$$

  - Use initial conditions at the onset of the Green Revolution (from the 1971 REDS) as instruments
  - Because these are fixed characteristics, we no longer need to impute incomes

Table 8: FE-IV Participation, Out-Marriage, and Network Participation Estimates

Dependent variable:	migration	out-marriage	participation
	(1)	(2)	(3)
Household income	0.262	0.166	-0.520
	(0.172)	(0.074)	(0.680)
Caste income	-0.110	-0.111	0.327
	(0.045)	(0.066)	(0.139)
Time trend	0.059	0.026	0.014
	(0.022)	(0.018)	(0.127)
Kleibergen-Paap F-statistic	10.52	8.05	2.91
Hansen J-statistic	2.62	6.74	4.17
	[0.62]	[0.15]	[0.38]
Number of observations	1,049	998	2,335

Source: REDS Panel, 1982 and 1999

Conclusion

- Why does India have migration rates that are so much lower than comparable developing economies?
  - Formal insurance is particularly weak in India [no evidence]
  - Informal insurance works particularly well there [high levels of risk-sharing have been documented throughout the developing world]
- There is, however, more to consumption-smoothing than risk-sharing
  - The size, scope, and connectedness of caste networks may be exceptional
  - Recent genetic evidence indicates that strict endogamy emerged 1900 years ago

Conclusion

- Can policies be implemented to increase mobility in this economy?
- We perform two counter-factual experiments with the estimated model
    1. Provision of credit to wealthy households
    2. Government safety net for poor households

Figure 7: Reducing Risk in Higher Income-classes

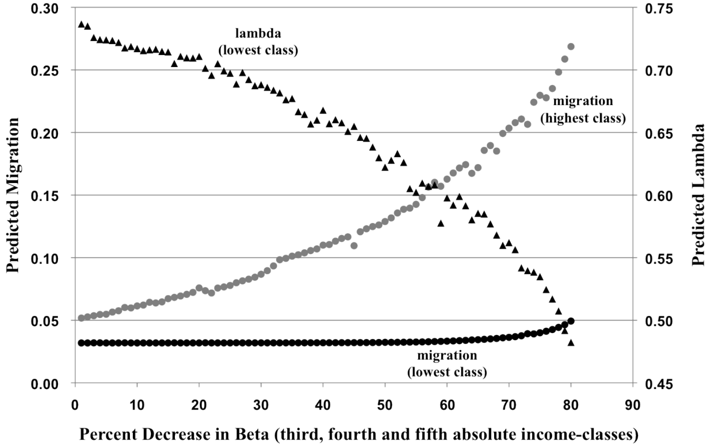


Figure 8: Reducing Risk in Lower Income-classes